Energy-Aware Optimal Clustering in Wireless Sensor Network Using Integer/Linear Programming

Zahra Eskandari
Dept. of Comp. Eng.,
Quchan University of Advanced Technology,
Iran,
z.eskandari@qiet.ac.ir

Seyed Amin Hosseini Seno
N/W Research Lab.,
Ferdowsi University of Mashhad, Iran,
hosseini@um.ac.ir

Deris Stiawan
Dept. of Comp. Eng.,
Faculty of Comp. Sc.,
Universitas Sriwijaya
Palembang, Indonesia
deris@unsri.ac.id

Bedine Kerim
Dept. of IT
College of CSIT,
Albaha University,
Saudi Arabia
bkerim@bu.edu.sa

Rahmat Budiarto
Dept. of CIS
College of CSIT,
Albaha University,
Saudi Arabia
rahmat@bu.edu.sa

***Abstract*— clustering is a hierarchical method to data transmission in wireless sensor networks, which has a considerable effect on energy conservation. A balanced and efficient clustering has an important role in these networks. This paper discusses an optimal clustering method in wireless sensor network. Firstly, by considering energy and distance parameters, we model the clustering problem using two techniques, Integer Linear Programming and Linear Programming. Then we propose a clustering algorithm based on the optimal selected cluster heads. Experimental results shown that Linear Programming technique has better performance in energy consumption and network lifetime in comparison to the Integer Linear Programming technique.**

Keywords—wireless sensor network; clustering; energy efficient, integer linear programming; linear programming

I. INTRODUCTION

Wireless sensor network (WSN) consists of nodes with limited processing ability and energy resources [1]. These nodes which sense light, temperature, jitter and so on, are deployed in environment densely and randomly and then periodically transmit sensed data to sink according to the routing algorithms. Data transmission is the most costly operation in the network and is in quadratic relation with transmission distance, therefore transmitting in long distance leads to quick node's power drop [2]. To conserve low energy resources of the nodes, routing algorithms must regard some considerations to consume less energy amount and consequently guarantee longer network lifetime. One of the most prominent approaches is clustering algorithm [3].

Clustering is a method by which the sensor nodes in a network organize themselves in hierarchical structures. The cluster members are managed by a leader node called Cluster Head (CH) [4]. The CH communicates with all the nodes of its own cluster directly and then communicates with other CHs. Indeed, CH receives the data sent by its members, aggregates and compresses the data and then transmits aggregated data to other CHs until the data reaches to the sink. An appropriate clustering algorithm reduces energy consumption and enhances the network lifetime [5]. Numerous techniques for clustering have been proposed by researchers such as LEACH [6], LEACH-C [7], HEED [8], CEFL [9], LEACH-E [10] and so forth.

Recently some works focus on improved earlier proposed clustering algorithms or obtaining optimal clustering. In optimal approaches, which be considered in next section, the process of selecting the CHs is modeled as optimization problem with emphasis on minimum energy consumption or maximum network lifetime as objective. Then, by using different tools such as fuzzy logic, integer linear programming and so on, the mentioned problem is solved and optimal or near optimal set of CHs is selected.

In this paper we try to use Linear Programming and Integer Linear Programming to solve the clustering problem. The important different between this work and former works which were using these techniques, is that our proposed model has high scalability. In fact, in this work, by using most important parameters in WSN such as energy and transmission distance, we formulate clustering problem as optimization problem.

The rest of the paper is organized as follows: In the next section we study some optimization works which done in this field. In Section 3, we briefly describe mathematical concepts of Linear Programming and Integer Linear Programming followed by explanations of two models of the clustering problem based on these techniques, and detail discussion on our proposed algorithm. In Section 4 we evaluate the mentioned models and present experimental results and lastly we conclude the work in Section 5.

II. BACKGROUND

Until now some clustering algorithms have been presented which in this section we review them briefly. In [11], the authors proposed a new static clustering method, which is based on the location of sensor nodes, the communication efficiency and network connectivity. They used fuzzy logic tools to select the best CHs and their goal was the network lifetime optimization. They simulated their proposed algorithm and can improve lifetime of the sensor network significantly compared to other approaches. Indeed they employed the Hausdorff distance, and presented the Hausdorff clustering algorithm. In [12] another fuzzy approach was proposed, Fuzzy c-Means (FCM) clustering approach, to determine the optimal number of clusters. They conclude that by optimal selection of number of clusters, the energy consumption decreased considerably.

In related to optimization problems, the earliest contribution [13] formulated a non-ILP algorithm clustering mechanism called Virtual Grid Architecture (VGA) and a corresponding ILP formulation. The objective was to find the minimum set of connected CHs. However, due to the limitation of the solver, maximum number of nodes (*N*) is restricted to 30 nodes. Furthermore, in [14] the authors proposed three different ILP formulations, which not only focused on the minimum number of CHs, indeed they specified number of CHs and then after formulating the clustering problem such that the maximum possible network lifetime was obtained, they selected optimal CHs. However, due to the complexity of the ILP formulations, only up to 9 nodes could be solved in their modeling. In [15] the authors proposed an ILP formulation of the clustering problem, building on the model presented in [14], which improved its weaknesses. Indeed, in [15] an improved ILP formulation for the Clustering Problem was presented and then the performance of it was evaluated.

Additionally, some optimal or near optimal works were presented too, although they do not use the same technique as our work. In [16] a new efficient CH selection algorithm based on the maximum weight of five different parameters was proposed and showed that the proposed algorithm is better than the other weighted clustering algorithms. In [17] a clustering algorithm of cluster head optimization based on energy (LEACH-E) was proposed to optimize the selection of cluster head. Their work avoid the node with low energy to be selected as a CH, and so they could balance the energy consumption of nodes. Simulation results showed that their algorithm had better results compared to LEACH protocol.

III. Proposed Method

In this section, having done an introduction of Linear Programming and Integer Linear Programming, we present our contribution which consists of two phases; phase one we present two types of cluster problem modelings, Linear programming (LP) model and Integer Linear Programming (ILP) model. Then phase 2 we present our clustering algorithm.

A. Linear programing and Integer Linear programming

At first we briefly describe Linear Programming and Integer Linear Programming which both deal with the problem of minimizing a linear function of several variables subject to equality and inequality linear constraints. In addition to mentioned constraints, ILP has some integrality restrictions on some or all of the variables. In fact, the problem is of the form:

$$min\{Cx \mid Ax \leq b\} \quad (1)$$

The set S: $S = \{ x \in \mathbb{R}^l, Ax \leq b\}$ in LP and set S: $S = \{ x \in \mathbb{Z}^k \times \mathbb{R}^l, Ax \leq b\}$ in ILP is called a feasible set and the element x is called a feasible solution. LP solver usually use simplex algorithm [18] to solve this problem, however some solvers such as CPLEX [19], SCIP [20] and so forth, use other efficient methods. MIP solvers use some efficient techniques such as Gomory's Cuts, branch and bound algorithm. Many of these methods use LP-relaxations [21], in the processes of finding an optimum point.

B. Proposed LP and ILP model

Due to the simplicity of ILP model, we pay attention to this modeling first, and then the LP model is considered.

1) ILP model

In this model, the minimum number of CHs is considered as objective function subject to a condition that all nodes must be covered by CHs. By defining the variable x_i as in (2):

$$x_i = \begin{cases} 1 & if\ node_i\ is\ selected\ as\ CH \\ 0 & else \end{cases} \quad (2)$$

The clustering problem is modeled as follows:

$$min \sum xi \quad (3)$$

subject to:

$$\sum yij * xi \geq 1 \qquad for\ j = 1..N, i = 1..N \quad (4)$$

If the variable x_i equals to 1, the node *i* is selected as CH. And the value of y_{ij} is determined in (5).

$$y_{ij} = \begin{cases} 1 & if\ node_i\ is\ covered\ by\ CH_j \\ 0 & else \end{cases} \quad (5)$$

and the constraints mean that each node *j* must be covered by at least one CH.

2) LP model

In the previous modeling, we only focused the node's coverage without considering the distance between each node and its CH. Since the energy consumption is related to transmission distance and it was not considered, so the previous model is not energy efficient. Thus, in this modeling the objective function is altered as in (6).

$$\min \sum \sum distance(node_i, node_j)\, x_i \quad (6)$$

subject to:

$$\sum y_{ij} * x_i \geq 1 \qquad for\ j = 1..N, i = 1..N \quad (7)$$

By considering the distance between each node with its CH, objective function is defined as minimum value for the sum of distance of each cluster members to their CHs. In this model, the definition for y_{ij} does not change but we remove the constraint that the x_i must be 0 or 1. Therefore in solving phase, positive value for each variable x_i demonstrates that node *i* is selected as CH.

C. Proposed Clustering Algorithm

The details of the proposed algorithm is presented in Fig. 1. To do clustering in the network, as a CH's energy reaches a predefined threshold, all the nodes in the network transmit their information to sink and then sink performs centralized phase of clustering algorithm.

Initially the sink calculates the average energy of the alive nodes (E_{ave}) then by considering the nodes which have higher energy than E_{ave} as candidate CHs, sink forms candidate of CH (CCH) set. Then, matrix *M* is formed as in (8).

$$M = \begin{pmatrix} y_{1,1} & \cdots & y_{1,N} \\ \vdots & \ddots & \vdots \\ y_{N,1} & \cdots & y_{N,N} \end{pmatrix} \quad (8)$$

The sink sets up the objective function for each model based on the LP or ILP model, and then solves the optimization problem and the CHs are selected. It must be mentioned that N is the number of nodes in the network.

After this step, sink performs Inter CHs routing and creates the spanning tree between CHs based on the energy and distance to sink parameters. Indeed CHs select a parent with more energy and less distance.

```
Proposed algorithm () {
Centralized-clustering-phase ();
Distributed-clustering-phase ();
While (nodes is alive) {
Data-packet-transmission-to-sink ();
If (check CH'energy decrease to threshold () ==
true) {
        Control-packet-transmission-to-sink ();
        Centralized-clustering-phase ();
        Distributed-clustering-phase () ;}
Centralized clustering phase () {
Calculate Eave;
Form CCH set;
Form M matrix;
Model and Solve ILP/LP problem;
Form-INTER-CH-Spanning-Tree ();
Broadcast-CH-spanning-Tree () ;}
```

Fig. 1. Proposed clustering algorithm

Then distibution phase starts. The sink broadcasts this information in the network, and each CH broadcasts in its surrounding, then each node selects the nearest CH with higher energy as its CH. It is notable that to achieving efficient transmission, the node adjusts transmission signal strength based on the distance between himself and its parent.

IV. Experimental Results and Discussion

We use commercial solver IBM Ilog Cplex 12.1 available under academic license to evaluate the proposed method. The simulation environment is a 25m*25m square area which the number of nodes (N) is varied between 100 and 500. The initial energy of each node is a random variable less than 200 Joule.

We suppose the variables ER_i , ET_{ij}, and EA_i represent the required energy for data receiving, transmission and data aggregation, respectively. Using the energy model in [22], the following equations are obtained, where c_1, c_2 and c_3 are constants:

$$ER_i = c_1 * K \tag{9}$$

$$ET_{ij} = c_2 * K * dist_{ij}^2 \tag{10}$$

$$EA_i = c_3 * K \tag{11}$$

Some events in the network occurred and then by detecting some events, a node forwards sensed data to the sink using spanning tree. All nodes that sense an event and forward their data to the sinks, consumes energy for only data transmission (ET_{ij}). CH nodes consume energy for both data receiving and data transmission (ER_i and ET_{ij}, respectively). Furthermore, they consume energy for data aggregation (EA_i).

As shown in Fig. 2, we set the value of N, the number of network nodes, to 300 and in the end of each round we sum the energy of alive nodes and calculate total energy of the network. As the LP modeling considers distance in addition to coverage of the nodes, so the solver try to minimize the distance between cluster members and their CH, in CHs selection process. In distribution phase the cluster members also select nearest CH. So, transmission distance decreases which lead to less energy consumption in data transmission. As a result, in LP model the energy of the network is saved better than ILP model.

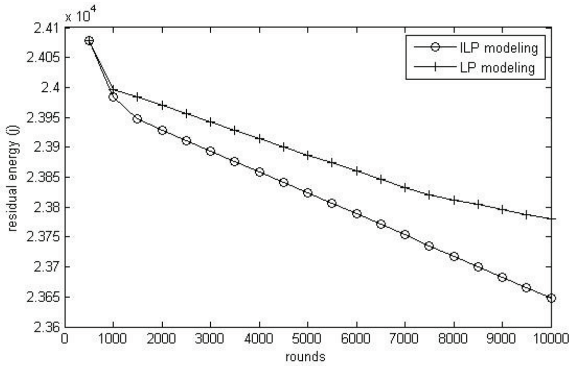

Fig. 2. Residual energy of the network at N=300

We also consider another proper parameter to be evaluated which is the of alive nodes. After each simulation round, we count the alive nodes in the network, as in LP model, each node selects the CH, by considering its energy and distance, a suitable CH is selected which led to less overall energy consumption. The rational is because nodes with low energy are never being selected as CH, so they spend less energy and remain longer time in the network. Therefore, the number of alive nodes are more than that in ILP model. The result of our comparison is shown in Fig. 3.

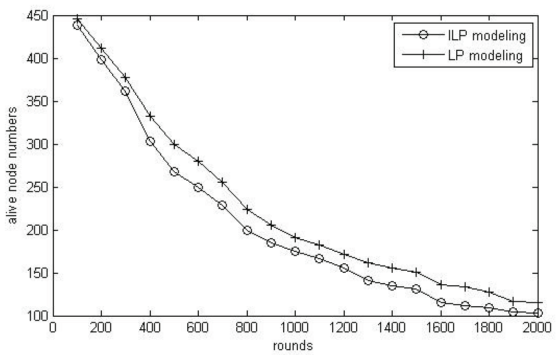

Fig. 3. Remaining alive nodes in the network

In Fig. 4, the lifetime for two models are plotted versus number of nodes. As the LP model consumes less energy and has more alive nodes, so its lifetime is higher than ILP model. Network lifetime is defined as the longevity of network in term of the number of rounds performed before the failure of 10% of the total nodes.

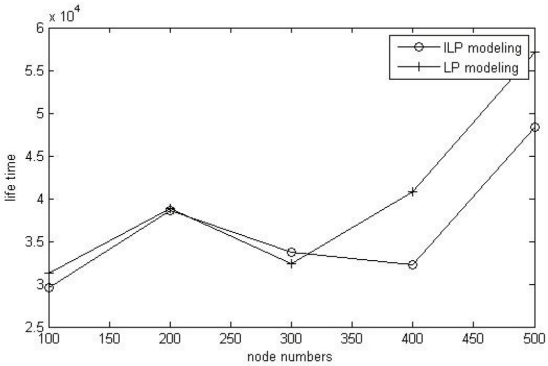

Fig. 4. Life time of the networks

V. CONCLUSION AND FUTURE WORKS

In this paper we model the clustering procedure problem as an optimization problem. We formulate the clustering problem as selecting minimum number of CHs to comply coverage of the network. By modeling the problem under LP and ILP techniques, we obtain better result in LP modeling because of considering distance parameter in objective function. Taking consideration the distance between nodes and their CH affects the energy used to transmit and in turn decrease overall energy consumption and longer alive nodes.

As a future work, we can model this problem as a multi objective function by considering important goals, minimum transmission energy consumption and maximum lifetime.

REFERENCES

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, "Wireless sensor networks: a survey," Computer Networks Journal, vol. 40, no. 8, pp. 102-114, 2002.

[2] S. Lindsey, C.S. Raghavendra, "PEGASIS: Power-Efficient GAthering in Sensor Information Systems", Proceeding of the IEEE Aerospace Conference, vol. 3, pp. 1125-1130, 2002

[3] C-P. Liu, R-F. Li, B. Fu, "Distributed clustering algorithm in dynamic topology network based on WSN," Chinese Journal of Scientific Instrument, 30 (2009), pp.2652-2657.

[4] K. Hussain, A.H. Abdullah, K.M. Awan, F. Ahsan and A. Hussain, "Cluster head election schemes for WSN and MANET: a survey, " World Applied Sciences Journal, vol. 23, no. 5, pp.611-620, 2013.

[5] S. Chinara and S.K. Rath, "A Survey on One-Hop clustering algorithms in Mobile Ad Hoc networks," Journal of Network and Systems Management, vol. 17, no. 1-2, pp. 183–207, 2009.

[6] J. Hoebeke, I. Moerman, B. Dhoedt and P Demeester, "An overview of Mobile Ad Hoc networks: applications and challenges," Journal of the Communications Network, 3 (July 2004): 60-66."

[7] M. Negnevitsky, Artificial intelligence: A guide to intelligent systems, Addison-Wesley, Reading, MA, 2001.

[8] H. S. Lee, K. T. Kim, and H. Y. Youn, "A new cluster head selection scheme for long lifetime of wireless sensor networks," in Computational Science and Its Applications—ICCSA 2006, vol. 3983 of Lecture Notes in Computer Science, pp. 519–528, Springer, 2006.

[9] T.L Saaty, Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process, RWS Publications, U.S.A. 2000.

[10] H. Yang and B. Sikdar, "Optimal cluster head selection in the LEACH architecture," Proceedings of International Performance, Computing and Communication Conference, 2007, New Orleans, LA, pp.93-100.

[11] X. Zhu and L. Shen, "Near optimal cluster-head selection for wireless sensor networks," China Journal Of Electronics, November 2007, vol.24, no.6, pp.721-725.

[12] A.S. Raghuvanshi, S. Tiwari, R. Tripathi and N. Kishor, "Optimal number of clusters in wireless sensor networks: An FCM approach," IEEE International Conference on Computer and Communicaton Technology (ICCCT'10), 2010, pp.817-823.

[13] J.N. Al-Karaki, A.E. Kamal, and R. Ul-Mustafa, "On the optimal clustering in Mobile Ad hoc Networks," Consumer Communications and Networking Conference, Las Vegas, 2004, pp. 71-76.

[14] W. El-Hajj, D. Kountanis, A. Al-Fuqaha, and H. Harbi, "Optimal hierarchical Energy efficient design for MANETs," International Conference on Communications and Mobile Computing, 2006, pp.287-292.

[15] S. Zahidi, " Solving the clustering problem in MANETs using SAT & ILP techniques," The 11th International IEEE Conference on Information Sciences, Signal Processing and their Applications: Main Tracks, 2012, pp. 624-629.

[16] P. Sivaprakasam,"An efficient clusterhead election algorithm based on maximum weight for MANET," Third International IEEE Conference on Advanced Computing (ICoAC 2011), Chennai, 2011, pp.315-320.

[17] M.S. Fareed, N. Javaid, M. Akbar, S. Rehman, U. Qasim, Z.A. Khan. "Optimal number of cluster head selection for efficient distribution of sources in WSNs," Seventh International IEEE Conference on Broadband, Wireless Computing, Communication and Applications (BWCCA), Victoria BC, 2012, pp.632-637.

[18] A. Schrijver, Theory of Linear and Integer Programming, John Wiley & Sons Ltd., New York, 1999.

[19] IBM ILOG CPLEX Optimizer, 2011. [Online]. http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/, (last accessed: 1 July 2016).

[20] SCIP: An open source MIP solver and constraint Integer Programming framework, 2011. [Online]. http://scip.zib.de/, (last accessed: 1 July 2016).

[21] L.A. Wolsey, Integer programming, John Wiley & Son, New York, 1998.

[22] X-W. Zhang, X-H Yu, X-Q. Chen, "Inter-query data aggregation in wireless sensor networks," International IEEE Conference on Wireless Communications, Networking and Mobile Computing, 2005, vol. 2, pp. 1472 – 1476.