http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 5 (1): 188-191, 2006
ISSN 1812-5638
© 2006 Asian Network for Scientific Information

Multi-level Clustering Architecture for Wireless Sensor Networks

Yulin He, Won-Sik Yoon and Jai-Hoon Kim
Department of Electrical and Computer Engineering, Ajou University, Suwon, Korea

Abstract: Recently wireless sensor networks have received increasing attention from the research community. One of the key design challenges in wireless sensor network is conserving the sensor nodes' energies, so as to maximize their lifetime. Clustering technique has proved to be a fundamental mechanism to increases network lifetime and scalability. This study was proposed a novel distributed clustering approach to organize the sensors in a wireless sensor network into clusters. Based on this algorithm, generate a multi-level clustering architecture in a top-down fashion and observe that the energy consumption decrease with the number of levels in the hierarchy. Numerical analysis was also given on the basis of the result from stochastic geometry.

Key words: Sensor networks, clustering, energy efficiency, voronoi tessellation

INTRODUCTION

In recent years, wireless sensor networks have received significant interest from the research Community and emerged as a platform for several important surveillance and control applications in both military as well as civilian domains[1,2]. Generally, sensor networks are formed by hundreds of even thousands of low-cost, low-power and tiny nodes that equipped with sensing, processing and wireless communication capabilities. These sensor nodes can be rapidly deployed in physical environments to collect useful information (e.g. seismic, acoustic, medical and surveillance data). Sensors in such environments typically do not have renewable source of energy and are expected to last until their energy drains. Therefore, energy is a very scarce resource and has to be managed carefully in order to extend the life of the sensors for the duration of a particular application.

Previous research has showed that the energy dissipation in transmitting a bit of data is higher than that in computation[3] and hence it might be energy-efficient to organize the sensors into clusters. In the clustered environment, the data collected by the sensor nodes is communicated to their cluster heads and then is forwarded to the Base Station (BS) through a hierarchy of cluster heads. Since the sensors are now communicating data over smaller distances in the clustered environment, the energy consumed in the network will be much lower than the energy consumed when every sensor communicates directly to the BS. In addition, the strong correlation between data from nodes located closely allows all data from nodes within the cluster to be processed locally, reducing the data set that needs to be transmitted to the end user. In particular, data aggregation techniques can be performed to combine several correlated data into a smaller set of information that maintains the effective data of the original data. Therefore, much less actual data is transmitted from the cluster to the BS.

Several clustering algorithms have been proposed in the context of wireless ad-hoc networks and sensor networks. In WCA[4], a weight-based distributed clustering algorithm has been proposed for mobile ad-hoc network, which combines several parameters including the ideal node degree, transmission power, mobility and battery power of mobile nodes when electing the cluster heads. ACE[5] clusters the sensor network using the node degree as the main parameter. The approach in[6] selects a d-hop dominating set to cluster the network based on node ID, while the approach in[7] selects a dominating set using linear programming relaxation techniques. However, in these researches the energy efficiency is not the main consideration when designing the clustering algorithm. In LEACH[8], the authors have proposed a distributed clustering algorithm for micro-sensor networks. The main advantages include local coordination for cluster formation among sensors, randomized rotation of cluster heads for improved energy utilization. However LEACH protocol does not guarantee good cluster head distribution and has to assume uniform energy consumption for cluster heads.

When designing the clustering algorithm, the essential operation is to identify a set of cluster heads from the set of nodes in the network and then cluster the

Corresponding Author: Won-Sik Yoon, Telecom Research Laboratory, Department of Electrical an Computer Engineering, Ajou University, 5 Woncheon-Dong, Youngtong-Gu, Suwon, 443-749, Korea
Tel: 82 31 219 2371

remaining nodes with these cluster heads. Cluster heads are responsible for coordination among the nodes within their clusters and aggregation of their data and communication with each other and/or with external observers on behalf of their clusters. This study presented a distributed clustering algorithm which the selection of cluster heads is probabilistically depended on the node's residual energy. Based on this approach, present an energy efficient data aggregation clustering hierarchy for wireless sensor networks. We also use the result in stochastic geometry[9,10] to evaluate the performance of the proposed clustering algorithm.

PROBLEM STATEMENT

System model: For the development of the clustering algorithm, made some assumptions about the sensor nodes and the underlying network.

The sensor nodes are distributed as per a homogeneous spatial Poisson process of intensity λ in 2-dimensional space. The locations of the sensor nodes are fixed after deployment. This is typical for sensor network applications. All nodes are assumed to have the ability to transmit their packets to any other sensor in the network or directly to the BS if needed. The nodes can use power control to adjust the amount of transmit power. All nodes have similar processing and communication capabilities and equal significance. All communication is over wireless links. A wireless link is established between two nodes only if they are in range of each other. One unit of energy is used by each node to transmit one unit of data over one unit of squared distance. All nodes have data to send to the base station and data aggregation is performed at cluster heads in each level of the clustering hierarchy, which combines all data transmitted from its members into one unit of data. Furthermore, the network is error-free; therefore sensor nodes do not have to retransmit any data.

Problem statement: Assume that n nodes are dispersed in a field and the above assumptions hold. Our goal is to identify a set of cluster heads which cover the entire field. Each node must be mapped to exactly one cluster. A node must be able to directly communicate with its cluster head. Cluster heads of multi-levels can form a clustering hierarchy which can collect sensed data from nodes and send back to the BS using an underlying routing protocol.

Assume that there are total h levels in the clustering hierarchy with level 1 being the highest level and level h being the lowest. In this clustering hierarchy, the sensors communicate the sensed data to level-h Cluster Heads (CHs). The level-h CHs then aggregate this data and communicate the result to level-(h-1) CHs and so on. Finally, the level-1 CHs communicate the aggregated data to the BS. The total energy cost of communicating the information from the sensors to the sink is the energy spent by the sensors to communicate the information to level-h CHs, plus the energy spent by the level-h CHs to communicate the aggregated information to level-(h-1) CHs, …, plus the energy spent by the level-1 CHs to communicate the aggregated information to the BS. Objective was to create an energy-efficient clustering hierarchy that can efficiently reduce the total energy cost.

CLUSTERING ALGORITHM

The Clustering Algorithm works in a top-down way, that is, the level-j cluster heads will be elected before the level-(j+1) cluster heads. The clustering procedure is initiated by the BS broadcasting a short message indicated it will recruit a new set of level-1 cluster heads. After receiving the message, each sensor node i set its probability of becoming a cluster head, denoted as $P_{CH}(i)$, by using the following formula:

$$P_{CH}(i) = P_{ini}(j) \times \frac{E_{res}(i)}{E_{max}(i)} \quad (1)$$

where, $P_{ini}(j)$ is the initial percentage of level-j cluster heads among all nodes, $E_{res}(i)$ is the estimated current residual energy in the node i and $E_{max}(i)$ is a maximum energy in the node i with a fully charged battery. According to the calculated probability, those nodes with higher residual energy are more likely to become cluster heads.

Once some sensor nodes have elected themselves to be the tentative cluster heads according to the probabilities calculated by the above equation (1), they inform the BS by a short message consisting of their nodes' ID. Note that $P_{ini}(j)$ is used to limit the number of tentative cluster heads. The BS then can randomly choose the final level-1 cluster heads among these tentative cluster heads according to the probability of becoming a cluster head in level-1. As indicated in[8], much better clusters can be produced by dispersing the cluster head nodes throughout the network if the additional information such as the physical location of each node is known as a prior.

Then BS informs those nodes that have been selected as the final level-1 cluster heads with a short message. After knowing themselves as a level-1 cluster heads, all cluster heads in level-1 must broadcast an advertisement message to all the other nodes in the network. This message is also a small message containing only the cluster head's ID and a message header. Each non-cluster

head node determines its cluster by choosing the cluster head that requires the minimum communication energy, based on the received signal strength of the advertisement from each cluster head and responses to its cluster head with a joining-cluster message shows that it will be a member of the cluster. In this way, the level-1 cluster heads and associated clusters are formed. If we need to select the level-2, level-3, …, level-h cluster heads continually, a similar procedure can be followed. The difference from the case in level-1 lies in the fact the cluster heads election procedure is now limited in the range of each cluster for the lower levels in the hierarchy.

NUMERICAL ANALYSIS

As mentioned in the problem statement section, the Total Energy Cost is used as the performance metric when evaluating the proposed algorithm.

By present assumptions, the sensor nodes are distributed according a homogeneous spatial Poisson process of intensity λ and hence, the number of nodes in a given region is a Poisson random variable N with mean λA, where is the area of that region. Without loss of generality, assume that for a fixed realization of the process there are n sensors deployed in a circle area of radius R ($A-\pi R^2$) and the BS is located in the center of the circle. Let D_i be a random variable that denotes the squared distance from sensor node i located in (x_i, y_i), I=1, 2, …, n to the BS, then the expectation of D_i can be given by:

$$E(D_i) = \int_A (x_i^2 + y_i^2)\left(\frac{1}{\pi R^2}\right) dA = \frac{R^2}{2} \tag{2}$$

Also assume that the probability of becoming a cluster head in level-j is p_j (j=1, 2, …, h, apparently, $p_1 < p_2 < \cdots < p_h$), the sum of squared distances from members of level-j cluster to their CHs is denoted by D_j and NCH_j presents the total numbers of CHs in level j.

Since there are on an average np_1 level-1 CHs communicating with the BS and the location of any CH is independent of the locations of other CHs, the energy cost of the communication from all the level-1 CHs to the BS is $np_1R^2/2$.

According to the algorithm, each non-cluster head joins the cluster with the minimum transmission energy, which typically means it joins the closest cluster. Such an organization is introduced by the Voronoi tessellation[9] generated by the locations of the cluster heads. The definition of a Voronoi tessellation with respect to a point set Π is given as follows:

Definition: The Voronoi tessellation with respect to the point set Π, denoted by $V(\Pi)$, is a collection of cells V_{yi} for $y_i \in \Pi$ such that $V_{yi}(\Pi) = \{ z \in R^2 \mid \| y_i - z \| < \| y_j - z \|, \forall y_j \in \Pi \}$ where the Euclidean distance from point y to z is given by $\| y-z \|$. i.e., all points in the plane are closer to y_i than to any other point in Π.

Firstly let us consider a single-level case for the simplicity of description. Since the sensor nodes become cluster heads with a certain probability and the non-clusterheads join the closest cluster, the cluster heads and the non-clusterhead nodes are distributed as per independent homogeneous spatial Poisson processes Π_c and Π_n of intensity λ_c and λ_n, respectively. Seen from the Π_c-points, the plane is divided into zones- Voronoi cells, each cell corresponding to a Π_c process point called its nucleus. If N_v is the random variable denoting the number of Π_n process points in each Voronoi cell and D_v is the sum of squared distance of all segments from the Π_n process points to the nucleus in a Voronoi cell, we have

$$E[D_v] = \lambda_n / \pi\lambda_c^2 \tag{3}$$

Now let us consider the case of multiple levels. Since the sensor nodes obey a homogeneous spatial Poisson process of intensity λ and the cluster heads in level-j are chosen according to the probability p_j, by properties of the Poisson process, level-j CHs, j=1, 2, …, h are governed by homogeneous Poisson processes of intensities $\lambda_j = \lambda p_j$. By using the Eq. 3, the sum of squared distance of a level-j CH from level-(j+1) CHs, j = h-1, …, 2, 1 in a typical level-j cluster is given by:

$$E[D_j] = \frac{\lambda_{j+1}}{\pi\lambda_j^2} = \frac{\lambda p_{j+1}}{\pi(\lambda p_j)^2} = \frac{p_{j+1}}{\pi\lambda p_j^2}. \tag{4}$$

And the expectation of the sum of squared distance of a level-h CH from the sensor nodes is:

$$E[D_h] = \frac{\lambda_h}{\pi\lambda^2} = \frac{p_h}{\pi\lambda}. \tag{5}$$

The expectation of the total numbers of level-j CHs is

$$E[NCH_j] = np_j,\ j = 1, 2, \ldots, h. \tag{6}$$

Let TEC be the total energy cost of communicating information from the sensors to the base station through the hierarchy of cluster heads generated by the clustering algorithms. After some manipulations, the expected value of TEC is given by:

$$\begin{aligned} E[TEC] &= \frac{np_1R^2}{2} + \sum_{j=1}^{h} \left\{ E[NCH_j]\ E[D_j] \right\} \\ &= \frac{np_1R^2}{2} + np_h \frac{p_h}{\pi\lambda} + \sum_{j=1}^{h-1} np_j \frac{p_{j+1}}{\pi\lambda p_j^2} \\ &= \frac{np_1R^2}{2} + \frac{np_h^2}{\pi\lambda} + \sum_{j=1}^{h-1} \frac{np_{j+1}}{\pi\lambda p_j}. \end{aligned} \tag{7}$$

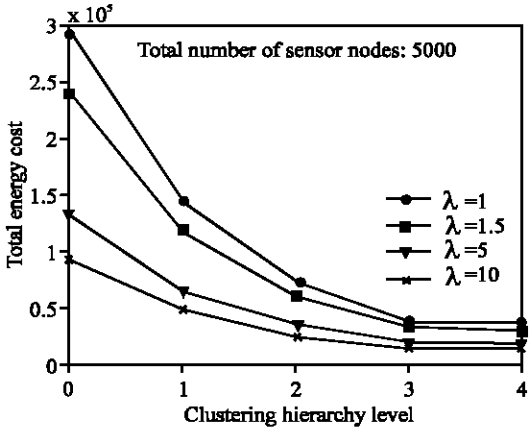

Fig. 1: Total energy cost vs. clustering hierarchy level

We obtain the performance of the proposed algorithm by generating a multi-level clustering architecture with different number of levels in it on sensor networks distributed uniformly with various spatial densities.

Figure 1 shows that the energy cost in the network decreases with the number of levels of clusters in the hierarchy. It was observe that, among the networks of same number of sensor nodes, the total energy cost was higher in the lower density network than in the higher density network. This was due to the fact that the average distance between a node and its cluster head in a lower density network was larger than the distance in a higher density network. In addition, the energy saving with increase in the number of levels in the hierarchy is not so obvious when the number of levels in the hierarchy is over 3, independent of the density of the sensor nodes. This can be explained by the reason that the more levels are in the hierarchy, the more internal nodes the data has to travel through the hierarchy of cluster heads to reach the base station.

CONCLUSIONS

In this study, we have proposed a distributed clustering algorithm for organizing sensors into a hierarchy of clusters. Based on this algorithm, we then generate a multi-level clustering architecture in a top-down way. Numerical analysis using the result from the stochastic geometry shows that the total energy cost in the system to communicate the data collected by the sensor nodes to the base station decrease with the number of levels in the hierarchy.

REFERENCES

1. Cerpa, A. *et al.*, 2001. Habitat monitoring: Application driver for wireless communications technology. In ACM SIGCOMM Workshop on Data Communications in Latin America and the Caribbean.
2. Kahn, J.M. *et al.*, 1999. Next century challenges: Mobile networking for smart dust. In Proceeding MOBICOM, Seattle, WA, pp: 271-278.
3. Pottie, G.J. and W.J. Kaiser, 2000. Wireless integrated network sensors. Communication of the ACM., 435: 51-58.
4. Chatterjee, M. *et al.*, 2002. WCA: A weighted clustering algorithm for mobile Ad-hoc networks. J. Cluster Computing, Special Issue on Mobile Ad-hoc Networking, 5: 193-204.
5. Chan, H. and A. Perrig, 2004. ACE: An emergent algorithm for highly uniform cluster formation. Proceeding of 1st Eur. Workshop Sensor Networks.
6. Amis, A.D. *et al.*, 2000. Max-min d-cluster formation in wireless ad-hoc networks. Proceeding of IEEE INFOCOM.
7. Kuhn, F. and R. Wattenhofer, 2003. Constant-time distributed dominating set approximation. Proc. ACM Symp. Principles of Distributed Computing (PODC).
8. Heinzelman, W.R. *et al.*, 2000. Energy-efficient communication protocol for wireless micro-sensor networks. In Proceedings of IEEE HICSS.
9. Baccelli, F. and S.A. Zuyev, 1996. Poisson voronoi spanning trees with applications to the optimization of communication networks. INRIA, Sophia Antipolis, France, Res. Rep., pp: 3040.
10. Foss, S.G. and S.A. Zuyev, 1993. On a certain segment process with voronoi clustering. INRIA, Rapport de Recherche.