AN INNOVATIVE APPROACH TO IMPROVE LIFETIME IN SENSOR NETWORKS USING OPTIMAL NUMBER OF SINKS

[1]Santhosh B D, [2]C. P. Sameerana, [3]Adarsha Sagar H V [4]Muniraju M
[1]Asst Professor, dept of CSE ,APSCE Bangalore[2]Associate Professor, HOD, Dept of CSE, APSCE Bangalore, [3]Asst Professor, dept of CSE [4]Asst Professor, dept of CSE Bangalore APSCE Bangalore
Email: [1]santhoshd8@gmail.com, [2]hodcse.apsce@gmail.com, [3]adarsha.sagar183@gmail.com, [4]m.raju077@gmail.com

Abstract—In this paper we investigate the benefits of placing optimal number of sinks for a wireless sensor network (WSN) to prolong the network lifetime, provided that the number of hops from each sensor to its nearest sink is no more than $h \geq 1$ and the sink location space is given in advance. We formulate this problem as a joint optimization problem, which consists of finding the optimal number of sinks for placement and devising an energy-efficient routing protocol for data collection. Due to the NP-hardness of the problem, we propose a novel heuristic by decomposing the problem into two sub-problems and solving them separately. As a result, the proposed optimization framework improves network performance from several aspects, including the network lifetime prolongation, network scalability improvement, average data delivery delay reduction and also the network robustness. We apply the maximum flow algorithm to find the flow of data and to deliver a feasible solution. We finally conduct extensive experiments by simulations to evaluate the performance of the proposed algorithm. The experimental results demonstrate that the proposed algorithm outperforms another heuristic and increases the network lifetime.

Keywords: Multiple Sink, Wireless Sensor Networks, Energy efficiency, Routing, Graph theory.

I. INTRODUCTION

A sensor network is a static ad hoc network consisting of hundreds of sensor nodes deployed on the fly for unattended operation. Each sensor node is equipped with a sensing device, a low computational capacity processor, a short-range wireless transmitter-receiver and a limited battery-supplied energy. Sensor nodes monitor some surrounding environmental phenomenon including multi-media(*e.g.*, video, audio) and scalar data (*e.g.*, temperature, pressure, light, infrared) [1], [10] which forward sensed data towards a base station for processing which is located on the periphery of the sensor network. Base station(s) collect the data from the sensor nodes and transmit this data to some remote control station. Although there have been significant progress in sensor fabrications including processing design and computing, advances of battery technology still lag behind, making energy resource the fundamental constraint in WSNs. To maximize the network lifetime, energy conservation in such networks thus is of paramount importance.

In traditional sensor networks, there is a single static sink (also referred to as the base station) with unlimited power supply, which serves as a gateway between the network and users. The sink functionalities typically include gathering sensing data from sensors in the network via multi-hop relays, performing data processing, and responding to user queries. The sink is often placed in a strategic location in the monitoring region to enable the network to operate as long as possible. As sensors are usually powered by batteries, it is difficult or even impossible to replace or recharge them especially when the network is deployed in human inaccessible or hostile environments. Energy is consumed every time sensors send or receive packets. Once its onboard energy supply drains, a sensor becomes dysfunctional. Therefore, preserving the sensors' energy is a key for keeping the network operational for longer periods of time. Therefore, utilizing sensor energy efficiently to prolong the network lifetime has become the main research focus in this area.

Most previous studies in WSNs focused on the improvement of network performance assuming that there is a single stationary sink. This traditional single sink paradigm however suffers the following main drawbacks which degrade the network performance. One is the so called *single sink neighborhood problem*, where the sensors within one-hop distance from the sink have to relay the sensing data for the other sensors that cannot reach the sink directly. As a result, these sensors consume much more energy than the others. Once they deplete their energy, the network will be partitioned and the sink will be disconnected from the rest of sensors even if those sensors are still fully operational with sufficient residual energy. In particular, with the increase of network size, the single sink neighborhood problem becomes worse.

The other is the network connectivity issue. It is compulsory that the network consisting of the sink and sensors should be connected. Otherwise, the data generated by the sensors in a fragment different from the fragment in which the sink is located cannot be collected ultimately. However, in some sparse sensor deployment scenarios, it is difficult to ensure that each sensor and the sink are in the same fragment due to the restriction of physical obstacles or other geographic constraints. To cope with the single sink neighborhood problem, the sink multiplicity strategy has been exploited and demonstrated to improve various network performance including network lifetime [2], [9], average data delivery latency [14], and system throughput [13]. Under such a paradigm, multiple sinks are placed in the monitoring region with each being used to gather sensing data of the sensors within a certain number of hops from the sink. Consequently, the relay workload of each sensor will be decreased, the average data delivery delay will be shortened, and the network lifetime will be prolonged.

In this paper, we aim to find the optimal number of sinks and their locations in a monitoring region for data gathering such that the network lifetime is maximized, subject to the following two constraints. One is that each sink can only be placed at one of the given potential sink locations. In practice, the sink location is determined by several factors. For instance, it is inappropriate to place a sink at a barrier location that obstructs the wireless communication between sensors and the sink or at a water pond. Therefore, instead of assuming any location in the monitoring region is appropriate for sink placement, we only place the sinks to some anchor locations, which are referred to as the *potential sink locations*. Such information is usually given by users a priori.

The other is a given controllable parameter h, which is the upper bound on the number of hops from each sensor to its nearest sink. It quantifies the extent of multi-hop routing. The choice of h trade-offs between the network lifetime to be delivered and the number of sinks to be placed. A larger h may result in a smaller number of sinks but will cause more energy consumption on data relay. As a result a shorter network lifetime will follow. On the contrary, the ideal situation for maximizing the network lifetime is $h = 1$ since each sensor can transmit its data to a sink directly and there is no relay required. Apparently, such an improvement on network lifetime is at the cost of using a prohibitively large number of sinks if the monitoring region is large and the transmission range of sensors is small. This paper strives to find a fine trade-off between the network lifetime and the number of sinks needed to meet the prescribed constraints.

The main contributions of this paper are as follows. We first formulate a joint optimization problem with the objective to find the optimal number of sinks for placement and find a routing protocol for data gathering such that the network lifetime is maximized, under the constraints that the potential sink locations are pre-defined and the number of hops from each sensor to its nearest sink is no more than h. Due to its NP-hardness, we then propose a novel heuristic by decomposing the problem into two sub-problems: finding the optimal number of sinks for placement meeting the given h-hop constraint and devising a tree-based routing protocol for data collection to maximize the network lifetime. The proposed algorithm exhibits low computational complexity and high scalability. The proposed optimization framework improves network performance including the network lifetime prolongation, network scalability improvement, and the average data delivery delay reduction. Furthermore, it also enhances the network robustness substantially, since sensing data from all sensors can be collected by multiple sinks regardless of the network connectivity. We apply the maximum flow algorithm to find the flow of data and to deliver a feasible solution. We finally conduct extensive experiments by simulations to evaluate the performance of the proposed algorithm. The experimental results demonstrate that the proposed algorithm outperforms another heuristic significantly in terms of network lifetime prolongation.

The remainder of the paper is organized as follows. Section II provides the literature survey on multiple sink placement. Section III introduces the system model and defines the problem precisely. Section IV proposes a novel heuristic algorithm for the problem, and Section V evaluates the performance of the proposed algorithm through experimental simulations. Section VI concludes the paper.

II. RELATED WORK

Multiple sink deployment has been extensively studied in previous works and most of them focused on the network lifetime maximization, assuming that the number of sinks is given [2], [11], [6], [13]. For example, Bogdanov *et al.* in [2] considered the multiple sink placement problem under different data generation rates. They dealt with several special communication graphs by proposing heuristic algorithms. With the objective to maximize the network lifetime, they aimed to find the optimal positioning of multiple sinks through the network flow technique. Qiu *et al.* [13] discussed the sink placement problem in multi-hop wireless networks. They proposed two linear programming solutions by incorporating an interference model and a fault tolerance model into the problem formulation. The routing structures in both [2] and [13] are based on flow techniques while this paper develops a tree-based routing protocol for data gathering considering the difficulty of flow control at each sensor at each time instance in practice. Gandham *et al.* [6] solved how to place $k \geq 1$ mobile sinks to collect sensing data in a monitoring region. They partitioned the network lifetime into a number of rounds and proposed a ILP (integer linear program) model with the objective to minimize either the maximum energy consumption among the nodes or the total energy consumption within each round. Kim *et al.* [9] studied the problem by employing linear program and *k*-mean clustering techniques, under the constraints including the residual energy of sensors, the data generation rate, and potential sink locations. There are also studies on multiple sink placement with other optimization objectives. For example, Youssef *et al.* [14] considered minimizing the data delivery latency between a sensor and a sink. They proposed several genetic algorithms to find the best locations for sink positioning through minimizing the number of hops between each sensor and one of the sinks, assuming that the number of sinks is given. They formulated the problem as a clustering problem, where the number of clusters is equal to the number of sinks. Note that they assumed that the sensors and sinks in the network can communicate directly with each other regardless of their physical distance. Xu Xu and Weifa Liang[15] proposed a novel heuristic for multiple sink placing and used a BFS search.

To the best of our knowledge, we are not aware of prior works that jointly determine the optimal number of sinks for placement and devise the tree-based routing protocol for network lifetime maximization, subject to the given potential sink locations and the bounded number of hops from each sensor to its nearest sink. This paper will provide a joint optimization framework for this problem.

III. SYSTEM MODEL

In this paper, we consider a wireless sensor network $G(V \cup S, E)$ consisting of n stationary sensors, where V is the set of sensors, S is the set of potential sink locations, and E is the set of links. There is a link between two sensors, or a sensor and a sink if they are within the transmission range of each other. The locations of sensors are fixed and known a priori. Each sensor equipped with an omni-directional antenna has a fixed, identical transmission range. Assume that each sensor $vi \in V$ has identical data generation rate ra. We also assume that sinks have unlimited energy supplies and after being deployed to certain locations in the network, they gather data from sensors via tree-based routing structure. Without loss of generality, we only consider the energy consumption on data transmission and reception [12]. The network lifetime is defined as the time of the first sensor's failure due to the depletion of its energy [3].

The *h-hop constrained multiple sink placement problem* in a wireless sensor network *G*(*VUS*,*E*) is to place the optimal number of sinks at some locations in *S* such that the number of hops between each sensor and one of the sinks is no more than $h \geq 1$, meanwhile, under this sink deployment, the network lifetime is maximized.

Let $S = \{s_1, s_2, \ldots, s_{|S|}\}$ be the set of potential sink locations and its subset $S' \subseteq S$ be the set of chosen locations for sink placement, that is, the set of chosen sinks, where $k = |S'|$. Denote by *Ts* the tree rooted at sink $s \in S'$ and $dt_{T_s}(v_j)$ the number of descendants of *Vj* in *Ts*. Recall that the data generation rate of each sensor is *ra*, then the energy consumption of sensor *Vj* on wireless communication per time unit is

$eCT_{s=} r_a[(dtTs(vj)+1)et+dtTs(vj)er]$----1

where *et* and *er* are the amounts of energy consumption on transmitting and receiving a bit of data, assuming that there is no data aggregation at each relay node when proceeding data routing. It can be seen that the value of *ecTs* (*v*) is greatly related to the value of *dtTs* (*v*). Denote by *CT* (*s*) the set of children of *s* in *Ts* in which sensors are within the transmission range of *s*. Sensors in *CT* (*s*) have to relay data for remote sensors and will consume energy faster thus they are often referred to as the *bottleneck sensors*. The maximum energy consumption among the sensors in *Ts* per time unit is

$$\max_{s \in S'}\{ecTs(v)|v \in CT(s)\}$$

and the network lifetime thus is

$$L=\min_{s \in S'}\{IE/eCT_s(V) \text{ where } \mathbf{v \in C_T(s)}\}\text{----2}$$

where IE is the initial energy capacity of each sensor. To maximize the network lifetime is equivalent to minimize the maximum number of descendants of bottleneck sensors in the routing tree rooted at each $s \in S'$. In other words, the joint optimization problem is then to identify a $S' \subseteq S$ with the minimum cardinality $k = |S'|$ such that each sensor is no more than *h* hops to one of the sinks in S' and meanwhile the maximum number of descendants of bottleneck sensors in the routing tree rooted at each $s \in S'$ is minimized.

However, the *h*-hop constrained multiple sink placement problem is NP-hard. Considering one of its special cases where *h* = 1 and there is no restriction on the potential sink locations, the problem becomes the unit disk covering problem (UDCP) that aims to find the minimum number of disks to cover all sensors in the network [7]. Assume that the radius of the disk (sink) is identical to the transmission range of sensors and a sensor is covered by a disk if they are within the transmission range of each other. Since the decision version of the minimum disk covering problem is NP-complete [7], the problem of concern in this paper is NP-complete, too.

IV. HEURISTIC ALGORITHM

Due to the difficulty of jointly determining the optimal number of sinks and devising a routing protocol to maximize the network lifetime, in this section we propose a heuristic for it. We decompose the problem into two sub-problems: finding the optimal number of sinks and their locations such that each sensor can reach a sink with no more than *h* hops; and constructing a load-balanced forest to maximize the network lifetime, in which each sink is the root of a routing tree with the depth no more than *h*. Each sensor belongs to only one of these trees, i.e., each sensor can reach a sink with no more than *h* hops. Specifically, the heuristic first calculates the set of sensors covered by a sink at each potential location subject to the given *h*-hop constraint. It then identifies a subset of sinks and their locations with minimum cardinality, covering all sensors in the network. It finally constructs load-balanced routing trees rooted at each chosen sink for efficient data gathering such that the network lifetime can be maximized. We now describe the proposed heuristic in detail.

A. Placing the optimal number of sinks meeting the h-hop Constraint

Given a potential sink location $s \in S$, let $N1(s) = \{u \mid (u, s) \in E, u \in V\}$ be the set of neighboring sensors of sink s and $Nh(s)$ be the set of sensors within h hops from sink s. i.e., $Nh(s) = \{v \mid$ the number of hops from v to s is no greater than $h\}$. The calculation of $Nh(s)$ for each $s \in S$ is as follows. A Breadth-First-Search (BFS) tree rooted at s is constructed, which is expanded layer by layer. The expansion will terminate when it reaches layer h. The set of sensors contained in this BFS tree is referred to as $Nh(s)$

Let $C = \{ Nh(s) \mid s \in S\}$ be the collection of sets derived by the set S of potential sink locations. The problem of placing the optimal number of sinks at locations in S such that each sensor can reach one of the chosen sinks with no more than h hops is equivalent to finding a sub-collection $S' \subseteq S$ such that $|S'|$ is minimized and $\cup_{s\in S'} Nh(s)=V$. It is a set cover problem, which is NP-complete [4]. Instead, a greedy heuristic will be employed and it delivers an approximate solution to the problem with the approximation ratio of $O(\log B)$, where $B = \max_{s\in S'}\{|N_h(s)|\} \leq n$.

For convenience, a sensor v is referred to be *covered* by a sink s if the number of hops from v to s is no more than h; otherwise, v is *uncovered* by s. If a given sensor v cannot be covered by any sink in S', then the sensor is *uncovered*. The proposed algorithm proceeds iteratively. Initially, all sensors in V are *uncovered* and the set of chosen sinks S' is empty. The algorithm iteratively selects a sink s such that the set $Nh(s)$ from C covering as many uncovered sensors as possible. Once a set $Nh(s)$ is chosen, it will be removed from C. The sink s and its current location will be added to set S'. The algorithm continues until all sensors in V are covered by the sinks in S'. The detailed description of the proposed algorithm for finding the optimal number of sinks is given.

Algorithm Find_Optimal_Sink can be easily implemented in polynomial time of $n = |V|$, $m = |E|$, and $|S|$. Since the number of iterations of the proposed algorithm is bounded by $\min(n, |S|)$, and the loop body of the algorithm can be implemented in time $O(n|S|)$, the time complexity of algorithm Find_Optimal_Sink thus is $O(n^2|S| + n|S|^2 + m|S|)$.

B. Routing protocol design for data gathering

Having identified the set of chosen sinks S', we now devise an energy efficient tree-based routing protocol for data gathering to maximize the network lifetime. Following Eq. (2), to maximize the network lifetime is equivalent to

```
Find_Optimal_Sink(V, S)
begin
    U ← V;
    C ← {N_h(s) | s ∈ S};
    S' ← ∅; /* the set of chosen sinks */
    while U ≠ ∅ do
        select a set N_h(s) ∈ C such that
        |U ∩ N_h(s)| is maximized;
        S' ← S' ∪ {s};
        U ← U − N_h(s);
        C ← C − {N_h(s)};
    endwhile
    return S';
end
```

minimize the maximum energy consumption among bottleneck sensors, while the energy consumption of each bottleneck sensor is related to the number of its descendants in the routing tree rooted at a chosen sink. In other words, the optimization objective is to group sensors into different clusters headed at different sinks and make each sensor belong to one cluster only. For each cluster, a load balanced routing tree rooted at the cluster head (a sink) will be built, such that (i) the number of hops from each sensor to its tree root is no more than h; and (ii) the maximum number of descendants among the bottleneck sensors is minimized. We refer to this clustering problem as the *load-balanced forest problem*, which can be approximately solved by the following three steps.

1) Partition sensors into h disjoint subsets: The $k = |S'|$ chosen sinks are compressed into a *virtual node r*, and every neighbor of a chosen sink in the original network now becomes a neighbor of the virtual node. A BFS tree rooted at r in the modified network is then constructed and as a result, the sensors in the network are partitioned into h disjoint subsets, according to the number of hops of each sensor to r. Let Vi be the set of sensors in layer i, then $\bigcup_{i=1}^{h} V_i = V$ and $V_i \cap V_j = \emptyset$, where $1 \leq i, j \leq h, i \neq j$. Note that V0 contains only the root *r* and *V*1 contains only the bottleneck sensors.

2) Find a load-balanced tree: However, finding an optimal load-balanced tree has been shown NP-complete [8]. We here adopt a heuristic proposed in [8] for the load-balanced tree construction. The heuristic is a greedy algorithm, which expands the tree layer by layer in a top-down fashion. Assuming a partial load-balanced tree spanning the sensors from layer 0 to layer $1 \leq l < h$ has been constructed, we now expand the tree by including the nodes in layer *l* + 1 as follows.

We first construct a node-weighted bipartite graph *Gl* =
(*X*, *Y*,*El*, *w*), where the nodes in *Vl* are grouped into different
subsets according to their ancestors in *V*1, i.e., nodes in the
same subset are the descendants of the same node in *V*1. Let
$X = \{x_1, x_2, \ldots, x_{|X|}\} \subseteq V_1$ be the set of ancestors of the nodes in *Vl* that are incident to nodes in layer *l* + 1, and *Y* be the set of nodes in layer *l* + 1, i.e., *Y* = *Vl*+1. For each *x* ∈ *X*, its weight *w*(*x*) is the number of descendants of *x* in the current tree. And each node *y* ∈ *Y* is assigned a weight *w*(*y*) = 1. *El* is the set of edges consisting of (*x*, *y*) if *x* ∈ *V*1 is the ancestor of a node *v* ∈¸ *Vl* and (*v*, *y*) ∈ *E*. The load balanced tree problem then is to choose a node *x* ∈ *X* as the ancestor for every node *y* ∈ *Y* such that the maximum number of descendants among the nodes in *V*1 in the resulting tree is minimized.

We then transform the problem into a maximum flow problem in an auxiliary flow network *N* by assigning its links
with different capacities dynamically by using Ford-Fulkerson algorithm[16], where assuming *s* is a source node and *t* is a destination node. Directed edges from *s* to *y* ∈ *Y* and *s* to *x* ∈ *X* are associated with capacity *c*(*s*, *y*) = 1 and *c*(*s*, *x*) = *w*(*x*) respectively. The directed edge from *y* to *x* $\langle y, x \rangle \in E'_l$ has capacity *c*(*y*, *x*) = 1 if edge (*x*, *y*) ∈ *El*. The capacity of the directed edge from each *x* ∈ *X* to *t*, *L*, is the maximum load among the nodes in *X*, that is *c*(*x*, *t*) = *L*. The value range of *L* is within the interval $[\max_{1 \leq i \leq |X|}\{w(x_i) \mid x_i \in X\},$ $\max_{1 \leq i \leq |X|}\{w(x_i) \mid x_i \in$ *X*} + |*Y* |]. Given a value of *L*, we apply the maximum flow algorithm to *Nl* to find a flow *f* from *s* to *t* and check whether
$|f| = \sum_{1 \leq i \leq |X|}\{w(x_i) \mid x_i \in X\} + |Y|$.

```
FORD-FULKERSON(G, s, t)
for each edge (u, v) ∈ E[G]
    do f[u, v] ← 0
       f[v, u] ← 0
while there exists a path p from s to t in the residual network G_f
    do c_f(p) ← min {c_f(u, v) : (u, v) is in p}
       for each edge (u, v) in p
           do f[u, v] ← f[u, v] + c_f(p)
              f[v, u] ← −f[u, v]
```

If yes, it delivers a feasible solution, we will check whether it still has a feasible solution by decreasing the value of *L*; otherwise, the value of *L* needs to be increased. The optimal value *Lopt* of *L* can be found through binary search. In the end, every node *y* in layer *l*+1 will be assigned an ancestor *x* ∈ *X* if *f*(*y*, *x*) = 1. The proposed maximum flow algorithm thus can be applied at most log |*Vl*+1| times to find the optimal load *Lopt* for the current tree expansion.

The partial load-balanced tree is then expanded by including the nodes in layer *l*+1 as follows. For each sensor *y* ∈ *Vl*+1, sensor *v* ∈ *Vl* becomes its parent if *v* is a descendant of *x* ∈ *X*, (*v*, *y*) ∈ *E*, and *f*(*y*, *x*) = 1. As a result, the partial load balanced tree is expanded upto layer *h*. It is straightforward that the approximate load-balanced tree rooted at the virtual node *r* is no more than *h* layers.

3) Find a load-balanced forest: The load-balanced forest consisting of *k* load-balanced trees rooted at the *k* chosen sinks is constructed as follows. A bipartite graph
$G_B = (S', V_1, E')$ is constructed, where
$V_1 = \cup_{s \in S'} C_T(s)$ and an edge (*s*, *v*) ∈ E' if sink $s \in S'$ is within the transmission range of sensor
v ∈ *V*1. A maximum matching in *GB* is then found. For each
matched *v* ∈ *V*1, there is a matched edge with $s \in S'$ as the other endpoint. For each unmatched sensor *v* ∈ *V*1, if there are multiple edges in E' incident to *v*, one of the edges is arbitrarily chosen and the other endpoint of the chosen edge

is a sink $s \in S'$. For both cases, *s* is the root of a load-balanced tree and the subtree rooted at *v* in the original load balanced tree will be part of this new tree. As a result, the sensors in the network have been partitioned into *k* load-balanced trees rooted at the *k* chosen sinks, and each sensor can reach its root (a sink) within *h* hops.

In summary, the time complexity of the proposed algorithm for finding a load-balanced forest is $O(mn\log n)$, where $n = |V|$, $m = |E|$. Sensor partitioning takes $O(m+n)$ time, using the Breadth-First-Search technique while the load-balanced tree algorithm takes $O(mn\log n)$ [8]. The complexity of finding a maximum matching in *GB* is $O(mn)$ [4] and it takes $O(n)$ time to construct the *k* load-balanced trees. For convenience, in the rest of the paper we refer to the proposed heuristic for the *h*-hop constrained multiple sink placement problem as algorithm Heuristic Opt Multisink Place, or HOMP for short.

IV. PERFORMANCE EVALUATION

In this section, we evaluate the performance of the proposed heuristic algorithm for *h*-hop constrained multiple sink placement problem and investigate the impacts of several parameters on the network lifetime through experimental simulations.

A. Simulation environment

We consider a wireless sensor network consisting of 80 to 240 sensors which are randomly deployed in a $100m \times 100m$ square region in the default setting. The potential sink locations in *S* are also randomly generated with the default setting $|S| = 100$. The transmission range *R* of each sensor is fixed to be 10 meters and its initial energy capacity IE is 100*Jules*. In all our experiments we adopt the energy consumption parameters of real sensors - MICA2 motes [5], where $e_t = 14.4 \times 10^{-6} J/bit$ and $e_r = 5.76 \times 10^{-6} J/bit$. We assume that the data generation rate of each sensor is $ra = 1$*bits/s*. Together with these parameters, the network lifetime can be calculated using Eq. (2). The values in figures is the mean of the results by applying each mentioned algorithm to 50 different network topologies of the same size.

B. Impact of the number of hops h and network size n on network performance

We evaluate the impact of parameters *h* - the upper bound on the number of hops from each sensor to its nearest sink, and network size *n* on the optimal number of sinks required and network lifetime.

1) Impact of h and n on the optimal number of sinks:

We first investigate the optimal number of sinks needed by varying the values of *h* and *n*. Fig. 1 plots the number of chosen locations for multiple sink placement under different

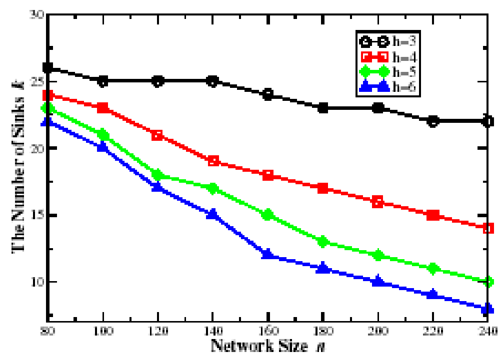

Fig. 1. Impact of h and n on the number of sinks k

constraints of *h* and *n*. It indicates that the number of sinks needed heavily relies on the value of *h*. With the same network size, a larger *h* will result in a smaller *k*. This is because with the increase of the number of hops, each sink can cover more sensors and less number of sinks are needed to cover all the sensors. Moreover, it is also shown that when *h* is fixed, the number of sinks decreases with the increase of network size *n*, since a single sink now can cover more sensors with the increase of sensor density. Thus, fewer sinks are needed.

2) Impact of h and n on the network lifetime:

We then study the impact of the number of hops *h* and the network size *n* on the network lifetime. Fig. 2 demonstrates that smaller *h* and *n* will result in a longer network lifetime. When the network size *n* is fixed, a larger *h* implies a smaller *k* by Fig. 1, which indicates that each sink can cover more sensors in a routing tree rooted at the sink. In other words, each child of the sink will bear a heavier load in comparison with the one with a smaller *h*. Consequently, it will lead to a shorter network lifetime. Similarly, by fixing the number of hops *h*, the network lifetime drops, with the increase of the network size n.

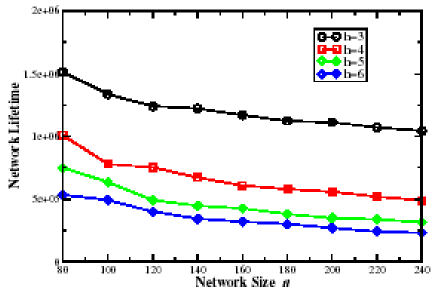

Fig. 2. Impacts of h and n on network lifetime

3) Impact of monitoring regions on network performance:

We finally evaluate the number of sinks *k* and the network lifetime by varying the monitoring area, while keeping the sensor density unchanged. We fix *h* = 5 and vary the monitoring area from 60*m* × 60*m* to 140*m* × 140*m*, which means that the network size *n* and the number of potential sink locations |*S*| will increase accordingly at the same rate.

Table I illustrates that the larger the monitoring region, the greater the value of *k* will be. That means, with the same sensor density, more sinks are required to meet the *h*-hop constraint. It also shows that the network lifetime decreases with the increase of the monitoring area. The reason behind is that in spite of the increase of the number of sinks, each bottleneck sensor still undertakes more relay workload, which causes a shorter network lifetime.

C. Performance evaluation on different heuristics

We compare the performance of the proposed heuristic against the BFS tree-based heuristic in terms of network lifetime. Recall that the results delivered by algorithm HOMP are the *k* locations in *S* for sink placement and load-balanced trees rooted at the *k* chosen sinks. For the BFS-tree based heuristic, we assume that its first two stages are identical to algorithm HOMP. The only difference lies in the routing protocol design, instead of building a load-balanced tree, a BFS tree rooted at the virtual node will be built, we refer to this variant as algorithm BFS Heuristic Opt MultiSink Place or BFS HOMP for short. Note that a

TABLE I
THE NUMBER OF SINKS *k* AND NETWORK LIFETIME WITH DIFFERENT MONITORING REGIONS DELIVERED BY HOMP

monitoring area(m^2)	60×60	80×80	100×100	120×120	140×140
n	36	64	100	144	196
$\|S\|$	36	64	100	144	196
k	9	14	20	29	38
network lifetime ($10^5 s$)	7.76	6.88	5.84	5.42	5.18

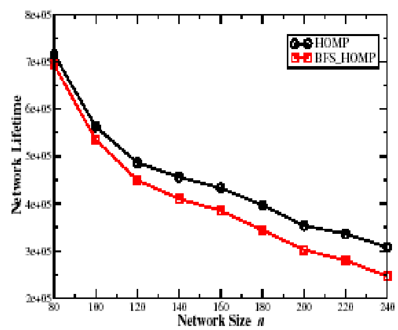

Fig. 3. Performance evaluation between algorithms HOMP and BFS_HOMP

To evaluate the performance of algorithms HOMP an BFS_HOMP, we vary *n* from 80 to 240 while fixing *h* = 5 and = 100. Fig. 3 implies that with the increase of the network size *n*, the network lifetime delivered by either algorithm HOMP or algorithm BFS_HOMP decreases, because the bottleneck sensors have to relay more data for other remote sensors. It is also shown that in terms of network lifetime, algorithm HOMP always outperforms algorithm BFS_HOMP since the former distributes the load among bottleneck sensors more evenly. With the increase of network size *n*, the gap between the network lifetime delivered by these two algorithms becomes larger. In general, algorithm HOMP performs 13% better than algorithm BFS_HOMP on average.

V. CONCLUSION

In this paper we have studied the problem of placing optimal number of sinks for network lifetime maximization, subject to the following constraints: all potential sink locations are given in advance and the maximum number of hops from each sensor to its nearest sink is bounded by a constant *h*. We formulated this problem as a joint optimization problem. Due to its NP hardness, we then devised a novel heuristic for it. We finally conducted extensive experiments by simulations to evaluate the performance of the proposed algorithm against the BFS tree-based heuristic. The experimental results demonstrate that the former outperforms the latter significantly in terms of network lifetime prolongation.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: a survey. *Computer Networks*, Vol. 38, 2002.

[2] A. Bogdanov, E. Maneva, and S. Riesenfeld. Power-aware base station positioning for sensor networks. *Proc. INFOCOM'04*, IEEE, 2004.

[3] J-H Chang and L. Tassiulas. Energy conserving routing in wireless ad hoc networks. *Proc. INFOCOM'00*, IEEE, 2000.

[4] T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein. *Introduction to Algorithms*. 3rd Edition, MIT Press, 2009.

[5] Crossbow Inc. MPR-Mote Processor Radio Board User's Manual.

[6] S. R. Gandham, M. Dawande, R. Prakask, and S. Venkatesan. Energy efficient schemes for wireless sensor networks with multiple mobile base stations. *Proc. Globecom'03*, IEEE, 2003.

[7] D. S. Johnson. Approximation algorithms for combinatorial problems. *JCSS*, Vol. 9, pp. 256–278, 1974.

[8] W. Liang, J. Luo. Load-balanced data gathering for network lifetime maximization in wireless sensor networks. *manuscript*, submitted for possible publication, Sept., 2009.

[9] H. Kim, T. Kwon, and P. Mah. Multiple sink positioning and routing to maximize the lifetime of sensor networks. *IEICE Trans. on Communications*, Vol. E91-B, pp 3499–3506, 2008.

[10] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson. Wireless sensor networks for habitat monitoring. *The 1st Int'l Workshop on Wireless Sensor Networks and Applications (WSNA)*, ACM, 2002.

[11] E. I. Oyman, and C. Ersoy. Multiple sink network design problem in large scale wireless sensor networks. *Proc. ICC'04*, IEEE, 2004.

[12] G. J. Pottie. Wireless sensor networks. *Proc. of Information Theory Workshop*, pp 139–140, 1998.

[13] L. Qiu, R. Chandra, K. Jian, and M. Mahdian. Optimizing the placement of integration points in multi-hop wireless sensor networks. *Int'l Conf. on Network Protocols (ICNP)*, 2004.

[14] W. Youssef and M. Younis. Intelligent gateways placement for reduced data latency in wireless sensor networks. *Proc. ICC'07*, IEEE, 2007.

[15] Xu Xu and Weifa Liang, Placing optimal number of sinks in sensor networks for network lifetime maximization. Proc ICC '11,IEEE 2011.

[16] Thomas S Cormen, Charles E Leiserson, Ronald L Rivest, Clifford Stein Introduction to algorithms 3rd edition.