Applied Mechanics and Materials Vol. 203 (2012) pp 257-262 *Online: 2012-10-26*
© (2012) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.203.257

Grid-Based Clustering Algorithm for Sensing

Haiyang Zhang[1], Xun Li[1], Xiangke Wang[1], Meng Zhang[2]

[1]School of Mechatronics Engineering and Automation
National University of Defense Technology
Changsha, Hunan Province, China
[2]Qinghe Building Zi 9, Beijing, China

zhyzhyzhy001@126.com

Keywords: sensing cluster, GAF, sampling

Abstract. Energy efficiency is considered as a challenge in Wireless Sense Networks because of the limited energy. In this paper a novel grid-clustering sensing algorithm, the SCA (the sensing clustering algorithm) is proposed in order to minimize energy expenditure and maximize network lifetime. Different to all conventional methods, the proposed algorithm clusters nodes depending on the sensing ability, and forms a comprehensive covered and fully connected network. Both of the theoretical analyses and the simulation indicate that the SCA reduces the energy consumption effectively.

Introduction

Wireless Sense Networks (WSNs) have been applied in many areas such as healthcare, target tracking, monitoring, smart homes, surveillance and intrusion detection [1]. Sensor devices in a wireless sensor network are resource constrained [2, 3]. In order to make the most efficient use of the limited resources and to extend the life of the whole system as far as possible, clustering algorithms were proposed. Traditional clustering algorithms were proposed based on the nodes' wireless communication capacity. A great of considerations are concerning with the analyses of the algorithms in terms of energy saving and using efficiency, and validated that the cluster algorithms perform better than the uncluster algorithms.

Conventional view assumed that the sensing and computing in WSNs consume less energy than the communication. Recently some experiments on complex phenomenon monitoring show that the above assumption might not be entirely true. For some real-time applications in monitoring area, especially for biological monitoring, high-power sensors are required [4]. If considering the fact that working time of sensors is often longer than that of wireless communicating, we can consequentially assume that sometimes data gathering consumes more energy than the communication. So for the energy management, we should not only care for the wireless communication module, but also focus on the sensing module.

GAF algorithm, as a kind of geography based clustering algorithm, was proposed in literature [5]. It divides the whole area where nodes are distributed into some small "virtual grids", and then selects a node from each grid to act as the cluster head (CH) in every cycle. The CH keeps in active state while others is in sleep state. This method reduces the energy consumption and extends the network lifetime. However the nodes in this algorithm are not treated equivalently. CH may die out early because of the heavy load on communication, which may break the network structure. So an advanced GAF algorithm was proposed in reference [6], which balances energy load by clustering dynamically and selecting CH in optimal position. Other grid-based clustering protocols, such as GROUP, a grid-clustering routing protocol that provides scalable and efficient packet routing for large-scale wireless sensor networks was proposed in [7]. GROUP can distribute the energy load among the sensors in the network, and provide in-network processing support to reduce the amount of information that must be transmitted to the sink.

WSNs are data-centric systems, in which sensing data and reducing the data redundancy effectively are the important evaluation indices. GAF algorithm is a routing algorithm in ad hoc networks, but it does not make any optimization for sensing application. Once an event happened in the observation region, all nodes closed to the event would participate in sampling signal, which produces a lot related data and results a waste of energy.

In this paper we proposed a novel clustering algorithm that is optimized for sensing, the sensing clustering algorithm (SCA). The basic idea is to divide the observation region into small grids based on sensor range R **.** The known geography information is helpfull to locate the event, so that related grids will rapidly form a sensing cluster (SC) dynamically. All nodes in the SC can sample signal effectively. Compared with the GAF algorithm, the proposed SCA algorithm requires less nodes, reduces the data redundancy and saves energy effectively.

Sensing Grid

In the clustering protocols, the network was divided into clusters. In clustering topological structure, a cluster is composed of a CH and several CMs (cluster members). The lower-level CH cluster is the members of higher-level network, and the top CH communicates with the base station directly. This kind of method will divide the whole network into regions that could connect with each other. According to a certain selection algorithm, a node in a cluster was chosen as the CH, and it was responsible to manage or control the whole cluster, distribute tasks among member nodes, collect information, fuse data in cluster and transmit data among clusters. The communication module of the member nodes can be turned off at most time, and the cluster heads are responsible for long distance communication. Clustering algorithm can not only ensure the data communication, but also, to a great extent, save the energy.

We propose a novel clustering algorithm based on the sensing ability of the nodes in WSNs. The algorithm draws the ideas from the GAF algorithm, whose network model based on the following assumptions:

1. All sensor nodes are with the same configuration: the same hardware and the battery capacity.

2. All sensor nodes are stationary and ware of their own locations through location techniques such as the GPS system.

3. Each sensor node is able to adjust its wireless transceiver's power consumption, so that it can adjust its communication radius.

Nodes are deployed in the observation region (a 2-dimensional area Q measures L on each side, that is Q is $L \times L$ in size). Drawing on the GAF's ideas, Q is divided into virtual grids with the same size, and thus nodes in one grid belong to a cluster. Virtual grid was sized based on the sensing range R of the node. The grid is a square with r units on each side as shown in the Fig.1.

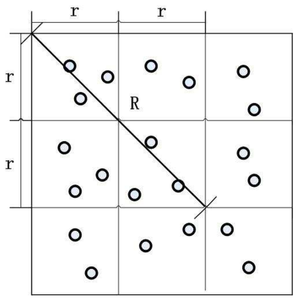

Fig.1 Grid structure

The distance between two possible farthest nodes in any two adjacent grids must not larger than R , it means that any node in a grid can sense event signal in the eight grids around it. So that it can satisfy the requirements of both the connectivity of adjacent grids and the observation of event signal. Therefore, we get:

$$(2r)^2 + (2r)^2 \le R^2; \tag{1}$$

which implies $R \ge \sqrt{8}r$.

Sensing clusters and nodes will be configured after all of the nodes are clustered. The corner of the rectangular sensing area near the SINK node is chosen as the origin, according to which 2-dimensional coordinate system is established. At the same time each grid is numbered by the coordinates. Then numbered all the nodes in grids, and in each grid all the nodes obtain their ID based on their coordinates. Initially NO.1 acts as the node GH (grid head) and works for period T , other nodes in the grid store their grid ID and GH node ID. In order to balance the energy consumption among the sensors in the grid, all nodes in the grid shall rotate as GH. Each node act as the GH for time T , and the grid selects its GH according to the following mechanism: each node records its residual energy E_R and sets a cycle counter C with the initial value being the ID of each node in the grid. After a cycle T , each counter will decrease 1; after the operation the node whose counter value is 1 will be the new GH for the next cycle T . There is a ratio $\eta = E_R / \overline{E}_R$, where E_R stands for residual energy of a node and $\overline{E_R}$ is the average residual energy of all the nodes in the grid. Once the value is less than a constant value α , the node will broadcast a retire message that contains its ID and quit the rotating GH, then the nodes whose IDs are ranked after it will renumber their IDs. This mechanism can avoid energy exhausted too early for a node, which may lead to the situation that the number of valid nodes decreases rapidly. When $\eta = E_R / \overline{E}_R$ is larger than β , the node will renumber itself and join the rotation again. All these operations are processed at the end of each cycle, so that the grid can enter working state rapidly in the next sampling cycle.

Node works in two modes by adjusting its wireless transceiver's power:

1. Maximum mode, in which a node sends wireless signal by maximum power $P_{\max}$, and its communication radius is the largest $R_{\max}$.
2. Energy-saving mode, in which the communication radius is $\frac{1}{2}R_{\max}$, while the power is $\frac{1}{4}P_{\max}$.

Mode 1 is used for long distance communication, while Mode 2 is used for communication in cluster.

Sensing Cluster Forming Protocol

Once an event occurred in the region Q, some nodes could detect the intrusion. Then the protocol will work.

Competition: firstly the node will wake up other nodes in its grid by broadcasting a message in energy-saving mode. Nodes in the grid sample the event signal immediately after being waken, and samples will be send to the GH; After analyzing and averaging these samples, GH will send a Sig-A message. The message contains

- the grid ID;
- the average signal sample strength;

At the same time it will receive other grids' messages, GH compares other grids' signal strength with its own, and the strongest signal grid become the center grid. Center grid will declare its center state by sending a message, and constitute a 3×3 SC with the surrounding eight grids.

Once the SG is formed, the center grid will be the head of sensing cluster (SCH), and other grids will be the members of sensing cluster (SCM). It will synchronize time with the SINK node, and then synchronizes time with the whole SC.

Working: besides sampling, the SCH also is responsible for data fusion and communication, and nodes in the grid will fulfill this function in turn. The rest eight grids are responsible for sampling, and reporting sensed data to the SCH regularly, while the SCH will finish preliminary data fusion, and sent to SINK node.

SCH switch: once the event moves, or some nodes are died, the corresponding SG and GH will change. In every sampling cycle, the SCH will broadcast an IN-message, which contains

- the average of signal strength;
- the top three grids' ID whose signal strength are the strongest among all the eight grids;

Usually the three grids are adjacent, and the top 1 grid is chosen as the probable region for the future forecast. This is a way used for event prediction. Once a grid is sensing that the strength of a signal is stronger than that of the SCH, there will be a new cycle of competition, and a new SC will be formed. The rest of the grids will be in sleep state accordingly.

The algorithm is illustrated as follows:

Step1：Initialize the configuration and parameters;

Step2：$Node\ i = active \quad S_i > 0, Node\ i \in Gi$

For $Node\ j \in Gi$， $Node\ j = active$;

Step3：For $Node\ j \in Gi \quad \overline{S}_i = \sum_{j=1}^{n} S_j$;

Step4：If $\overline{S}_i > \overline{S}_k$， $Gk \subset G$

$Gi = SCH$;

Step5：Select new SCH

If $\overline{S}_k > \overline{S}_{SCH}$

$Gk = SCH$;

Where G stands for the set of all the grids, S_j is the signal strength of node j sensed in the i th grid, $\overline{S}_{SCH}$ is the average signal strength in the SCH.

Performance Evaluation

This part analyzes the energy efficiency of GAF and SCA. The sensing radius of each node is R, the communication radius is R_{MAX}, and the nodes density is ρ. Therefore there are $R_{MAX}^2 \times \rho / 5$ nodes in a GAF cluster, and $R^2 \times \rho / 8$ node in a SCA cluster. Once an event is happened, SCA uses 9 grids to sample the signal, and only one node participates in the sampling in every grid, which means that about 9 nodes are active. For the GAF algorithm, once an event is happened in the region, all $R_{MAX}^2 \times \rho / 5$ nodes in a cluster will be waken and sense the event.

Suppose each node will produce $m(bit)$ data in every sample cycle. Then SCH will fuse all the data and send them to the SINK node. There is a parameter A, which stands for the relative between all the data produced by SCM nodes. The data D that SCH sends to SINK can be calculated by $D = M \times A \times m(bit)$, where $M = R_{MAX}^2 \times \rho$, $A = [\frac{1}{M}, 1]$. Two extreme values of A are analyzed as follows:

- $A = \dfrac{1}{M}$. In this situation all data are highly related, all nodes sample the same data, and SCH could fuse all the data into one;
- $A = 1$, which means all data are completely not related, and SCH will send M data to the SINK node.

According to above analyses, the energy consumption on communication will be calculated as follows:

For GAF algorithm, the inner cluster communication energy consumption is

$$P_{SND} = (M-1) \times m \times P,$$

the energy consumption from the SCH to SINK is

$$P_{SND2} = (M-1) \times m \times A \times n \times P,$$

Where P is the power of send $1bit$ data in a hop, and n is the hops from SCH to SINK.

For SCA algorithm, we obtained accordingly that

$$P_{SND} = 9 \times m \times \frac{1}{4} P_{MIN}, \quad P_{SND2} = 9 \times m \times A \times n \times \frac{R_{MAX}}{R} \times \frac{1}{4} P_{MIN}, \text{ and } \quad P_{TAL} = P_{SND} + P_{SND2};$$

Simulation Results

We use Matlab to simulate SCA and compare it with the GAF algorithm in energy consumption. Nodes are deployed in a $100 \times 100 m^2$ rectangle region evenly. The grid size for GAF is $10 \times 10 m^2$, and $2.5 \times 2.5 m^2$ for SCA. The power consumption of CC2420 for transmission is $35mW$, and for reception is $38mW$, respectively. Power of a accelerometer (3 axis) of ADI is $30mW$ [9]. In every cycle an event is assumbed to happen in the region randomly, whose signal radius varies from $6m$ to $10m$. we use these data for a 25-cycle simulation, and the following two plots show the comparisons between the two algorithm in energy consumption. The results show that the SCA consumes much less energy than the GAF. To be more specifical, when the signal radius is 6, SCA consume 37.5% energy of the GAF in sensing signal, and 62.5% of the GAF in other aspect such as data fusion and competition for the GH. When the signal radius is 9, SCA only consumes 19.3% energy of the GAF in sensing signal, 58.8% of the GAF in other aspects.

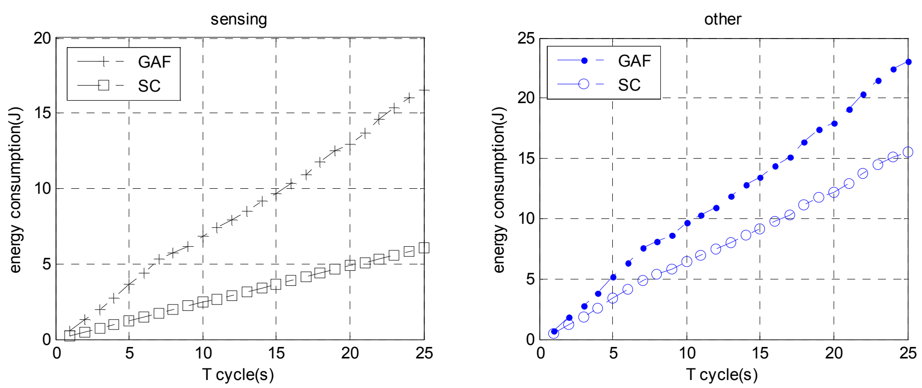

Fig. 2 Simulation results (signal radius= 6 ± 1)

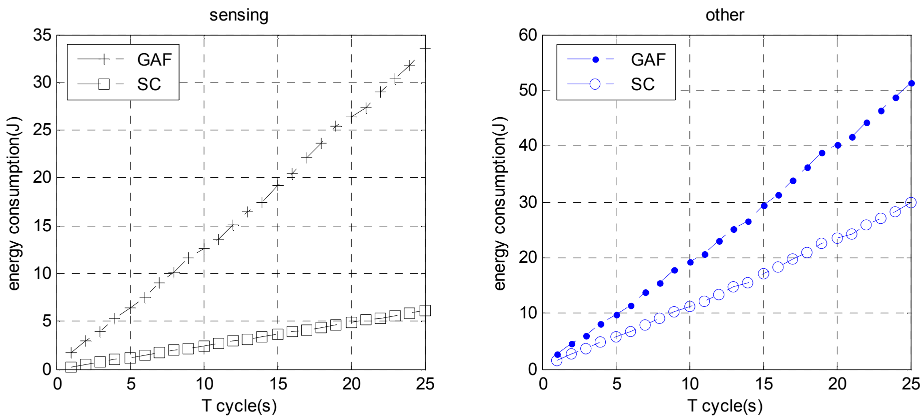

Fig. 3 Simulation results (signal radius=9±1)

Conclusion

In this paper, we described SCA, a grid-clustering sensing algorithm for WSNs. This algorithm clusters nodes into grids based on their sensing capacity. Combined with the known geographic information, it is much more simple and its complication of computation is quite low. So it adapts to the resource-limited situation in sensor node. Simulations validate that SCA is a feasibility and efficiency cluster algorithm for large-scale WSNs, and compared with the GAF it is much more energy-efficient.

Reference

[1] Akyildiz, I.F., Su W., Sankarasubramaniam Y., Cayirci E.: Wireless sensor networks: a survey, Comput. Netw., 2002, 38, (4), pp. 393-442

[2] Wieselthier, J.E., Nguyen, G.D., Ephremides, A.: Resource management in energy-limited, bandwidth-limited, transceiver-limited wireless networks for session-based multicasting, Comput. Netw.: Int. J. Comput. Telecommun. Netw., 2002, 39, (5), pp. 113–131

[3] Potdar, V., Sharif, A., Chang, E.: Wireless sensor networks: a survey. 2009 Int. Conf. on Advanced Information Networking and Applications Workshops, 2009, pp. 636–641

[4]Cesare Alippi, Giuseppe Anastasi and Mario Di Francesco: Manuel Roveri, Energy Management in Wireless Sensor Networks with Energy-hungry Sensors[J], Abstract IEEE Instrumentation and Measurement Magazine Vol. 12, N. 2, April 2009: 16-23

[5] Xu Ya，Heidemann and John Estrin D.: Geography-informed Energy Conservation for Ad Hoc Routing[C], In Proceedings of the Seventh ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM 2001) July 16-21, 2001, 70-84,Rome, Italy.

[6] QI Xiao-gang, QIU Chen-xi: An Improvement of GAF for Lifetime Elongation in Wireless Sensor Networks[J], Journal of Convergence Information Technology Vol. 5, N.7, September 2010;

[7] Liyang Yu, Neng Wang, Wei Zhang and Chunlei Zheng: GROUP: a Grid-clustering Routing Protocol for Wireless Sensor Networks Wireless Communications[C], Networking and Mobile Computing, 2006. (WiCOM2006). International Conference Sept 22-24. 2006, 1-5, Wuhan, China

[8]Li Xun, Wang Jianwen: A Cooperate Framework for Target Tracking in Wireless Sensor Networks[C], The International Conference on Automatic Control and Artificial Intelligence (ACAI2012), Nov.14, 2011, Xiamen, China

[9] Cesare Alippi, Giuseppe Anastasi, Mario Di Francesco and Manuel Roveri: Energy Management in Wireless Sensor Networks with Energy-hungry Sensors[J], Abstract IEEE Instrumentation and Measurement Magazine Vol. 12, N. 2, April 2009. 16-23

Review of Modern Engineering Solutions for the Industry
10.4028/www.scientific.net/AMM.203

Grid-Based Clustering Algorithm for Sensing
10.4028/www.scientific.net/AMM.203.257