Journal of the Operations Research
Society of Japan
2009, Vol. 52, No. 4, 366-376

A NEW CLUSTERING MODEL OF WIRELESS SENSOR NETWORKS USING FACILITY LOCATION THEORY

Takehiro Furuta — *Tokyo University of Science*
Mihiro Sasaki, Fumio Ishizaki — *Nanzan University*
Atsuo Suzuki, Hajime Miyazawa — *Nanzan University*

(Received November 20, 2008; Revised March 24, 2009)

Abstract In this paper, we study mathematical formulations for clustering problems which arise in wireless sensor networks as examined from the standpoint of *facility location theory*. Following facility location theory, LEACH-C, one of the principal studies on cluster-based network organizations, formulates the clustering problem as a p-median problem. In this paper, we examine some drawbacks to the formulation put forward in LEACH-C. We then formulate the problem as an uncapacitated facility location problem to overcome these drawbacks. Computational experiments show that compared to LEACH-C, the proposed algorithm based on our formulation can extend the total lifetime of sensor networks.

Keywords: Mathematical modeling, facility location theory, p-median problem, UFLP, wireless sensor network

1. Introduction

Wireless sensor networks receive significant attention due to their rich applications in scientific, medical, commercial and military domains. A wireless sensor network is formed by tens to thousands of sensor nodes randomly deployed in a target field. Sensor nodes are organized into ad hoc networks and send information about monitored events to a data sink or to a remote base station (BS) through the organized network.

One of the crucial challenges in the organization of sensor networks is energy efficiency. This need for energy efficiency arises because sensor node battery capacities are severely limited and battery replacement is impractical. The sensor node battery constraint limits the functional life of the network. The functional life of each individual node varies based upon the demands placed on its battery. Thus, an important characteristic in the design of sensor networks is their robustness in face of the demise of individual sensor nodes. Various network architectures and protocols to save energy consumption and to extend sensor network lifetimes have been studied (e.g., see [4] and the references therein). Among these architectures and protocols, cluster-based network organizations are considered to be the most favorable approach in terms of energy efficiency. In this approach, sensor nodes are organized into clusters, and one sensor node in each cluster is selected as the cluster head (CH) which then plays a special role as a transfer point (see Figure 1). Additionally, each CH creates a transmission schedule for the sensor nodes within the cluster. This schedule allows the radio components of each non-CH-node to be powered down except during scheduled transmit times.

Rotation of CH duties among the sensor nodes within the cluster becomes an important factor in organizing sensor networks. Since the BS is generally far away from the sensor field, data transmission to the BS consumes significant energy from the CH batteries. Therefore, if the same node continuously works as a CH, it will die quickly. In response to this need to not to drain the battery power of a single sensor, clustering algorithms introduce a CH duty rotation among the sensor nodes in a cluster.

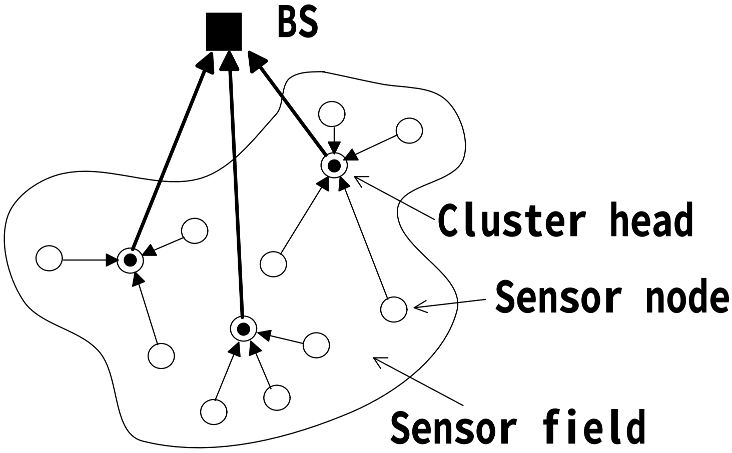

Figure 1: Cluster-based sensor network

The foundational studies on cluster-based network organizations are LEACH (Low Energy Adaptive Clustering Hierarchy) [2] and LEACH-C (LEACH-Centralized) [3]. LEACH is based on self-organizing networks. LEACH-C is a centralized cluster formation version of LEACH, wherein the BS organizes and controls the network. More precisely, the LEACH-C protocol provides for centralized cluster formation, for local processing for the aggregation of sensing data, and for the rotation of CHs for each round. These activities are designed to achieve uniform energy consumption among sensor nodes and to maximize network lifetime. Since the BS usually does not have any energy constraints, centralized cluster formation methods can be attractive alternatives to non-centralized methods. Heinzelman *et al.* demonstrated LEACH-C to be more efficient than LEACH in terms of energy consumption based on computational results [3].

Recently, researchers have applied facility location theory and mathematical programming approaches to sensor network problems. Patel *et al.* [7] introduced a variation of a maximal expected covering location model considering the relocation costs of CHs. Alfieri *et al.* [1] proposed a mixed integer linear programming model to determine the active set of sensor nodes so as to cover a given ratio of the sensor area. Krivitski *et al.* [5] proposed a heuristic algorithm called local algorithm for solving the clustering problem. They solved large size p-median problems by the algorithm without considering battery expense of sensors.

In this paper, we study centralized clustering algorithms for wireless sensor networks from the standpoint of *facility location theory* and using a mathematical programming approach. From this standpoint, we consider LEACH-C as formulating the clustering problem as a p-median problem [6], one of the well-known facility location problems. We point out drawbacks of the formulation used in LEACH-C. For example, LEACH-C does not consider critical factors to decide CHs such as the distance from CHs to the BS, energy consumption caused by receiving and aggregating data, the number of CHs to be selected and so on. To overcome the drawbacks of LEACH-C, we formulate the clustering problem as an uncapacitated facility location problem (UFLP) [6] incorporating sensor node battery levels and energy consumptions that have not been taken into account in LEACH-C. The UFLP finds the optimal number of facilities and their locations to minimize the sum of fixed costs to locate the facilities and transportation costs under the assumption that the facilities have no capacity limit. This paper relies on the basic network setting assumptions inherent in LEACH-C to provide for a formulation improvement for networks which could utilize LEACH-C. These basic network setting assumptions include the assumptions that all sensor nodes can connect to the BS and to other sensor nodes.

The remainder of this paper is organized as follows. In Section 2, we point out drawbacks in the clustering algorithm of LEACH-C, and offer some basic ideas for overcoming these drawbacks. In Section 3, we formulate the clustering problem as a UFLP. Based on this formulation, we propose a new clustering algorithm for wireless sensor networks. In Section 4, we show computational results comparing the useful lifetimes of sensor networks. Finally, in Section 5 we offer concluding remarks.

2. Centralized Cluster-based Sensor Networks

The operation of cluster-based sensor networks is usually divided into *rounds*. Each round has two phases, which are the clustering phase and the data transmission phase. Rounds are repeated to continuously monitor events. We briefly describe the centralized cluster formation algorithm employed in LEACH-C below.

In the clustering phase of LEACH-C, each sensor node first reports its location (this information may be obtained by a GPS receiver) and its battery level to the BS. The BS computes the average battery level of the nodes and selects sensor nodes which have above-average battery levels as CH candidates. Finally, the BS determines a cluster formation by solving a p-median problem where the objective is to minimize the sum of the squared distances from each sensor node to the nearest CH and then assigns the CHs.

In the data transmission phase of LEACH-C, each non-CH node sends monitored data to its CH. After the CH receives all the data from the sensor nodes within its cluster, a *data aggregation* process can be carried out by the CH. Since the data monitored by each sensor node within a cluster are often correlated or redundant, the BS does not require all the data. The data aggregation process removes redundant data and reduces the size of data sent to the BS. The data aggregation process consequently reduces the energy consumption used for data transmission by the CHs.

Heinzelman *et al.* [3] reported that LEACH-C performs better than LEACH in their simulation. Although LEACH-C performs better than LEACH, there are drawbacks in the clustering algorithm of LEACH-C. Those drawbacks are listed below.

- The LEACH-C algorithm only considers the energy consumed by data transmission, although CHs also expend battery power while receiving and aggregating data.
- Although energy consumption of the sensor nodes is important consideration, the LEACH-C algorithm in data transmission phase does not take this into account directly. The algorithm uses the squared distances as a measure to evaluate the energy consumption. Since energy consumption is proportional to the squared distance, using squared distances could be possible. However, the LEACH-C algorithm only focuses on the squared distances between sensor nodes and the nearest CHs, and does not focus on between the CHs and the BS that significantly impact energy consumption due to much longer distances.
- In the LEACH-C algorithm, the number of CHs is predetermined and fixed throughout the rounds. As the rounds continue and battery levels fall, the number of CH candidates

may be less than the fixed number required by the algorithm. As a result, although many sensor nodes are still functional, the p-median problem will become infeasible due to a shortage of CH candidates. When this happens, it is no longer possible to continue rounds using this algorithm unless CHs are selected in an alternative way.

To overcome the drawbacks mentioned above, we formulate the clustering problem as a UFLP where the objective is to maximize the total battery level of all sensor nodes by taking into account all types of energy consumption. The facilities correspond to the CHs and the fixed cost for each sensor node corresponds to the energy consumption required to transmit data to the BS as a CH. By formulating the problem as a UFLP, the number of CHs selected in each round is more flexible than it is using the p-median problem. This alternative algorithm can extend the useful life of the network.

3. Formulation

As mentioned in the previous section, each round has two phases. The first phase is the clustering phase, in which the clustering problem is solved. In this section, we formulate the clustering problem as a UFLP, and show the p-median problem formulation used in LEACH-C. The following notations are employed:

N: the set of sensor nodes,
d_{ij}: the distance from sensor node $i \in N$ to sensor node $j \in N$ (m),
f_i: the distance from sensor node $i \in N$ to the BS (m),
b_i: the battery level of sensor node $i \in N$ (J),
l: the data size sent by a sensor node (bit),
E: the coefficient for the radio dissipate to run the transmitter or receiver circuitry (J/bit),
E_{DA}: the coefficient for data aggregation (J/bit),
n: the number of sensor nodes which have positive battery level,
α: the parameter to determine CH candidates ($0 < \alpha \leq 1$),
S_i: 0 if sensor node i has a positive battery level and 1 otherwise.

For the purpose of this study, we assume that every sensor node sends a fixed length message (l bits) in each round. Regarding data aggregation, we adopt perfect data aggregation as LEACH-C does, wherein the received messages are aggregated into a single message at the CHs regardless of the number of messages received. As a result, every CH sends l bits of data to the BS. The parameter α is introduced to allow more flexible CH candidate selection. Note that $\alpha = 1$ throughout in LEACH-C, which means that battery levels of CH candidates are greater than or equal to the average battery level of all sensor nodes. We determine CH candidates using the value of α times the average battery level. For example, if $\alpha = 0.5$, then battery levels of CH candidates are greater than or equal to 0.5 times the average battery level. As mentioned in Section 1, CHs perform not only the role of transfer points but also other roles in the control of sensor networks. However, for the purpose of comparison with LEACH-C, we omit these factors and assume that CHs perform only the role of transfer points.

Again, for the purpose of comparison with LEACH-C, we assume the same models of energy consumption for transmitting and receiving data as LEACH-C [3]. Amplifier energy used for data transmission is defined by two models depending on the distance between the sensor nodes. If the distance is less than the threshold distance d_0, we use the free space model [8, 9]. In all other instances, we use the multi-path model [8, 9]. In the former model, energy consumption is proportional to the squared distance, and in the latter model, energy consumption is proportional to biquadrate distance. Amplifier energy used for data transmission from sensor node i to j is given by,

$$D_{ij} = \begin{cases} E + \epsilon_{fs} {d_{ij}}^2 & (\text{if } d_{ij} < d_0) \\ E + \epsilon_{mp} {d_{ij}}^4 & (\text{if } d_{ij} \geq d_0), \end{cases}$$

and that from sensor node i to the BS is given by,

$$F_i = \begin{cases} E + \epsilon_{fs} {f_i}^2 & (\text{if } f_i < d_0) \\ E + \epsilon_{mp} {f_i}^4 & (\text{if } f_i \geq d_0), \end{cases}$$

where ϵ_{fs} (pJ/bit/m^2) and ϵ_{mp} (pJ/bit/m^4) are the coefficients for the two models, respectively. Amplifier energy used for data reception from a sensor node is lE. Note that E is a fixed energy consumption. Since every sensor node including those serving as CHs sends l bits of data, it consumes lD_{ij} or lF_i joules of energy.

We further introduce the following decision variables.

x_i: binary variable such that $x_i = 1$ if sensor node $i \in N$ is selected as a CH, and $x_i = 0$ otherwise.

y_{ij}: binary variable such that $y_{ij} = 1$ if sensor node $i \in N$ belongs to the cluster where sensor node $j \in N$ is a CH, and $y_{ij} = 0$ otherwise.

To improve network efficiency, we propose a new formulation for the clustering problem of sensor networks. The clustering problem is formulated as the following integer programming problem:

$$\text{maximize} \quad \sum_{i \in N} \left\{ b_i - \left(l \sum_{j \in N} D_{ij} y_{ij} + lF_i x_i \right) - lE \sum_{j \in N} y_{ji} - lE_{DA} \sum_{j \in N} y_{ji} \right\} \tag{3.1}$$

$$\text{s.t.} \quad x_i + \sum_{j \in N} y_{ij} + S_i = 1, \quad i \in N, \tag{3.2}$$

$$\left(b_i - \frac{\alpha}{n} \sum_{k \in N} b_k \right) x_i \geq 0, \quad i \in N, \tag{3.3}$$

$$y_{ij} \leq x_j, \quad i, j \in N, \tag{3.4}$$

$$x_i \in \{0, 1\}, \quad i \in N, \tag{3.5}$$

$$y_{ij} \in \{0, 1\}, \quad i, j \in N. \tag{3.6}$$

The objective of this model is to maximize the sum of sensor node battery levels after each round. Each term in brackets of objective function (3.1) is described as follows. The second term enclosed within parentheses represents the total energy consumption of sensor node i used for data transmission. The third and the fourth terms represent the energy consumptions of sensor node i used for data reception and for data aggregation, respectively. From constraint (3.2), each sensor node either plays the role of a CH or sends data to the nearest CH as long as its battery level is positive. Constraints (3.3) ensure that each sensor node which has at least α times as much as the average battery level of all live sensor nodes will be a candidate to be a CH. Constraint (3.4) states that only CHs can receive data.

Note that objective (3.1) is the maximization of the total sum of the battery levels of sensor nodes and can be rewritten in the standard form of the objective in the UFLP:

$$\text{minimize} \sum_{i \in N} \sum_{j \in N} (E + D_{ij} + E_{DA}) y_{ij} + \sum_{i \in N} F_i x_i.$$

Hence, the problem can be regarded as a UFLP.

In a similar manner, we can formulate the clustering problem in LEACH-C as the following p-median problem:

$$\begin{aligned} \text{minimize} \quad & \sum_{i \in N} \sum_{j \in N} (d_{ij})^2 y_{ij} \\ \text{s.t.} \quad & \sum_{j \in N} x_j = p, \\ & (3.2), (3.3), (3.4), (3.5), (3.6). \end{aligned}$$

Note that in LEACH-C, the objective is to minimize the total sum of squared distances between the sensor nodes and the nearest CHs, and energy consumption is not directly considered. Also, the parameter α is fixed and equal to 1 in constraint (3.3), which means that only sensor nodes that have a battery level greater than the average level of all live sensor nodes can be CH candidates.

4. Computational Experiments

We made computational experiments to compare the performance of our clustering algorithm with that of LEACH-C. Although Heinzelman *et al.* [3] used heuristic solutions in the simulated annealing algorithm, we used exact solutions of the clustering problem in the experiments. Our reason for using exact solutions rather than heuristic solutions was that our objective was to examine how long we can extend the network lifetime by using our UFLP based formulation for the clustering problem. We used an optimization software Xpress-MP (2005B) to obtain exact solutions. All experiments were run on a PC with Intel Pentium 4 processor (2.53GHz) and 512MB RAM.

We used the same physical constants and parameters as LEACH-C [3] in the experiments: $b_i = 0.5$ J, $E = 50$ nJ/bit, $\epsilon_{fs} = 10$ pJ/bit/m^2, $\epsilon_{mp} = 0.0013$ pJ/bit/m^4, $E_{DA} = 5$ nJ/bit, $d_0 = 87$ m, and $l = 4200$ bits. We solve the UFLP based formulation with various values of α from 0.1 to 1.0. In addition, as is done in LEACH-C, we solve the p-median problems with $p = 5$ and $\alpha = 1$ for comparison purposes. We used two types of data sets. In the first data set, 100 sensor nodes were randomly deployed in a square of 100 by 100 meter (data 1, $\cdots$, data 5). Figure 2 shows sensor node locations of data 5. In the second data set, 100 sensor nodes were deployed in a square of 400 by 400 meter (data 6, $\cdots$, data 10). For convenience, we define the lower left of the squares as ($x = 0$, $y = 0$), and the upper right

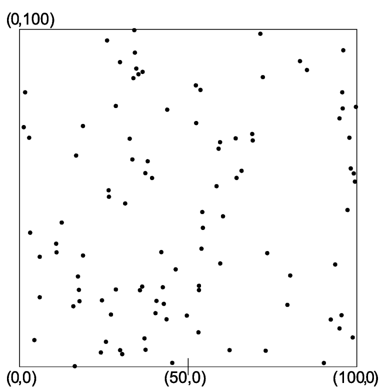

Figure 2: This figure shows sensor node locations of data 5. Black circles denote sensor nodes which are randomly deployed in a square of 100 by 100 meter

Table 1: Comparisons of the number of rounds for each survival rate (s.r.) between our UFLP-based formulation with $\alpha = 1.0$ (UFLP) and the p-median based formulation in LEACH-C ($p$med) in a 100 meters square

	data 1		data 2		data 3		data 4		data 5	
s.r.	$p$med	UFLP	$p$med	UFLP	$p$med	UFLP	$p$med	UFLP	$p$med	UFLP
99%	902	902	611 (100)	902	907	916	908	920	888	902
90%	907	922	N/A	919	923	933	919	940	896 (93)	915
70%	908 (90)	933	N/A	933	924 (88)	950	920 (89)	950	N/A	933
50%	N/A	939	N/A	939	N/A	957	N/A	959	N/A	943
30%	N/A	944	N/A	944	N/A	961	N/A	965	N/A	948
10%	N/A	952	N/A	954	N/A	969	N/A	972	N/A	957
0%	N/A	963	N/A	960	N/A	981	N/A	979	N/A	963

Table 2: Comparisons of the number of rounds for each survival rate (s.r.) between our UFLP-based formulation with $\alpha = 1.0$ (UFLP) and the p-median based formulation in LEACH-C ($p$med) in a 400 meters square

	data 6		data 7		data 8		data 9		data 10	
s.r.	$p$med	UFLP	$p$med	UFLP	$p$med	UFLP	$p$med	UFLP	$p$med	UFLP
99%	38	49	31	47	36	53	50	59	32	42
90%	53	70	61	74	50	68	76	81	62	74
70%	105	119	102 (78)	135	131 (78)	171	103	113	90 (76)	109
50%	235	245	N/A	180	N/A	247	179	201	N/A	151
30%	254 (44)	359	N/A	291	N/A	342	183 (50)	392	N/A	271
10%	N/A	441	N/A	423	N/A	463	N/A	545	N/A	502
0%	N/A	526	N/A	455	N/A	469	N/A	550	N/A	575

of the 100 meter square as ($x = 100$, $y = 100$) and of the 400 meter square as ($x = 400$, $y = 400$). We assign the BS location at ($x = 50$, $y = 175$) in the data sets for the 100 meter square and at ($x = 200$, $y = 475$) in the data sets for the 400 meter square. The average computational time of our clustering algorithm using Xpress-MP is 4.76 seconds per round.

For the sake of data thoroughness, we simulated data transmission from every node to the BS until all sensor nodes died even though most networks would deteriorate below useful levels before that. To evaluate the performance of the clustering algorithms, we introduce the survival rate, which is defined as the percentage of functional, or live, sensor nodes over all sensor nodes. Table 1 shows comparisons of the numbers of rounds between our UFLP based formulation with $\alpha = 1.0$ and the p-median based formulation in LEACH-C for the data sets for the 100 meter squares. Table 2 shows the comparisons for the data sets for the 400 meter squares. Each table shows the numbers of rounds operated until survival rates decrease to 99%, 90%, 70%, 50%, 30%, 10%, and 0%. We use 100 sensor nodes to develop the data in the experiments. Therefore a 99% survival rate means that the first sensor node has died. In other words, the numbers that appear in the row labeled "99%" show the number of rounds that are predicted to be run while all sensor nodes remain functional. Note that the numbers appearing in the row labeled "0%" express the theoretical maximum network lifetime. The number in the parenthesis that appears in the "$p$med" columns is

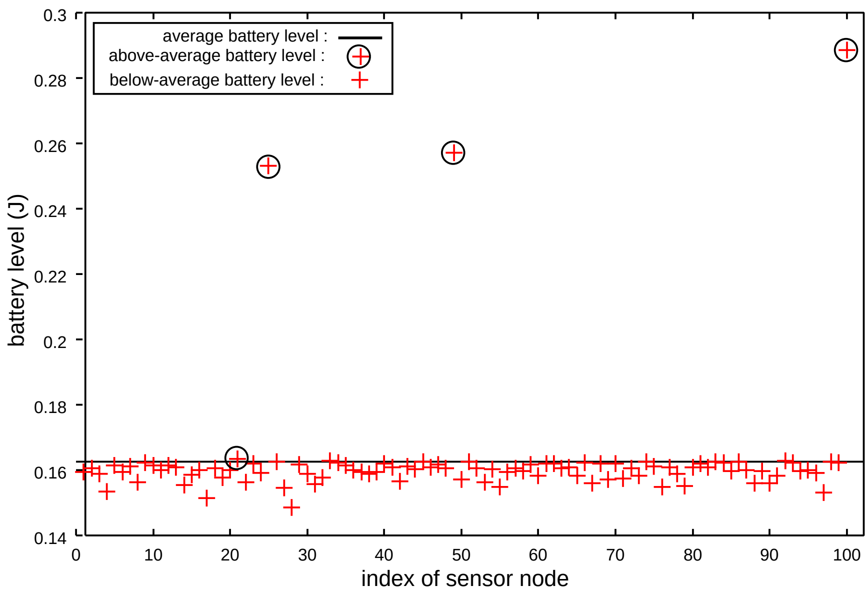

Figure 3: The battery level of each sensor node at the 611th round of data 2 (100×100)

the number of nodes still functional when the clustering problem becomes infeasible. For instance, 908(90) means that the problem becomes infeasible at the 908th round and that ninety of the sensor nodes are still functioning.

The p-median problem in LEACH-C becomes infeasible due to a lack of CH candidates. Tables 1 and 2 show that many sensor nodes are still functioning when the problem becomes infeasible. Of particular note in experimental data set 2, all sensor nodes are still functioning when the problem becomes infeasible at the 611th round. In other words, there would be less than five sensor nodes that have above-average energy among the 100 sensor nodes. This seems unlikely so we examine what happens at the 611th round.

Figure 3 shows the battery level for each sensor node at the 611th round in data set 2. The horizontal line denotes the average remaining battery level. The circled crosses indicate those battery levels that are above the average. The data indicates that only four sensor nodes would be available to serve as CH candidates. However, the p-median problem requires a minimum of five candidates in order to select five CHs. The data also shows that three of the four remaining candidates have significantly larger remaining battery levels than the other sensor nodes. Through further investigation, we found that three of those nodes with high battery levels are physically closer to the BS. As a result, they do not expend as large an amount of energy in sending data to the BS when they are selected to serve as CHs. However, in spite of the advantage that their location provides vis-a-vis data transmission efficiency, they are selected to serve as CHs only at about the same frequency as other sensor nodes that lack this location advantage. In short, though they have sufficient battery levels to take on the CH role more frequently, they actually play a CHs role less frequently than what would optimize the use of their energy resources within the network. In this case, a relatively few sensor nodes may have much larger remaining battery levels and may thereby skew the average battery level. This skewing of the average results in more of the sensor nodes being calculated as having below-average energy. This then leads to the shortage of CH candidates and makes the problem infeasible earlier than it could be when there are many live nodes remaining.

As an alternative, we propose a UFLP based formulation in which the number of CHs

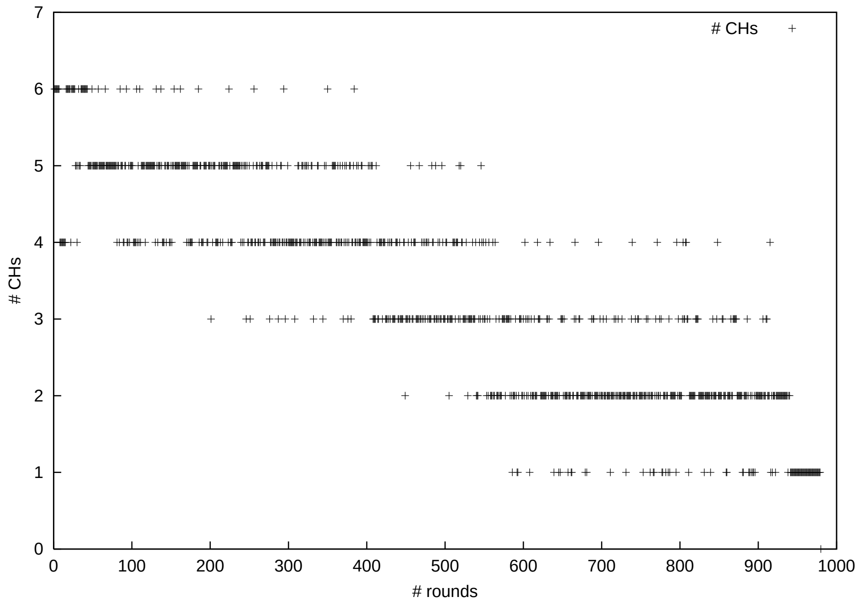

Figure 4: The number of CHs in each round with the case of data1 (100×100) and $\alpha = 0.9$

is dynamically optimized to maximize the total battery level in each round as shown in Figure 4. Using this formulation, the problem remains feasible even after quite a large number of rounds, although the sensor nodes do not uniformly expend their batteries.

In Tables 1 and 2, we observe that, in terms of total network lifetime, our UFLP based formulation is more suitable than the p-median based formulation proposed in LEACH-C. Using our formulation, we obtain exact solutions to the clustering problems without becoming infeasible until all sensor nodes die. The numbers of rounds using our formulation are greater than or equal to the number of rounds using the p-median based formulation at any of the survival rates where we made an observation.

To see how the number of rounds changes with various values of α using our formulation, we simulated data transmission from every node to the BS until all sensor nodes died. Table 3 shows the computational results for the data sets for the 100 meter square and Table 4 shows the results for the data sets for the 400 meter square. In Tables 3 and 4, the columns labeled "ave." indicate the average number of rounds run until survival rates decrease to 99%, 90%, 70%, 50%, 30%, 10%, and 0%. The columns labeled "std." indicate the standard deviation. The average numbers of rounds are calculated as the average over the five data sets in each type.

From Tables 3 and 4, we can see that the number of rounds is sensitive to the value of α. The number of CH candidates in each round increases as the value of α decreases. As the value of α decreases, the total network lifetime is extended, but the number of rounds it takes for the network to drop to a 70% survival rate decreases. Thus, when the value of α is small, many sensor nodes die quickly, but the total network lifetime tends to be extended. On the other hand, when the value of α is large, each sensor node tends to last longer, but many sensor nodes tend to die simultaneously and the total network lifetime tends to be shorter. From these observations, we can select an appropriate value for α according to required characteristics of the sensor networks. For example, if we need at least a 70% survival rate to monitor events in the sensor field, we would set a relatively large value for α. If the network lifetime is more important than the number of live nodes, we would set a relatively small value for α.

Table 3: The number of rounds vs. survival rate (s.r.) of nodes : 100×100

	α											
	0.1		0.3		0.5		0.7		0.9		1.0	
s.r.	ave.	std.	ave.	std.	ave.	std.	ave.	std.	ave.	std.	ave.	std.
99%	227.6	17.47	641.6	11.22	903.6	19.27	905.0	12.02	912.8	13.50	908.4	8.88
90%	390.8	25.16	706.0	4.53	922.6	13.94	925.6	14.67	930.4	12.07	925.8	10.38
70%	691.0	23.11	795.4	10.71	933.0	13.78	939.8	12.60	943.8	11.21	939.8	9.31
50%	969.6	14.12	892.6	12.18	943.8	12.21	948.6	13.74	951.2	11.30	947.4	9.84
30%	1222.8	19.41	1095.0	30.65	954.6	11.35	958.6	13.79	957.4	10.71	952.4	9.91
10%	1460.4	28.92	1276.0	29.33	985.8	14.86	977.2	9.68	967.8	11.21	960.8	9.09
0%	1501.0	29.84	1347.4	22.43	1123.6	12.24	1053.6	13.28	988.2	11.14	969.2	9.96

Table 4: The number of rounds vs. survival rate (s.r.) of nodes : 400×400

	α											
	0.1		0.3		0.5		0.7		0.9		1.0	
s.r.	ave.	std.	ave.	std.	ave.	std.	ave.	std	ave.	std.	ave.	std.
99%	20.4	7.70	60.8	14.17	61.0	9.30	57.8	8.90	51.6	7.13	50.0	6.40
90%	67.4	5.77	93.6	10.55	100.8	10.13	95.0	6.40	83.2	6.61	73.4	4.98
70%	118.8	23.15	144.2	29.39	148.4	30.26	145.0	29.01	136.6	28.37	129.4	25.27
50%	176.4	25.73	197.2	39.40	217.2	47.86	218.2	45.89	211.6	39.37	204.8	41.60
30%	278.2	51.06	312.6	52.66	334.4	42.34	342.4	45.40	335.6	51.29	331.0	49.56
10%	495.8	50.48	503.6	43.32	492.6	70.56	489.0	54.36	478.8	51.86	474.8	49.07
0%	813.0	92.43	728.8	92.88	691.8	76.76	615.6	81.86	567.2	79.75	515.0	51.63

5. Conclusion

In this paper, we formulate the clustering problem as a UFLP and develop a centralized method for determining good clustering with the objective of maximizing the lifetime of the total network. In LEACH-C, a foundational study on clustering problems, the number of cluster heads is predetermined but over time, sensor node battery levels and the number of live sensor nodes change. Depending on location of the sensor nodes and their battery levels, the p-median problems in LEACH-C may become infeasible relatively quickly because the number of cluster head candidates becomes smaller than the predetermined number of cluster heads.

In contrast, using our formulation as a UFLP, we can find an optimal number of cluster head and select the optimal cluster head candidates at the same time. Furthermore, using ten random data sets, the numbers of rounds at any survival rates in our computer experiments using our UFLP based formulation are greater than or equal to the number of rounds those using the p-median based formulation in LEACH-C. In these computational experiments, we use 100 sensor nodes examples and obtain exact solutions using Xpress-MP.

In this paper, we focused on the number of live sensor nodes and the network lifetime to evaluate our algorithm. Some factors to develop more practical clustering still remain unconsidered. For example, to develop a clustering method so that live sensor nodes cover the whole target area as long as possible is one of the most important future work.

Acknowledgments

The authors would like to thank two anonymous referees for their valuable suggestions and comments.

References

[1] A. Alfieri, A. Bianco, P. Brandimarte, and C. F. Chiasserini: Maximizing system lifetime in wireless sensor networks. *European Journal of Operational Research*, **181** (2007), 390–402.

[2] W.R. Heinzelman, A. Chandrakasan, and H. Balakrishnan: Energy-efficient communication protocol for wireless microsensor networks. *Proceedings of the 33rd Hawaii International Conference on System Sciences* (Maui, Hawaii, January 2000).

[3] W.B. Heinzelman, A.P. Chandrakasan, and H. Balakrishnan: An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, **1** (2002), 660 – 670.

[4] M. Kochhal, L. Schwiebert, and S. Gupta: Self-organizing of wireless sensor networks. In J. Wu (ed.): *Handbook on Theoretical and Algorithmic Aspects of Sensor, Ad Hoc Wireless, and Peer-to-Peer Networks* (Auerbach Publications, 2006), 369–392.

[5] D. Krivitski, A. Schuster, and R. Wolff: A local facility location algorithm for sensor networks. *Distributed Computing in Sensor System, First IEEE International Conference, DCOSS 2005* (Marina del Rey, CA, June 30 - July 1, 2005).

[6] P.B. Mirchandani and R.L. Francis: *Discrete Location Theory* (Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley & Sons, New York, 1990).

[7] D.J. Patel, R. Batta, and R. Nagi: Clustering sensors in wireless ad hoc networks operating in a threat environment. *Operations Research*, **53** (2005), 432–442.

[8] T. Rappaport: *Wireless Communications: Principles & Practice* (Prentice Hall, New Jersey, 1996).

[9] D. Tse and P. Viswanath: *Fundamentals of Wireless Communication* (Cambridge University Press, Cambridge, 2005).

Takehiro Furuta
Faculty of Engineering
Tokyo University of Science
1-14-6, Kudankita, Chiyoda,
Tokyo 102-0073, Japan
E-mail: takef@fw.ipsj.or.jp