2009 IEEE International Advance Computing Conference (IACC 2009)
Patiala, India, 6–7 March 2009

Multi-Hop Routing for Multi-Stationed Wireless Sensor Networks

S H Manjula*, Prabhavathi M**, Arjun A V**, Shaila K**, L Nalini**, K R Venugopal**, L M Patnaik***
**Department of Computer Science and Engineering
**University Visvesvaraya College of Engineering, Bangalore University, Bangalore 560 001
*Dr. M G R University, Chennai
***Vice Chancellor, Defence Institute of Advanced Technology(Deemed University), Pune, India
shmanjula@yahoo.com

***Abstract*— Wireless Sensor Networks (WSN) consists of tiny autonomous devices called sensors capable of sensing, processing and transmitting information. Energy consumption, routing and maximizing lifetime are the important challenges in the sensor network. In this paper, we propose Multiple Tree Construction (MTC) algorithm to address the problem of finding path from each node to its nearest base stations to maximize the network lifetime. The algorithm deals with having optimal number of base stations. Analytical and simulation results show that MTC performs better than existing algorithms involving single base station.**

***Index Terms*— Network Lifetime, Base Station, Query, Optimal Placement, Data Aggregation.**

I. INTRODUCTION

Wireless Sensor Networks consist of a large number of sensor nodes deployed over area of interest. A sensor node is a micro-electronic device which consists of a sensing unit, a processing unit, a tranceiver unit and a power unit. Sensing unit senses the data in the area under monitoring and send data to the base station. Processing unit consist of a small storage and a microprocessor. Transceiver unit connects the sensor node with the network. Power unit consists of a small battery which is the main source of power. Sensor networks find potential applications in wide variety of fields. WSN are used in environmental and habitual monitoring, indoor climate control, surveillance, treaty verification and intelligent alarms.

A sensor node is resource constrained having limited processing speed, storage capacity, processing power and communication bandwidth. The lifetime of a sensor node strongly depends on battery. Recent advances in the hardware technology helps to overcome constraints on processing speed and communication bandwidth. In most of the cases, the network must operate for longer durations and since the nodes are wireless the available energy resources limits the overall operation of network. Most of the time a sensor node must operate unattended and be adaptive to the environment.

Considering sensor network as a database, the lifetime of sensor network is defined to be the number of queries successfully answered by the network. Every sensor node sends the sensed data to the base station when requested. A query is defined to be one such request for data sensed by all the nodes in the network. The network is said to deteriorate even if one node in the network fails. So the problem of maximizing network lifetime reduces to maximizing the number of queries answered before the first node failure in the network.

A base station is a special type of a node in a sensor network which has high computing power and unlimited energy supply. This node is responsible for quering the sensor nodes. Processing of the sensed data and taking necessary action is the task of a base station. Multiple base stations co-ordinate with each other through satellite communication. Normally a base station is situated in the border of the surviellance area.

Motivation : A single base station network has a limitation in satisfing a number of queries. Since the distance from the base station to a sensor node is more, the sensed data is transmitted through multiple hops. This causes the heavy work load on the intermediate sensors as a result nodes near the base station will drain more energy. In this work, we introduce the concept of multiple base stations to solve this particular problem.

Contribution : We present the MTC algorithm to construct routing trees for different base stations resulting in energy efficient transfer of data. We propose a model for optimal placement of base stations for any network. Further, we compare the performance of the network for different number of base stations to validate our model.

Organization : In rest of the paper, Section II gives overview of literature survey. We formulate the problem of maximizing the number of queries in Section III. The design and analysis of MTC algorithm is presented in Section IV and implementation details in Section V. Analysis of algorithm and comparison of simulation results is given in Section VI. Section VII contains the conclusions.

II. LITERATURE SURVEY

Akyildiz et al., [1] explain the concept of design issues in WSN. Ranjan et al., [2] proposed a generic architecture for power aware routing to decouple the sources from sinks in a sensor network. This paper defines special type of nodes called core nodes where data and query aggregation takes place. Maintaining the clusters is of major concern in this work.Braginsky et al., [3] present a scheme called Rumor routing which allows for queries to be delivered to events in the network. Rumor routing is intended to fill the gap between query flooding and event flooding. Shiwen Mao et al., [4] propose an edge-based routing protocol called Beamstar Algorithm which exploits the properties associated with directional antenna and power control at the base station. It is designed so that each sensor node can determine the location information passively with minimum control overhead. Further, the concept is extended to multiple base stations, where in, upon an event a sensor node simply chooses the closest base station to send the request.

Long Gan et al., [5] use amorphous computing for pattern formation, to tackle the energy efficient routing problem in sensor networks. An autonomous data agent carries data packet sent from a sensor node and make its own routing decision based on local information. The data agent considers both the routing cost and the remaining energy of the node to choose the next hop. This work does not guarantee successful delivery of data packets and achieveing an optimal solution. Hong Luo et al., [6] propose a routing algorithm called Minimum Fusion Steiner Tree for energy efficient data gathering with aggregation in WSN. This algorithm not only optimises the data transmission cost, but also incorporates the fusion which reduces the amount of data resulting in significant energy savings. The disavantage of this paper is that it fails to detect situations where fusion is more energy consuming than forwarding the whole amount of incoming data as it is.

Sabbineni et al., [7] present new dissemination protocol for WSN. This protocol uses location information to reduce redundant transmissions, there by saving energy. It discuss virtual grid formation to achieve location aided flooding. The limitation of this work is because of its uniform grid sizes. Mathew et al., [8] propose bootstrapping as a possible phase for energy which aims at saving energy by reducing the number of collisions and turning off radio. The diadvantage of this paper is that it requires nodes to be highly synchronized.Weifa Liang et al., [9] present an Maximizing Network Lifetime (MNL) algorithm for energy efficient routing of data in sensor network having single base station. MNL is considered as the benchmark for evaluating MTC algorithm. MNL algorithm reduces the energy consumption of node but does not relate to the distance of a node from base station.

III. PROBLEM FORMULATION

A. Problem Statement

Given a surveillance area as two dimensional space XxY, consisting of $N(S)$ sensors and $N(BS)$ base stations, the objectives are to

(i) Find optimum number of base stations.
(ii) Locate an optimal placement of base stations.
(iii) Maximize the number of queries satisfied by network.

B. Assumptions

In our formulation, we make a number of simplifying assumptions.

(i) Each node is aware of its location and the boundaries of the surveillance region through location services.
(ii) Base stations are resistant to computational failures.
(iii) Nodes are adaptive to environment and dies only because of draining power.
(iv) Sensors are static and has identical sensing and communication range.

Finding optimum number of base stations can be considered as a maximization problem. Every base station when added to the network maximizes the lifetime of a network by reducing the distance between any node and the base station but delays the action to be taken on the network and incurs communication overhead. This overhead forces the base stations to have an agreement policy for stability of the network. Number of base stations considered would be optimum iff adding another base station to the network results in more overhead than increasing the efficiency of the network.

Placing a base station at an optimal position in the border of surveillance area reduces the distance between the base station and a sensor node. For multiple base stations we propose optimal placement positions which result in reduction of the distance between a base station and the farthest node. Since most of the simulations consider surveillance area to be square, we suggest two models for the placement of base stations in next section. The third objective is formulated as the problem of reduction in the energy consumption of a node to send the sensed data. The model for energy consumption for a node based on its distance from the nearest base station and amount of data to be sent is proposed in the later part of this work.

C. Base Station Model

It defines different possibilies for placing base station so that the distance from any node to the base station is less. When two base stations are used we find that the optimal placement of base stations is on opposite edges of surveillance area as shown in Fig. 1. We formulate the distance between a node and a base station using euclidian distance. As shown in Fig. 1., there are two base stations and many sensor nodes in surveillance area.

If base station grid position is BS_{Xj} and BS_{Yj} for j^{th} base station, then for any node i placed in the co-ordinate position S_{Xi} and S_{Yi} of a surveillance area of length and breadth l, we prove the distance to any node from base station is minimum for both two base station model and four base

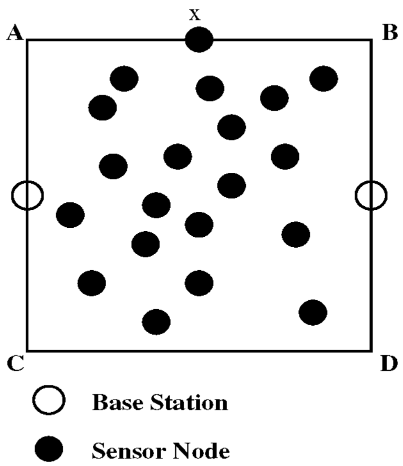

Fig. 1. Optimal placement of base stations in 2 - Base Station Model

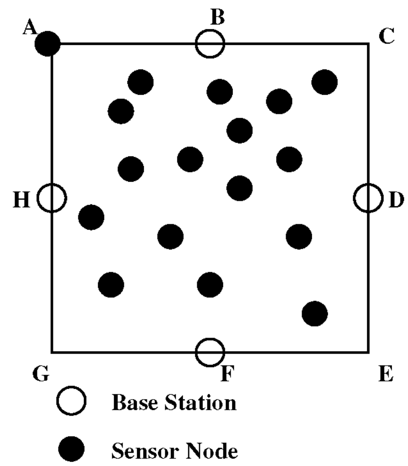

Fig. 2. Optimal placement of base stations in 4 - Base Station Model

station model considering optimal placement for base stations.

Theorem 1: If there are two base stations placed symmetrically and optimally, then the distance to every node from any base station is given by,

$$\sqrt{(BS_{X1} - S_{Xi})^2 + (BS_{Y1} - S_{Yi})^2} \leq \frac{l}{\sqrt{2}} \quad (1)$$

OR

$$\sqrt{(BS_{X2} - S_{Xi})^2 + (BS_{Y2} - S_{Yi})^2} \leq \frac{l}{\sqrt{2}} \quad (2)$$

Proof: We have the first base station placed at the center of the edge AC and the second base station at the center of the edge BD with the cordinates (BS_{X1}, BS_{Y1}) and (BS_{X2}, BS_{Y2}) respectively. Base stations are deployed symmetrically because sensor nodes are deployed uniformly in the surveillance area. The distance for any node from a base station increases as the node moves diagonally to base stations. As shown in Fig. 1., node x is the farthest node from both the base stations.

The distance between node x and the first base station is calculated as follows. Since the node x is placed at the center of AD,

$$Ax = \frac{l}{2} \quad (3)$$

Even the distance between the base station and A is $\frac{l}{2}$. So the euclidian distance from base station to the node x is given by,

$$\sqrt{(BS_{X1} - S_{X1})^2 + (BS_{y1} - S_{Y1})^2} = BS_1x \quad (4)$$

BS_1x is the distance from base station to node x and is computed as follows,
From the right angled triangle ABS_1x

$$BS_1x^2 = Ax^2 + ABS_1^2 \quad (5)$$

$$BS_1x^2 = (\frac{l}{2})^2 + (\frac{l}{2})^2 \quad (6)$$

$$BS_1x = \frac{l}{\sqrt{2}} \quad (7)$$

Substituting value of Equation. 7. in Equation. 4. we get,

$$\sqrt{(BS_{X1} - S_{X1})^2 + (BS_{Y1} - S_{Y1})^2)} = \frac{l}{\sqrt{2}} \quad (8)$$

Since the node x is the farthest node from the base station, the distance from base station to any other node in the surveillance area is,

$$\sqrt{(BS_{X1} - S_{Xi})^2 + (BS_{Y1} - S_{Yi})^2} \leq \frac{l}{\sqrt{2}} \quad (9)$$

Because of the symmetry we have a similar equation for BS_2

$$\sqrt{(BS_{X2} - S_{Xi})^2 + (BS_{Y2} - S_{Yi})^2} \leq \frac{l}{\sqrt{2}} \quad (10)$$

■

When four base stations are used, the optimal placement of the base stations are at the center of each boundary line of the surviellance area. Fig. 2. indicates the placement of base stations.

Theorem 2: If there are four base stations placed symmetrically and optimally, then the distance to every node from any base station is given by,

$$\sqrt{(BS_{Xj} - S_{Xi})^2 + (BS_{Yj} - S_{Yi})^2} \leq \frac{l}{2} \quad (11)$$

where j = 1,2,3,4.

Proof: Each of the four base stations are placed at the center of each boundary edge of the surveillance area as shown in Fig. 2. A node when placed in the center of the square or at any corner becomes the farthest node. In either case the distance is evaluated as,

$$\sqrt{(BS_{X1} - S_{X1})^2 + (BS_{Y1} - S_{Y1})^2} = BS_1A \quad (12)$$

Since node is at A and we have base station BS_1 placed exactly at the middle of AG we have,

$$BS_1A = \frac{l}{2} \tag{13}$$

hence the distance from the base station to the farthest node is

$$\sqrt{(BS_{X1} - S_{X1})^2 + (BS_{Y1} - S_{Y1})^2} = \frac{l}{2} \tag{14}$$

For any other node in the surveillance area, the distance will be,

$$\sqrt{(BS_{X1} - S_{Xi})^2 + (BS_{Y1} - S_{Yi})^2} \leq \frac{l}{2} \tag{15}$$

Because of the symmetry, the distance to any node from any base station will be

$$\sqrt{(BS_{Xj} - S_{Xi})^2 + (BS_{Yj} - S_{Yi})^2} \leq \frac{l}{2} \tag{16}$$

for j = 1,2,3,4.

■

D. Energy Consumption Analysis

When there is a query each node will have some data to be transmitted to the base station. For all nodes in the network, amount of data transmitted η depends on the sensed data at the node as well as aggregation of the data transmitted by its children. The transmission cost is proportional to amount of data and the distance. The energy spent by a node is given by the equation,

$$ES(v) = TE(v) + SE(v) \tag{17}$$

where $ES(v)$ is the energy spent by a node v, $TE(v)$ and $SE(v)$ are the energy spent by the node for transmission and sensing. $TE(v)$ is directly proportional to square of the distance between a node to its parent. $TE(v)$ is given by the equation,

$$TE(v) = \eta * d^2 \tag{18}$$

$$\eta = \eta_0 + \eta_1 + \eta_2 + ... + \eta_m \tag{19}$$

where η_0 is the amount of sensed data in bytes for the current node and η_1 to η_m are the amounts of data transmitted by its m-children respectively. This work aims at reducing the distance from a node to its parent d by adding more base stations as explained in the previous section.

IV. ALGORITHM

We propose an algorithm for Multiple Tree Construction in the surveillance area keeping the base stations as the root nodes. There is an edge between two nodes when they both are in communication range of each other. Since multiple base stations are used, we get different trees for each base station. MTC algrithm has three phases: Node Deployment, Query Issue and Tree Construction. The algorithm works in distributed fashion.

First phase of the algorithm deals with deployment of nodes in surveillance area and base stations in the border. Second phase of the algorithm discusses the way base station queries a sensor node. The algorithm is designed for time driven network. In the third phase of algorithm the routing path is found by sensor nodes.

A. Node Deployment

Deployment of nodes and base stations varies the performance of the algorithm to a large extent. The random deployment of nodes affects tree construction. We suggested a method for the placement of base stations in the previous section. Random deployment of nodes makes analysis complex, we reduce the complexity by making the deployment uniform over surveillance area. From $Theorem1$ we know that having two base stations optimally placed restricts the distance between the nodes and base station to be $\frac{l}{\sqrt{2}}$ and it further reduces to $\frac{l}{2}$ when four base stations are optimally placed as in $Theorem2$. We simulate the performance of the algorithm for different placements of the base stations.

B. Query Issue

Query is considered as message sent by a base station requesting sensed data from every node. The request message is broadcasted to every node in the network. Every node replies to the query by sending the sensed data. Multiple request message conflicts are resolved by the timestamp used in the request message. The time gap between the consecutive queries should be sufficiently large so that the nodes have replied for the previous requests. The time interval between two consecutive queries t_q is given by,

$$t_q = 2 * (h_l * t_h + t_c) \tag{20}$$

where t_h is time required by a node to send data per hop. This value remains constant because the transmission range for each node in a network is same. h_l is number of hops from a farthest node to the base station. t_c is computation time required for constructing a tree. We have a multiplication by a constant term 2 since the request has to reach the farthest node first and then the data is sent.

C. Tree Construction

Routing the sensed data to the base station holds the key for the performance of MTC algorithm. Constructing a tree gives a single and optimal path for routing the sensed data. After the first two phases of the algorithm every node is ready with the data that has to be sent to one of the base stations. Before sending the data, every node has to decide on the base station to which data has to be sent and the path to be followed.

Initially all the nodes will be in a set called Not_in_Tree. For each base station a tree is constructed taking the base

TABLE I
ALGORITHM: MULTIPLE TREE CONSTRUCTION (MTC)

```
tree_construction()
 begin
 NodeDeployment (Uniform)
 if(∀vRE(v) > 0)
   begin
   waits till QueryIssue
   Intially all nodes in Not_in_Tree set
   for every node v in Not_in_Tree set
     begin
     for every base station j in the network
       begin
       for every neighbor i which belong to the tree
       having base station j as root
         begin
         Simulate by transmitting data through
         node i
         Update residual energy of all nodes in the
         path
         Find most sensitive node (having minimum
         residual energy) in path through i
         end
       LSN = Maximum_Energy{all sensitive
       nodes}
       end
     LSN_max = Maximum_Energy{all LSN}
     add the node to correspondng Base Station set
     end
   end
 end
```

station as the root. For each node v in Not_in_Tree set find all the neighbors already in each of the partially constructed trees. For each such neighbor i simulate the act of data transmission assuming path via i. Upadate the residual energy of the nodes in the path as it will be after data transmission is done. Find the most sensitive node having least residual energy in the path. Among all such sensitive nodes select the node having maximum residual energy, call it as LSN (Least Sensitive Node). Likewise we get a LSN for each tree having base station j as root. Among all these LSNs choose one having maximum energy LSN_{max}. Now node v is added to the corresponding tree in which LSN_{max} is present and attached to the neighbor which is the leaf node in the path containing LSN_{max}.

V. IMPLEMENTATION

In order to implement MTC algorithm we have designed two messages; $Neighbor$ message and $Request$ message. Both messages are broadcasted in the network at different phases of the algorithm.

$Neighbor$ message is used to build a neighbor table at every node since the algorithm is executed in distributed fashion. Every node just after the first phase of the algorithm sends a $Neighbor$ message giving its location information. This location information is further used by the algorithm to evaluate the distance between the nodes.

$Request$ message is initiated at the Base station when there is a query. This message contains a timestamp to avoid the conflict among multiple request messages. The timestamp is represented as (t, BS_i), where t is the sequence number for the $Request$ message and BS_i is considered as base station id of i^{th} base station. When any node in a network receives two messages, conflict is resolved as $(t_1, BS_i) < (t_2, BS_j)$ iff $t_1 < t_2$ or $t_1 = t_2$ and $BS_i < BS_j$. First condition would be true when a single base station sends more than one request and second condition would be true when multiple base stations send requests at the same time.

The message length of the message transmitted by a node depends on the length of the message sensed by itself and the length of the messages received by all its children. The message length takes a random number in the range 1 to 8 so the message length sensed by each node varies from 1KB to 8KB. Suppose the node has j children then the message length transmitted by the node to its parent is the sum of its sensed message length and the message lengths of all the messages received by all its children.

All the nodes in the network have same communication and sensing ranges. We deploy the nodes using a uniform bi-variate and placement of the base station is calculated analytically.

VI. PERFORMANCE ANALYSIS

We measure the performance of our algorithm based on the number of queries satisfied by the network. MTC algorithm is shown to perform better than MNL algorithm given in [9] through the simulation studies.

A. Simulation setup

The simulation is performed using ns-2.31 simulator on Fedora core 8.0. The surveillance area is considered to be $100m$X$100m$. We deploy nodes in numbers of 10, 20, 30, 40 and 50. The deployment is random and uniformly distributed. Two models are considered having two and four base stations respectively. Every node in a network is provided with 2x10^6 units of initial energy. Node failure occurs when node's residual energy has reached zero which in turn results in network failure. We run simulations for 20 times and take the result as the average value to nullify the effect of randomness.

B. Simulation results

We evaluate the performance of the network in terms of number of queries satisfied. In Fig. 4. we compare the results of MNL and MTC algorithm for random deployment of base stations. The improvement is consistent for different network sizes. The improvement shows that the MTC algorithm satisfies almost more than double the number of queires satisfied by MNL algorithm. The difference in number of queries satisfied by both the algorithms is huge and there is obvious improvement in the network lifetime.

In Fig. 5., MTC algorithm with four base stations is compared with MNL algorithm and improvement is also found to be better than MNL algrithm. We compare the results for different number of nodes deployed. Compared

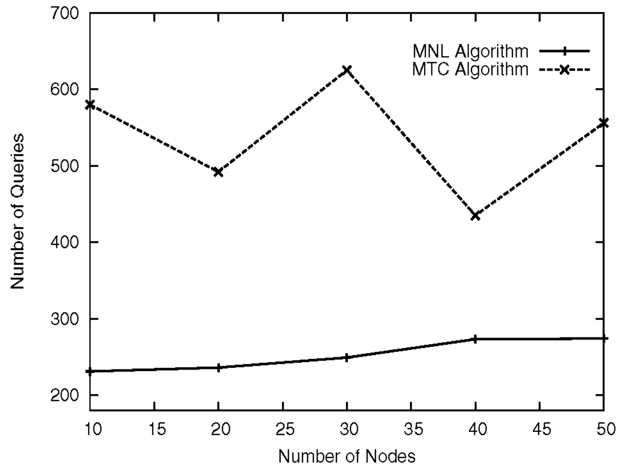

Fig. 3. MNL Algorithm Vs MTC Algorithm for Two Base Station Model

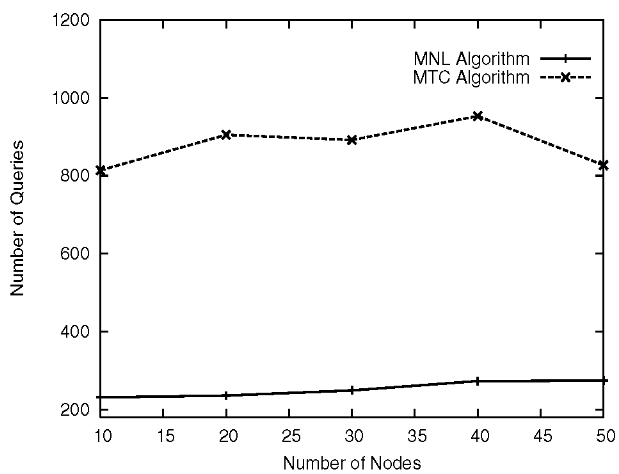

Fig. 4. MNL Algorithm Vs MTC Algorithm for Four Base Station Model

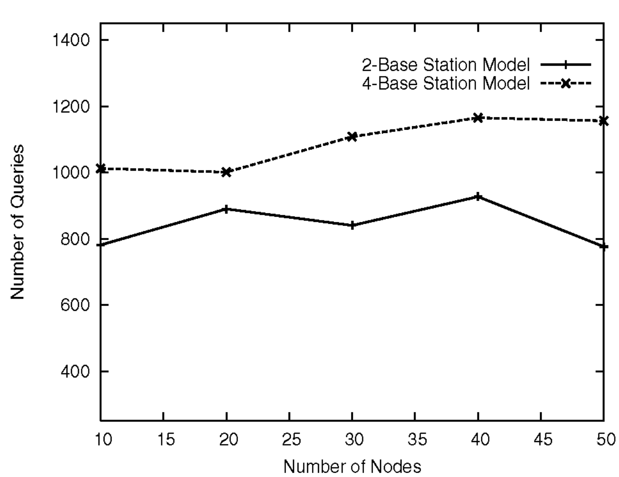

Fig. 5. Two Base Station Model Vs Four Base Station Model

to Two Base Station model, Four Base Station model works better. MTC algorithm has almost 800 queries satisfied for the same time as MNL algorithm satisfies around 230 queries.

Finally, we compare the two models described for different number of node deployment in Fig. 6. The lifetime is increased with large value for two base station model, further increases slightly for four base station model. We use the optimal values of both models for comparision.

VII. CONCLUSIONS

In this paper, we consider the deployment of multiple base stations to solve data gathering problem. We present the MTC algorithm as an energy efficient routing algorithm to transmit the sensed data to base station. We propose two and four base station models with optimal deployment. MTC algorithm's performance is compared with the MNL algorithm through extensive simulations and found that MTC outperforms MNL.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "Wireless Sensor Networks: A Survey," *IEEE Communication Magazine*, vol. 40, no. 8, pp. 102–114, 2002.

[2] Rishi Ranjan and Ashraf Saad, "Generic Architechture for Power Aware Routing in Wireless Sensor Networks," in *IEEE International Conference on Local Computer Networks*, pp. 575–576, 2004.

[3] David Braginsky and Deborah Estrin, "Rumor Routing Algorithm for Sensor Networks," in *ACM International Conference on Mobile Computing*, pp. 22–31, September 2002.

[4] Shiwen Mao and Y. Thomas Hou, "Beam Star: An Edge-Based Approach to Routing in Wireless Sensor Networks," in *IEEE Transactions on Mobile Computing*, pp. 1–13, 2007.

[5] Long Gan, Jiming Liu and Xiaolong Jin, "Agent-Based Energy Efficient Routing in Sensor Networks," in *ACM Transaction*, July 2004.

[6] Hong Luo, Yonghe Lin and Sajal K Das, "Routing Correlated Data with Fusion Cost in Wireless Sensor Networks," in *IEEE Transactions on Mobile Computing*, vol. 5, pp. 1620–1632, November 2006.

[7] Harshavardhan Sabbineni and Krishnendu Chakrabarty, "Location-Aided Flooding: An Energy-Efficient Data Dissemination Protocol for Wireless Sensor Networks," in *IEEE Transactions on Computers*, vol. 54, pp. 36–46, January 2005.

[8] Rajesh Mathew, Mohemed Younis and Sameh M. Elsharkawy, "Energy-Efficient Bootstrapping for Wireless Sensor Networks," in *Innovations System Software Engineering*, 2005.

[9] Weifa Liang and Yuzhen Lin, "Online Data Gathering for Maximizing Network Lifetime in Sensor Networks," in *IEEE Transactions on Mobile Computing*, vol. 6, pp. 2–11, January 2007.