NANZAN-TR-2006-01

A New Cluster Formation Method for Sensor Networks Using Facility Location Theory

Takehiro Furuta, Mihiro Sasaki, Fumio Ishizaki,
Atsuo Suzuki, Hajime Miyazawa

August 2006

Technical Report of the Nanzan Academic Society
Mathematical Sciences and Information Engineering

A New Cluster Formation Method for Sensor Networks Using Facility Location Theory

Takehiro Furuta
Nanzan University
takef@fw.ipsj.or.jp

Mihiro Sasaki
Nanzan University
mihiro@nanzan-u.ac.jp

Fumio Ishizaki
Korea University
fumio@ieee.org

Atsuo Suzuki
Nanzan University
atsuo@nanzan-u.ac.jp

Hajime Miyazawa
Nanzan University
miyazawa@nanzan-u.ac.jp

Abstract

In this letter, we study cluster formation methods for wireless sensor networks from a view of *facility location theory*. From a view of facility location theory, we can consider that LEACH-C, which is one of the preliminary studies on cluster-based network organization, formulate the cluster formation as a p-median problem. We point out difficulties in the formulation of the p-median problem. To overcome the difficulties, we formulate the cluster formation as an uncapacitated facility location problem (UFLP). Based on the formulation, we propose a new cluster formation method for sensor networks. Computational experiments exhibit that the proposed method can extend the network lifetime of sensor networks, compared to LEACH-C.

Keyword: Wireless sensor network, location theory, LEACH-C, mathematical programming

1 Introduction

Wireless sensor networks have been paid much attention due to their rich applications in the scientific, medical, commercial and military domain. A wireless sensor network is formed by tens to thousands sensor nodes randomly deployed in a target field. Sensor nodes should be first organized into an ad hoc network. Sensor nodes then send information on monitored event to a data sink or a remote base station (BS) through the organized network.

One of crucial challenges in the organization of sensor networks is energy efficiency, because battery capacities of sensor nodes is severely limited and replacing the batteries is not practical. Various network architectures and protocols to efficiently organize sensor networks have been studied (e.g., see [1] and references therein). Among them, cluster-based network organization is considered as the most favorable approach to the problem of energy efficiency. In this approach, sensor nodes are organized into clusters, and in each cluster, one sensor node called cluster head (CH) is selected to play a special role (see Figure 1). Each CH creates a schedule for the nodes in its cluster. This allows the radio components of each non-CH-node to be turned off all times except during its transmit time. Once each CH has all data from the nodes in its cluster, it aggregates the data and then transmits the compressed data to BS. Since the BS is far away, CHs will die quickly, because CHs exhaust much amount of energy for the transmission of the compressed data to the BS. Thus, in order not to drain the battery power of a single sensor, some cluster formation algorithms include the rotation of CH position among sensor nodes.

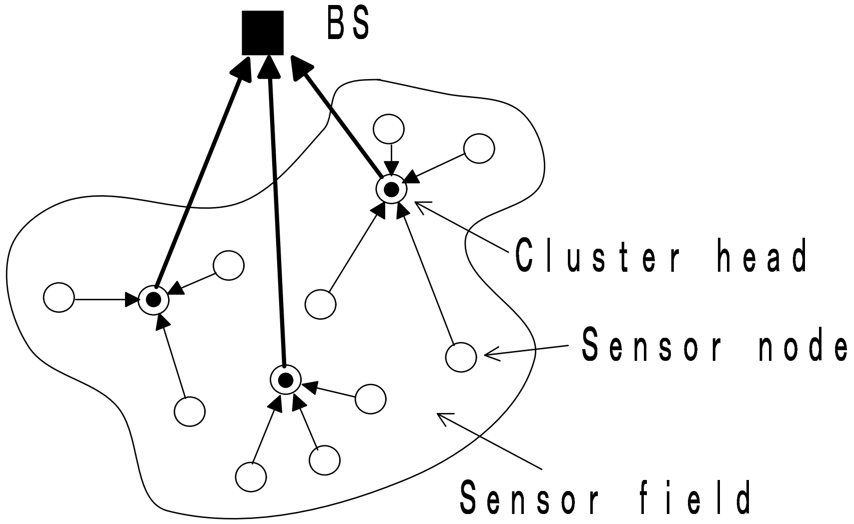

Figure 1: Cluster based sensor network

One of the preliminary studies on cluster-based network organization is LEACH (Low Energy Adaptive Clustering Hierarchy) [2] and LEACH-C (LEACH-centralized) [3]. LEACH-C is a centralized cluster formation version of LEACH, where the BS organizes the network. Since the BS usually does not have energy constraint as sensor nodes, centralized cluster formation methods can be attractive alternatives if the BS can find a good cluster formation in a reasonable time.

In this letter, we study cluster formation methods for wireless sensor networks from a view of *facility location theory*. From a view of facility location theory, we can consider that in LEACH-C, the cluster formation is formulated as a p-median problem [4], which is one of well-known facility location problems. We point out difficulties in the formulation of the p-median problem. To overcome the difficulties, we formulate the cluster formation as an uncapacitated facility location problem (UFLP) [4]. Based on the formulation, we propose a new cluster formation method for sensor networks.

2 Cluster formation in LEACH-C

The centralized cluster formation in LEACH-C operates as follows [3]: In the cluster formation phase of LEACH-C, each sensor node first reports information about its current 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 nodes, and as the candidates of CH, the BS selects sensor nodes whose battery level is greater than or equal to the average battery level. Finally the BS determines CHs and finds a cluster formation by solving a p-median problem where the objective is to minimize the sum of the squared distances from each sensor nodes to its nearest CH. (for the formulation of the p-median problem, see Section III.) The p-median problem is solved in each round.

We point out that the formulation of the p-median problem in LEACH-C has the following difficulties:

- Although it is important to consider energy consumption or remaining battery level of each node in order to achieve a long network lifetime, LEACH-C does not directly consider them. Instead, LEACH-C considers only the distances from a sensor node to its CH to determine CHs. However, it does not consider the distances from CHs to the BS. Since usually the distances from CHs to the BS is much longer than those from sensor nodes to their CHs, these distances are critical factor governing the network lifetime.
- LEACH-C considers only the energy consumption of data transmission to select CHs, although CHs also use their battery powers for receiving and aggregating data.
- LEACH-C predetermines the number of CHs and fixes it through all rounds. This can cause that the p-median problem becomes infeasible due to the lack of CH candidates, even though many sensor nodes are still alive.

To overcome the above difficulties, we formulate the cluster formation as the UFLP where the objective is to maximize the total battery level of all sensor nodes and all kinds of energy consumption are taken into account.

3 Formulation

In this section, we provide the mathematical formulation of the UFLP in our method. For comparison, we also provide the mathematical formulation of the p-median problem in LEACH-C. We employ the following notations:

N**:** the index set of sensor nodes,

d_{ij}**:** the distance from sensor node $i \in N$ to sensor node $j \in N$,

f_i**:** the distance from sensor node $i \in N$ to the BS,

b_i**:** the battery level of sensor node $i \in N$ (J),

l**:** the data amount sent by each sensor node (bit),

E**:** coefficient for the radio dissipate (J/bit),

E_{DA}**:** coefficient for data aggregation (J/bit/signal),

n**:** the number of sensor nodes which have a positive battery level,

α**:** parameter to determine CH candidates ($0 < \alpha \le 1$),

S_i**:** 0 if sensor node i has a positive battery level and 1 otherwise.

We assume that the model of the energy consumption of transmitting and recieving the data is the same as LEACH-C [3]. Amplifier energy used for data transmission is defined by two models depending on the distance between the sensor nodes [3]. If the distance is less than the distance threshold d_0, we use the free space model, otherwise, we use the multi-path model. Amplifier energy used for data transmission from sensor node i to j (D_{ij}) and that from sensor node i to the BS (F_i) are given by

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

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

where ϵ_{fs} and ϵ_{mp} are coefficients for the two models, respectively.

We further introduce the following decision variables for $i, j \in N$.

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

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

We now propose a new formulation for cluster formation in sensor networks. The cluster formation is formulated as the following integer programming problem:

$$\max \sum_{i \in N} \left\{ b_i - (lE + l \sum_{j \in N} D_{ij} y_{ij} + l F_i x_i) - lE \sum_{j \in N} y_{ji} - E_{DA} \sum_{j \in N} y_{ji} \right\} \tag{1}$$

$$\text{s.t.} \quad x_i + \sum_{j \in N} y_{ij} + S_i = 1, \quad i \in N, \tag{2}$$

$$\left(b_i - \frac{\alpha}{n} \sum_{i \in N} b_i \right) x_i \ge 0, \quad i \in N, \tag{3}$$

$$y_{ij} \le x_j, \quad i, j \in N, \tag{4}$$

$$x_i \in \{0, 1\}, \quad i \in N, \tag{5}$$

$$y_{ij} \in \{0, 1\}, \quad i \in N, j \in N. \tag{6}$$

From constraint (2), each sensor node plays a CH or sends the data to a CH as far as its battery level is positive. Constraint (3) ensures that sensor node whose battery level is at least α times as much as the average battery level can be the candidates of CH. Constraint (4) means that only CHs can receive the data.

Note that the objective (1) is to maximize the total sum of battery level of sensor nodes and it can be rewritten to the standard form of the objective in the UFLP as

$$\min \sum_{i \in N} \sum_{j \in N} (lE + lD_{ij} + E_{DA}) y_{ij} + l \sum_{i \in N} F_i x_i.$$

Hence, the problem can be regarded as the UFLP.

In a similar manner, we can formulate the cluster formation in LEACH-C as the following integer programming problem:

$$\begin{aligned} \min \quad & \sum_{i \in N} \sum_{j \in N} (d_{ij})^2 y_{ij} \\ \text{s.t.} \quad & \sum_{j \in N} x_j = p, \\ & (2), (3), (4), (5). \end{aligned}$$

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 the energy consumption is not directly considered. Also, the parameter α is fixed and equal to 1 in constraint (3).

4 Computational Experiments

Table 1: Average number of rounds vs. survival rate of nodes

α	0.1		0.3		0.5		0.7		0.9		1.0	
	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

To fairly compare the performance of our cluster formation method with that of LEACH-C, we provide some computational experiments. We used the following physical constants and parameters in the experiments: $b_i = 0.5$ nJ, $E = 50$ nJ/bit, $\epsilon_{fs} = 10$ pJ/bit/m^2, $\epsilon_{mp} = 0.0013$ pJ/bit/m^4, $E_{DA} = 5$ nJ/bit/signal, $d_0 = 87$ m, $\alpha = 0.1 - 1.0$, $l = 4200$bit, $p = 5$ (for LEACH-C [3]). We used 5 data sets (data1,...,data5), where sensor nodes are randomly located between ($x = 0$, $y = 0$) and ($x = 100$, $y = 100$). We simulated data transmission from every node to the BS at ($x = 50$, $y = 175$) until the battery levels of all sensor nodes reach to 0. To compare the performance of our method with that of LEACH-C in terms of network lifetime, the decisions of CHs were conduced based on the exact solutions[1]. To get the exact solutions, we use the solver tool Xpress-MP (2005B) on PC with Intel Pentium 4 processor (2.53GHz) and 512MB RAM. The average computational time of our method is 4.76 (s) per round.

Table 1 shows the average number of rounds (ave.) until the survival rate of sensor nodes reaches to 99%, 90%, 70%, 50%, 30%, 10%, and 0%, and the standard deviation (std.) for various values of α. The 99% means that the first sensor node died because we use 100 sensor nodes data

[1] In [3], the authors use the simulated annealing algorithm to find heuristic solutions.

Table 2: LEACH-C : The number of rounds per the percentage of nodes alive

	data1	data2	data3	data4	data5
99%	902	611(100)	907	908	888
90%	907		923	919	896(93)
70%	906(90)		924(88)	920(89)	

in the experiments. In other words, the numbers of the row of 99% show the number of rounds which all sensor nodes can be used. The average numbers of rounds are averages over the 5 data sets. For each data set, Table 2 shows the number of rounds until the survival rate of sensor nodes reaches to 99%, 90%, 70% in LEACH-C. In Table 2, the numbers in parenthesis are the numbers of nodes alive when the problem becomes infeasible. For instance, 906(90) means that the problem becomes infeasible at the 906th rounds and 90 sensor nodes are alive. We observe in Table 1 and Table 2 that our method with parameter $\alpha = 0.1 - 1.0$ is superior to LEACH-C. We also observe that the p-median problem in LEACH-C became infeasible, although many sensor nodes are still alive. Here the infeasibility means that p CHs cannot be selected in the p-median problem. The infeasibility at early stage of rounds in LEACH-C comes from the fact that the number of CHs is fixed through all rounds in LEACH-C. On the other hand, as shown in Figure 2, the number of CHs is dynamically changed in each round in our method. This keeps the problem being feasible even after a quite large number of rounds. From Table 1, we see that the optimal value of the parameter α depends on the required survival rate of sensor nodes. For example, when the required survival rate is 70%, the optimal value of α is 0.9. When the required survival rate is 10%, the optimal value of α is 0.1.

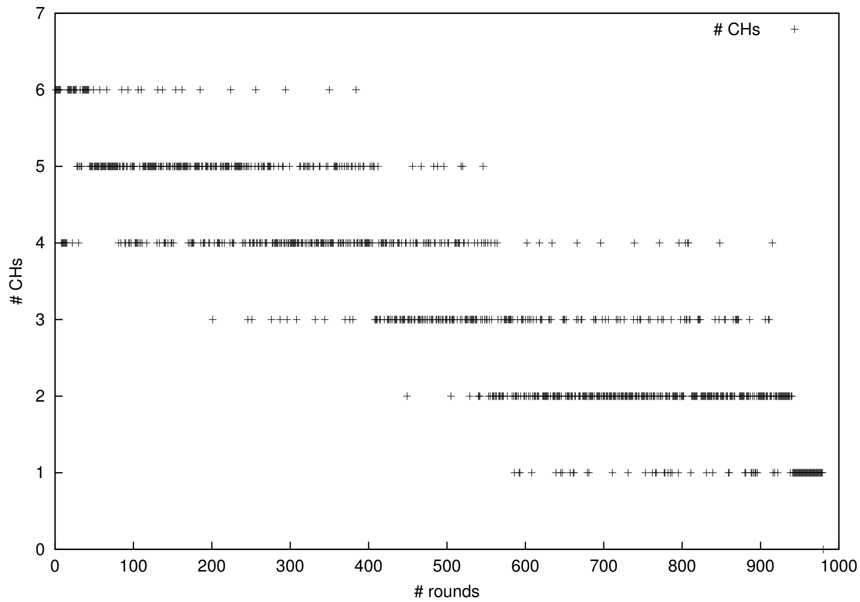

Figure 2: The number of CHs in each round with the case of data1 and $\alpha = 0.9$

5 Conclusion

We formulate the cluster formation problem as the UFLP and develop a centralized method to find a good cluster formation maximizing the network lifetime. Our method is superior to the cluster formation method of LEACH-C for the whole range of survival rates. We use 100 sensor nodes examples for computational experiments and get exact solutions using solver tool. Suitable heuristics for the UFLP [5] may enable us to deal with more practical size of sensor network problems.

References

[1] M. Kochhal, L. Schwiebert, and S. Gupta, "Self-organizing of wireless sensor networks," in *Handbook on Theoretical and Algorithmic Aspects of Sensor, Ad Hoc Wireless, and Peer-to-Peer Networks*, J. Wu, Ed. Auerbach Publications, 2006, pp. 369–392.

[2] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *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*, vol. 1, no. 4, pp. 660 – 670, October 2002.

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

[5] D. Ghosh, "Neighborhood search heuristics for the uncapacitated facility location problem," *European Journal of Operational Research*, vol. 150, pp. 150–162, 2003.