An Improved Method for Data Gathering in Wireless Sensor Networks Using Mobile Collector

[I]R. Kingsly Stephen, [II]P.C. Sarah Prithvika, [III]P.S.K. Patra
[I]Research Scholar, Anna University, Chennai, Tamil Nadu, India
[II]Dept. of CSE, Agni College of Technology, Anna University, Chennai, Tamil Nadu, India
[III]Head of the Dept., Dept. of CSE, Agni College of Tech., Anna University, Chennai, Tamil Nadu, India

Abstract
In large-scale wireless sensor networks, a new mechanism for gathering data is proposed by incorporating mobility into the network. A mobile collector is used to gather data from the sensor nodes.A robot or vehicle with a powerful battery and transceiver can be used as a mobile collector. The lifetime of the sensor nodes in wireless sensor is maximized. This is achieved by minimizing the length of each data-gathering tour. This method improves scalability and ensures that energy consumption among sensors is uniform.

Keywords
Data gathering, mobile data collector, movement planning, wireless sensor networks

I. Introduction

Wireless Sensor Networks are used in a large variety of applications. Sensor nodes are usually deployed in large scale networks without prior pre-configuration. The energy of sensor nodes is used for two tasks, sensing the environment and uploading data to the data sink. Energy utilization of sensors during sensing is stable, but the data-gathering mechanism is the significant factor that determines network lifetime. The data packets are aggregated at a point and then sent to the data sink. Sensor nodes that are close to the sink lose more energy as they have to relay the sensed data to the data sink. If these sensor nodes fail, then the other nodes cannot reach the data sink. So for large-scale sensor networks it is inefficient to have a single static data sink to gather data from all the sensors. A mobile collector can act as data transporter that connects disconnected networks. The path of the moving mobile collector acts as virtual links connecting separated sub-networks. The assumption in this paper is that the data rate is less and that the data is usually delay insensitive. The mobile collector starts its tour from the data sink and traverses through the field and collects data. The path of the mobile collector can be well-planned to improve the lifetime of the network.

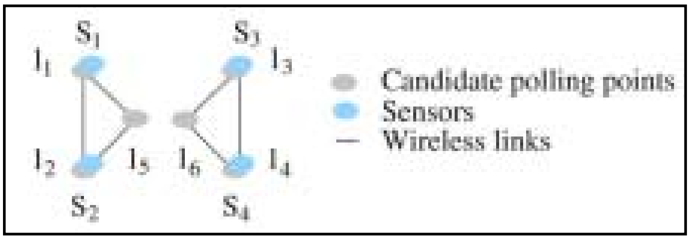
II. Related Works

In hierarchical networks, clusters are formed and form the lower layer of the network. At higher levels, cluster heads collect sensing data from sensors and forward the data to the data sink. Scalability and energy-efficiency is usually improved for such two-layered hybrid networks. Cluster head is not only used to aggregate data but also to make scheduling and routing decisions. Cluster heads fail faster than other nodes. To avoid the problem of cluster heads failing faster than other nodes, cluster heads can be rotated. Each node must have the capacity to be chosen as cluster head, as any node can be chosen to be the cluster head. Changing cluster heads frequently results in high overhead due to high information exchange among sensor nodes. This disadvantage can be avoided by using heterogeneous networks. The network consists of a few resource-rich nodes and many resource-limited nodes. The resource-rich nodes act as cluster heads, and the network is organized into a two-layeredhierarchical network. Sensor nodes can be attached to animals [1], [2]. Radio-tagged zebras and whales were used as mobile nodes to sense data in the wild environment. The animals wander in the environment and they exchange data only when they move close to each other. The mobility of the animals is unpredictable and so the delay is not predictable. Sensor nodes can be placed in public transportation vehicles such as trains and buses which move along fixed paths, but it is very restrictive because it depends on the existing routes and schedules of trains and buses.In [3], [4] a number of mobile observers, called data mules, collect data from the sensors directly when they are in close range, temporarily store the data, and send the collected data to wired access points.The movement of mules is modeled as 2-D random walk.In [5], mobile observers move along the sensing field along straight lines that are parallel and fetch data from sensors. To decrease latency, data sent by some sensors are allowed to be passed to other sensors to reachthe mobile observers. This scheme works well ina uniformly distributed large-scale sensor network. But,in practice, data mules may not always be able to move alongstraight lines, for example, obstacles or barriers may blockthe moving paths of data mules. When not all sensors are connected and only a small number of data mules are available, the datamules may not cover all the sensors nodes in the network if theyonly move along straight lines.In [6], the tradeoff between energy-saving and data-gathering latency in mobile data gathering wasanalyzed by exploring a balance between the relay hop count of the local data aggregation and the moving tour length of the mobile collector. Two algorithms were proposed,in which a subset of sensors is selected as polling points tostore the locally aggregated data and then to upload the data to the mobile collector when it arrives.Meanwhile, when sensors are associated with these polling points, it is assured that any packet relay is limited within a given number of hops. In [7], a heterogeneous and hierarchical architecture for large scale monitoring was proposed for thedeployment of WSNs with mobile sinks, where the sensors transmit their sensing data to the gateway nodes for storage temporarily through multi-hop relays andthe mobile sinks travel along pre-determined paths to collect data from close by gateway nodes. In this data-gatheringmodel, the capacitated minimum forest problem was studied, and approximate algorithms were devised for instanceswhere all gateways have uniform and arbitrary capacities, respectively.

So based on the above discussions it is clear that to be suitable for various network topologies it is essential to have a mechanism where the data collection is done by a mobile data collector that moves along the field and collects data.

III. Terminology

The M-collector can visit the transmission range of every static sensor, such that sensing data can be collected by a single-hop communication without any relay.



Sensors : S_1, S_2, S_3 and S_4
Neighbor set of l_1, l_2 and l_5 : { S_1, S_2 }
Neighbor set of l_3, l_4 and l_6 : { S_3, S_4 }
Candidate polling point set : { l_1, l_2, l_3, l_4, l_5, l_6 }

Fig. 1: Examples of Polling Points, Neighbor Sets, and Candidate Polling Point Set

1. Polling Points – Polling points are the positions where the M-collector polls sensors.
2. Neighbor Set – The neighbor set of a point in the plane is the set of sensors that can upload data to the M-collector directly without relay, if the M-collector polls sensors at this point.
3. Candidate Polling Point Set –Since it is only possible to test a certain number of points and their corresponding neighbor sets in the plane, and we must select polling points from this set of points, which we refer to as the candidate polling point set.

While an M-collector is moving, it can poll nearby sensors one by one to collect the data. On receiving the polling message, a sensor simply uploads the data to the M-collector directly without relay. M-Collector polls sensors from polling points. When an M-collector moves to apolling point, it polls nearby sensors with the same transmissionpower as sensors, such that sensors that receive the pollingmessages can also upload packets to the M-collector in a single hop.After collecting data from sensors around the polling point, theM-collector moves directly to the next polling point in the tour.

Thus, each data-gathering tour of an M-collector consists of anumber of polling points and the straight line segments connecting them.Thus, the problem of devising the optimal tourcan be considered as the problem of determining the locationsof polling points and the order to visit them. Before an M-collector starts a data-gathering tour, it needs to determine thepositions of all polling points and which sensors it can poll ateach polling point.Since the M-collector can only collect data atpolling points, each sensor must be in the neighbor set of at leastone polling point to upload data without relay. In other words,the union of all neighbor sets of all polling points must cover allthe sensors in the field.

In some existing work, the transmission range of an omnidirectional antenna was simply assumed to be a disk-shaped areaaround the transceiver. Based on this assumption, given a pointin the plane, the neighbor set of this point consists of all sensorswithin the disk-shaped area around this point. However, dueto the uncertainties of a wireless environment, such as signalfading, reflection from walls and obstacles, and interference,it is hard to determine the boundary of the transmission rangewithout actually measuring [8], [9]. Therefore, in practice, itis almost impossible to obtain the neighbor set of an unknownpoint, unless the M-collector has moved to this point andtested wireless links between it and its one-hop neighbors, ora sensor has been placed at this point and acquired all its one-hop neighbors during the neighbor discovering phase.

In fig. 1, we illustrate the polling points, neighbor set, andcandidate polling point set by an example, where there are foursensors s_1, s_2, s_3, and s_4 deployed at positions l_1,l_2,l_3, and l_4, respectively. During the exploration phase, the M-collectordiscovers the neighbor sets of l_5 and l_6 by broadcasting “Hello”messages at these points. Thus, l_5 and l_6 can be added into thecandidate polling point set. Since sensors s_1, s_2, s_3, and s4alsoreport their one-hop neighbors to the M-collector by sending“ACK” to the M-collector, l_1, l_2, l_3, and l_4 also become candidate polling points.After the discovering phase, we assume that each sensor hasknowledge and information of all its one-hop neighbors and the M-collectoracquires the information about the neighbor set of each polling point.

IV. Single Hop Data-Gathering Problem

Here, we consider the problem of determining the shortest movingtour of an M-collector that visits the transmission range of eachsensor.For the sake of simplicity, weassume that M-collectors move at a constant speed and ignore thetime for making turns and data transmission, such that we canroughly estimate the time of a data-gathering tour by the tourlength. By moving through the shortest tour, data canbe collected in the shortest time such that the users will havethe most up-to-date data. This is called as the single-hop data-gathering problem (SHDGP).

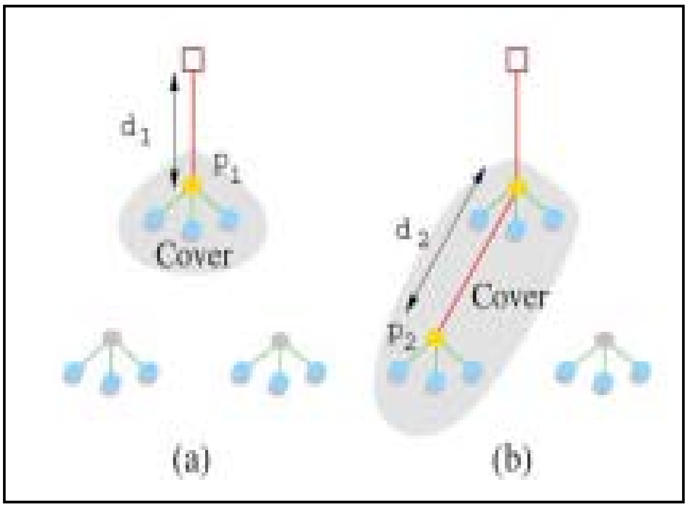
The SHDGP can be formalized as follows, after the candidate polling point set is obtained. Given a set of sensornodes S = { s_1, s_2,..., s_{ns}}, a set of candidate polling pointset L = { l_0, l_1, l_2 ,...,l_{nl}}, where l_0 denotes the starting andthe ending point of the tour, and the neighbor set nb(l_i) ofeach candidate polling point l_i (i = 1,2,...,n_l) find a set ofpolling points and determine the sequence to visit them, suchthat every sensor in S belongs to the neighbor set of at least onepolling point, and the total length of line segments connectingall polling points is minimized. We define a complete directedgraph G = (L, A) and associate a nonnegative cost c_{ij} witheach arc $a_{ij} \in$ A, where c_{ij} is equal to the cost of the distancebetween the candidate polling points l_i and l_j.

The SHDGP canbe formulated as an MIP. In this formulation, the objective function (1) minimizes the total cost (distance) of the data-gathering tour. x_{ij} is an indicatorvariable denoting whether arc a_{ij} from candidate polling points l_i to l_j belongs to the optimal tour. Binary variable I_i indicateswhether candidate polling point l_i is on the optimal tour. Constraints (2) and (3) make sure of the fact that every node in the tourmust have one arc pointing toward it and the other arc pointing away from it. Constraint (4) enforces that every sensor must bein the neighbor set of at least one polling point belonging tothe tour, such that every sensor can communicate with the M-collector directly. Constraints (5)–(7) exclude the solutions withsub-tours, which is similar to that in [12] proposed by Gavishfor the capacitated minimal directed tree problem. The Constraint (5) restricts that flow can only take place in an arc if itis on the tour. Constraint (6) specifies that the number of units offlow entering vertex l_0 is equal to the number of polling pointsin the tour. Finally, constraint (7) enforces that one unit flowsout of each of the other points in the tour. In a way similar to that in [12], it can be shown that constraints (5)–(7) can prohibita tour that does not include the starting and ending point l_0.

Minimize
$$\sum_{i,j\in L, i\neq j} c_{ij}x_{ij} \qquad (1)$$
Subject to
$$\sum_{i\in L, i\neq j} x_{ij} = I_j \qquad \forall j\in L \qquad (2)$$
$$\sum_{j\in L, j\neq i} x_{ij} = I_i \qquad \forall i\in L \qquad (3)$$
$$\sum_{j\in nb(l_i)} I_i \geq 1 \qquad \forall j\in S \qquad (4)$$
$$y_{ij} \leq |L| x_{ij} \qquad \forall i,j\in L \qquad (5)$$
$$\sum_{j\in L\setminus\{l_0\}} y_{jl_0} = \sum_{j\in L\setminus\{l_0\}} I_j \qquad (6)$$
$$\sum_{i} y_{ji} - \sum_{k} y_{kj} = I_j \qquad \forall j\in L\setminus\{l_0\} \qquad (7)$$
where
$$x_{ij} = \begin{cases} 1, & \text{if data-gathering tour contains arc } a_{ij} \\ 0, & \text{otherwise} \end{cases}$$
$$I_i = \begin{cases} 1, & \text{if data-gathering tour contains} \\ & \text{candidate polling point } l_i \\ 0, & \text{otherwise} \end{cases}$$
y_{ij} : flow value from l_i to l_j on arc a_{ij}.

V. An Algorithm for Single Hop Data- Gathering Problem

It is interesting to compare the SHDGP with a similar problem,i.e., the covering salesman problem (CSP). The problem that is considered is obtaining the shortest tour ofa subset of all cities such that every city not on the tour iswithin some predetermined distance dist of a city that is onthe tour. If the transmission range of each sensor could bemodeled as a disk-shaped area, the SHDGP can be simplifiedto the CSP by setting dist in the CSP equal to the transmissionrange of sensors.Note that if the transmission range of each sensor canbe modeled as a unit circle, the covering line algorithm for theCSP could also approximate the optimal tour for the SHDGP.However, as discussed earlier, in practice, transmission rangesof sensors are far from unit circles and are difficult to estimate.



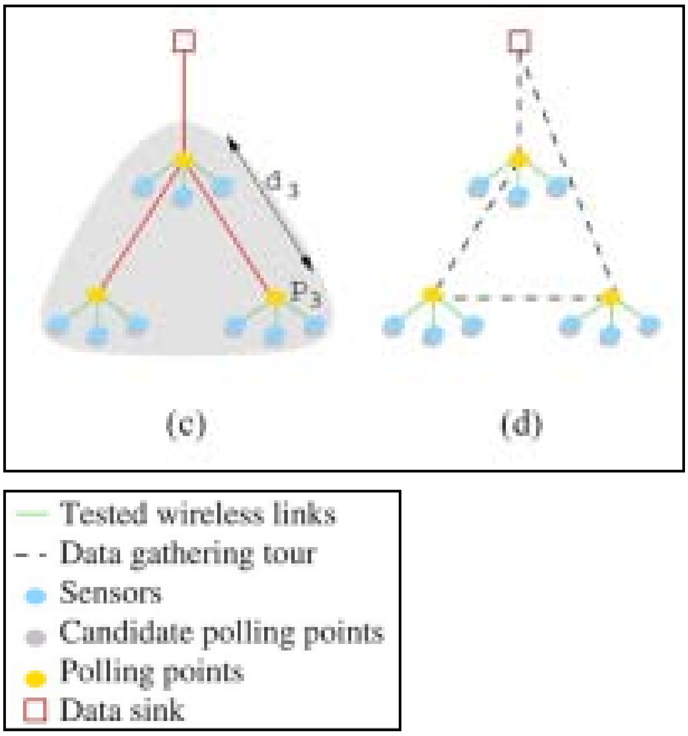


Fig. 2: Spanning tree covering algorithm. (a) Neighbor set of p1 is covered with the average cost d1/3. (b) Neighbor set of p1 is covered with the average cost d2/3. (c) Neighbor set of p1 is covered with the average cost d3/3. (d) Data-gathering tour obtained by the spanning tree covering algorithm.

We must choose a subset of points from the candidate polling pointset, each of which corresponds to a neighbor set of sensors. At each stage of the algorithm, a neighbor set of sensors canbe covered when its corresponding candidate polling pointis chosen as a polling point in the data-gathering tour. Thealgorithm will terminate after all sensors are covered. Thealgorithm tries to cover each uncovered neighbor set of sensorswith the minimum average cost at each stage In Fig. 3, an example of thespanning tree covering algorithm is illustrated. In the Fig. 2, starting from the data sink, the M-collector chooses p_1 as thefirst polling point, since compared with other candidate pollingpoints, the uncovered sensors in the neighbor set of p_1 can becovered with the smallest average cost $d_1/3$. Next, p_2 and p3will be picked as the second and third polling points with the average cost $d_2/3$ and $d_3/3$, respectively. After that, the shortesttour can be approximated on all chosen polling points and thedata sink, as shown in Fig. 3(d).

VI. Simulation Results

It has been shown by simulation that the throughput and energy-utilized using M-Collector is better than without one [11] as depicted in Fig. 3 and Fig. 4.

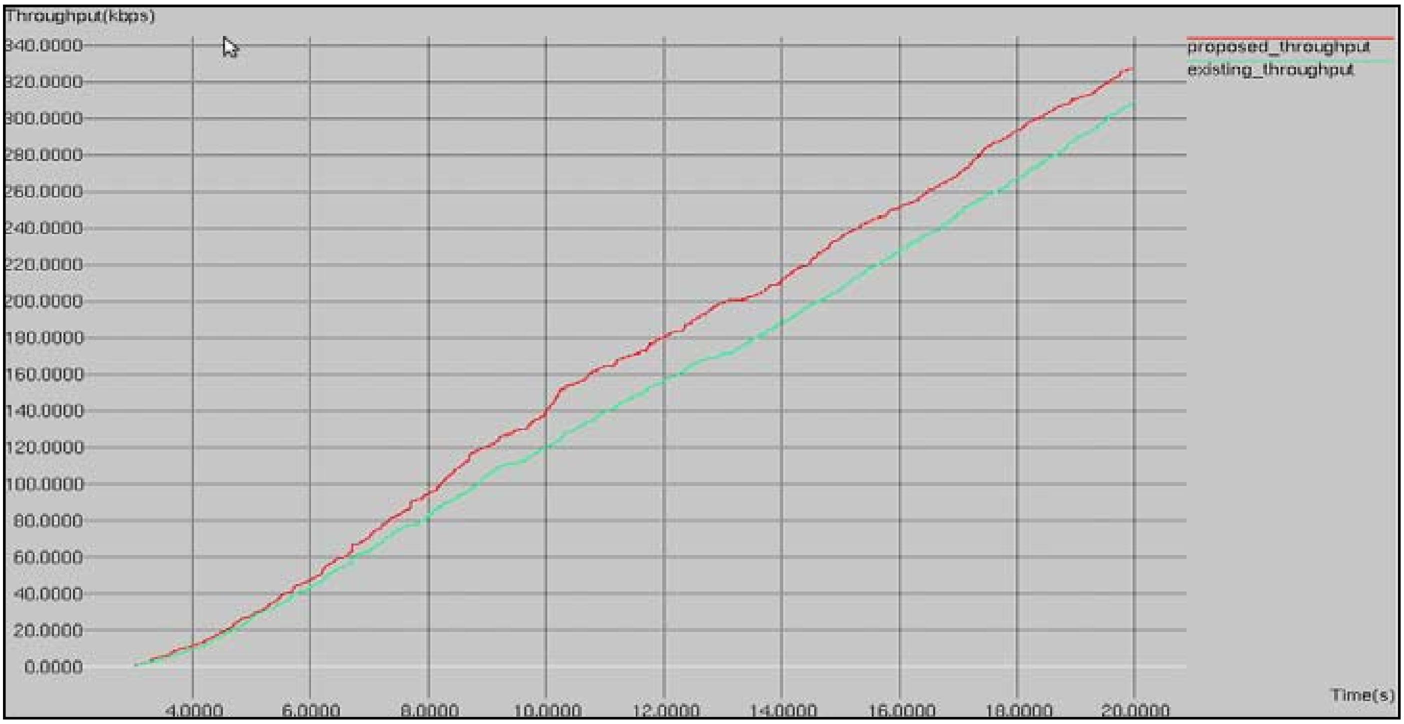


Fig. 3: Comparison of throughput in existing and proposed systems

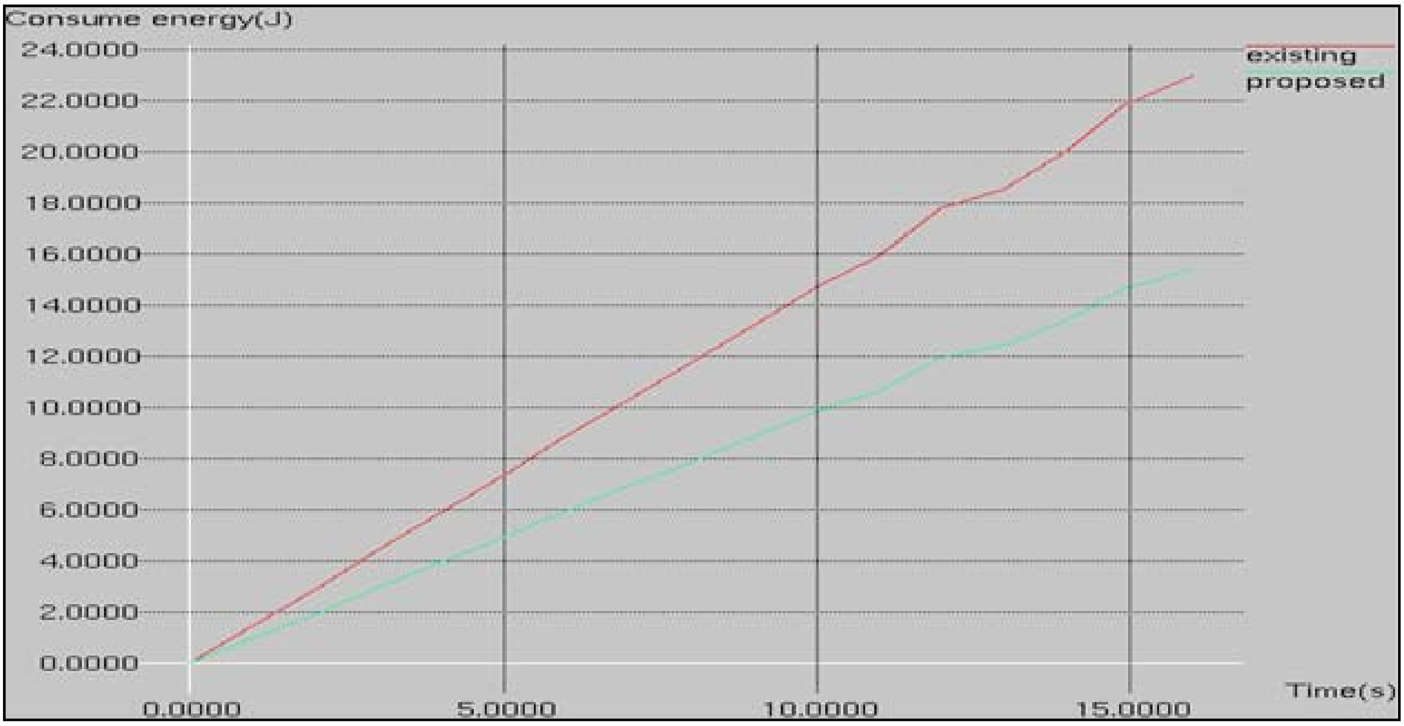


Fig. 4: Comparison of energy consumed in existing and proposed systems

VII. Conclusion and Future Work

In this paper we have discussed the techniques used to effectively collect data in large scale networks by introducing mobility into the network. In order to satisfy time constraints,we can improve this mechanism byusing multiple mobile collectors, byletting each of them move through a shorter sub-tour. In [10] it has been assumed that the data collected is not critical, but we also have to consider the case where emergency data has to be transmitted. The emergency data can be transmitted directly to the data sink, instead of waiting for the M-Collector to collect the data.

References

[1] *P. Juang, H. Oki, Y. Wang, M. Martonosi, L. Peh, andD. Rubenstein, "Energy-efficient computing for wildlife tracking: Design tradeoffs and early experiences with ZebraNet," in Proc. ASPLOS, 2002, pp. 96–107.*

[2] *T. Small and Z. Haas, "The shared wireless infostation model—A new adhoc networking paradigm (or where there is a whale, there is a way)," in Proc. ACM MobiHoc, 2003, pp. 233–244.*

[3] *R. C. Shah, S. Roy, S. Jain, and W. Brunette, "Data MULEs: Modeling athree-tier architecture for sparse sensor networks," in Proc. IEEE Workshop Sens. Netw. Protocols*

Appl., 2003, pp. 30–41.

[4] S. Jain, R. C. Shah, W. Brunette, G. Borriello, and S. Roy, "Exploitingmobility for energy efficient data collection in wireless sensor networks".Norwell, MA: Kluwer, 2005.

[5] D. Jea, A. A. Somasundara, and M. B. Srivastava, "Multiple controlled mobile elements (data mules) for data collection in sensor networks," in Proc. IEEE/ACM Int. Conf. DCOSS, Jun. 2005, pp. 244–257.

[6] M. Zhao and Y. Yang, "Bounded relay hop mobile data gathering in wireless sensor networks," IEEE Trans. Comput., vol. 61, no. 2, pp. 265–277, Feb. 2012.

[7] W. Liang, P. Schweitzer, and Z. Xu, "Approximation algorithms for capacitated minimum forest problems in wireless sensor networks with a mobile sink," IEEE Trans. Comput., to be published.

[8] G. Zhou, T. He, J. Stankovic, and T. Abdelzaher, "RID: Radio interference detection in wireless sensor networks," in Proc. IEEE INFOCOM, 2005,pp. 891–901.

[9] G. Zhou, T. He, and J. Stankovic, "Impact of radio irregularity on wireless sensor networks," in Proc. 2nd Int. Conf. MobiSys, Jun. 2004, pp. 125–138.

[10] M. Ma, Y. Yang and M. Zhao, "Tour Planning for Mobile Data-Gathering Mechanisms in Wireless Sensor Networks", IEEE TRANSACTIONS ON VEHICULARTECHNOLOGY, VOL. 62, NO. 4, MAY 2013.

[11] B.Manimekala1 and M. Kayalvizhi, "A Data Transfer in Wireless Sensor Networks Using AODV Protocol", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, No 1, January 2012

[12] B. Gavish, "Formulations and algorithms for the capacitated minimal directed tree problem," J. ACM, vol. 30, no. 1, pp. 118–132, Jan. 1983.