A Tree-based Energy-Efficient Algorithm for Data-Centric Wireless Sensor Networks

Yean-Fu Wen, Frank Yeong-Sung Lin, and Wen-Cheng Kuo
Department of Information Management, National Taiwan University
Department of Information Management, China University of Technology
{d89002,yslin,r93039}@im.ntu.edu.tw

Abstract

The nature of wireless sensor networks make them suitable for a great variety of applications, especially over wide areas, or in remote or hostile locations; however, such environments make battery capacity an especially important concern, where replacing or recharging of batteries is infeasible for one reason or another. Battery capacity restrictions on highly energy-constrained sensor networks can be mitigated, by adopting data-aggregation techniques and by managing the scheduling of nodes. These effectively reduce the overall amount of data transmitted, thereby conserving energy. In this paper, we address the construction of energy-efficient data-aggregation trees, an NP-problem, in different rounds of communication, seeking to maximize the lifetime of heterogeneous sensor networks. This problem is subject to constraints on such networks: battery capacity, data-sensing scheduling, and round calculation. We derive a near-optimal primal feasible solution using Lagrangean Relaxation. The experimental results show that our proposed algorithm outperforms similar algorithms.

1. Introduction

Wireless sensor networks (WSNs) provide many advantages to researchers who seek to monitor environmental data. Data collection is facilitated by these networks, and is possible in even remote and hostile environments where humans cannot venture.

Sensor nodes in a WSN are designated to monitor some events of interest. Sensor nodes are typically deployed to constitute a field of sensors. Environmental data, such as temperature, is sensed by a sensor node. The data is collected and then transmitted to a sink node, for further processing, as shown in Figure 1. Each sensor node, or source node, within the sensing range produces data.

Several factors should be taken into account when designing wireless sensor networks (WSNs), such as coverage [4] and lifetime [2] [3] [8] [10] [12] [17] [18]. The battery capacity of sensors is fixed, and it is often infeasible to recharge the batteries due to remote deployment, so the power consumption is rapidly by operations. Therefore, an issue fundamental to WSNs is extending the network lifetime. In this paper, the lifetime is defined as the number of rounds until the information about the occurrence of any of events cannot be delivered to the sink node.

WSNs are primarily used for two applications, event-driven and query-based. In a periodic WSN, all source nodes periodically sense the events and report the sensed data to the sink node; in a query-based WSN, users can at any time request information from a sensor node. In both event-driven and query-based models, when a specific event occurs, the raw data is collected and processed before relaying it to the sink node with query instructions. During the processing, duplicate and useless data is abandoned. The source nodes are responsible for collecting the raw data, and the intermediate nodes are responsible for aggregating and transmitting it to the sink node through a reverse multicast tree where multiple source nodes transmit data to the sink node. This process is called data-centric routing [11].

Data aggregation is the key to efficient data-centric routing. By aggregating the data from different source nodes, duplicate information can be eliminated; thus, the total number of transmissions can be reduced. By adopting data aggregation, we can achieve energy-efficient transmission [2] [5] [8] [9] [18]. However, the construction of this type of data- aggregation tree has been proven to be NP-complete [2] [6], which signifies that general algorithms cannot provide optimal solution to the problem. Krishnamachari *et al.* [2] devised three aggregation heuristics, namely, the Shortest Paths Tree (SPT), Center at Nearest Source (CNS), and the Greedy Incremental Tree (GIT) to sub-optimally solve the problem. In this paper, we have proposed a means, based on an approach determined by Lagrangean Relaxation (LR), by which more efficient feasible solutions using a tree-based algorithm may be gained.

21st International Conference on Advanced Networking and Applications(AINA'07)
0-7695-2846-5/07 $20.00

Power consumption has a great impact on the system lifetime of WSNs. Therefore, many researchers have focused on designing routing algorithms to reduce power consumption. For more about query-based WSNs where the sensor nodes transmit data traffic through the aggregation tree to the sink node in a round, see [5] [10] [12] [16]. In order to maximize the total number of rounds, the source nodes in these WSNs construct appropriate aggregation trees that send the sensed data to the sink node each round.

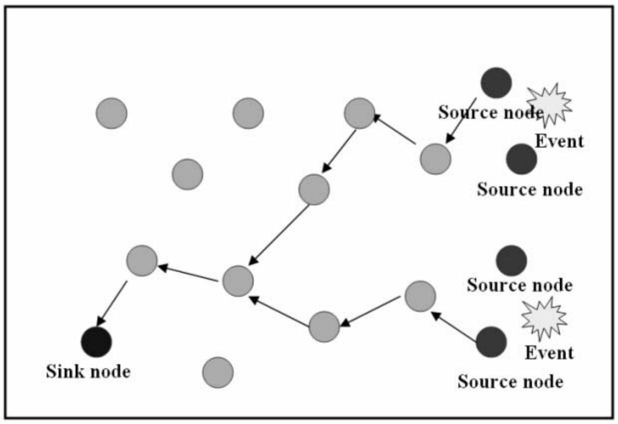

Figure 1. An example in which event data collected by sensor nodes is transmitted along routing paths to a sink node.

In [12], linear programming was used to solve the problem, whereas in [10], Prim's algorithm was adopted, which determines as the data-aggregation tree a minimum cost spanning tree. From these and our previous work [7], the basic tree-based concept is extended. After a predefined number of rounds, the source nodes reconstruct, as a new data-aggregation tree, the spanning tree with the minimum cost. When constructing the aggregation tree, the cost, or weight, is the remaining battery power of sensor nodes. For a sensor, the greater its weight, the greater the probability that it will be selected as an intermediate node. By adopting the remaining battery power concept, the sensitivity of sensor nodes to the spanning tree is improved.

Relevant to our work, in [16], the authors proposed a chain-based protocol, Power-Efficient Gathering in Sensor Information System (PEGASIS), in order to extend system lifetime. This algorithm, which operates by means of a greedy method, constructs a chain by which information is transmitted.

In this paper, a meaningful mathematical formulation of WSNs is essential. Through the mathematical model, we can understand the theoretical bounds of the performance, and the impact of different input parameters, such as the number of nodes, topology, and battery capacity level [1]. We implement an optimization-based model for WSNs in which a network topology is constructed and aggregation routing is assigned. Seeking to maximize the numbers of executing rounds of the trees, we use an LR-based approach that obtains a near-optimal solution. For the purposes of this model, we take into consideration data sensing scheduling, data aggregation and battery capacity constraints.

The remainder of the paper is organized as follows. Presented in the next section is a mixed-integer nonlinear programming formulation of the tree-based data centric routing problem, which includes the assignment of data sensing schedules and data-aggregation routing. In Section 3, our LR-based approach is outlined. In Section 4, the heuristic for obtaining primal feasible solutions to the problem is addressed. In Section 5, computational results are reported. Finally, in Section 6, our conclusions are presented.

2. Problem formulation

In this paper, we consider heterogeneous sensor networks consisting of data source nodes and communication nodes in which sensors are randomly scattered over the area of interest. The locations of the sensors are fixed and known by the sink node a priori. The sink node is assumed to have an inexhaustible supply of energy. The battery capacity of sensors is given to be highly energy-constrained. Environmental events are periodically sensed by the source nodes. Each of these events is sensed by only one source node, in order to decrease energy consumption.

At the beginning of each round, the sink node instructs the source nodes to sense the environmental conditions as well as broadcasting the routing information to all nodes. In a round of communication, environmental data aggregated by the communication nodes is relayed to the sink node. Subsequently, the data is transmitted by the source nodes to the sink node over the data-aggregation tree.

We seek to determine a routing scheme which maximizes the number of successful data transmission. The following decision variables are relevant to determination of such routing schemes: (i) decisions about which source nodes wake up to sense the environment in each round, (ii) selection of the routing for data-aggregation for each round, (iii) the number of times that each tree may be used. The problem description is summarized on next page.

The objective problem of Z_{IP} is the maximization of the total number of rounds for each given data-aggregation tree. The number of rounds that tree t is used is denoted as θ_t ($\theta_t \in \{0, 1, 2 \cdots, M_t\}$. In each round, the data is transmitted via a tree t.

$$Z_{IP} = \max \sum_{t \in T} \theta_t \qquad \text{(IP)}$$

Problem Assumption:
· Heterogeneous WSNs;
· Fixed sensing range and fixed transmission range;
· Bi-directional links;
· Error-free transmission within the transmission radius; and
· The location of each sensor node is known by the sink node.
Given:
· The network topology, which includes the set of nodes V and the set of links L.
· The set of source nodes $W \in V$.
· The set of communication nodes $U \in V$.
· The sink node q.
· The set of data-aggregation trees T, rooted at the sink node.
· The set of candidate paths P_w from the source node w to the sink node. Note that we are interested in determining only one path for each source node w. That path can be determined by the proposed algorithm. Thus, it is not necessary for the set of all candidate paths to be determined or initially listed.
· The initial battery level C_w for each node w, evaluated by residual power (nJ).
· The sensing, transmission, and receiving costs for each link with respect to the energy consumption rates E_m, E_s, and E_r (nJ).
· The energy required for the broadcast of the routing information from the sink nodes to all nodes, denoted as E_q (nJ).
· The total number of rounds R.
· The set of events Ξ.
Object:
To maximize the number of event transmission rounds.
Subject to the following constraints:
· Data sensing scheduling: each event must be monitored by at least one source node. That event data must be transmitted to the sink node;
· Tree: the routing paths for the source nodes and links on the paths are constructed as an aggregation tree;
· The battery capacity: the total energy consumption of a node cannot exceed its initial energy level; and
· The maximum number of rounds: must be calculated based on the battery capacity.
To determine:
· The wake up time of each source node and communication node;
· Which tree t is selected to transmit sensed data to the sink node in the round r;
· A routing path and links from a source node to that sink node; and
· The maximum number of rounds that the network may continue functioning.

The objective function is subjected to (1)-(12), described as the following constraints.

(a) Backhaul selection constraints

The indicator function $e_{\vartheta w}$ is 1 if the event ϑ is in the coverage of the source node w, otherwise 0. For each event in the coverage of source node w, it is necessary to be sensed by one waking source node w in round r (denoted by 0-1 decision variable $\pi_{wr} = 1$) as shown in (1).

$$\sum_{w \in W} \pi_{wr} e_{w\vartheta} = 1, \ \forall \vartheta \in \Xi; r \in R \tag{1}$$

(b) The path constraint

For each waked source node w in round r, it has to select a path to transmit sensed data to the sink node. Let x_{pr} be a 0-1 decision variable to denote whether a path is selected in the round r. x_{pr} is equal to 1 if the path p is selected to transmit sensed data; otherwise 0. Then, the constraint is formulated as:

$$\sum_{p \in P_w} x_{pr} = \pi_{wr}, \ \forall w \in W; r \in R \tag{2}$$

It requires that if one path is selected for the source node w in round r, it must also be on the sub-tree adopted by the source node w. Here, the 0-1 decision variable ϕ_{ur} is set to 1 if the communication node u is awake in the round r, and 0 otherwise.

$$\sum_{p \in P_w} x_{pr} \gamma_{pu} \leq \phi_{ur}, \ \forall u \in V; w \in W; r \in R \tag{3}$$

where γ_{pu} is the indicator function which is 1 if relay node u is on the path p and 0 otherwise.

(c) The link constraints

It requires that if one path is selected for the source node w in round r, it must also be on the sub-tree adopted by the source node w. An 0-1 decision variable $y_{r(u,v)}$ is set to 1 if link (u, v) is used in round r, and 0 otherwise.

$$\sum_{p \in P_w} x_{pr} \delta_{p(u,v)} \leq y_{r(u,v)}, \ \forall (u, v) \in L; w \in W; r \in R \tag{4}$$

where $\delta_{p(u,v)}$ denotes the indicator function which is 1 if the link (u, v) is on the path p and 0 otherwise.

We select at most one outgoing link of node in round r.

$$\sum_{v \in V} y_{r(u,v)} \leq 1, \ \forall u \in V; r \in R \tag{5}$$

We select exactly one outgoing link of the source node in round r. Here, the decision variable π_{wr} equal to 1 if the data source node w is awake in the round r, and 0 otherwise.

$$\pi_{wr} \leq \sum_{v \in U} y_{r(w,v)}, \ \forall w \in W; r \in R \tag{6}$$

The number of selected incoming links (u, q) of the sink node in round r must be larger than 1.

$$\sum_{u \in V} y_{r(u,q)} \geq 1, \ \forall r \in R \tag{7}$$

(d) The battery capacity constraints

For each communication node w, the total receiving and communication power consumption can not exceed its initial energy level.

$$\sum_{r\in R} E_q + \sum_{r\in R} \pi_{wr}\left(E_m + E_s\right) \le C_w, \ \ \forall w \in W \qquad (8)$$

For each source node w, the total sensing and communication power consumption can not exceed its initial energy level.

$$\sum_{r\in R} E_q + \sum_{r\in R} \phi_{ur}\left(E_r + E_s\right) \le C_u, \ \ \forall u \in U \qquad (9)$$

(e) The numbers of trees calculation constraints

It requires that the sub-tree adopted by any source node w be a subset of the shared spanning tree. This shared spanning tree is selected to be shared by the source nodes. Here, the 0-1 decision variable z_{tr} is denoted whether a tree t is selected or not. It is equal to 1 if the tree t is selected to transmit the sensed data to the sink node in the round r, and 0 otherwise.

$$y_{r(u,v)} \le \sum_{t\in T} \sigma_{t(u,v)} z_{tr} \ \ \forall (u,v)\in L; r \in R \qquad (10)$$

where $\sigma_{t(u,v)}$ is an indicator function which is 1 if the link (u, v) is on the tree t, and 0 otherwise.

In round r, a tree must be selected to transmit data to the sink node. The constraint is formulated as:

$$\sum_{t\in T} z_{tr} = 1 \ \ \forall r \in R \qquad (11)$$

Accordingly, the total number of rounds θ_t that tree t is used to transmit data to the sink node is calculated as:

$$\sum_{r\in R} z_{tr} \le \theta_t, \ \ \forall t \in T \qquad (12)$$

3. The LR-based Approach

Adopting a LR-based approach [13] [14] [15], by relaxing Constraints (3), (4), (6), (8), (9), and (10) into the objective function (IP), it is transformed into a LR problem. As a convention, we first multiply the objective function of the primal problem by negative one and transform it into a minimization problem. For a vector of non-negative multipliers, the resulting LR problem is as follows:

$$\begin{aligned}
&Z_{LR}\left(\mu^1_{rwu}, \mu^2_{rwuv}, \mu^3_{rw}, \mu^4_{w}, \mu^5_{u}, \mu^6_{ruv}\right) \\
&= \min \Bigg\{ -\sum_{t\in T}(\theta_t) + \\
&\sum_{r\in R}\sum_{w\in W}\sum_{u\in U} \mu^1_{rwu}\Big(\sum_{p\in P_w} x_{pr}\gamma_{pu} - \phi_{ur}\Big) + \\
&\sum_{r\in R}\sum_{w\in W}\sum_{(u,v)\in L} \mu^2_{rwuv}\left(\sum_{p\in P_w} x_{pr}\delta_{p(u,v)} - y_{r(u,v)}\right) + \\
&\sum_{r\in R}\sum_{w\in W} \mu^3_{rw}\Big(\pi_{wr} - \sum_{v\in U} y_{r(w,v)}\Big) + \\
&\sum_{w\in W} \mu^4_{w}\Big(\sum_{r\in R} E_q + \sum_{r\in R}\pi_{wr}\left(E_m + E_s\right) - C_w\Big) + \\
&\sum_{u\in U} \mu^5_{u}\Big(\sum_{r\in R} E_q + \sum_{r\in R}\phi_{ur}\left(E_R + E_s\right) - C_u\Big) + \\
&\sum_{r\in R}\sum_{(u,v)\in L} \mu^6_{ruv}\left(y_{r(u,v)} - \sum_{t\in T}\sigma_{t(u,v)} z_{tr}\right)\Bigg\}
\end{aligned}$$

Subject to: (1), (2), (5), (7), (11), and (12).

This LR problem can be decomposed into four sub-problems, each of which is independent, related to decision variables. Descriptions of and algorithms for each of these four sub-problems are presented in the Appendix. Once a solution has been determined for each of the sub-problems, the solution to the LR problem may be determined, which establishes a lower bound (LB). The value of the follow-up primal feasible solution gives us an upper bound (UB). The distance between the tightest LB and the UB, computed by $(UB - LB)/LB * 100\%$, gives the degree of optimality of the problem solution.

We employ the subgradient method [13], one of the most popular methods used to solve the Lagrangean dual problem (D). Let the decision variable vectors ($x_{pr}, \theta_{ur}, y_{r(u,v)}, \pi_{wr}, z_{tr}$, and θ_t) be the subgradients of Z_D. Iteration $k+1$, determined by updating multiplier vector, $\pi^k = \left(\mu^1_{rwu},\ \mu^2_{rwuv},\ \mu^3_{rw},\ \mu^4_{w},\ \mu^5_{u},\ \mu^6_{ruv}\right)$ using $\pi^{k+1} = \pi^k + t^k g^k$. The step size t^k is determined by $t^k = \kappa \cdot (Z^h_{IP} - Z_D(\pi^k)) \Big/ \left\|g^k\right\|^2$, where Z^h_{IP} is the primal objective function value for a heuristic solution (an upper bound of Z_{IP}), and is a constant, $0 < \kappa \le 2$.
The Lagrangean dual problem (D) is

$$Z_D = \max Z_{LR}\left(\mu^1_{rwu}, \mu^2_{rwuv}, \mu^3_{rw}, \mu^4_{w}, \mu^5_{u}, \mu^6_{ruv}\right) \qquad (D)$$

subject to

$$\mu^1_{rwu}, \mu^2_{rwuv}, \mu^3_{rw}, \mu^4_{w}, \mu^5_{u}, \mu^6_{ruv} \ge 0 \qquad (13)$$

4. Obtaining the Primal Feasible Solution

Data-aggregation trees, which aggregate and transmit data to sink nodes, are determined by an LR-based algorithm (LRA), is constructed by means of the solutions to the sub-problems (SUB1) and (SUB3) presented in the Appendix. Of the candidate data-aggregation trees, designated z_{tr} and π_{wr}, redundant trees are eliminated from the set of trees. The total amount of energy consumed by a data-aggregation tree gives its cost; initially, the aggregation tree with the lowest cost is used. After the battery in a node has become entirely depleted of energy, to ensure the delivery of all data by an alternate route, a substitute tree is selected from the set of candidate trees. When all candidate trees have become exhausted, a data-aggregation tree based on the remaining energy of nodes is reconstructed to transmit data. The procedures of this heuristic are shown as follows. The complete LRA is shown in Algorithm 1.

Step 1 We construct the candidate trees according to the decision variables z_{tr} and π_{wr}.

Step 2 We sort the candidate trees in ascending order according to the costs of the trees, determined by energy consumption.

Step 3 We select an aggregation tree, which satisfied all constraints, from the set of candidate trees and calculate the energy consumption of nodes. Repeat **Step 3** until no trees are connected cause by any node energy exhausted.

Step 4 We reconstruct an aggregation tree based on the remaining energy of nodes and use it to transmit information. Then, we calculate the energy consumption of nodes. Repeat **Step 4** until we can not find any aggregation tree.

Step 5 Finally, we obtain the maximum number of communication rounds.

5. Evaluation and Experimental Results

We conducted several experiments, controlling for different parameters to evaluate the solution quality of our approach. Each event is covered by some source nodes and one of them senses and transmits the information to the sink node in each round of communication. We varied the number of source nodes and communication nodes that were used in the construction of a network. In the grid network experimental conditions, the nodes were deployed in a grid, a regularly spaced topology. In the random network experimental conditions, the nodes were deployed randomly, i.e. the sensors were randomly scattered over the area of interest.

Using LR, we determined a gap (%) that allowed us to compare our approach with Power Efficient Data gathering and Aggregation Protocol (PEDAP) algorithm and Power Efficient Data gathering and Aggregation Protocol-Power Aware (PEDAP-PA) algorithm.

We simulated the networks using a program written in C++. The test platform was a PC with Pentium 4 2.4G CPU and 512MB DRAM. We executed our program on Windows 2000. Table 1 shows the experimental results for grid topologies; Table 2, for random topologies. In these two tables, "Zdu" denotes the Lagrangean dual solution value. "ZIP" denotes the objective value obtained from LRA. "Gap", calculated by $(ZIP - Zdu)/ZIP * 100\%$, is used to evaluate our solution quality. "Imp. Ratio" is used to compare the LRA with the existing algorithms. "Imp.

Algorithm 1 LRA

Initialize Lagrangean multiplier vector $(\mu^1_{rwu}, \mu^2_{rwuv}, \mu^3_{rw}, \mu^4_{w}, \mu^5_{u}, \mu^6_{ruv})$ to be zero vectors;
UB = an arbitrary large number;
$LB = 0$;
$improve\_counter = 0$; {used to count the number of iterations in which LB or UB has not improved.}
$\kappa = 2$; {κ is a constant, described in Section 3.}
for $iteration = 1$ to $Max\_Iteration\_Number$ {a bound on the number of iterations of the LR procedures.} **do**
  Run subproblem $(SUB1) - (SUB4)$ {shown in Appendix.}
  Calculate $Z_{LR}()$ to get LB, denoted by Z_{LR}; {Z_D denotes the current maximum LB.}
  if $Z_{LR} > Z_D$ **then**
    $Z_D = Z_{LR}$;
    $improve\_counter = 0$; {LB is updated if LB is greater than Z_D.}
  else
    $improve\_counter = improve\_counter + 1$;
  end if
  if $improve\_counter = Improve\_Threshold$ **then**
    $improve_counter = 0$;
    $\kappa = \kappa/2$; {the parameters of Lagrangean dual procedures are adjusted}
    RUN Getting primal feasible heuristic to get ub; {shown in Section 4.}
  end if
  if $ub < UB$ **then**
    $UB = ub$;
  end if
  RUN update-step-size();
  RUN update-Lagrangean-multiplier();
end for

21st International Conference on Advanced Networking and Applications(AINA'07)
0-7695-2846-5/07 $20.00 © 2007 IEEE

Table 1. Evaluation (gap and improvement ratios) of proposed algorithm in grid networks with different number of nodes and data source nodes.

No. of Nodes	Events	No. of Source Nodes	Zdu	ZIP	Gap(%)	PEDAP	Impr. Ratio 1(%)	PEDAP-PA	Imp. Ratio 2(%)
16	1	2	90.49	86	5.22	49	43.02	47	45.35
	2	5	81.35	80	1.68	43	46.25	43	46.25
25	1	2	98.77	92	7.36	47	48.91	46	50.00
	2	5	86.54	82	5.54	42	48.78	39	52.44
	4	10	88.25	82	7.63	42	48.78	33	59.76
36	1	2	112.67	108	4.33	61	43.52	46	57.41
	2	5	89.53	84	6.58	49	41.67	46	45.24
	4	10	91.99	84	9.51	27	67.86	45	46.43
49	1	2	85.77	81	5.89	48	40.74	48	40.74
	2	5	86.87	81	7.24	48	40.74	48	40.74
	4	10	63.04	63	0.06	36	42.86	36	42.86
	8	20	48.19	44	9.52	36	18.18	36	18.18
64	1	2	89.98	85	5.86	45	47.06	39	54.12
	2	5	90.64	85	6.63	45	47.06	39	54.12
	4	10	92.62	85	8.97	45	47.06	40	52.94
	8	20	46.28	45	2.83	41	8.89	41	8.89
81	1	2	85.51	80	6.89	53	33.75	54	32.50
	2	5	87.00	79	10.13	53	32.91	53	32.91
	4	10	83.27	79	5.41	52	34.18	53	32.91
	8	20	69.90	61	14.59	52	14.75	53	13.11

Ratio 1" $= (ZIP - PEDAP)/ZIP * 100\%$ and "Imp. Ratio 2" $= (ZIP - PEDAP - PA)/ZIP * 100\%$.

In Tables 1 and 2, the small gap for our solution is less than 14.59%. This implies that our solution quality is near optimal. Moreover, our LR-based algorithm outperforms other algorithms in different network sizes. We adopted a dynamic source node concept and the LRA provided us with a good mechanism with which to find near-optimal selected trees. The improvement ratios indicate significantly greater energy efficiency for our proposed heuristic, which exceeded others by $8.89\% - 72.64\%$.

With a fixed network size, as the number of source nodes increases, the system lifetime shortens. Since the energy consumers increase, the total energy consumption increases. However, we observe that the number of rounds declining trend is not a linear function but a step function, as shown in Figure 2. That is because some nodes, or articulation points, play an important role in the routing. Once these nodes are drained of their energy, the network becomes disconnected. The lifetime of the system declines if the load of the articulation points is exceeded, such as occurs when there is an increase in traffic or a greater number of nodes.

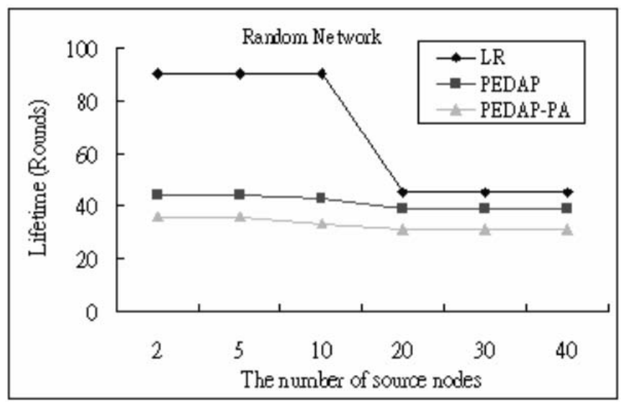

Figure 2. Evaluation of lifetime (Rounds) in random network (Number of Nodes = 81)

6. Conclusion

In this paper, we address the construction of a data-aggregation tree in heterogeneous sensor networks, seeking to maximize the number of event data transmissions in successive rounds of communication. In order to maximize the lifetime of the system, we proposed an approach based on Lagrangean Relaxation to construct candidate aggregation trees, carefully allocating traffic according to the round of communication. We conducted several experiments involving different network topologies. According to our experimental results, we can claim that our LR-based algorithm is superior to the PEDAP algorithm by 8.89% to 72.64%, and to the PEDAP-PA algorithm by 8.89% to 71.7%.

21st International Conference on Advanced Networking and Applications(AINA'07)
0-7695-2846-5/07 $20.00 © 2007 IEEE

Table 2. Evaluation (gap and improvement ratios) of proposed algorithm in random networks with different number of nodes and data source nodes.

No. of Nodes	Events	No. of Source Nodes	Zdu	ZIP	Gap (%)	PEDAP	Impr. Ratio 1 (%)	PEDAP-PA	Imp. Ratio 2 (%)
16	1	2	75.42	71	6.23	34	52.11	39	45.07
	2	5	76.54	71	7.80	34	52.11	39	45.07
25	1	2	93.31	89	4.84	50	43.82	51	42.70
	2	5	93.52	89	5.08	48	46.07	49	44.94
	4	10	87.16	81	7.60	48	40.74	48	40.74
36	1	2	111.20	106	4.91	34	67.92	31	70.75
	2	5	111.93	106	5.60	29	72.64	30	71.70
	4	10	83.65	79	5.89	29	63.29	30	62.03
49	1	2	83.15	79	5.25	39	50.63	41	48.10
	2	5	83.68	79	5.93	39	50.63	41	48.10
	4	10	84.92	79	7.49	39	50.63	41	48.10
	8	20	74.34	68	9.32	37	45.59	37	45.59
64	1	2	100.22	96	4.40	40	58.33	43	55.21
	2	5	99.66	94	6.02	40	57.45	45	52.13
	4	10	94.71	89	6.41	40	55.06	43	51.69
	8	20	85.15	79	7.79	40	49.37	41	48.10
	16	30	86.95	79	10.06	42	46.84	40	49.37
81	1	2	93.16	90	3.52	44	51.11	36	60.00
	2	5	93.42	90	3.80	44	51.11	36	60.00
	4	10	94.51	90	5.01	43	52.22	33	63.33
	8	20	47.04	45	4.52	39	13.33	31	31.11
	16	30	47.81	45	6.23	39	13.33	31	31.11
	24	40	46.82	45	4.04	39	13.33	31	31.11

References

[1] B. Krishnamachari and F. Ordonez, "Analysis of Energy-Efficient, Fair Routing in Wireless Sensor Networks through Non-Linear Optimization", Workshop on Wireless Ad hoc, Sensor, and Wearable Networks, in Proc. IEEE VTC, Orlando, Florida, USA, October 2003.

[2] B. Krishnamachari, D. Estrin, and S. Wicker, "Modelling Data-Centric Routing in Wireless Sensor Networks", in *USC Computer Engineering Technical Report CENG 02-14*, 2002.

[3] C.W. Shiou, Frank Y.S. Lin, H.C. Cheng, Y.F. Wen, "Optimal Energy-Efficient Routing for Wireless Sensor Networks", in Proc. IEEE AINA, Taipei Taiwan, March 2005.

[4] Frank Y.S. Lin, and P.L. Chiu, "A Near-Optimal Sensor Placement Algorithm to Achieve Complete Coverage- Discrimination In Sensor Networks", IEEE Communications Letters, 9(1):43-45, January 2005.

[5] Frank Y.S. Lin and Y.F. Wen, "Multi-sink Data Aggregation Routing and Scheduling with Dynamic Radii in WSNs", IEEE Communications Letters, 10(10):692-694, October 2006.

[6] Garey, M.R., D.S. Johnson, "Computers and Intractability: A Guide to the Theory of NP-completeness", Freeman San Francisco, 1979.

[7] H.H. Yen, Frank Y.S. Lin, "Near-Optimal Tree-Based Access Network Design", Computer Communication, 28(2):236-245, 2005.

[8] H.H. Yen, Frank Y.S. Lin and S.P. Lin, "Efficient Data-centric Routing in Wireless Sensor Networks", In Proc. IEEE ICC, Seoul Korea, May 2005.

[9] H.H.Yen, Frank Y.S. Lin, and S.P. Lin, "Energy-Efficient Data-Centric Routing in Wireless Sensor Networks", IEICE Trans. on Communications, E88-B(15):4470-4480, 2005.

[10] H. O. Tan and I. Korpeoglu, "Power Efficient Data Gathering and Aggregation in Wireless Sensor Networks", ACM SIGMOD Record, 32(4):66-71, 2003.

[11] Jamal N. Al-Karaki and Ahmed E. Kamal, "Routing Techniques in Wireless Sensor Networks: A Survey", IEEE Wireless Communication, 11(6):6-28, 2004.

[12] K. Kalpakis, K. Dasgupta and P. Namjoshi. "Efficient Algorithms for Maximum Lifetime Data Gathering and Aggregation in Wireless Sensor Networks", Computer Networks, 42(6):697-716, August. 2003.

[13] M.L. Fisher, "The Lagrangean Relaxation Method for Solving Integer Programming Problems," Management Science, 27(1):1-18, January 1981.

[14] M.L. Fisher, "An Application Oriented Guide to Lagrangean Relaxation," Interfaces, 15(2):10-21, April 1985.

[15] Bazaraa M.S., H.D. Sherali, and C.M. Shetty, "Lagrangian Duality and Saddle Point Optimality Conditions", Nonlinear Programming: Theory and Algorithms, 2nd Edition, pp. 199-242, John Wiley & Sons, Inc, Singapore, 1993.

[16] S. Lindsey and C. S. Raghavendra, "PEGASIS: Power-Efficient Gathering in Sensor Information Systems", In Proc. IEEE Aerospace Conference, Big Sky, MT, Mar. 2002.

[17] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks", In Proc. 33rd Hawaii International Conference on System Sciences, Hawaii, USA, Jan. 2000.

[18] Y.F. Wen and Frank Y.S. Lin, "Cross-Layer Duty Cycle Scheduling with Data Aggregation Routing in Wireless Sensor Networks", In Proc. IFIP EUC 2006, Seoul Korea, August 2006, LNCS 4096.

Appendix: Algorithms for the sub-problems

Each of the following four sub-problems, generated from the problem (LR), is related to decision variables.

Sub-problem (SUB1) is related to the decision variables x_{pr} and π_{wr}. The objective problem

$$Z_{SUB1} = \min \left\{ \begin{array}{c} \sum\limits_{r\in R}\sum\limits_{w\in W}\sum\limits_{p\in P_w} \left[\begin{array}{c} \sum\limits_{u\in U} \mu^1_{rwu}\gamma_{pu}+ \\ \sum\limits_{(u,v)\in L} \mu^2_{rwuv}\delta_{p(u,v)} \end{array} \right] x_{pr}+ \\ \sum\limits_{w\in W}\sum\limits_{r\in R} \left[\mu^3_{rw} + \mu^4_w\left(E_m+E_s\right)\right]\pi_{wr} \end{array} \right\} \tag{SUB1}$$

subject to (1) and (2).

The proposed algorithm for (SUB1) is as follows:

Step 1 For each source node w, compute the coefficient $\left(\mu^3_{rw} + \mu^4_w\left(E_m+E_s\right)\right)$ for each π_{wr}.

Step 2 If the corresponding coefficient is negative, set π_{wr} to 1; otherwise, set π_{wr} to 0.

Step 3 If π_{wr} is equal to 1, find the shortest path with nonnegative arc weights, given by $(\sum\limits_{u\in U}\mu^1_{rwu}\gamma_{pu} + \sum\limits_{(u,v)\in L}\mu^2_{rwuv}\delta_{p(u,v)})$ by Dijkstra's algorithm. **Step 3** is repeated for all π_{wr}.

The computational complexity of the above algorithm is $O(|V|^2)$.

Sub-problem (SUB2) is related to decision variable $y_{r(u,v)}$. The objective problem,

$$Z_{SUB2} = \min \sum_{r\in R} \left[\begin{array}{l} \sum\limits_{(u,v)\in L}\left(\mu^6_{ruv} - \sum\limits_{w\in W}\mu^2_{rwuv}\right) y_{r(u,v)} \\ -\sum\limits_{v\in U}\sum\limits_{w\in W}\mu^3_{rw} y_{r(w,v)} \end{array} \right] \tag{SUB2}$$

subject to (5) and (7).

The proposed algorithm for solving (SUB2) is as follows:

Step 1 For each link, we compute the coefficient $(\mu^6_{ruv} - \sum_{w\in W}\mu^2_{rwuv})$. If the link is the outgoing link of the source node, the corresponding coefficient is set to $(\mu^6_{ruv} - \sum_{w\in W}\mu^2_{rwuv} - \sum_{w\in W}\mu^3_{wr})$.

Step 2 For all outgoing links of node, we find the smallest coefficient. If the smallest coefficient is negative then we set the corresponding $y_{r(u,v)}$ to 1 and the other outgoing links $y_{r(u,v)}$ to 0, otherwise we set all outgoing link $y_{r(u,v)}$ to 0. Repeat **Step 2** for all nodes.

Step 3 For all incoming links of the sink node, we find the smallest coefficient and set the corresponding $y_{r(u,v)}$ to 1 and the other incoming links $y_{r(u,v)}$ to 0.

The computational complexity of the above algorithm is $O(|R||V|^2)$ for each link.

Subproblem (SUB3) (related to decision variables z_{tr} and θ_t). The objective function

$$Z_{SUB3} = \min -\sum_{t\in T}\theta_t - \sum_{r\in R}\sum_{t\in T}\left(\sum_{(u,v)\in L}\mu^6_{ruv}\sigma_{t(u,v)}z_{tr}\right) \tag{SUB3}$$

subject to (11) and (12).

The proposed algorithm for solving (SUB3) is described as follows:

Step 1 For each z_{tr}, we solve the maximum cost spanning tree problem with nonnegative arc weights by Prim's algorithm.

Step 2 We compute the sum of the number of the rounds that each tree t is used.

The computational complexity of the above algorithm is $O(|V|^2)$ for each tree.

Sub-problem (SUB4) (related to decision variable ϕ_{ur})

$$Z_{SUB4} = \min\sum_{u\in U}\sum_{r\in R}\left(\mu^5_u\left(E_r+E_s\right) - \sum_{w\in W}\mu^1_{rwu}\right)\phi_{ur} \tag{SUB4}$$

Subject to $\phi_{ur} \in 0, 1$.

This problem can be further decomposed into $|U||R|$ independent subproblems, each of which may be solved optimally by a simple algorithm. For each communication node n in round r, the coefficient $(\mu^5_u\left(E_r+E_s\right) - \sum_{w\in W}\mu^1_{rwu})$ is computed. Then, if the coefficient is negative, ϕ_{nr} is set to 1; otherwise, ϕ_{nr} is set to 0. The computational complexity is $O(|W|)$ for each communication node.