Energy-Aware Clustering and Routing Scheme in Wireless Sensor Network

Chang Lou, Xiaofeng $\operatorname{Gao}^{(\boxtimes)}$, Fan Wu, and Guihai Chen

Department of Computer Science and Engineering, Shanghai Key Laboratory of Scalable Computing and Systems, Shanghai Jiao Tong University, Shanghai 200240, China mcfatealan@sjtu.edu.cn, {gao-xf,fwu,gchen}@cs.sjtu.edu.cn

Abstract. Maximizing the network lifetime is always a main challenge ahead of wireless sensor network (WSN). Clustering and routing has been proved to be energy-efficient strategies for extending the network lifetime. In this paper, we put forward an energy-consumption model for sensors in WSNs and calculate network energy consumption in a short period for any given network configuration, including sensor state scheduling, clustering, and routing information. Then we address an energy-aware optimal planning problem with area coverage and connectivity constraints, seeking for the best sensor scheduling scheme to extend the network lifetime. We formulate it as an Integer Linear Programming (ILP) model, and add some extra constraints to reduce the scale of the model. We use Gurobi to compute this model and compare the basic model, scalereduced model, with a previous work OPT-ALL-RCC [6]. The simulation results prove that the reduction is necessary and our model have better performance than the previous model.

Keywords: Wireless sensor network \cdot Clustering \cdot Routing \cdot Energy-efficient

1 Introduction

A large collection of densely deployed, spatially distributed, and autonomous devices (or nodes) that communicate via wireless signals and cooperatively monitor physical or environmental conditions is called wireless sensor network (WSN) [1]. In surveillance applications, sensors are deployed in a certain field

© Springer International Publishing Switzerland 2015

DOI: 10.1007/978-3-319-21837-3_38

This work was supported in part by the State Key Development Program for Basic Research of China (973 project 2012CB316201), in part by China NSF grant 61422208, 61202024, 61472252, 61272443 and 61133006, CCF-Intel Young Faculty Researcher Program and CCF-Tencent Open Fund, the Shanghai NSF grant 12ZR1445000, Shanghai Chenguang Grant 12CG09, Shanghai Pujiang Grant 13PJ1403900, and in part by Jiangsu Future Network Research Project No. BY2013095-1-10. The opinions, findings, conclusions, and recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies or the government.

K. Xu and J. Zhu (Eds.): WASA 2015, LNCS 9204, pp. 386-395, 2015.

to detect and report events like presence, movement, or intrusion in the monitored area [2]. Generally speaking, the battery of sensors is unchangeable or the uncertainty of the deployed area makes recharging impossible. In consequence, maximizing the lifetime of WSNs is always a vital issue of this topic. In this paper, we mainly conserve the network energy from the following two steps [2,3]:

- 1. Energy-efficient scheduling of sensor states (active or sleep);
- 2. Energy-efficient clustering and routing among sensors;

We assume that the sensor deployment is dense enough that in any certain area there is more than one sensor. That allows sensors take turns to monitor the overlapped area. By scheduling some redundant nodes into sleep states, the network lifetime is extended while the coverage and connectivity is preserved.

Clustering has been admitted as one of the energy-efficient way for WSNs [4, 14–17]. In a cluster-based WSN, some sensors are elected as CHs (cluster head). The duty of CHs is to receive the information collected from non-CHs and send to the sink node (a special node which has infinite energy and all the information feedbacks to this node). One main advantage of clustering is that it could solve some potential problems like bandwidth limits thus increasing the capacity of the system [5]. Furthermore, CHs could function as routers. Since a long-distance transmitting costs much more energy than a short one [7], the nodes from the relatively far location could send the information through a node chain.

Our target is to seek an optimal planning to extend the network lifetime. While maintaining the coverage and connectivity of the whole network, we try to find a synthesize scheme including sensor state scheduling, clustering, and routing to maximize the lifetime of the network. Compared with Chamam's work in [6], a similar design to schedule sensors acting for different roles, we further consider all possible energy consumption in a WSN, and further introduce a reduction method to reduce the problem scale. Thus our designs are more accurate and time-efficient. We also provide various numerical experiments and comparisons to validate the efficiency of our design.

The rest of this paper is organized as follows: In Sect. 2, we introduce some related work about energy-efficient planning of sensor network. In Sect. 3, we give necessary definition, address energy consumption model for single node, put forward the optimal planning problem, and state all assumptions. In Sect. 4, we show how to formulate this model into an integer linear programming problem. In Sect. 5, we describe how to reduce the problem scale to improve the computing efficiency. In Sect. 6, we display the simulation results and comparisons with previous literature. In Sect. 7, we make the conclusion.

2 Related Work

Much progress has been made on the study of sensor state scheduling, clustering, and routing. For sensor state scheduling, Xing et al. [8] proposed Coverage Configuration Protocol (CCP), which can provide different degrees of coverage requested by applications and nodes' states can be alternated according to the



Fig. 1. Multi-hop clustering sensor network

coverage degree. Heinzelman et al. [7] proposed LEACH (Low-Energy Adaptive Clustering Hierarchy), a clustering-based protocol that put forward the concept of cluster head and utilizes randomized rotation of CHs to balance the energy load among the sensors in WSN. The protocol proved to be energy-efficient and convenient to adapt to practical situations. Lindsey et al. [9] proposed PEGA-SIS (power-efficient gathering in sensor information systems), an improved, near optimal chain-based protocol. In PEGASIS, each node communicates only with a close neighbor and takes turns transmitting to the base station, thus reducing the amount of energy spent per round. However, the transmission distance is limited to only one hop and routing is not considered. Manjeshwar et al. [12,13] proposed TEEN (Threshold-sensitive Energy Efficient sensor Network protocol) and APTEEN (Adaptive Periodic Threshold-sensitive Energy Efficient sensor Network protocol). In TEEN, the data transmission is not transmitted as frequently as sensing activities. A cluster head sensor sends its members the threshold value of the sensed attribute and a small change in the value of the sensed attribute that triggers the node to switch on its transmitter and transmit. Rodoplu et al. [10] proposed Small Minimum Energy Communication Network (MECN). The protocol identifies a relay region for every node, which consisting of nodes in a surrounding area where it is more energy efficient to transmit through those nodes than direct transmission. But still it does not prove its routing strategy is the optimal strategy (Fig. 1).

Most works concentrate on raising new mechanisms or protocols to conserve network's energy, while only a few considered trying to address a linear programming model to formulate the problem and to find the optimal solution. Chamam et al. [6] put forward an optimal configuration problem. The authors classified nodes into three levels (cluster, active, sleep), and set a standard energy constant for each state. The target is to determine the state of each sensor in the next round. Nevertheless, the energy consumption rate in the real situation varies significantly among CHs and active nodes. The number of connected nodes and link ways could largely affect the energy consumption rate, and the communication cost is overly underestimated. The authors suggest that the balance of retained energy among nodes are critical and use balance as the target function of ILP, without giving any proof or reference. Also, the model ensures the existence of the spanning tree but the specific optimal routing scheme is not considered.

3 Problem Statement and Assumptions

3.1 Network Configuration

Consider a cluster-based WSN deployed in a certain flat area, nodes can collect information from the environment, transmit data to others, receive data from others, or sleep. It is important to alternate sensors that are redundant to sleep. According to [3], the power consumption in sleeping mode is generally about 5 % ~ 10 % of active mode.

Besides, sensors in our network are divided into clusters. Sensors connect to cluster heads (CHs) to send message to sink. The election of CHs is critical since CHs form the backbone of the network. CHs usually dissipate energy much faster than non-CHs, for its frequent data exchange with others. Thus usually we choose those who have more energy left to become CHs. On the other hand sensors should also choose wisely about which cluster to join. A cluster of more nodes can have more potential CHs but after deaths of some nodes on the border, the whole area could be cut off from the network leaving a coverage hole.

Directly sending data from sensors to sink has been proved inefficient. Instead of simple sink-to-CH communication, we seek some intermediate CHs to act as routers. Message will be transferred from CH to CH and the delivering path follows some rules. Each time there is only a very short distance to pass data so the energy consumption could be lowered. Still there is a tradeoff between transferring times and transferring distance because compared to sending data directly, intermediate CHs need to process this package and spend extra energy. Thus it is unwise to use routers as much as it can. Also, we need to ensure that for each CH there will be a path for it to communicate with sink. To achieve this, there should be a spanning tree in the topology.

Now we give the definition of coverage. We use the basic Disk model to judge if an area is covered by a set of sensors, which means a point can only be covered or not. The covered area is a disk with the sensor as its center. We assume all sensors having the same sensing range. In this article we require the monitored area fully covered.

Definition 1. Sensor set S covers the area A if and only if \forall point $P \in A, \exists$ sensor $i \in S$, the distance between P and i is less than the sensing range of sensor i.

Then we define the connection between sensors. To make sure communication between sensors is stable and reliable, we assume only sensors within safe distance can establish connection.

Definition 2. Sensor *i* can reach to sensor *j* in one hop if and only if the distance between *i* and *j* is less than the communication range of sensor *i*.

Our optimal solution consisting three layers (state scheduling, clustering, routing) above, and it has to meet the following demands:

- 1. The energy consumption is minimized.
- 2. The whole area should be fully covered.

- 3. Every active sensor is connected to one and only one CH.
- 4. There must be a spanning tree including all CHs.

We made the following assumptions:

- 1. All sensors are homogenous and fixed. Sensors can collect data from all directions.
- 2. The location and energy information of all sensors are known to sink. Each sensor is unique and can be recognized by sink. We do not consider the initial step of the network construction and neither its energy consumption.
- 3. Sensors are randomly deployed and the network is dense enough.
- 4. Sensors can function well as long as retained energy is not used up. The lifetime of their components is not considered.
- 5. Only CHs can act as routers. CHs receive and transmit data to other CHs or sink.
- 6. The amount of data a sensor collects from the environment in a period is steady. Here we treat the size of message package every sensor needs to pass to sink as a constant.

3.2 Energy Consumption Estimation

We address the SC model to estimate the energy consumption.

In the SC model, we only consider two major energy dissipations: Standalone cost and Communication cost. The standalone cost includes the energy a sensor uses for gathering data from environment and data processing. It is relatively stable and in the most common situation it will not change. While the communication cost relate to the amount of transmitted data and transmitting range. For example, to transmit a k-bit message a distance d needs:

$$E_{Tx}(k,d) = E_{elec} \cdot k + \epsilon_{amp} \cdot k \cdot d^2.$$

 E_{elec} and ϵ_{amp} are constants. We assume $E_{elec} = 50nJ/bit$, $\epsilon_{amp} = 100pJ/bit \cdot m^2$ [5]. d means the distance between communicating nodes.

To receive this message, the radio expends:

$$E_{Rx}(k) = E_{elec} \cdot k.$$

During the period δt , the energy a sensor consumes:

$$\delta E = (E_S + \sum kt_j \cdot (E_{elec} + \epsilon_{elec} \cdot d_j^2) + \sum kr_j \cdot E_{elec}) \cdot \delta t.$$

 E_S is a constant, which means standalone cost rate. kt and kr mean bits transmitted and received in a time unit.

Now we can calculate the energy consumption of any node in the next period if we already have the sensor scheduling, in other words, clustering and routing scheme for the next round. If a node is active but not a CH in the next period, the energy consumption includes standalone cost and transmitting cost for sending collected data to CH. If a node is a CH, the energy consumption should not only include the above two costs, but also the receiving cost for gathering data from the cluster, the receiving and transmitting cost for being a routing node.

Undoubtedly we want the total sum of δE as small as possible, which means the whole energy consumption is minimized. But this does not necessarily lead to the network lifetime maximization. For example, some sensors die out quickly while others are nearly as healthy as newborns. In this situation the network is becoming equivalently sparse, which leads to the coverage holes and disagrees with our assumption. Keeping the balance of retained energy among the sensors is of vital importance. Thus, instead to minimize total sum of δE , we choose to minimize the sum of ratio between predicted consumed energy and retained energy of sensors $\delta E/E_i$.

4 Modeling and Definitions

Our work is on the basis of research of Chamam et al. [6]. After taking communication costs into consideration, we formulate the optimal planning problem into an integer linear programming problem.

First we define all the sets, constants, variables we will use later in Table 1. Here S, C, d_{ij}, ρ_{ic} are constants, while $X_i, Y_i, Z_{ij}, V_{ij}^{kl}$ are variables we want to compute. The domain of variables i, j, k, l, c is $i, j, k, l \in \{1...N\}, c \in \{1...M\}$.

Symbol	Meaning
S	the set of sensors, $S \in \{1, \cdots, N\}$
C	the set of cells composing the monitored area, $C \in \{1, \cdots, M\}$
d_{ij}	$d_{ij} = 1$ if sensor <i>i</i> can reach sensor <i>j</i> in one hop, otherwise 0
$ ho_{ic}$	$ \rho_{ic} = 1 $ if sensor <i>i</i> covers cell <i>c</i> , otherwise 0
X_i	$X_i = 1$ if sensor <i>i</i> is Active, otherwise 0
Y_i	$Y_i = 1$ if sensor <i>i</i> is a CH, otherwise 0
Z_{ij}	$Z_{ij} = 1$ if sensor <i>i</i> is connected to CH <i>j</i> , otherwise 0
V_{ij}^{kl}	$V_{ij}^{kl} = 1$ if flow (i,j) passes through the link kl , otherwise 0

Table 1. Set table

The problem is modeled below. The meaning of the target function has been described in the last section, namely the sum of the predicted consumption rate of all sensors, including standalone cost and communication cost. Constraint (1) is coverage requirement that the monitored area should be fully covered. Constraint (2) means a sensor has to be active if it is a CH. Constraint (3) guarantees that there must be one sensor that can reach the sink (though we did not include sink in the sensor table, we can regard sink as a special node). Constraint (4) guarantees that a sensor has to be active if it is connected to CHs, a sensor has to be a CH if there are other nodes connecting to it and the distance requirement for the connection between CHs, which we talked in the

definition of connection. Constraint (5) guarantees that one sensor can only join one cluster to avoid redundancy and disorder, and the CHs have the maximum connection limits taking bandwidth of sensors into consideration. Constraint (6) means this is a 0-1 programming problem. Constraints (7) (8) guarantee that the network contains a spanning tree, which ensures the connectivity of network. However, Constraint (7) is not linearized. Thus we use Constraints (9) to help to linearize it. In the end, we get the linear programming model.

$$\forall i = 1..|S|, \ Y_i \le X_i \tag{2}$$

$$\sum_{i=1}^{|S|} Y_i \cdot d_{i0} \ge 1 \tag{3}$$

$$\forall i, j = 1..|S|, \ j \neq i, \ Z_{ij} \le X_i - Y_i, \ Z_{ij} \le Y_j, \ Z_{ij} \le d_{ij}$$
(4)

$$\forall i = 1..|S|, \sum_{\substack{i=1, i \neq i}}^{|S|} Z_{ij} + Y_i = X_i, \ Z_{ij} \le N_{max}$$
(5)

$$X, Y \in \{0, 1\}^{|S|}, \ Z \in \{0, 1\}^{|S|^2}, \ V \in \{0, 1\}^{|S|^4}$$
(6)

$$\forall i, j, k = 1..|S|, \ j \neq i, \sum_{l \in |S|, l \neq k} V_{ij}^{kl} - \sum_{l \in |S|, l \neq k} V_{ij}^{lk} = \begin{cases} 0 & \text{if } k \neq i \text{ and } k \neq j \\ Y_i \cdot Y_j & \text{if } k = i \\ -Y_i \cdot Y_j & \text{if } k = j \end{cases}$$
(7)

$$\forall \ i, j, k, l = 1..|S|, \ j \neq i, \ k \neq l, \ V_{ij}^{kl} \le Y_i, \ V_{ij}^{kl} \le Y_j, \ V_{ij}^{kl} \le Y_k, \ V_{ij}^{kl} \le Y_l, \ V_{ij}^{kl} \le d_{kl} \qquad \left(8\right)$$

$$\forall i, j = 1..|S|, \ Y_i \cdot Y_j \le Y_i, \ Y_i \cdot Y_j \le Y_j, \ Y_i \cdot Y_j \ge Y_i + Y_j - 1 \tag{9}$$

5 Problem Scale Reduction

Though the size of the problem grows with the number of sensors in a polynomial speed, for 10^2 sensors there should at least be 10^8 variables to optimize since $V \in \{0,1\}^{|S|^4}$. We can eliminate some computing redundancy to make the problem smaller. For V_{ij}^{kl} , it makes sense only if sensor k and l can reach each other, which is written as : $d_{ij} = 1$.

Also, we can strictly control the use of routing, for example, sensor i,k,l,j should gradually approach the sink in sequence. The constraint is written as:

$$d_{i0} \ge d_{k0} > d_{l0} \ge d_{j0}.$$

6 Simulation Results

We perform our simulation in Gurobi [11], a commercial optimization solver, which has a superior performance dealing with large-scale constraints and variables. For convenience we use "EAS" to refer to our model.

	Reduced		Not reduced	
Sensor num	Variables	Constraints	Variables	Constraints
10	164	10598	8220	25230
20	2376	24732	144840	268060
30	10039	76223	757860	1395970
40	32899	233293	2435280	4233000
50	74310	519420	6005100	10523050

Table 2. Number of variables andconstraints vs number of sensors

Table 3. Number of variables and con-straints vs communicating ranges

	Reduced		Not reduced	
Comm. range	Variables	Constraints	Variables	Constraints
10	1970	19740	757860	920950
20	5122	41804	757860	1067110
30	10039	76223	757860	1395970
40	10749	81193	757860	1829230
50	19286	140952	757860	1928410

6.1 Performance Evaluation: Impact of Problem Scale Reduction

Variation of the Number of Variables and Constraints with the Number of Sensors: In a settled map, the larger the number of sensors is, the denser the network becomes and the diversity of network structure has more potential. We fix the map size as $100 \cdot 100$, the sensing and communication ranges as 30, and change the number of sensors to see the impact. The result is shown in Table 2. The number of variables and constraints is increasing w.r.t. sensor number, but the reduced model has gained a small-scale advantage and a slower growing speed.

Variation of the Number of Variables and Constraints with the Communicating Ranges: The communicating ranges has a great impression on network planning. A large communicating range means a cluster could potentially contain more nodes, while a small communicating range needs routing technique more. We fix the size of the map as $100 \cdot 100$, the number of sensors as 30. We change the communicating ranges to see the impact. The result is shown in Table 3. Since our reduction technique takes the connecting possibility between nodes into consideration, the communicating ranges affect the degree of reduction and the smaller the communicating ranges become, the better the reduction behaves.

Variation of the computation time with the problem scale: Let us take a closer look at the impact of the problem scale on the computation time. We fix the size of the map as $100 \cdot 100$, the sensing range and communication ranges 50, and change the number of sensors to see the impact. The result is shown in Table 4. We can see that there is a positive correlation between the computation time and the problem scale. Also, when we tried to solve the unreduced model, we run out of memory resources on our machine, which reveals the necessity of reduction.

Variation of Objective Function Values Obtained by Two Models: Objective function values are positively correlate to energy consumption in the whole network. We fix the size of the map as $100 \cdot 100$, the sensing range and communication ranges as 30, and change the number of sensors to see the impact. The result is shown in Table 5. We can see that after reduction, the obtained objective function value will not inevitably cause the decrease of solution quality. On the contrary, it helped to find a near-optimal solution.

	Computation time		
Sensor num	Reduced	Not reduced	
10	5	5	
20	8	141	
30	41	406	
40	70	1031	
50	138	Out of memory	

Table 4. Variation of the compu-	
tation time with the problem scale	

Table 5.	Variation of	the objective func	-
tion with	the problem	scale	

	Objective function		
Sensor num	Reduced	Not reduced	
10	4.18	6.2	
20	0.4	6.158	
30	5.92	5.318	
40	4.157	5.679	
50	3.85	Out of memory	

6.2 Comparison Between Reduced OPT-ALL-RCC and EAS

Here we choose to compare our model to the reduced OPT-ALL-RCC(generated by doing the same reduction to OPT-ALL-RCC as to our EAS model) instead of OPT-ALL-RCC for sake of time efficiency. We fix the size of the map as 100·100, the sensing range and communication ranges as 50, and change the number of sensors to see the impact.



Fig. 2. Computation time using reduced OPT-ALL-RCC and EAS



Fig. 3. Network lifetime using reduced OPT-ALL-RCC and EAS

Computation Time: The result is shown in Fig. 2. For the same problem set, the reduced OPT-ALL-RCC and EAS spend similar time to find a near-optimal solution because after reduction these two models have the exact same problem scale. In other words EAS needs to spend less time than the original versional of OPT-ALL-RCC.

Network Lifetime: The result is shown in Fig. 3. There is an obvious network lifetime lift from OPT-ALL-RCC to EAS. Especially when the scale of network gets large, our model could effectively lengthen the network lifetime.

7 Conclusion

In this article we propose a model that based on communication cost estimation to calculate the energy consumption of sensor network. On top of this we put forward a mechanism to find an optimal or near-optimal configuration for sensor network to maintain coverage and connectivity and extend lifetime of sensor network. The simulation shows that our mechanism could prolong the lifetime of network with the computation time remains nearly constant. As future research directions, we plan to further decrease the problem scale and improve the energy consumption model.

References

- 1. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., et al.: Wireless sensor networks: a survey. Comput. Netw. **38**(4), 393–422 (2002)
- Chong, C.Y., Kumar, S.P.: Sensor networks: evolution opportunities, and challenges. Proc. IEEE 91(8), 1247–1256 (2003)
- Raghunathan, V., Schurgers, C., Park, S., et al.: Energy-aware wireless microsensor networks. IEEE Signal Process. Mag. 19(2), 40–50 (2002)
- Ahmed, A.A., Younis, M.: A survey on clustering algorithms for qireless sensor networks. Comput. Commun. 30(14), 2826–2841 (2007)
- Yao, Y., Giannakis, G.B.: Energy-efficient scheduling for wireless sensor networks. IEEE Trans. Commun. 53(8), 1333–1342 (2005)
- Chamam, A., Pierre, S.: On the planning of wireless sensor networks: energyefficient clustering under the joint routing and coverage constraint. IEEE Trans. Mob. Comput. 8(8), 1077–1086 (2009)
- Heinzelman, W.R., Chandrakasan, A., Balakrishnan, H.: Energy-efficient communication protocol for wireless microsensor networks. In: IEEE International Conference on System Sciences, vol. 2, pp. 1–10 (2000)
- Xing, G., Wang, X., Zhang, Y., et al.: Integrated coverage and connectivity configuration for energy conservation in sensor networks. ACM Trans. Sens. Netw. (TOSN) 1(1), 36–72 (2005)
- 9. Lindsey, S., Raghavendra, C.S.: PEGASIS: power efficient gathering in sensor information systems. In: IEEE Aerospace Conference Proceedings, vol. 3 (2002)
- Rodoplu, V., Meng, T.H.: Minimum energy mobile wireless networks. IEEE J. Sel. Areas Commun. 17(8), 1333–1344 (1999)
- 11. Gurobi Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Gurobi
- Manjeshwar, A., Agarwal, D.P.: TEEN: a routing protocol for enhanced efficiency in wireless sensor networks. In: IEEE Parallel and Distributed Processing Symposium(IPDPS), vol. 3, pp. 30189a–30189a (2001)
- Manjeshwar, A., Agarwal, D.P.: APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks. In: IEEE Parallel and Distributed Processing Symposium(IPDPS), pp. 195–202 (2002)
- Akkaya, K., Younis, M.: A survey of routing protocols in wireless sensor networks. Ad Hoc Netw. 3(3), 325–349 (2005)
- Arboleda, L.M.C., Nasser, N.: Comparison of clustering algorithms and protocols for wireless sensor networks. In: IEEE Canadian Conference on Electrical and Computer Engineering(CCECE), pp. 1787–1792 (2006)
- Younis, M., Munshi, P., Gupta, G., Elsharkawy, S.M.: On efficient clustering of wireless sensor networks. In: IEEE Workshop Dependability and Security in Sensor Networks and Systems, pp. 78–91 (2006)
- Karl, H., Wiling, A.: Protocols and Architectures for Wireless Sensor Networks. Wiley, New York (2005)