INFORMATION
TECHNOLOGY JOURNAL

Traffic Aware Relay Based Routing and Deployment Algorithms in Wireless Sensor Networks

Bin Zeng, Lu Yao and Rui Wang
Research Center of Information Management, Naval University of Engineering,
Wuhan, Hubei, 430033, China

Abstract: For energy efficient routing scheme in the wireless sensor networks, it is important to use sub-optimal paths supported by relay nodes to increase the survivability and lifetime of networks. However, through routing algorithms play an important role, they will not be able to make good performance without an appropriate deployment of relays. First, the co-design problem, relay routing problem and relay deployment problem are formulated together with the end to end network energy consumption. Second, an optimal routing Tree algorithm considering network traffic, are proposed in order to significantly reduce the network energy consumption through the efficient utilization of the deployed relay nodes. Third, the relay deployment problem is considered to find the optimal locations for a given number. A traffic-aware relay deployment algorithm is proposed by adopting the network routing and traffic information. Our algorithms has been evaluated through a series of simulations and compared with state-of-the-art approaches. The results show that they lead to significant improvement on the network energy consumption over the existing traffic-unaware strategies.

Key words: Sensor networks, routing algorithm, relay node deployment, energy efficient algorithms

INTRODUCTION

There has been significant research interest by using higher energy relay nodes as cluster heads in wireless sensor networks in order to achieve balanced data gathering, reduction of transmission range and fault tolerance. In such networks, sensor nodes lie in the lower tier and transmit their data to their respective relays or cluster heads. Therefore, it is not necessary for sensor nodes to route the packets which reduces the energy consumption of the nodes. Each relay node is responsible to route the data from the sensor nodes to the sink as a cluster head. Furthermore, a relay node can not only forward the data from the sensor nodes in the same cluster but also forward data transferred from other relay nodes by using multi-hop communication paths.

In the last few years, the research of data communication in sensor networks has begun to focus on the deployment and routing based on relay nodes. A greedy geographic and energy-aware routing protocol (Chen *et al.*, 2009) is designed to reduce the energy consumption of relay communication without increasing the energy cost of relay nodes. A tree-based routing protocol called SN-MPR (Faheem and Boudjit, 2010) is proposed to reduce energy consumption of the network without affecting data dissemination efficiency. In MERR (Minimum Energy Relay Routing) (Zimmerling *et al.*, 2007) relay nodes are used to route data generated from sensor nodes to the sink.

It is well known that the network performance is impacted not only by routing protocol but also by the network structure.

Therefore, some sensor node and relay node deployment methods are proposed to prolong the lifetime of sensor networks. To create a network backbone the most efficient way is to construct Minimum Spanning Tree (MST) out of sensor nodes. MST backbone involves large number of nodes that must remain active to ensure network connectivity. For example, an approximation algorithm is proposed to maintain at least a connection path between sensor nodes with the minimum number of relay nodes by Efrat *et al.* (2008). A computational fluid dynamics simulation model was established to decide how to place sensor in the greenhouse in order to increase the accuracy of information monitor for measuring the microclimate environment of greenhouse (Zhi *et al.*, 2009).

Bari *et al.* (2008) have analyzed the jointly impact of routing algorithms and relay node placement

Corresponding Author: Bin Zeng, Department of Management Engineering, Research Center of Information Managemen, Naval University of Engineering, No. 717, JieFang Road, Wuhan, Hubei, 430033, China

schemes on network performance in wireless sensor networks powered by ambient energy harvesting. However, existing deployment and routing algorithms are decoupled in the design process. Firstly the positions of the relay nodes are computed and then an appropriate routing protocol is developed based on this predetermined position information. In this paper we try to jointly optimize both placement and routing of relay nodes. An integrated Integer Linear Program (ILP) formulation (Feng *et al.*, 2009) is proposed to compute the minimum number of relay nodes by deploying them in suitable locations to satisfy the requirements of data communication. However, this algorithm can work well only in centralized approach and hundreds of sensors without considering the impact of message traffic flow. A Traffic-aware relay node deployment problem is transformed into Euclidean Steiner Minimum Tree problem (ESMT) and some algorithms are developed for discrete relay node assignment by Yuzhen and Weifa (2008). But the optimal solutions are only suitable for the simple case of one source node, both with single and multiple traffic flows without considering the impact of routing algorithms.

In this study, we will study the basic sensor relay deployment or selection problem. The goal is to design the methodology and algorithms for the relay deployment problem as well as its association with the relay based routing problem. In particular, our contributions are:

- We formulate the relay based routing and deployment problems by the performance metric of energy consumption
- We introduce a relay based routing algorithms including some advanced algorithms for optimal relay deployment
- We conduct experiments for performance of our integrated routing and deployment algorithms which increases the performance of data gathering

This study will not focus on the practical protocol such as joining and leaving clusters in the routing trees and how to form distributed algorithms. Our goal instead is to provide a well-formulated optimization problem and design some heuristics to solve it.

PROBLEM DESCRIPTION

The relay-based data gathering is implemented as a two-tier network. An upper tier network consists of a set of relays placed on an edge nodes network. The sensor

edge nodes network is a collection of sensor nodes connected by wireless links. Then a routing tree of relays is built with a routing algorithm.

We first present a model of the edge nodes network and the relay nodes network before the objective functions are defined. The edge nodes network is represented by an undirected graph G, consisting of a set of N nodes $G = \{N_1, N_2, \ldots n_N\}$. There are two parameters associated with link $l_{i,j}$ between nodes n_i and n_j, $d(i, j)$ for distance and $b(i, j)$ for data traffic. If there is no link $l_{i,j}$ of the edge nodes network between nodes n_i and n_j, let $d(i, j) = b(i, j) = 8$.

A relay nodes network consists of a set of R relays $\{R_1, R_2, \ldots, r_R\}$. These relays are to be placed on R nodes. A routing algorithm connects the relay nodes to a routing tree based on the criteria as minimal transmission energy consumption. An edge node is routed to a relay according to the same criteria. More specifically, a routing tree T can be defined by $\{X, Y, Z\}$. Here, X is a mapping function of relay deployment:

$$x(i) = \begin{cases} 1 & \text{if } n_i \text{ is a relay node} \\ 0 & \text{otherwise} \end{cases} \qquad (1)$$

Z is a function for each node to identify its immediate corresponding relay node in the tree:

$Z(i) = k$, if $X(k) = 1$ and n_k relays the data from n_i and Y defines which ENs network links are used when $n_{z(i)}$ relays the data from n_i:

$$y(i,u,v) = \begin{cases} 1 & \text{if } l_{u,v} \text{ is used for path from } n_{z(i)} \text{ to } n_i, \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

Now we define two basic metrics, the aggregate distance and the aggregate data traffic flow. For each node n_i, the end-to-end distance from node n_i to the sink n_s is:

$$d_A(i) = \underbrace{\sum_{u=1}^{N} \sum_{v=1}^{N} d(u,v) \cdot y(i,u,v)}_{d'} + \underbrace{d_A(z(i))}_{d''} \qquad (3)$$

where, d' is the distance from node n_i to relay node $n_{z(i)}$ and d'' is recursively defined as the distance from $n_{z(i)}$ to sink node n_s.

For node n_i, the data volume of delivering one message to its relay node $n_{z(i)}$ is:

$$l_A(i) = \sum_{u=1}^{N} \sum_{v=1}^{N} b(u,v) \cdot y(i,u,v) \qquad (4)$$

We use a simplified model (Yuzhen and Weifa, 2008) for the communication energy consumption. The energy

spent for transmission of an l-bit packet over distance d_A between node n_i and sink node n_s raised to the power of γ, we have:

$$E_A(i) = l_A \left(\alpha_{tx} + \varepsilon d_A(i)^\gamma \right) \qquad (5)$$

where, α_{tx} is the energy/bit consumed by the transmitter of electronics at node n_i and γ accounts for the energy dissipated in the transmit amplifier. The aggregate energy consumption is the sum of transmission energy cost of all nodes:

$$EA = \sum_{i=1}^{N} E_A(i) \qquad (6)$$

In general, a routing tree T defined by {X,Y,Z} is relevant to both the relay deployment and routing, where X defines the relay deployment and Z and Y define the routing. Therefore, the problem can be formulated as:

- Routing Problem. Given the relay deployment X, design a routing algorithm to construct the routing tree RT, Z and Y, subject to minimizing energy consumption EA
- Deployment problem. By limiting the number of relays to be deployed, design a deployment algorithm to generate a mapping function X, subject to optimizing EA with a routing algorithm RT

Two design problems have been formulated for a two-tier sensor network, the relay-based routing problem and the relay deployment problem, together with the important performance metrics of energy consumption. Solving deployment problem usually is more difficult than routing problem. Routing algorithms are very important role, but without an appropriate deployment of relays, the performance of routing algorithms will not make full use.

ROUTING ALGORITHM

Given placement X, a routing algorithm will build a routing tree RT for minimizing energy consumption of nodes. In order to minimize EA, a sub-graph with all relay nodes is constructed by utilizing the shortest path among them. Relay-based Energy Efficient algorithm (RBEE) routes relays one by one as follows, starting from the nearest relay to the sink. The first relay connects to the sink node through the shortest path. The following relays connect to the sink or a placed relay whichever minimizes EA. After the routing tree is established, each edge node connects to the relay node that generates the smallest EA value. The complexity of this algorithm is $O(RN^2)$.

Algorithm: Given X, generate Z and Y

Objective: minimize EA

Step 1: Apply the Dijkstra's algorithm to the entire network graph G to generate all-pair SPTS with: $sp(i,j) = sp(j,i)$ representing the shortest path between n_i and n_j, nextNodeOnSPT(i,j) representing the next node on the shortest path toward n_s

Step 2: Sort all relay nodes in ascending order of length of their shortest path towards n_s

reset mark[i] = false for all n_i and mark[s] = true

Step 3: Establish an init routing with all n_i following the SPT towards n_s compute the init value $EA_{cur}|Z(i) = s$

Step 4: For each node n_i with x(i) = 1, mark[i] = false, according to their order in step 2. Set u = i and Z(i) = q be an already routed relay such that gain(i,q)>gain(i,j), where mark[q] = mark[j] = true, gain(i,j) = $(EA_{new}|Z(i) = j)$-EA_{cur}

do v = nextNodeOnSPT(k_1, q);

Y(i, u, v) = 1; u = *v*,

  while (v≠q)

  update $EA_{cur}|Z(i) = s \rightarrow z(i) = q$

Step 5: For each node n_i with X(i) = 0

Let k_1 = i and Z(i) = q be a relay such that gain(i,q)>gain(i,j), where X[q] = x[j] = 1, gain(i,j) = $(EA_{new}|Z(i) = j)$-EA_{cur}

do v = nextNodeOnSPT(u, q);

Y(i, u, v) = 1; u = v,

  while (v≠q)

  update $EA_{cur}|Z(i) = s \rightarrow z(i) = q$

Usually, routing algorithms are applied after relays are deployed. Sometimes, routing algorithms can help the decision of relay deployment, as discussed in the following section.

RELAY DEPLOYMENT ALGORITHM

A relay deployment algorithm finds the optimal locations for relays. The objective functions may be an optimization of energy consumption, network lifework, connectivity or coverage. In general, the deployment problem can be reduced to an NP-hard problem (Pourahmadi *et al.*, 2011). Thus, heuristics are used for the deployment algorithms. Almost all traditional deployment algorithms are traffic independent, therefore, relays are not reassigned or relocated when the traffic pattern changes. Therefore, their performance is not satisfactory in general. Our algorithms, including OPT-EA and IRD, produce good performance, but depend on the traffic pattern. Traffic-aware algorithms can work well when the traffic pattern can be predicted in advance and keep relatively stable. It also can help to relocate the position of relays when they needed to be redeployed or moved since traffic changes much.

Opt-EA algorithm: The traffic-aware deployment algorithms find a better position of relays by utilizing the routing information. First, a routing algorithm is applied to produce the routing paths by building a routing tree assuming all nodes are relay nodes. Once we have the routing tree, traffic T of a node is defined as data volume

passing through the node plus the traffic transmitted at the node assuming no relay is deployed. If T of a node is more than 1, it cab be seen as a branching node. The number of branching nodes $M \leq N$. If the number of relays to be placed $R = M$, there is no relayed message on every link. When $R < M$, a deployment algorithm selects nodes to place R relays.

The thought of the algorithm is described as follow. Since the relay is the higher-power node and even rechargeable, in order to maximize the network lifetime, the relays should be deployed into some places with high energy consumption. Transmission energy consumption is dependent upon the product of the node traffic and the distance between the node n_i and its assigned relay $n_{z(i)}$ The first relay is placed on the node with the highest energy consumption. Then the value of $p(i)$ is recomputed for each edge node and the next relay is placed on the node with the highest energy consumption. This process continues until R relays are placed. The complexity of this algorithm is $O(N^2)+O(routing)$.

Algorithm: For a routing algorithm, output deployment X

Step 1: Set $X(i) = 1$ for all i, apply routing algorithm to generate a initial routing tree with nextNodeOnTree(i,Z(i)) representing the next node on the path towards $n_{z(i)}$.
Step 2: Set $X(i) = 0$ for all i and $X(s) = 1$
Step 3: For $j = 0$ to R, set $p(i) = 0$ for all i
Step 4: For each n_i, set $u = i$
    do $v = nextNodeOnTree(i,Z(i))$
    if $X(i) = 1$ then $p(v) = p(v)+1$
    $u = v$
    while $(v \neq p(i))$
Step 5: find a node N such that $p(N)*EA(N) > p(j)*EA(j)$, where $X(N) = X(j) = 0$
Step 6: set $X(N) = 1$

Traffic-aware deployment algorithms utilize traffic information. Therefore, they are able to produce better relay positions. However, the deployment will change when the traffic pattern changes.

Incremental relay deployment algorithm: Incremental Relay Deployment (IRD) algorithm is designed for a near optimal solution when exhaustive search for the optimal solution is not feasible. It is often used for a baseline measurement.

Suppose R relays are to be deployed. We choose one location at a time. In the first iteration, we assume a relay is to be deployed on each of the N locations and evaluate their objective functions, respectively. The location that can produce the best value of the objective function is selected to place the first relay. In the second iteration, we search for the next best location, considering both the sink and the relays already deployed, produces the smallest value of the objective function. This process continues until all R relays are deployed. Its complexity is $O(N^2) \cdot O(routing)$ too.

Algorithm: For a routing algorithm, output deployment X

Priority: Minimize EA
Step 1: Set $X(i) = 0$ for all i, applying routing algorithm to generate a initial routing tree and compute the initial value of $EA_{cur}|Z(i) = s$
Step 2: Set $X(s) = 1$
Step 3: For $i = 0$ to R
Set n_k be a node with $X(k) = 0$
such that gain(k)>gain(j) where $X(k) = X(j) = 0$ and
$gain(i,j) = EA_{cur}-EA_{new}|X(j) = 1$ by applying routing algorithm
Set $X(k) = 1$
update $EA_{cur}|Z(k) = 0 \rightarrow z(k) = 1$

SIMULATION RESULTS

In this section, relay deployment algorithms with different routing strategies are evaluated by NS2. We deploy $N = 3000$ edge sensor nodes by uniform distribution in a field of 10000×10000 m with the sink centered in the deployment area. The data rate of each sensor node is randomly picked from (0, 10 kbps] in each time period. We use 5 topologies in our simulation study generated by TopoGenerator. The number of relays ranges R from 0 to 100. α_{tx} is set to be 50e-9(J/bit), ϵ is set to be 10e-12 (J/bit/m^2), γ is set to be 2.

Comparison of routing algorithm: The performance of our routing algorithm (RBEE) is compared to the following two approaches; the first algorithm produces minimum degree spanning tree (MDST) as the routing tree such that the total energy consumption is minimized. The second commonly built routing tree is Shortest Path Tree (SPT). In SPT every node sends data along the shortest path to the sink node. To build a SPT every node selects the neighbor relay as the return path that minimizes the routing cost to the sink.

We first evaluated routing trees built by various routing algorithms, in terms of the performance metrics of energy consumption (EA) of all the edge nodes. We selected two deployment algorithms, K-MEANS and IRD, to generate placement for this experiment. Three routing algorithms, SPT, MDST and RBEE, are evaluated.

Figure 1 and 2 illustrates the energy consumption for the given K-MEANS and IRD deployment, respectively. Evidently, the IRD deployment has produced a better placement than the K-MEANS deployment. With a good relay deployment, the difference between three routing algorithms is not significant. The difference of the energy consumption value is within 8%. Given a Random placement, good routing algorithms can help to some extent. RBEE results in lower energy consumption, particularly for the K-MEANS deployment, where it consumes up to 20% less energy than MDST does. The energy consumption of RBEE routing is only slightly lower than SPT routing. RBEE exhibits a much better performance than MDST. The difference of the energy consumption can be as large as 40%. RBEE is the best in
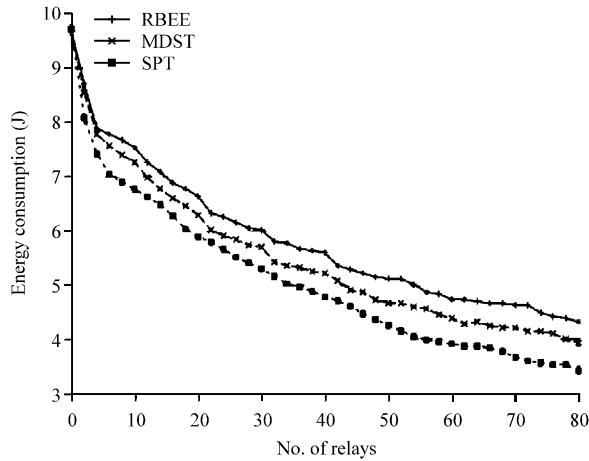
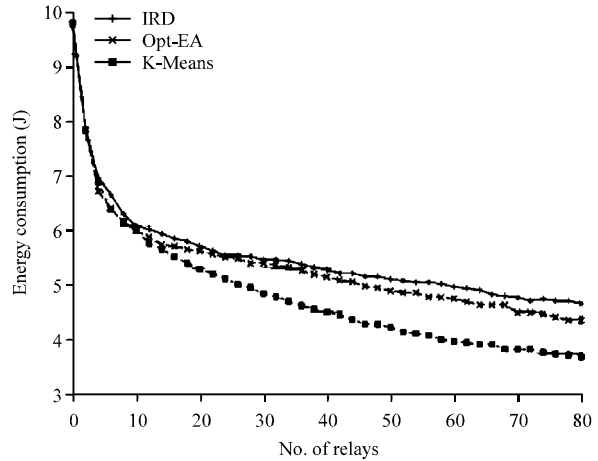Fig. 1: Energy consumption comparison with K-MEANS
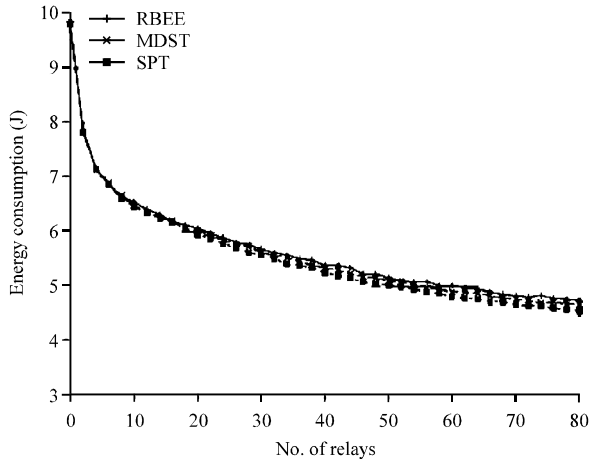deploy

Fig. 2: Energy consumption comparison with IRD deploy

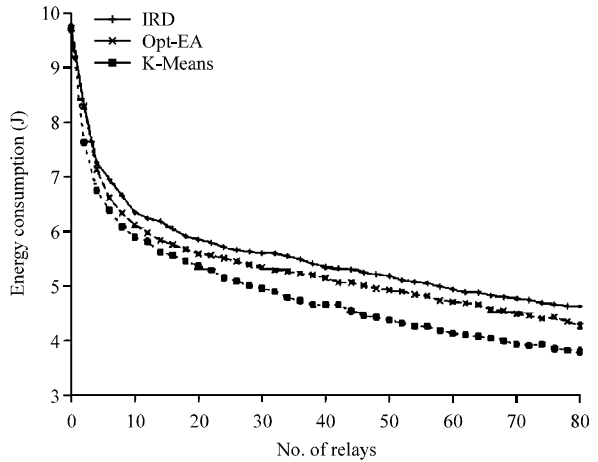Fig. 3: Energy consumption comparison with SPT route

Fig. 4: Energy consumption comparison with RBEE route

terms of its energy consumption which is always the lowest. However, it is the slowest algorithm among the three routing algorithms since it re-computes the EA metric with every iteration.

Comparison of deployment algorithm: Here, the deployment structure resulting from two deployment algorithms are compared. The first approach called by Lifetime Oriented Deployment (LOD) tries to solve an optimization problem to achieve a maximal lifetime without any restriction on the locations of relay nodes. The second random algorithm (K-MEANS) determines relays' locations by the K-means algorithm.

After the relay placement is generated, two routing algorithms, SPT and RBEE, are used to construct the routing trees. The energy consumption is shown in Fig. 3 and 4. Most deployment algorithms significantly reduce

the energy consumption when the first ten relays are placed. The energy consumption with ten relays is about two times that of 100 relays. The IRD deployment produces the best performance, as expected. The K-MEANS deployment does not give satisfactory results when SPT routing is used. However, traffic-aware routing can improve the performance of the K-MEANS deployment, but the energy consumption of the K-MEANS deployment is about 40% larger than that of the IRD deployment. Again, the energy consumption of Opt-EA is very close to that of the IRD algorithm. Since Opt-EA is faster, it can serve in place of the IRD algorithm. The LOD deployment algorithm exhibits higher energy consumption when a few relays are deployed and it outperforms the K-Means deployment algorithm as more relays are placed.
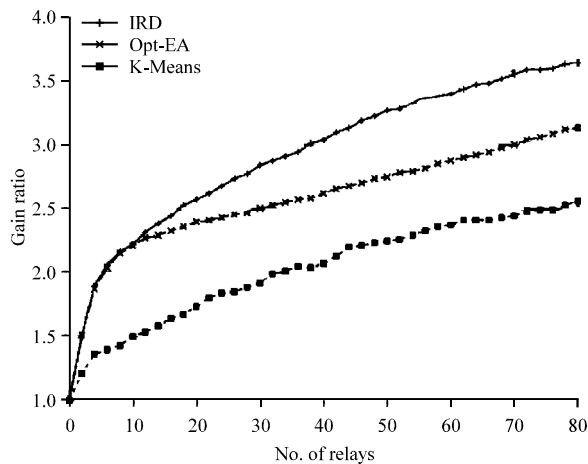
Fig. 5: Gain ratio with RBEE route

The benefit of placing relays can be measured by the metric of gain ratio which is defined as the ratio of the energy consumption with no relay deployed to the energy consumption with R relays deployed. Figure 5 illustrates the gain of various deployment algorithms with RBEE routing. The gain of K-MEANS deployment is very low, especially for the RBEE routing. The heuristic deployment such as Opt-EA performs well for a few relays, but not as good as LOD deployment algorithms for more relays. IRD and Opt-EA produce the highest gains.

Most existing schemes of data gathering in wireless sensor networks use heuristic deployment with SPT routing. It is a practical method and simple to implement. However, comparing IRD deployment with RBEE routing, the energy consumption of heuristic deployment with SPT can be 60% larger. Furthermore, many two-tier sensor networks do not implement true SPT routing due to practical considerations, so the aggregate distance can be very larger. With a good relay deployment, even the simplest SPT routing can perform well.

CONCLUSION

This study addressed the relay deployment problem in sensor networks. Firstly the problem was formulated and objective functions were proposed. Based on the performance metric of energy consumption, a heuristic routing and deployment algorithms have been presented. The performance study showed that the relay deployment has more impact than routing on network performance. From our study, the results below can be obtained:

- With a good relay deployment, the performance difference between routing algorithms is not significant. Within a cluster, an optimal deployment can be practical even by using the simple SPT routing
- From our experiment, the performance of random deployment is hardly acceptable. However, in some situations where relays cannot be deployed at proper locations, such as that in an bad environment, a good routing algorithm such as RBEE can improve the performance
- The IRD deployment algorithm is the best and Opt-EA is very close to IRD. A combination of Opt-EA deployment integrated with RBEE routing is able to give satisfactory performance, with relatively low complexity

There still exist more problems to be addressed in this area. We are designing the distributed versions of routing and deployment algorithms which may make near-optimal performance.

ACKNOWLEDGMENTS

REFERENCES

Bari, A., X. Yufei and A. Jaekel, 2008. Integrated placement and routing of relay nodes for fault-tolerant hierarchical sensor networks. Proceedings of 17th International Conference on Computer Communications and Networks, August 3-7, 2008, St. Thomas, US Virgin Islands, pp: 1-6.

Chen, G., C. Li, M. Ye and J. Wu, 2009. An unequal cluster-based routing protocol in wireless sensor networks. J. Wireless Networks, 15: 141-161.

Efrat, A., S.P. Fekete, P.R. Gaddehosur, J.S. Mitchell, V. Polishchuk and J. Suomela, 2008. Improved approximation algorithms for relay placement. Proceedings of the 16th Annual European Symposium on Algorithms, September 15-17, 2008, Karlsruhe, Germany, pp: 356-367.

Faheem, Y. and S. Boudjit, 2010. SN-MPR: A multi-point relay based routing protocol for wireless sensor networks. Proceedings of 2010 IEEE/ACM International Conference on Cyber, Physical and Social Computing, December 18-20, 2010, Hangzhou, pp: 761-767.

Feng, W., W. Dan and L. Jiangchuan, 2009. Traffic-aware relay node deployment for data collection in wireless sensor networks. Proceedings of 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, June 22-26, 2009, Rome, Italy, pp: 12-18.

Pourahmadi, V., S. Fashandi, A. Saleh and A.K. Khandani, 2011. Relay placement in wireless networks: A study of the underlying tradeoffs. IEEE Trans. Wireless Commun., 10: 1383-1388.

Yuzhen, L. and L. Weifa, 2008. Energy-efficient multiple routing trees for aggregate query evaluation in sensor networks. Proceedings of 6th International Conference on Wired/wireless Internet Communications, May 28-30, 2008, Tampere, Finland, pp: 201-212.

Zhi, A.E., T. Hwee-Pink and W.K.G. Seah, 2009. Routing and relay node placement in wireless sensor networks powered by ambient energy harvesting. Proceedings of IEEE Wireless Communications and Networking Conference, April 5-8, 2009, Budapest, Hungary, pp: 1-6.

Zimmerling, M., W. Dargie and J.M. Reason, 2007. Energy-efficient routing in linear wireless sensor networks. Proceedings of IEEE International Conference on Mobile Adhoc and Sensor Systems, October 8-11, 2007, Pisa, pp: 1-3.