社団法人 情報処理学会 研究報告
IPSJ SIG Technical Reports

2006－ＭＢＬ－37（９）
2006／5／18

クラスタリングと内部ネットワーク処理を利用した経路制御アルゴリズムに基づくデータ収集の仕組み

秋 恩炯† 峯 恒憲‡ 雨宮 真人‡

†九州大学大学院システム情報科学府知能システム学専攻 〒816-8580 春日市春日公園 6-1
‡九州大学大学院システム情報科学研究院知能システム学部門 〒816-8580 春日市春日公園 6-1

あらまし ネットワークの生存期間を延ばすため，センサーネットワークでは，エネルギーに敏感で，効率的なデータ収集手法が要求される．この論文では，エネルギーを効率的に扱うデータ収集手法を提案する．この手法では，まずセンサーノードをクラスタリングし，作成された各クラスタ毎に，分散されたデータ経路木を，ネットワーク内部でのデータ統合手法に基づいて，形成する．この手法では，センサーノードの ID を利用して，順番に，１つのクラスタヘッドを選択する．更に，クラスタの形成とデータ経路木の構成を同時に実行する．これにより，観測されたデータの経路木を構成するのに従来必要であったエネルギーを削減できる。また、各ノードでデータの集約を行うことで、転送するデータの総量も削減する。実験結果により、提案手法が、ダイレクトスキームや LEACH と比べて、ネットワークの生存期間の観点で性能を上回ることがわかった。

キーワード センサーネットワーク、データ収集、クラスタリング、データ集約、データ経路木

A Data Gathering Mechanism based on Clustering and In-Network processing Routing Algorithm

Eunhyoung Chu † Tsunenori Mine ‡ Makoto Amamiya ‡

Department of Intelligent Systems
{ † Graduate School， ‡ Faculty} of Information Science and Electrical Engineering
Kyushu University
6-1 Kasuga-Koen, Kasuga, Fukuoka 816-8580 JAPAN
Phone: +81-92-583-7615 FAX: +81-92-583-1338
E-mail: †chueh@al.is.kyushu-u.ac.jp ‡{mine, amamiya}@is.kyushu-u.ac.jp

Abstract Sensor networks require energy-aware, efficient data-collecting methods to extend network lifetime. In this paper, we propose an energy-efficient data gathering mechanism which clusters sensor nodes and forms a distributed data-routing tree based on in-network data fusion within each cluster. In our mechanism, the cluster formation and the data-routing tree construction are simultaneously carried out so that they reduce their energy required to organize a multi-hop routing tree of sensed data. The mechanism also performs data aggregation at each member node to reduce the amount of transmission data. Moreover our work distributes energy load to each node to avoid the intensive energy consumption of a cluster-head. Experimental results show that our data gathering mechanism outperforms the direct scheme protocol and the LEACH protocol on the point of view of the network lifetime.

A list of keywords Sensor networks, data gathering, clustering, data aggregation, data-routing tree

1. Instruction

Recent advances in computing technology have led to the development of a new computing device: the wireless, battery-powered, smart sensors. Sensors which are capable of sensing, computing and communicating may be deployed in ad hoc network without infrastructure and centralized control. Self-organizing and self-configuring capability are requisite for such sensor networks. In addition, activities of sensors are severely constrained by limited resources such as battery power, memory and computing capability available, which require sensor networks to be energy-efficient.

Direct transmission that sensors directly transfer data to the base station, may not be adequate to the sensor networks because of their limited power. Communication may occur via intermediaries in a multi-hop fashion. Moreover because adjacent sensor nodes obtain similar or identical data, using in-network aggregation in a multi-hop communication is useful to reduce the volume of transmission data. A clustering technique which gathers data from several representative sensor nodes by grouping sensors provides scalability for the sensor networks that are composed of hundreds or thousands of nodes. Clustering is essential for applications requiring efficient data aggregation. Another advantage of clustering technique is to reduce energy consumption of the network [2-12].

We present an energy-efficient data gathering mechanism which employs a hierarchical clustering algorithm and in-network processing. We call the mechanism CIPRA (data gathering mechanism based on Clustering and In-network Processing Routing Algorithm). CIPRA prolongs network lifetime by distributing energy consumption, is able to self-organize a data-routing tree and is self-configuring with local information of each sensor. Sensors should dynamically adjust radio transmission energy to adapt to the change of a network topology caused by disappearance of nodes. CIPRA distribute the energy load of the cluster-head to member sensors so that energy of each sensor equally decreases over the whole network. In CIPRA, after sensors senses data, each node sends the data to its neighbor node instead of its cluster-head. Neighbor nodes aggregate data to reduce the amount of data and transfer the aggregated data to their neighbor nodes or their cluster-head. Using local communication among neighbor nodes lessens the communication distance. In-network processing at each member node distributes the energy load of cluster-heads to the member nodes.

The rest of the paper is organized as follows. Section 2 discusses related work, and Section 3 describes our sensor networks and a radio model. Section 4 describes the proposed algorithm in detail. And then in Section 5 experimental results show energy-effectiveness of our algorithm. Then we conclude in Section 6 and present the future work in Section 7.

2. Related work

Many protocols have been proposed and designed in order to extend the lifetime of sensor networks with constrained resources. For example, Directed Diffusion is data-centric in its network view and performs all routing decisions through local, neighbor-to-neighbor interactions [8]. It provides a reactive routing technique, discovering routes between information sources and the base station. TAG (Tiny Aggregation) is a generic aggregation service for ad hoc networks of TinyOS motes. TAG aggregates data in the network using piggybacking aggregation queries on the existing ad hoc network protocol [9].

As a representative clustering protocol LEACH protocol was proposed [4-5]. It was a solution using randomized cluster-head selection and data aggregation at cluster heads. In LEACH, a pre-determined percentage of sensor nodes become cluster-heads per round. After clusters are formed, their cluster-heads gather and aggregate sensor data from their members in their vicinity, and transfer the aggregated data to the base station. LEACH employs the cluster-head rotation for balancing the energy load of cluster-heads. However LEACH does not guarantee to make a good cluster head distribution and select the pre-determined optimal number of cluster-heads per round. ACE makes clusters of the senor networks using the node degree as the main parameter [12]. In PEGASIS, each node aggregates data over a chain routing path after forming it with the closest neighbor and only a cluster head transmits the aggregated data to the base station [6]. HEED selects cluster-heads according to nodes' energy and a secondary parameter, such as node proximity to nodes' neighbors or degree of the node [7].

3. Sensor networks

We assume several properties about sensor networks.

1) A set of sensors is dispersed on a rectangular field.
2) The sensors are located in fixed but unknown place the network topology does not change.
3) The sensors use the symmetric communication channel.
4) The sensors make all decisions without a centralized controller.
5) The sensors enable to adjust their radio radius.

In our work, we use the same radio energy dissipation model as LEACH. To transmit an l bit message a distance d, the radio expends

$$E_{TX}(l,d)=E_{TX\text{-}elec}+E_{TXamp}(l,d)$$
$$=l*E_{elec}+l*\varepsilon_{amp}*d^k \quad (1)$$

and to receive this message, the radio expends

$$E_{RX}(l)=E_{RX\text{-}elec}=l*E_{elce} \quad (2)$$

In figure 1, there are some communication energy parameters: the electronics energy (E_{elec}), the amplifier energy ε_{amp} and the energy for data aggregation (E_{DA}). ε_{amp} varies according to distance d between a sender and a receiver : $\varepsilon_{amp}=E_{fs}$ assuming a free space model when $d < d_0$ and k=2, while $E_{amp}=E_{mp}$ assuming a multipath model when $d \leq d_0$ and k=4, when d_0 is a constant distance that depends on the environment. Each parameter is set as follows: E_{elec}=50nJ/bit, E_{fs}=10nJ/bit, E_{mp}=0.0013pJ/bit, E_{DA}= 5nJ/bit/signal and d_0=75m.

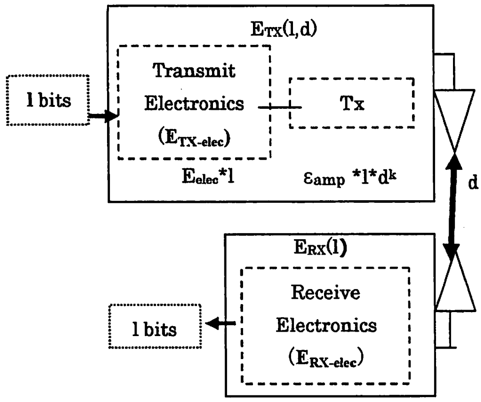

Figure 1. The radio model

4. Algorithm

4.1 Protocol outline

In this section, we describe CIPRA, a novel data gathering algorithm that employs a clustering architecture based on in-network aggregation at each sensor node. The problems we address are to design how to select the optimal cluster head, how to form clusters and how to construct an energy-aware data-routing tree path for the in-network processing.

A round of CIPRA is composed of three phases: (1) cluster head selection, (2) cluster formation and the data-routing tree construction and (3) data transmission.

4.2 Cluster head selection: A cluster head

Sensors consume considerable energy to send data to a remote base station located in far away distance. To lessen the number of sensors transferring data to the base station may prolong the network lifetime because of reducing energy consumption. It is most energy-efficient to have only their single cluster head. We use ID of a node. Using sensor ID, each sensor takes turns a role of the cluster head.

4.3 Routing-tree construction for data transmission and in-network aggregation

CIPRA organizes a topology of a doughnut form based on a tree structure, in which the cluster head is located at the center. During constructing a data-routing tree, each node selects a parent node which receives and fuses data transmitted from its child nodes. Also, sensors determine their tree value which is a depth of the data-routing tree. The tree value also presents the order of their data transmission to their parent nodes after the data-routing tree is constructed. The value of the root is zero and the value of other nodes increases one by one towards the leaves of the tree.

A data-routing tree organizing method starts from the cluster-head which is the root of the tree. A sensor broadcasts in turns to neighbors in its radio range which is a range of a doughnut ring. Then sensors which receive a message of routing-tree construction, select their parent and set their tree value and broadcast a message to their neighbors.

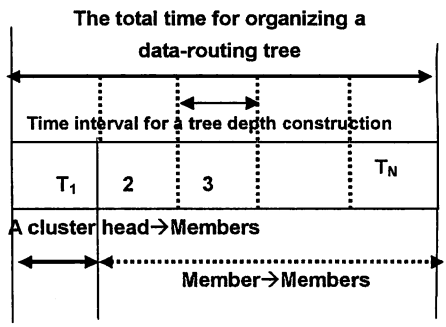

Figure 2. The time of organizing a data-routing tree

Figure 2 shows the time interval of data-routing tree construction. The total time of construction a data-routing tree is calculated considering the longest length of a field, especially, a diagonal length of a field. The total time is divided into N by a time interval of constructing the tree. T_N is a pre-calculated number which is the diagonal length of a field divided by the radio range which is a ring diameter of the doughnut form. During each time interval, the processes of broadcasting and receiving data-routing messages and of selecting parents are done. In other words, a sensor which receives the data-routing messages from other sensors selects a parent and sets tree value. Then the sensor broadcasts the data-routing message.

Figure 3 depicts a process of the data-tree construction in Figure 2. First, as shown in figure 3.1, cluster head node0 broadcasts a message to its neighbors within its radio range. The message is composed of the sender sensor's ID, a tree level value of the sensor. Sensors that receive the message from a cluster head set their parent to the cluster head and tree level to 1. After that, they broadcast their message to their neighbor sensors in their radio range. As presented in figure 3.2, each sensor that receives the message calculates its own tree value by adding 1 to the tree level in the message selects its own parent node and broadcasts its message to neighbors.

(1) At time interval T_1

1.1) A cluster head, node0, broadcasts to members

Message <ID: 0,Tree level value: 0 >

1.2) If sensors receive the message from a cluster head

1.2.1) Set parent to the cluster head and its tree level value to 1

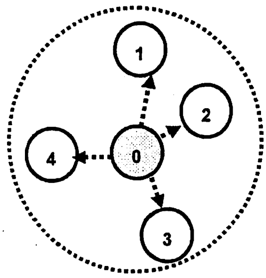

Figure 3.1 Message broadcast from a cluster head to members

(2) At time interval from T_2 to T_N

2.1) Sensors that have received the message broadcast to their neighbors

Message < id, the tree level value >

2.2) If sensors receive the message

2.2.1) Select a parent among the received sensors

2.2.2) Set a sensor that is located in the minimum distance to a parent

2.2.3) Set its tree level valve to one added by 1 to the tree-level value in the message

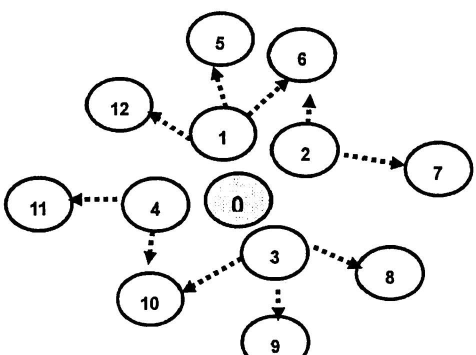

Figure 3.2 From members to members

(3) Repeat step 2 until N time

Figure 3. The CIPRA algorithm constructs the routing tree

In PEGASIS[6], authors assumed that all nodes have global knowledge of the network and employed the greedy algorithm. A chain-based path using the greedy algorithm was constructed to transmit data to the base station. However it is difficult or impossible that sensor nodes have global knowledge of the network due to the limitation of their memory capacity and processing power. CIPRA does not require each node to know the network topology. CIPRA performs clustering and data-routing tree formation using local information. When comparing with LEACH, the transmission distance of each sensor in CIPRA is shorter than that in LEACH because each sensor in LEACH directly transmits data to a cluster head by single hop.

4.4 Data transmission

In this phase, we describe data flow over a data –routing tree to the base station via a single cluster head. Each sensor waits for data from all of its child sensors before sending up its aggregation to its parent node. The total data transmission time is divided into N intervals for transmitting and receiving data between parent nodes and child nodes. Each interval is called communication slot. The number of slot is equal to the pre-calculated depth of routing tree. During a communication slot, the sensor delivers its packet to its parent. The parent node aggregates the received data from its child sensors and transmits data to its parent node. These processes are repeated until the data reach the cluster-head. Cluster-head aggregates the gathered data and transfer it to the base station.

For instance, as shown in figure 4, while node 1 and node 2 transmit their data, node 11 waits for receiving them. After receiving the data, node 11 aggregates and transfers the data. The slot mechanism gives energy consumption advantage. To preserve sensor energy, sensors are put into sleeping mode. This is done during idle times after each node has finished sending its data.

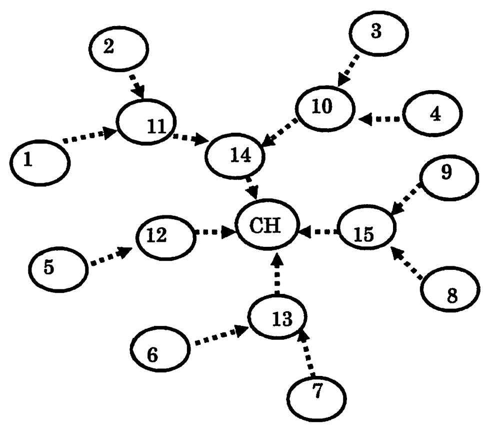

Figure 4. Data transmission over a data-routing tree

In LEACH, each member sensor transmits its data to its cluster head, and the cluster head only aggregates data. On the other hand, in CIPRA each node aggregates data using multi-hop communication across a tree path so that the amount of data is reduced.

5. Evaluation

In this section, we evaluate the performance of CIPRA with a simulator. The aim of the sensor networks is to transmit data of a monitored environment to the base station with prolonged lifetime. Namely, if sensor nodes can't cover the monitored environment, data quality of a system will rapidly decrease. We therefore pay attention to a fact that the satisfactory number of sensors must keep up in a monitored field to maintain high data quality. Therefore we examine the network lifetime that is defined as the number of rounds until energy of the first sensor node is run out.

We assume that 100 nodes are uniformly deployed in a monitored field with 100*100 m^2. The position of the base station is located at (50,175) and (50, 200) in each case of two simulations. Further we assume each node aggregates data received from nodes into a single packet whose size is 800 bits. The initial energy of nodes is 0.25J. The LEACH protocol uses five clusters for a network composed of 100 nodes according to its optimal cluster-head selection equation [4-5].

5.1 A parent selection; the minimum distance VS. the minimum energy

We consider the minimum distance and the minimum energy as the parameter for selecting a parent node. The minimum distance selects a parent that is located in the closest distance from the sender. The minimum energy is to select a parent which holds the least consumption of energy after data is transmitted. In other words, a child node calculates its residual energy after it transmits data to its parent and residual energy of the parent after a parent node receive data. A parent holds the most the residual energy of a child and a parent is selected. That is each node consumes the least amount energy. Therefore consumed energy is the least case.

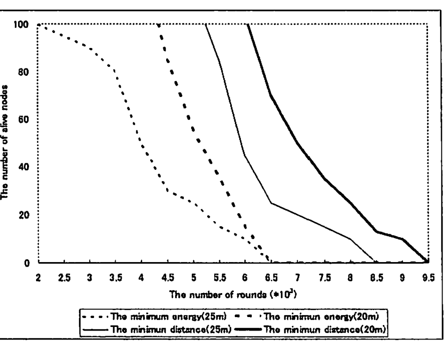

Figure 5. The minimum distance VS. The minimum energy

Figure 5 shows that the network lifetime of using the minimum distance parent selection lasts about 1.3 times as long as that of the minimum energy where both methods ware same radio range. When using the minimum energy, during organizing data-routing tree each node must broadcast its residual energy information. Several child sensors select a sensor which is located in the lower tree level range than it of them and in the edge of a range as their parent. The parent node receive data from several its child nodes. Therefore, compression ratio of data aggregation of the minimum distance is higher than it of the minimum energy.

5.2 Comparison with LEACH and Direct transmission

We compare CIPRA with LEACH and Direct transmission.
For LEACH, we specified that 5% of the nodes would be elected cluster heads.

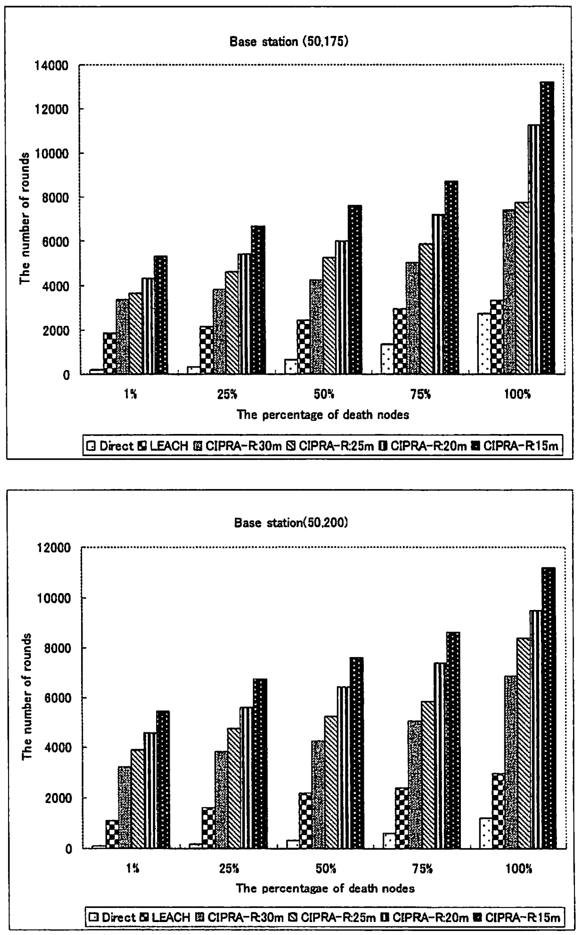

Figure 6. The number of rounds

Figure 6 shows the number of rounds according to the percentage of death nodes by varying the radio range. CIPRA has different network lifetime according to its radio range. For example, the radio range 15m has the longest lifetime of network it is and about 1.5 times than that of the radio range 30m. The lifetime becomes longer as more transmission energy is reduced.

As shown in Figure 5, CIPRA have a good improvement about the first death node compared with LEACH. Moreover the lifetime of the last node death is almost three-times longer than that in LEACH. In other words, CIPRA of 15m radio range lasts about 3 times as long as LEACH.

6. Conclusion

In this paper, we discussed energy-aware and efficient data gathering mechanism, CIPRA, which is based on clustering and a data-routing tree using data aggregation at nodes in clusters. Experimental results showed that our data gathering mechanism outperform the direct scheme protocol and the LEACH protocol from the point of view of the network lifetime. Where not only the time of the first node death, but also the lifetime of the last node became prolonged.

We also conducted an experiment of a parent selection comparing the minimum distance with the minimum energy. Experimental results showed that the minimum distance provides the energy-efficient tree construction than the minimum energy.

7. Future work

CIPRA reduces the energy required for transmitting data to the base station when selecting one cluster head. However when a network composed of several hundreds or thousands sensors requires it multiple cluster heads. In case of the above, we will need an algorithm selecting multiple cluster heads considering the residual energy of each node. Furthermore the increases of the death nodes form holes which break off the communication between nodes. To deal with this, we have to adaptively increase the radio range according to the number of the death nodes in the radio range of each sensor. We need to investigate a energy-efficient method to cope with the hold problems.

References

[1] Ian F. Akyidiz, Weilian Su, Yogesh Sankarasuramniam, "A survey on sensor network", IEEE Communication Magazine, Vol 40, NO.8, pp.102-114, Aug 2002

[2] S. Bandyopadhyay, E. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor network", IEEE Infocom, San Francisco, CA, 2003

[3] Mohamed Younis, Moustafa Youssef, Khaled Arisha, "Energy-Aware routing in Cluster-based sensor networks", Proceedings of the 10th IEEE/ACM International Symposium on Modeling, Analysis, Simulation of Computer and Telecommunication System, Fort Worth, Octorber, 2002

[4] W. R. Heinzelman, A. Chandrakasan, H.Balakrishnan, "Application-sepcific protocol architecture for wireless microsensor sensor networks", IEEE transaction on Wireless Communication 1(4), pp. 600-670, 2002

[5] W. R. Heinzelman, A. Chandrakasan, H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks", in 33rd Annual Hawaii International Conference on System Sciences, pp.3005-3014 , 2000

[6] Stemphanie Lindsey, and Cauligi S. Ranhavendrea, "PEGASIS: Power-Efficient Gathering in Sensor Information Systems", ICC 2001

[7] O. Younis and S. Fahmy, "Distributed clustering in ad-hoc sensor networks: A hybrid, Energy-efficient approach", In Proceeding of IEEE INFOCOM, August 2004

[8] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: A scalable and robust communication paradigm for sensor networks", In Proc, 6th Annual International Conference on MobiCom 2002. pp. 56-67, Boston, ACM press Augest 2000

[9] S. Madden, M. Franklin, J. Hellerstein, and W. Hong, "TAG: a tiny aggregation service for ad-hoc sensor networks", In Pro. of ACM WSNA'02, 2002

[10] Jonathan Beaver, Mohamed A. Sharaf, Alexandros Labrinidis, and Panos K. Chrysantihis, "Location-aware Routing for Data Aggregation in Sensor networks" In: Proceedings of Geo Sensor Networks Workshop 2003

[11] Vivek Mhatre, Catherine Rosenberg, "Design guidelines for wireless sensor networks: communication, clustering and aggregation" to appear in Ad Hoc. Networks Journal 2003.

[12] H. Chan and A. Perrig, "ACE: An Emergent Algorithm for Highly Uniform Cluster Formation", Proc. First European Workshop Sensor Networks, Jan. 2004

[13] M, Bhardwaj, A. Chandrakasan, and T. Garnett, "Upper bounds on the lifetime of sensor networks", in IEEE International Conference on communications, pp.5-67, 2001

[14] P. Levis, S. Madden, J. Polastre, R. Szewczyk, K. Whitehouse, A. Woo, D. Gay, J. Hill, M. Welsh, E. Brewer, and D. Culler, "Tinyos: An operating system for wireless sensor networks" Springer-Verlag, To Appear. 2004.

[15] L. M. S. C. of the IEEE Computer society, "Wireless lan medium access control (mac) and physical layer (phy) spec-ification," in IEEE Std 802.11, 1997.

[16] S. Ganeriwal, R. Kumar, and M. B. Srivastava, "Timing-sync protocol for sensor networks," in SenSys '03: Proceedings of the 1st international conference on Embedded networked sensor systems. New York, NY, USA: ACM Press, pp. 138–149, 2003