Aggregation Using the Concept of Dynamic-Sized Data Packet for Effective Energy Saving in Wireless Sensor Network

Smitha N. Pai[1(✉)], H. S. Mruthyunjaya[2], Aparna Nayak[1], and A. Smitha[1]

[1] Department of Information and Communication Technology, M.I.T., Manipal University, Manipal, India
{smitha.pai,aparna.nayak,smitha.a}@manipal.edu

[2] Department of Electronics and Communication Technology, M.I.T., Manipal University, Manipal, India
mruthyu.hs@manipal.edu

Abstract. Data aggregation process can extract relevant information from raw data obtained from various sources using certain mathematical functions. Aggregation reduces the transmission of redundant data. A protocol named DP_AODV is implemented in this paper. Aggregator nodes (cluster head) are identified using the positional information. Routes are established between these aggregator nodes using efficient routing techniques. Data is aggregated along the path to the destination conserving additional energy. The aggregation process involves averaging the data if it is within the threshold range, else, only the data part along with the positional information is appended to the payload. Size of the Data packet varies dynamically based on the number of nodes having co-related data at that particular instance. The common header occupies a substantial part of the packet. Avoiding multiple transmission of common part of the header saves energy.

Keywords: Aggregation · Wireless sensor network · Energy · Data packet

1 Introduction

Most applications using sensors are used to monitor, measure continuously varying physical parameter like humidity, temperature, light intensity, etc. Sensors require power to run the electronic circuitry. The source of power can be from the battery, solar panel or electrical grid lines. Applications like irrigation in agriculture need batteries running for one crop season of nearly six months. In the agricultural field, the moisture content in the soil, humidity and temperature is measured continuously. The main consumption of energy in this network is during transmission and reception of data. Measured data has to be sent from the location where it has sensed (source) to the main collection center called the base station (sink). If the distance between the source and sink is larger than the transmission range of the sensor, data is sent using multiple hops. Efficient route between source and sink is essential. Energy consumption is further reduced by aggregating data at some strategic location. DP_AODV, a Dynamic-Sized

N. Kumar and A. Thakre (Eds.): UBICNET 2017, LNICST 218, pp. 25–36, 2018.
https://doi.org/10.1007/978-3-319-73423-1_3

Data Packet Protocol is proposed in the current paper and it addresses all these issues. The ns2 simulator is used to strengthen the result. Co-related data is aggregated at the aggregator node based on the threshold value. If the difference in the collected data is beyond the threshold value, the data is appended to the payload part of the existing packet. The second section that follows substantiates the research work that is carried out in the relevant area. The third section addresses the methodology used to attain the required result. This is followed by the experimental results to reinforce the idea. Elaboration on the result follows. The methodology incorporated in this paper involves the process of clustering, routing, time synchronization between various clusters and packet handling.

2 Related Work

The current work is based on COMMON-Sense, a project which is still going-on at Pavagadh district, India [1]. The temperature and humidity of the surrounding area along with the moisture content in the soil is measured. In the paper emphasis is given to conserving energy at the MAC layer. Finally, data is sent via the gateway to the monitoring station. This paper [1], forms the basis for selecting the environmental parameters for the current project.

Literature survey is carried out along the lines of routing, co-relation, aggregation, data handling, energy consumption, connectivity, and coverage. Literature [2, 3], addresses intensively the various Routing techniques in wireless sensor networks. A delay-aware network structure for WSNs with in-network data fusion is proposed [4]. Study on optimized transmission and fusion cost is emphasized in work [5]. Spatial co-relation awareness in the dynamic and scalable tree is available in work [6]. Best response dynamics to local data is discussed in the literature [7]. Adaptive clustering without relying on exact sensor location information is dealt with in reference [8]. Whole integrated network situation for head node selection is studied in the paper [9]. Not only balancing the energy expenditure among sensors but also extending the network lifetime by equal usage of multiple optimal intermediate routers is addressed in the literature [10]. Route based on less time to reach the base station with aggregation taking place at the first level of the tree is dealt with [11]. Hierarchical Agglomerative Clustering where in, repeated merging of small clusters are carried out until all clusters scale to the satisfied threshold is accounted for [12]. Network division into unequal sized grid-shaped cluster is handled [13]. In [14], the grid whose cluster head consumes more energy takes part in cluster head rotation, shares energy load, balancing the energy dissipation. Some nodes send data in the form of a chain to the cluster head which again aggregates data at the sink. A method of clustering and sending data, based on the prediction of the value obtained is emphasized [15]. Data aggregation is carried along the spokes of the wheel [16]. Ring based data gathering is addressed [17]. In [18], the transmission range and the k-neighbor topology control strategy is used to prolong the network life. In [19], the network coverage and connectivity are achieved by dividing the sensors into sets of equivalent nodes and finding a better path. Paper [20] addresses that the connectivity is maintained even when some nodes have failed by dynamically adjusting the transmission power of the nodes. In [21], layered space-time directed graph is used to reduce the energy consumption and delay in the network. Reduction of energy consumption and data collection delay increases the accuracy [22]. In [23], the optimal size of the data packet is computed. This literature helps in identifying the maximum size of the data packet that can be supported in the network. In [24], the connectivity, is expressed by the probability that a node lies on a path to the sink, as a function of the probability that adjacent cells in a grid are connected. Survey paper [25], shows various ways of routing based on communication model, reliability, topology, and network structure for energy savings. Most of the literature surveyed gives the only emphasis on one or two parameters like routing, aggregation, low energy consumption, etc. In the current paper, actual value of the data is used for computation (though ns2 does not support data handling). Data transmission is efficiently handled by using techniques like routing and aggregation. The maximum packet size of 1024 bytes can be used for transmission.

3 Clustering and Routing

Clustering a set of sensor nodes help in localizing the co-related data in a randomly deployed network. Within each cluster, every node is in the hearing range of each other. Figure 1 depicts a network of 75 nodes. Selection of aggregator node is based on the positional information. Set of nodes within the cluster associate itself with the cluster head. Figure 2 shows the process of routing within and across the cluster. The green path shows the route followed to send data across the aggregator node. Blue lines shows the path to send data within the cluster. The number written in red, close to each of the node specifies the maximum number of hops that are required for any node to reach that aggregator node. Algorithm 1 shows the process of selecting the cluster head and Algorithm 2 to find a neighbor within the cluster.

Fig. 1. Selection of aggregator node

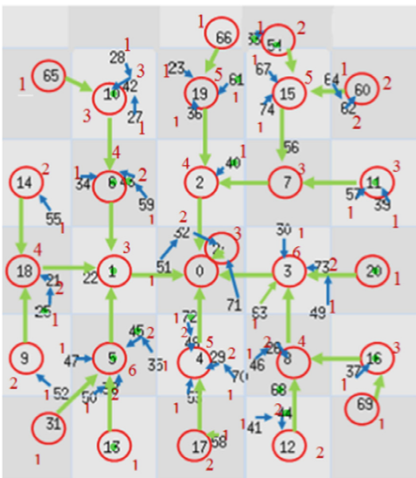


Fig. 2. Routing within and across clusters (Color figure online)

The Algorithm 1 finds the node closest to the center of a given rectangular cluster area. This node is assigned the status of cluster head.

```
Algorithm 1: To identify the cluster head in each cluster
  Input: N is the Node in the current cluster C
         Xi, Yi are the Coordinate of the node Ni in C
         Xc, Yc are the Coordinates of the topological
         Centre of the cluster C
  Output: Aggregator Node H
  begin
   H := Ø; MinDist := ∞; m := ∞; i:= 0
   While (Ni exists) repeat
       MinDist := DistanceBetween ((Xc, Yc), (Xi, Yi);
       If   (m > MinDist) m := MinDist;    H:= Ni
       i=i+1
   EndWhile.
  End.
```

In Algorithm 2, every node tries to find out which node is closest to itself within the given cluster. These nodes are associated with each other such that the node which sends the data is far from the center of the cluster and the one which receives data is closer to the center.

```
Algorithm 2: To find the neighbor during the routing
process within the cluster.
  Input: Ni is the Node id for Node i
         Na is the aggregator node (cluster head)
         Distia is the Distance between node Ni and Na
         Neighbor Nij  // Neighbor of i is j
  Output: Neighbor N
  Begin
    Distia: = ∞; Distja:= ∞; i:= 0; j:=1;
    For all Neighbor j, Vj= 0 // Node j is not visited
    If (Ni) N← Nia // If Ni exists Neighbor N of i is a
    While (Nj & !Vj) repeat
    If (Distia < Distij))
            N← Nia         // Neighbor N of i is j
      Else
            N← Nji         //Neighbor N of Nj is Ni
     End If
     Vj=1;
     j:=j+1
    End While
  End.
```

4 Time Synchronization Among Nodes to Send Data to Next Node

Synchronization is essential to carry out lose-less in-time data transmission between clusters. Algorithm 3 computes the waiting period that is required before forwarding the packet to the next hop neighbor. Figure 3 shows how many hop counts are required to reach the destination from any node in the network.

```
Algorithm 3: Waiting period and Elapsed Time
computation for forwarding of packets at any node N
Input: H is the Hop count to reach the current node from
the location where it has sensed the data
M :=1   // Hop count to itself set to 1
N is the No of nodes transferring data to that node
Output: Waiting period T
begin
  M:= Max of (H) from all neighbors    //for node N
  T:=The Time required to receive data from one hop
              neighbor x M
  ElapsedTime:= 0
  InitialTime:= CurrentMeasuredTime()
  While (M ≠ 1 && T > ElapsedTime) repeat
       If (packet received)
              M:= M-1
       EndIf
       ElapsedTime:=CurrentMeasuredTime()-InitialTime
   EndWhile
  Forward the Packet
 End.
```

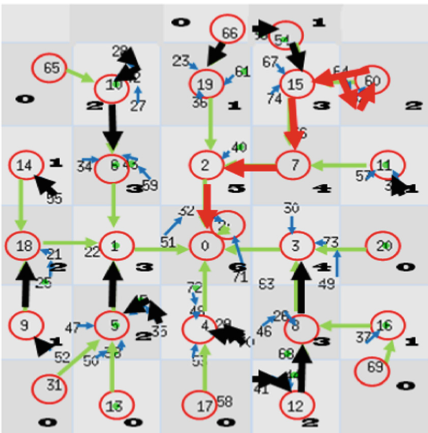


Fig. 3. Hop count computation to reach sink from any node in the network (Color figure online)

The black lines in the diagram represents how many hops are required to reach the base station from any of the node along its aggregating path. The red line represents the maximum path that is being taken to reach by any of the node in the whole network. In the current example this path length is six.

All the nodes sense the data and this information is forwarded to its neighbor. In order to carry out the aggregation process, all the data have to be available at the same time at the aggregator node. Some nodes take longer path (hops) to reach the aggregating cluster head nodes. In order to achieve the time synchronization, the path length for any node's data to reach the aggregator node is computed before-hand. At each aggregating location, the maximum among all the path length, for any of the nodes sending data to that node is computed. The waiting period is computed based on this maximum path length. The number of nodes which are aggregating their data at each aggregator is computed beforehand. As each data is received at the aggregator node the counter is decremented by one. As soon as the counter becomes one, data is sent to the next node without waiting. The Algorithm 3 computes the maximum hop counts required to reach the aggregator cluster head.

5 Aggregation of the Data at the Payload

Algorithm 4 is used to aggregate the data in the packet. If there is co-relation between data in the packet is within the threshold value (0.001 in this paper) then the data is aggregated else data is added to the payload along with the positional information. The choice of 0.001 is as per user requirement. The number of nodes contributing to the aggregation is based on this value. Smaller the value, smaller number of nodes will contribute to the aggregation. In a scenario like agriculture, the relative difference in the moisture content of the soil is low. Hence large area can be covered in one go.

Algorithm 4: Packet creation and forwarding towards Aggregator Node.

```
Input: D is the Data
       Ni is the Node I; Na the aggregator node a
Output: Packet P forwarded to next aggregator node
Count is the No of nodes whose data matches with the
current node's data computed so far.
begin
    Count:= 1
    Dp:= Ø; //Previous nodes data
    Dc:= Current node data
    StartTime := 0
While Ni ≠ Na repeat
ElapsedTime:= CurrentTime - StartTime
WaitTime := Max no of nodes sending data to this node
            multiplied by the time required for one hop
   For each Dp within the cluster
     If   (ElapsedTime > WaitTime)
         If | Dp - Dc | <= Threshold value
```

```
// computes average of the similar data computed so far
           Dp← (Dp x Count + Dc) / (Count +1)
           Count←Count +1
         Else
            Dp= Dp followed by Dc
 //New unexisting data which is beyond the threshold
             Count ←1
            EndIf
       EndIf
   EndForEach
 // Update in the packet format the values of Dp and the
 count corresponding to that Dp
 P=CMN Header appended with Dp, Count and current time
 Packet P forwarded to next hop node
 Endwhile
 End.  //The packet P available at the aggregator node
```

The algorithm for aggregation involves sensing the data and forwarding it to its next hop neighbor. At each of the nodes, there is a waiting period before aggregating and forwarding the data. This time is dependent on any of the nodes with a longest path (hops) which is incident at the aggregator node. After the waiting period all the data which have arrived are analyzed for co-relation. If there is a difference of less than 0.001 cm, then the average of the data which is incident at this node is appended to the current packet. Information on number of nodes having the similar data is available in the packet. Other data which is not co-related is simply appended to the current packet. The size of the data packet varies dynamically based on the co-related value. If the co-relation is large, smaller packet size is sent.

6 Simulation

Ns2 is used for carrying out the simulation. The agricultural application is taken up for designing the parameters [26]. The soil moisture content is noted once in every 5 min which is the input data to the simulator. The rate at which data is acquired is based on the rate in which the data changes at a specific location in the field. If it is raining, (or if the field is getting irrigated) then the moisture content value changes rapidly. During frequent changes the readings are taken often. One reading per day is taken during dry weather. Table 1 shows the simulation parameters that is used for the current work.

Table 2 shows the readings as obtained at the sink. *Time Now* specifies the time interval at which the data packet is obtained at the sink. *Cnt* refers to the number of nodes having *that specific* co-related value of *Data*. *Id* refers to the node which first sensed *that* value of data at the interval *Data time*. Data which was not able to synchronize within the specified time arrive late.

The first line in the Table 2 indicates that data is received at the sink node at the time interval 8.30 ms. The data at each of the nodes were sensed at the time 7.50 ms. The node with Id 51 has data of 2.094 cm. The count (Cnt) of *one* signifies only one neighbor exist with that value. This data is appended to the packet containing data obtained by node with id 32 and data 2.096 cm. This is follows data with Id 71 with data 2.098 cm. Here the count value is *two* signifying that there are two nodes which are close by and both are having their readings which are within a difference of 0.001. The average of the two readings is 2.098 cm. Each line represents one packet. The size of each of the packet is varying as noticed in Table 2.

Table 1. Simulation parameters used for carrying out the experiment

Radio parameters		Simulation parameter	
Radio frequency	868 MHz | 915MHz	Simulation interval	300secs.
Antenna Height	1m (min. reqd. height 0.0819m)	Topology	Random deployment
Antenna Type	Omnidirectional –Quarter wave	Channel	Wireless
Transmit Power	3.16 mW =5dBm	Network interface	Wireless Physical
Receive Power	−104dBm@ 5dBm=3.98e-14W	Queue length	50
Carrier Sense Threshold	−104dBm@ 5dBm	Network size	75 nodes
Capture Threshold	10 dBm	Transport protocol	UDP
Gain of transmitting/ receiving antenna	1	Application Traffic	CBR
Path loss	1	Data acquisition interval	150 | 300 | 600
Sensor parameters (Tiny Node)		Topology	Random deployment
Receive Power	0.042W=16.23dBm	Channel	Wireless
Idle Power	0.006W=7.78dBm	Network interface	Wireless Physical
Transmit Power	0.099W=19.95dBm	Queue length	50
Sleep Power	0.000003W=-25.2 dBm	Network size	75 nodes
		Transport protocol	UDP
		Application Traffic	CBR
		Data acquisition interval	150 | 300 | 600

Table 2. Readings as measured at the sink

Time now	Data time	Id	Data	Cnt	Data time	Id	Data	Cnt	Data time	Id	Data	Cnt	Data time	Id	Data	Cnt
8.60	7.50	51	2.094	1	7.50	32	2.096	1	7.50	71	2.098	2				
8.83	7.50	72	2.093	1	7.50	48	2.096	1	7.72	17	2.096	2	7.84	4	2.097	4
9.50	7.50	28	2.100	1	7.50	42	2.098	1	7.72	10	2.096	2	7.90	6	2.096	4
	7.65	14	2.098	2	7.52	13	2.098	1	7.74	9	2.095	2	7.56	18	2.099	3
	7.98	5	2.097	7	7.8	1	2.097	2								
10.00	7.52	19	2.098	5	7.74	60	2.095	3	7.72	54	2.098	2	7.56	15	2.099	3
	7.88	7	2.096	2	7.56	11	2.099	3	7.88	2	2.096	2				
11.0	7.5	63	2.097	1	7.86	3	2.095	4	7.50	12	2.100	3	7.50	8	2.098	4
	7.74	16	2.095	3	7.5	20	2.097	1								
11.16	7.50	27	2.100	1												

All timing measurements are in seconds. Data readings are in centimeters. *Cnt refers to the count.*

7 Result Analysis

The current protocol DP_AODV is compared with the protocol M_AODV [27]. M_AODV takes the efficient route to the destination with minimum hop count. Unnecessary broadcasting of path search messages are reduced in M_ADOV as compare to AODV protocol [28]. The result shows a relative comparison of the energy consumed by both the protocol for the same scenario.

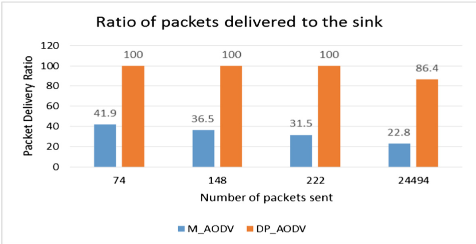
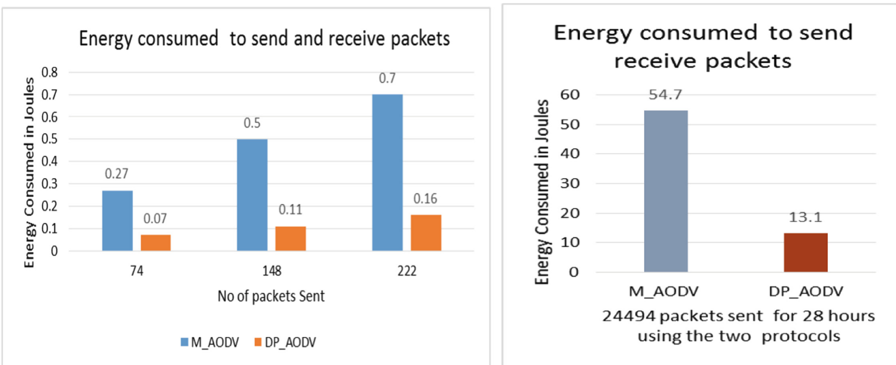


Fig. 4. Packet delivery ratio



a) Small number of packets delivered b) large number of packets delivered

Fig. 5. Energy consumption to send and receive packets

Figure 4 shows the packet delivery ratio for varying number of packets (74, 148, 222 and 24494) that are sent to the base station. Figure 5a shows the amount of energy consumed to send (74, 148, 222) packets and Fig. 5b to send 24494 packets.

The packet delivery ratio is high in the case of DP_AODV as compared to M_AODV. The dropping of packet is maximum in the case of M_AODV. This is because of huge traffic flow in the network. Packet delivery ratio is high in DP_AODV. When few (up to 222) packets are sent, the packet drop is low in M_AODV and almost nil in DP_AODV. When large number of packets are sent, due to non-synchronization of time between packets, queue size increases leading to dropping of packets in DP_AODV.

The energy consumption during transmitting and receiving of data is high in the case of M_AODV as compared to DP_AODV. With DP_AODV, packets are locally aggregated with regard to co-related data. If the difference in data is within of 0.001, data is aggregated, else the Id of the node with the measured data is appended to the existing payload.

8 Discussion and Conclusion

Wireless sensor network has limited energy. Hence energy conservation is a critical issue in the system design of WSN. In applications involving environmental, physical parameters measurement it is essential that the batteries run continuously for long intervals of time. To accomplish this, the number of transmissions and receptions should be reduced. Communication is the main cause for heavy energy consumption. In-network aggregation, clustering and combining of co-related data and combining data from multiple sensors and sending it as a combined packet, saves energy.

In the current paper, routing techniques are incorporated such that the route follows the co-related data within the clusters. Data from many packets are aggregated together and are sent simultaneously. Aggregation here refers to appending data to existing packet (if data is un-related) or averaging, if data is related. This results in Dynamic-Sized Data Packet. The common part of the header is retained, and only the data part is appended to the common header along the path. Data is sent across the aggregator nodes (cluster head) using minimum hop count to reach the base station.

Efficiency is improved using DP_AODV due to the following reasons. Firstly, number of packets reaching the base station is reduced. Due to this reason, the processing overhead at base station is reduced. Traffic is reduced. Packet collision and dropping of packets is reduced. Secondly, all data are sent to the same base station. This redundant information need not be specified for each data that is sent. Thirdly, the size of the data packet changes dynamically based on how many data are co-related. Common part of data is sent only once resulting in energy savings. The problem associated with aggregation is that the aggregator node has to wait for all nodes to send data. A large amount of delay is introduced in obtaining the value at the base station. Synchronization is of utmost importance. Loss of single packet can result in loss of large amount of information.

The current paper emphasis on how energy is saved using proper routing, clustering, aggregation and data combining techniques. This is especially useful for application as in agriculture where the soil moisture content measurement could be carried out for nearly six months when the information collected is highly related. The proposed method can also be incorporated for other applications as well. Monitoring the health of the coral reef, toxicity in water (for marine life), humidity and luminosity information for growth of crop are examples of some other application.

References

1. Panchard, J., Rao, S., Prabhakar, T.V., Jamadagni, H.S., Hubaux, J.: COMMON-sense net: improved water management for resource-poor farmers via sensor networks. In: International Conference on Information and Communication Technologies and Development (2006)
2. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirc, E.: Wireless sensor networks: a survey. J. Comput. Netw. **34**(4), 393–422 (2002)
3. Al-Karaki, J.N., Kamal, A.E., et al.: Routing techniques in wireless sensor networks: a survey. J. IEEE Wirel. Commun. **11**(6), 6–28 (2004)
4. Aschenbruck, N., Fuchs, C.: An integrated simulation environment for sensor data fusion applications in wireless mesh networks. In: Military Communication Conference, pp. 1778–1783 (2011)
5. Luo, H., Liu, Y., Das, S.K.: Routing correlated data with fusion cost in wireless sensor networks. IEEE Trans. Mob. Comput. **5**(11), 1620–1632 (2006)
6. Villas, L.A., et al.: A spatial correlation aware algorithm to perform efficient data collection in wireless sensor networks. Ad Hoc Netw. (2011)
7. Zeydan, E., et al.: Energy-efficient routing for correlated data in wireless sensor networks. Ad Hoc Netw. **10**(6), 962–975 (2012)
8. Ci, S., et al.: Adaptive clustering in wireless sensor networks by mining sensor energy data. Comput. Commun. **30**(14), 2968–2975 (2007)
9. Sha, C., Wang, R., Huang, H., Sun, L.: Energy efficient clustering algorithm for data aggregation in wireless sensor networks. J. China Univ. Posts Telecommun. **17**(2), 104–109 (2010)
10. Liu, T., Li, Q., Liang, P.: An energy-balancing clustering approach for gradient-based routing in wireless sensor networks. J. Comput. Commun. **35**, 2150–2161 (2012)
11. Misra, S., Thomasinous, D.: A simple, least-time, energy-efficient routing protocol with one-level data aggregation for wireless sensor networks. J. Syst. Softw. **83**(5), 852–860 (2010)
12. Du, T., Qu, S., Liu, F., Wang, Q.: An energy efficiency semi-static routing algorithm for WSNs based on HAC clustering method. J. Inf. Fusion **21**, 18–29 (2013)
13. Yuea, J., et al.: Energy efficient and balanced cluster-based data aggregation algorithm for wireless sensor networks. In: Proceedings of the International Workshop on Information and Electronics Engineering, Harbin, China, vol. 29, pp. 2009–2015 (2012)
14. Tang, F., et al.: A chain-cluster based routing algorithm for wireless sensor networks. J. Intell. Manuf. **23**(4), 1305–1313 (2012)
15. Jiang, H., Jin, S., Wang, C.: Prediction or not? An energy-efficient framework for clustering-based data collection in wireless sensor networks. IEEE Trans. Parallel Distrib. Syst. **22**(6), 1064–1071 (2011)
16. Sutagundar, A.V., Manvi, S.S.: Wheel based event-triggered data aggregation and routing in wireless sensor networks: agent-based approach. J. Wirel. Pers. Commun. **71**(1), 491–517 (2013)
17. Lu, K.H., et al.: Hierarchical ring-based data gathering for dense wireless sensor networks. J. Wirel. Pers. Commun. **64**(2), 347–367 (2012)
18. Guo, D., et al.: Prolonging network lifetime and balancing network energy in multi-domain WSNs. Int. J. Sens. Netw. **20**(4), 209–218 (2016)
19. Boucetta, C., et al.: Tree-based modeling of redundancy and paths in wireless sensor network. Int. J. Inf. Commun. Technol. **8**(2), 212–234 (2016)
20. Deniz, F., et al.: An adaptive energy-aware and distributed fault-tolerant topology-control algorithm for heterogeneous wireless sensor networks. J. Ad Hoc Netw. **44**, 104–117 (2016)

21. Chen, H., et al.: Spanning tree-based topology control for data collecting in predictable delay-tolerant networks. J. Ad Hoc Netw. **46**, 48–60 (2016)
22. Ardakani, S.P., et al.: CBA: a cluster-based client/server data aggregation routing protocol. J. Ad Hoc Netw. **5**, 1–33 (2016)
23. Sankarasubramaniam, Y., Akyildiz, I.E., Mchughlin, S.W.: Energy efficiency based packet size optimization in wireless sensor networks, pp. 1–8. IEEE (2013)
24. Legakis, H., Mehmet-Ali, M., Hayes, J.F.: Lifetime analysis for wireless sensor networks. J. Sens. Netw. **17**(1), 1–16 (2015)
25. Nikolaos, A., et al.: Energy efficient routing protocols in wireless sensor networks: a survey. IEEE Commun. Surv. Tutor. **15**(2), 551–591 (2013)
26. Pai, S.N., Shet, K.C., Mruthyunjaya, H.S.: Simulation environment for sensors placed in agricultural field. In: International Conference and Workshop on Emerging Trends in Technology, pp. 752–755 (2012)
27. Pai, S.N., Shet, K.C., Mruthyunjaya, H.S.: Efficient path finding algorithm for transmission of data in agricultural field using wireless sensor network. In: Das, V.V., Chaba, Y. (eds.) Mobile Communication and Power Engineering. CCIS, vol. 296, pp. 345–348. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-35864-7_50
28. Perkins, C., Belding-Royer, E., Das, S.: Ad hoc on-demand distance vector (AODV) routing. No. RFC 3561 (2003)