International Journal of Engineering & Technology, 7 (1.1) (2018) 426-430

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Quality based drip drag match data collection in wireless sensor network

V. Jayaraj [1] *, S. Alonshia [1]

[1] *Department of Computer Science and Engineering, Bharathidasan University, Tiruchirappalli, Tamilnadu, India*
**Corresponding author E-mail: cse@bdu.ac.in*

Abstract

Although data collection has received much attention by effectively minimizing delay, computational complexity and increasing the total data transmitted, the transience of sensor nodes for multiple data collection of sensed node in wireless sensor network (WSN) renders quality of service a great challenge. To circumvent transience of sensor nodes for multiple data collection, Quality based Drip-Drag-Match Data Collection (QDDM-DC) scheme have been proposed. In Drip-Drag-Match data collection scheme, initially dripping of data is done on the sink by applying Equidistant-based Optimum Communication Path from the sensor nodes which reduces the data loss. Next the drag operation pulls out the required sensed data using Neighbourhood-based model from multiple locations to reduce the delay for storage. Finally, the matching operation, compares the sensed data received by the dragging operation to that of the corresponding sender sensor node (drip stage) and stores the sensed data accurately which in turn improves the throughput and quality of data collection. Simulation is carried for the QDDM-DC scheme with multiple scenarios (size of data, number of sinks, storage capacity) in WSN with both random and deterministic models. Simulation results show that QDDM-DC provides better performance than other data collection schemes, especially with high throughput, ensuring minimum delay and data loss for effective multiple data collection of sensed data in WSN.

Keywords: *Wireless Sensor Network; Data Collection; Equidistant; Neighbourhood.*

1. Introduction

With the quality in wireless sensor network communication, collection of large data is bit difficult as sensor nodes have only limited power and minimal memory resources. Although there exist many different types of data collection methods designed for wireless sensor networks, quality of service has to be focused due to the high data intensive nature. Distributed Eigen-system Realization Algorithm (DERA) [1] was designed with the objective of improving the structural health data in wireless sensor network. However, different from health monitoring using WSN, snapshot and continuous data collection using Cell-based Path Scheduling [2] algorithm was designed aiming at improving the network capacity using order-optimal capacity. Recently substantial efforts have been spent for data collection in WSN. With the objective of reducing data collection latency, Combine Skip Substitute (CSS) [3] scheme was presented. Another approach to data collection in WSN utilized a multithreaded design [4] of coding-based data dissemination protocol to reduce the dissemination delay. In WSN network, false aggregation of data occurs frequently disturbing the network in a drastic manner. To solve this issue, Data Aggregation and Authentication (DAA) [5] protocol was designed to improve the early detection of false data. A vector quantization model based on hyperplane [6] was designed to reduce the distortion rate by applying optimal strategies. Improvement in data collection is also possible by extending the network lifetime. To improve the security and accuracy of data being collected, energy-efficient and high-accuracy model based on private sensor readings was presented in [9]. Extensive research has been concentrated on the area to reduce the energy consumption during data collection in WSN. In [10], comparison of energy efficient data collection techniques was presented. In this paper, we proposed a Quality based Drip-Drag-Match Data Collection (QDDM-DC) scheme with the objective of improving the throughput by minimizing the storage delay and data loss is presented. In the QDDM-DC scheme, we first propose an Equidistant-based Optimum Communication Path by dripping data on the sink from the sensor nodes is presented. Then we propose a Neighbourhood-based Drag model that efficiently pulls out the sensed data from multiple locations using Neighborhood drag algorithm. Finally, drip data sequence number and acknowledgement number is used to perform an efficient matching operation that in turn improves the throughput and quality of data collection. The QDDM-DC scheme achieves comparably better results of the counterpart using much less storage delay as well as data loss. The efficacy of the data collection scheme is demonstrated through both simulation and experiment.

2. Quality based drip-drag-match data collection

One of the main criteria in designing a WSN application is improving the Quality of Services and reducing data loss through efficient data collections. Quality of services (QoS) refers to network ability to perform large data communication. Wireless Sensor Network with sensor nodes sense target object and collect data from different locations. The continuous collection of data from wireless sensor network should possess minimum delay, high throughput and minimal data loss for efficient and quality data communication. With the quality in wireless sensor network communication, collection of large data is bit difficult as sensor nodes have only limited power and minimal memory resources. With this objective, Quality based Drip-Drag-Match Data Collection (QDDM-DC) scheme is investigated in this work by starting with a network model, followed by which elaborate descriptions of the proposed scheme is presented.

2.1. Network model for data collection

Let us consider a wireless sensor network '$G = (V, E)$' composed of a set 'V' of 'n' sensor nodes denoted as '$V = S_1, S_2, \ldots, S_n$' and a set '$E$' of edges denoted as '$E = e_1, e_2, \ldots, e_m$' in a rectangular area '$m * n$'with the nodes able to communicate up to range 'r'. Let '$Neigh(p)$' represents the set of neighbors of node 'p', where '$Neigh(p) = (p, q) \leq r$ ' and a sink 's' that performs the task of data collection. Now the problem is to design an efficient data collection scheme to not only reduce data loss but also storage and improve the throughput during data collection.

2.2. Equidistant-based optimum communication path

In Drip-Drag-Match data collection scheme, initially dripping of data is done on the sink from the sensor nodes through the communication path using equidistant rectangular model. The communication path through which the data is dripped is designed in such a way so as the optimum communication path must be equidistant from any furthest sensor node in the network. The maximum distance a data from a sensor node, on the circumference of network, before arriving at a sensor node within communication range 'r' of the sink node 'SN', must be the same as the maximum distance from a sensor node at the Centre of the network to a sensor node within communication range of the sink. If one sink is located in the Centre of the network, then for a rectangular model, the number of hops the data from sensor node at the farthest end to reach the sink is given as below.

$$H_{r\_number} = \left(\frac{m}{2*r}\right) + \left(\frac{n}{2*r}\right) \quad (1)$$

From (1), the count of hop data for a rectangular field 'H_{r_number}' is obtained based on the transmission range 'r' and width 'm' and height of the network 'n' respectively. Next, to ensure equidistance between sensor nodes at the Centre of the network and sensor nodes at the circumference of the network, the hop frequency is equal to '$H/2$'. Figure 1 shows the Equidistant-based Optimum Communication Path.

As shown in the figure, the sensor nodes in the network using equidistant rectangular model act as temporary nodes for any data destined for the sink node. With the movement of sink node along the communication path, these sensor nodes send the data to the sink node. The path through which the hops drips the data 'H_r' is measured as given below.

$$H_r = \left(\frac{m}{4*r}\right) + \left(\frac{n}{4*r}\right) \quad (2)$$

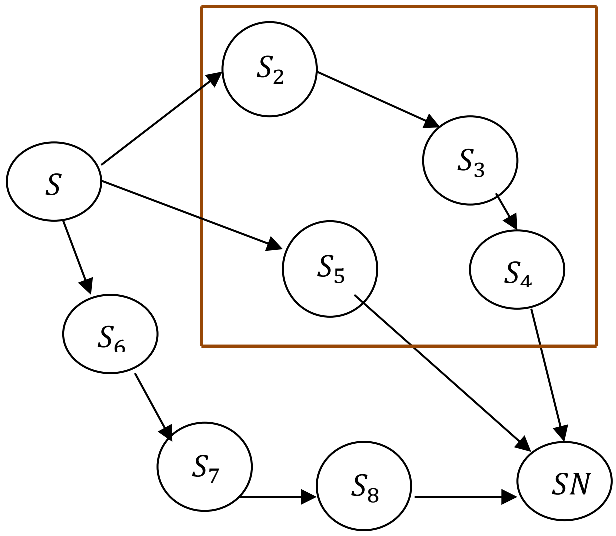

Fig. 1: Structure of Equidistant-Based Optimum Communication Path.

Where 'm, n' symbolizes the width and height of the network with a transmission range of 'r'. From (2), the optimal communication path is formulated as given below.

$$Optimal_p = r * \frac{H_r}{2} \quad (3)$$

From (3), the optimal path '$Optimal_p$' is obtained through which the data loss is reduced because of minimum delay between the time an event occurs and the time the sink receives the data. Finally, for each dripped data, a drip data sequence number is assigned. Figure 2 shows the equidistant rectangular algorithm.

Input: Sensor Nodes '$S = S_1, S_2, \ldots, S_n$', Sink Node '$SN$', communication range '$r$', width '$m$', height '$n$' drip data sequence number '$D = D_1, D_2, \ldots, D_n$'
Output: Reduces data loss
1: Begin 2: For each Sensor Nodes 'S' 3: Measure the hop number using (1) 4: Measure hops for transmission range 'r' using (2) 5: Measure optimal communication path using (3) 6: Assign 'D' drip data sequence number to the sender sensor node 7: End for 8: End

Fig. 2: Equidistant Rectangular Algorithm.

As shown in the figure, for each sensor node, dripping of data is performed from the sensor node through the sink using the equidistant rectangular algorithm. The algorithm initially, obtains the hop number for rectangular region. Followed by it, the hops for within the communication range are measured and finally, optimal communication path is obtained.

2.3. Neighbourhood-based drag model

Once the dripping of data is performed on the sink node, the drag operation pulls out the required sensed data from multiple locations to reduce the delay for storage by dividing the network optimally. The size of the rectangular model is mathematically calculated as given below.

$$a = \frac{A}{\sqrt{N_{SN}}}\ ; b = \frac{B}{\sqrt{N_{SN}}} \quad (4)$$

Whenever an event occurs, the drag operations pulls out the sensed data by aggregating the data by electing immediate neighbour nodes to forward the data to the sink. Figure 3 shows the structure of the Neighbourhood-based Drag model.

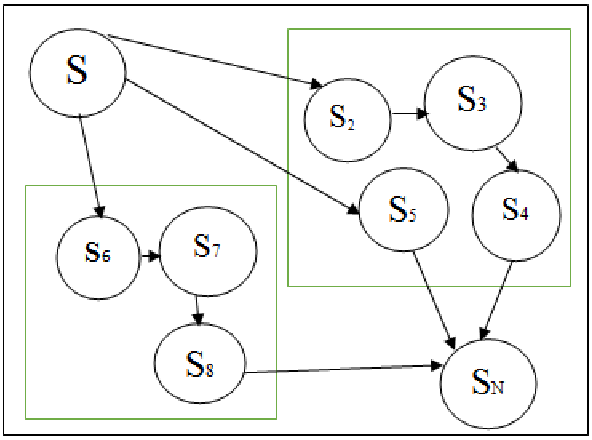

Fig. 3: Structure of the Neighbourhood-Based Drag Model.

As shown in the figure, the Neighbourhood-based Drag model pulls out the required sensed data until the data is received by an immediate neighbour node that is in direct communication range of the path of the sensor node and is formulated as given below.

$$\mathrm{Drag_i} = \sum_{i=1}^{n} \mathrm{MLoc_i DP_i} \quad (5)$$

From (5), '$\mathrm{Drag_i}$' symbolizes the drag operation performed using the sensed data '$\mathrm{DP_i}$' present in multiple locations '$\mathrm{MLoc_i}$' in the network. The data is stored and when the sensor node passes the immediate neighbour node, all stored data are transmitted to the sink node. The time taken for a sink node 'SN' to accelerate the data is given as below.

$$\mathrm{Acc_S} = \frac{\mathrm{Vel_{FS}} - \mathrm{Vel_{IS}}}{\mathrm{Total\_Vel(S_i)}} \quad (6)$$

From (6), given the initial '$\mathrm{Vel_{IS}}$' and final velocity '$\mathrm{Vel_{FS}}$' of the sensor node with respect to the total velocity of all the sensor nodes '$\mathrm{Total\_Vel(S_i)}$' in the network, the acceleration of the sink node '$\mathrm{Acc_S}$' is obtained. With this obtained sink node acceleration, the sensed data from multiple locations are obtained. Finally, a drag acknowledgement number is assigned to the sensed data during dragging operation. Figure 4 shows the Neighbourhood Drag algorithm.

Input: Sensor Nodes '$\mathrm{S} = \mathrm{S_1, S_2, \dots, S_n}$', Sink Node 'SN', Data Packets '$\mathrm{DP} = \mathrm{DP_1, DP_2, \dots, DP_n}$', Location '$\mathrm{Loc} = \mathrm{Loc_1, Loc_2, \dots, Loc_n}$', Drag Acknowledgement Number '$\mathrm{DAck} = \mathrm{DAck_1, DAck_2, \dots, DAck_n}$'
Output: Minimizes storage time delay
1: Begin 2: For each location 'Loc' 3: Measure size of rectangular model for efficient sensing using (4) 4: Perform drag operation using (5) 5: Measure time taken to accelerate data using (6) 6: Assign data acknowledgement number 'DAck' to the sensed data 6: End for 7: End

Fig. 4: Neighbourhood Drag Algorithm.

As shown in the figure, for each location Neighbourhood Drag algorithm pulls out the required sensed data that has a minimum of four immediate neighbours from multiple locations. This ensures that the sensor nodes on the perimeter of the network necessarily sense the data from multiple locations based on neighbourhood drag, reducing delay for storage.

2.4. Drip-drag-match data collection

One of the most important phase, that ensures the reliability in the sensor side and enhances throughput is the matching operation. In this matching operation, the sensed data received by the dragging operation to that of the corresponding sender sensor node (drip stage) is compared. The successful comparison refers to that the scheme has stored the sensed data accurately. In this phase, the drip data sequence number to the sender sensor node 'D' is compared with that of the data acknowledgement number 'DAck' to the sensed data. Upon successful matching, the sensed data is stored accurately in the sink. Therefore, we propose a novel Drip-Drag-Match algorithm in this section. Figure 5 shows the Drip-Drag-Match algorithm.

Input: Sensor Nodes '$\mathrm{S} = \mathrm{S_1, S_2, \dots, S_n}$', Sink Node 'SN', drip data sequence number '$\mathrm{D} = \mathrm{D_1, D_2, \dots, D_n}$', Data Packets '$\mathrm{DP} = \mathrm{DP_1, DP_2, \dots, DP_n}$', Location '$\mathrm{Loc} = \mathrm{Loc_1, Loc_2, \dots, Loc_n}$', Drag Acknowledgement Number '$\mathrm{DAck} = \mathrm{DAck_1, DAck_2, \dots, DAck_n}$'
Output: Optimizes throughput
1: Begin 2: For each Data Packets 'DP' and Sensor Nodes 'S' 3: Obtain 'D' drip data sequence number 4: Obtain data acknowledgement number 'DAck' 5: If 'D' = 'DAck' 6: Store sensed data in the sink node 7: End if 8: If 'D' ≠ 'DAck' 9: Do not store sensed data in the sink node 10: End if 11: End for 12: End

Fig. 5: Drip-Drag-Match algorithm.

As shown in the figure, for each sensed data 'DP' and sender sensor node 'S' the Drip-Drag-Match algorithm cross verifies the drip data sequence number with the data acknowledgement number. If both the numbers are matched, then the sensed data is stored in the sink node. Otherwise, the sensed data is retransmitted and the process is continued with the next sensed data. Theoretical analysis shows that the Drip-Drag-Match algorithm can improve data collection significantly.

3. Experimental settings

In this section, we present the numerical data obtained as a result of applying Quality based Drip-Drag-Match Data Collection (QDDM-DC) scheme.

Table 1: Simulation Setup

Parameters	Values
Network area '$m * n$'	1200m * 1200m
Transmission range 'r'	300m
Node density	8, 16, 24, 32, 40, 48, 56
Data packet density	5, 10, 15, 20, 25, 30, 35
Data packet size	312 bytes
Data rate	250kbps
Simulation time	1200s
Number of runs	7

In Table 1 the set of input parameter is tabulated and performance evaluation for QDDM-DC scheme is performed via simulation. Our example WSN consists of 56 sensor nodes deployed in a rectangular area of 1200m * 1200m placed in a random manner in the wireless sensor network. The nodes are distributed in an area using Random Way point model for simulation, whereas the link layer provides the link between two nodes and the design of link is multi direction. The radio ranges are dynamically adjusted within a network area of 1200m*1200m to maintain network connectivity. The sink node collects the data packets of range 5 – 35 and forwards the data to the sink node with each data packet size differing from 50bytes to 312bytes. The simulation time varies from 50 simulation seconds to 1200 simulation seconds. Omni directional antenna is used for simulation and at any instant of time only single process is performed (i.e., either packet transmission or packet reception).

Performance evaluation is conducted for the proposed QDDM-DC scheme with metrics such as data packet density, node density, throughput, storage delay and data loss during data collection at the sink node in WSN. The metrics were compared with the state-of-the-art schemes namely, methods namely, Distributed Eigen-system Realization Algorithm (DERA) [1] and Cell-based Path Scheduling (CPS) [2] for data collection in WSN.

4. Discussion

In this section, we present the performance evaluation of our proposed scheme QDDM-DC through NS2 simulation and compared with two existing schemes, Distributed Eigen-system Realization Algorithm (DERA) [1] and Cell-based Path Scheduling (CPS) [2] for efficient data collection in Wireless Sensor Networks.

The nodes in SM-ADS framework are positioned in uniform topology. To evaluate the efficiency of QDDM-DC scheme, the following metrics like data size, number of sink, node density, throughput, delay for storage, data loss in Wireless Sensor Network is measured.

Scenario 1: Throughput

Throughput measures the percentage of the data packets sent by the source nodes and received by the sink.

$$T = \frac{P_r}{P_s} * 100 \tag{7}$$

From (7), 'T' symbolizes the throughput which is measured on the basis of the packets received 'P_r' by the sink node to the packets sent 'P_s' by the source sensor node in network. Higher the throughput, more efficient the quality of data collection is said to be. The results of seven simulation runs conducted to measure the throughput rate are listed in table 2. As listed in table 3, the QDDM-DC, DERA and CPS measures the rate of throughput in terms of percentage (%). The rate of throughput obtained using our scheme QDDM-DC offer comparable values than the state-of-the-art schemes.

Scenario 2: Delay for storage

Storage delay is the delay time taken to store the data packets obtained from multiple locations at the sink node.

$$SD = DP_i * \text{Time}\ (MLoc_i DP_i) \tag{8}$$

Table 2: Tabulation for Throughput

Data packet density	Throughput (%)		
	QDDM-DC	DERA	CPS
5	90.23	81.52	79.37
10	92.14	84.89	82.15
15	93.18	86.24	83.34
20	90.27	85.43	82.53
25	91.49	86.56	81.66
30	93.15	88.34	83.44
35	95.89	90.47	87.57

From (8), the storage delay 'SD' is obtained with respect to the data packets sensed 'DP_i' from multiple locations '$MLoc_i$' by the sink node. In table 3 we further compare the storage delay consumed by different data packet density for data collection at the sink in WSN. The experiments were conducted using thirty five sensor nodes and the storage delay consumed is measured in terms of milliseconds (ms).

Table 3: Tabulation for Storage Delay

Data packet density	Storage delay (ms)		
	QDDM-DC	DERA	CPS
5	6.72	7.25	7.86
10	8.41	10.12	11.32
15	13.21	15.43	16.73
20	15.89	17.23	18.53
25	19.32	21.31	23.61
30	24.12	27.23	29.43
35	29.35	32.43	35.83

Scenario 3: Data loss

Data loss refers to the amount of loss in data while data is collected from the sensor node on the sink and is measured in terms of megabytes (MB).

$$DL = [S_i * P_{size}] - \text{Data obtained} \tag{9}$$

From (9), the data loss is obtained is measuring the difference between the sensor nodes that sends the packet and the data obtained at the sink node. Lower the difference or data loss, the efficiency of data collected at the sink node is improved. Table 4 shows the data loss on the sink node with respect to 35 sensor nodes with a moving speed of 25 m/s. To better perceive the efficacy of the proposed QDDM-DC scheme, substantial experimental results are presented in table 4 and compared against the existing schemes DERA [1] and CPS [2] respectively.

Table 4: Tabulation for Data Loss

Node density	Data loss (MB)		
	QDDM-DC	DERA	CPS
8	42	57	71
16	53	68	74
l24	74	89	98
32	89	104	114
40	97	112	122
48	101	116	126
56	105	120	130

5. Conclusion

In this paper, it is proposed that a Quality based Drip-Drag-Match Data Collection (QDDM-DC) scheme and analyze the performance of quality of service and throughput using Equidistant-based Optimum Communication and Neighbourhood-based model to perform efficient data collection in WSN. The efficient data collection scheme is formulated as an optimal path model where the communication path is identified based on the equidistant rectangular model.

By measuring the optimal communication path, we propose a Neighbourhood Drag algorithm based on the time taken to accelerate the data to obtain the required sensed data from multiple locations. Finally, a math operation ensures throughput and quality of data collection at the sink node. Extensive simulation is carried out to evaluate the proposed scheme and data delivery method. The results validate the effectiveness of the proposed scheme and show that it significantly outperforms two traditional data collection methods.

References

[1] Liu X, Cao J, Song WZ, Guo P & He Z, "Distributed Sensing for High-Quality Structural Health Monitoring Using WSNs", *IEEE Transactions on Parallel and Distributed Systems*, Vol.26, No.3, (2015), pp.738-747. https://doi.org/10.1109/TPDS.2014.2312911.

[2] Ji S, Beyah R & Cai Z, "Snapshot and Continuous Data Collection in Probabilistic Wireless Sensor Networks", *IEEE Transactions on Mobile Computing*, Vol.13, No.3, (2014), pp.626-637. https://doi.org/10.1109/TMC.2013.30.

[3] He L, Pan J & Xu J, "A Progressive Approach to Reducing Data Collection Latency in Wireless Sensor Networks with Mobile Elements", *IEEE Transactions on Mobile Computing*, Vol.12, No.7, (2013), pp.1308-1320. https://doi.org/10.1109/TMC.2012.105.

[4] Gao Y, Bu J, Dong W, Chen C, Rao L & Liu X, "Exploiting Concurrency for Efficient Dissemination in Wireless Sensor Networks", *IEEE Transactions on Parallel and Distributed Systems*, Vol.24, No.4, (2013), pp.691-670. https://doi.org/10.1109/TPDS.2012.195.

[5] Ozdemir S & Çam H, "Integration of False Data Detection With Data Aggregation and Confidential Transmission in Wireless Sensor Networks", *IEEE/ACM Transactions on Networking*, Vol.18, No.3, (2010), pp.736-749. https://doi.org/10.1109/TNET.2009.2032910.

[6] Fang J & Li H, "Hyperplane-Based Vector Quantization for Distributed Estimation in Wireless Sensor Networks", *IEEE Transactions on Information Theory*, Vol.55, No.12, (2009), pp.5682-5699. https://doi.org/10.1109/TIT.2009.2032856.

[7] Zhu YH, Wu WD, Pan J & Tang YP, "An energy-efficient data gathering algorithm to prolong lifetime of wireless sensor

networks", *Computer Communications*, Vol.33, No.5, (2010), pp.639-647. https://doi.org/10.1016/j.comcom.2009.11.008.

[8] Jiang H, Jin S & Wang C, "Prediction or Not? An Energy-Efficient Framework for Clustering-based Data Collection in Wireless Sensor Networks", *IEEE Transactions on Parallel and Distributed Systems*, Vol.22, No.6, (2011), pp.1064-1071. https://doi.org/10.1109/TPDS.2010.174.

[9] Li H, Lin K & Li K, "Energy-efficient and high-accuracy secure data aggregation in wireless sensor networks", *Elsevier, Computer Communications*, Vol.34, No.4, (2011), pp.591–597. https://doi.org/10.1016/j.comcom.2010.02.026.

[10] Rohankar R, Katti CP & Kumar S, "Comparison of Energy Efficient Data Collection Techniques in Wireless Sensor Network", *Elsevier, Procedia Computer Science*, Vol.57, (2015), pp.146–151. https://doi.org/10.1016/j.procs.2015.07.399.

[11] Rohankara R, "Agent Based Predictive Data Collection in Opportunistic Wireless Sensor Network", *Elsevier, Procedia Computer Science*, Vol.57, (2015), pp.33–40. https://doi.org/10.1016/j.procs.2015.07.361.

[12] Liu Y, Zeng QA & Wang YH, "Energy-Efficient Data Fusion Technique and Applications in Wireless Sensor Networks", *Hindawi Publishing Corporation, Journal of Sensors*, (2015), pp.1-3.

[13] Liu A, Cai LX, Luan TH & Ranabahu A, "QoS-Aware Data Collection in Wireless Sensor Networks", *Hindawi Publishing Corporation, International Journal of Distributed Sensor Networks*, (2015), pp.1-4.

[14] Alhmiedat T, "An Adaptive Energy-Efficient Data Collection System for Zig Bee Wireless Sensor Networks", *Hindawi Publishing Corporation, International Journal of Distributed Sensor Networks*, (2015), pp.1-13.

[15] Palanisamy T & Krishnasamy KN, "Bayes Node Energy Polynomial Distribution to Improve Routing in Wireless Sensor Network", *Plos One*, Vol.10, No.10, (2015), pp.1-15. https://doi.org/10.1371/journal.pone.0138932.

[16] Di Francesco M, Shah K, Kumar M & Anastasi G, "An adaptive strategy for energy-efficient data collection in sparse wireless sensor networks", *European Conference on Wireless Sensor Networks*, (2010), pp.322-337. https://doi.org/10.1007/978-3-642-11917-0_21.

[17] Gao S, Zhang H & Das SK, "Efficient Data Collection in Wireless Sensor Networks With Path-constrained Mobile Sinks", *IEEE Transactions on Mobile Computing*, Vol.10, No.4, (2011), pp.592-608. https://doi.org/10.1109/TMC.2010.193.

[18] Liu CX, Liu Y & Zhang ZJ, "Improved reliable trust-based and energy-efficient data aggregation for wireless sensor networks", *International Journal of Distributed Sensor Networks*, Vol.9, No.5, (2013). https://doi.org/10.1155/2013/652495.

[19] Almi'ani K, Almi'ani M, Al-ghonmein A & Al-Moghrabi K, "Data Collection Scheme for Wireless Sensor Network With Mobile Collector", *International Journal of Wireless & Mobile Networks*, Vol.6, No.4, (2014), pp.19-25. https://doi.org/10.5121/ijwmn.2014.6402.