



Design Issues and Parameters for Cluster Head Selection in Energy Efficient Wireless Sensor Networks

Rajkaran Singh, Dr. Amit Sharma

Punjab Technical University

Punjab, India

Abstract— *Wireless sensor networks comprise a large number of small sensor nodes scattered across limited geographical areas. The nodes in such networks carry sources of limited and mainly unchangeable energy. Therefore, it is required to operate these networks under energy efficient protocols and structure. Energy efficient clustering algorithms have been developed to reduce the networks energy consumption and extends its life time. In this paper we discuss various design issues and certain criteria's that should be considered while selecting the best cluster head (CH) in order to improve the network lifetime.*

Keywords— *Wireless sensor network, Clustering protocol, Distributed, Network lifetime, Energy efficiency*

I. INTRODUCTION

Wireless sensor networks comprises of a large number of sensor nodes randomly scattered over an physical area that gather data from the environment. Since the location of sensor nodes does not necessarily need to be determined in advance, they can be deployed over an inaccessible area. In network clustering have been shown as effective technique in reducing energy consumption in WSNs. In a typical clustering algorithm, a number of nodes in a network will be selected as the cluster heads (CHs). The remaining nodes will be regarded as cluster members (CMs) and form connections with the CHs. A CH will collect information from its CMs. A Wireless sensor network (WSNs) consists of a set of sensor nodes which is not controlled by any outside sources. It will have little communication links, and then, collectively performs tasks without help from any central servers. Accurate data extraction is difficult in sensor networks. WSNs consist of large number of small autonomous wireless devices, called sensor nodes. Sensor nodes are battery powered devices [1]. The energy requirement for sensor nodes while sensing, communication, and computation, leads to energy consumption, when transmitting data. Thus, it is a great consequence to design an energy efficient routing scheme for reporting sensory data to achieve a high delivery ratio and enhancement of network lifetime.

The topology of the WSNs can vary from a simple star network to an advanced multi-hop wireless mesh network. Cluster-based routing algorithm has a better energy utilization rate compared with non-cluster routing algorithm. A wireless sensor network system usually includes sensor nodes, sink node or base station. The data monitored by sensor nodes is transmitted along other nodes one by one, that will reach the sink node after a multi-hop routing and finally reach the base station, through the wired and or wireless network. The energy, the storage capacity and communication capability of sensor nodes are very limited. Random distribution of the nodes in the sensing field makes battery recharge or exchange an impossible fact. Due to their energy constraints, wireless sensors usually have a limited range to transmit, so it makes multi hop data routing toward the cluster head more energy efficient than direct transmission. A primary design goal for wireless sensor networks is to use the energy efficiently. The idea of clustering is to use the information aggregation mechanism in the cluster head (CH) to reduce -the amount of data transmission, thereby, reduce the energy dissipation in communication. In the clustering routing algorithm for wireless sensor networks, LEACH is the well-known routing protocol in wireless sensor network because it is the simplest and first most efficient routing protocol. It enhanced the network lifetime a lot by using the clustering approach and multi hop communication.

II. RELATED WORK

W. R. Heinzelman, A. P. Chandrakasan and H. Balakrishnan [2] proposed Low Energy Adaptive Clustering Hierarchy (LEACH) protocol in 2000. It is one of the most popular hierarchical routing algorithms for sensor networks. The idea is to form clusters of the sensor nodes based on the received signal strength and use local cluster heads as routers to the sink. This will save energy since the transmissions will only be done by such cluster heads rather than all sensor nodes. Optimal number of cluster heads is estimated to be 5% of the total number of nodes. All the data processing such as data fusion and aggregation are local to the cluster. Cluster heads change randomly over time in order to balance the energy dissipation of nodes. This decision is made by the node choosing a random number between 0 and 1. The node becomes a cluster head for the current round if the number is less than the following threshold:

$$T(n) = \begin{cases} \frac{p}{1 - p \times (r \bmod 1/p)} & \text{if } n \in G \\ 0 & \text{Otherwise} \end{cases} \quad (1)$$

where p is the desired percentage of cluster heads (e.g. 0.05), r is = the current round, and G is the set of nodes that have not been cluster heads in the last $1/p$ rounds.

S. Lindsey and C. Raghavendra [3] introduced Power Efficient Gathering in Sensor Information Systems (PEGASIS) protocol in 2002. It is an improved version of LEACH. Instead of forming clusters, it is based on forming chains of sensor nodes. One node is responsible for routing the aggregated data to the sink. Each node aggregates the collected data with its own data, and then passes the aggregated data to the next ring. The difference from LEACH is to employ multi hop transmission and selecting only one node to transmit to the sink or base station. Since the overhead caused by dynamic cluster formation is eliminated,

multi hop transmission and data aggregation is employed, PEGASIS outperforms the LEACH. However excessive delay is introduced for distant nodes, especially for large networks and single leader can be a bottleneck.

In 2001, A. Manjeshwar and D. P. Agarwal [4] proposed Threshold sensitive Energy Efficient sensor network Protocol (TEEN) protocol. Closer nodes form clusters, with a cluster heads to transmit the collected data to one upper layer. Forming the clusters, cluster heads broadcast two threshold values. First one is hard threshold; it is minimum possible value of an attribute to trigger a sensor node. Hard threshold allows nodes transmit the event, if the event occurs in the range of interest. Therefore a significant reduction of the transmission delay occurs. Unless a change of minimum soft threshold occurs, the nodes don't send a new data packet. Employing soft threshold prevents from the redundant data transmission. Since the protocol is to be responsive to the sudden changes in the sensed attribute, it is suitable for time-critical applications.

A. Manjeshwar and D. P. Agarwal [5] proposed AdaPtive Threshold sensitive Energy Efficient sensor Network Protocol (APTEEN) protocol in 2002. The protocol is an extension of TEEN aiming to capture both time-critical events and periodic data collections. The network architecture is same as TEEN. After forming clusters the cluster heads broadcast attributes, the threshold values, and the transmission schedule to all nodes. Cluster heads are also responsible for data aggregation in order to decrease the size data transmitted so energy consumed. According to energy dissipation and network lifetime, TEEN gives better performance than LEACH and APTEEN because of the decreased number of transmissions. The main drawbacks of TEEN and APTEEN are overhead and complexity of forming clusters in multiple levels, implementing threshold-based functions and dealing with attribute based naming of queries.

O. Younis and S. Fahmy proposed [6] Hybrid Energy Efficient Distributed clustering Protocol (HEED) protocol in 2004. It extends the basic scheme of LEACH by using residual energy as primary parameter and network topology features (e.g. node degree, distances to neighbors) are only used as secondary parameters to break tie between candidate cluster heads, as a metric for cluster selection to achieve power balancing. The clustering process is divided into a number of iterations, and in each iterations, nodes which are not covered by any cluster head double their probability of becoming a cluster head. Since these energy-efficient clustering protocols enable every node to independently and probabilistically decide on its role in the clustered network, they cannot guarantee optimal elected set of cluster heads. Upon running HEED protocol on the nodes, each sensor node chooses a random number between 0 and 1. If the number is smaller than or equal to CH probability of the node, the node will become a CH on a trial basis and sends alert messages to its adjacent nodes. The probability of becoming a CH is calculated as:

$$E_{prob} = C_{prob} \times \frac{E_{residue}}{E_{max}} \quad (2)$$

where, $E_{residue}$ is the residual energy of the node, E_{max} is the maximum energy when the battery is full, and C_{prob} denotes the primary percentage of the number of CHs which is initially set to 5 %. If there are several candidates for the CH role, the ones with lower communication costs will be chosen

III. DESIGN ISSUES

The challenges posed by the deployment of sensor networks is a superset of those found in wireless ad hoc networks. Sensor nodes communicate over wireless, lossy lines with no infrastructure. An additional challenge is related to the limited, usually nonrenewable energy supply of the sensor nodes. In order to maximize the lifetime of the network, the protocols need to be designed from the beginning with the objective of efficient management of the energy resources [1]. Let us now discuss the individual design issues in greater detail.

Fault Tolerance. Sensor nodes are vulnerable and frequently deployed in dangerous environment. Nodes can fail due to hardware problems or physical damage or by exhausting their energy supply. We expect the node failures to be much higher than the one normally considered in wired or infrastructure-based wireless networks. The protocols deployed in a sensor network should be able to detect these failures as soon as possible and be robust enough to handle a relatively large number of failures while maintaining the overall functionality of the network. This is specially relevant to the routing protocol design, which has to ensure that alternate paths are available for rerouting of the packets. Different deployment environments pose different fault tolerance requirements.

Scalability. Sensor networks vary in scale from several nodes to potentially several hundred thousand. In addition, the deployment density is also variable. For collecting high-resolution data, the node density might reach the level where a node has several thousand neighbors in their transmission range. The protocols deployed in sensor networks need to be scalable to these levels and be able to maintain adequate performance.

Production Costs. Because many deployment models consider the sensor nodes to be disposable devices, sensor networks can compete with traditional information gathering approaches only if the individual sensor nodes can be produced very cheaply.

Hardware Constraints. At minimum, every sensor node needs to have a sensing unit, a processing unit, a transmission unit, and a power supply. Optionally, the nodes may have several built-in sensors or additional devices such as a localization system to enable location-aware routing. However every additional functionality comes with additional cost

and increases the power consumption and physical size of the node. Thus, additional functionality needs to be always balanced against cost and low-power requirements.

Transmission Media. The communication between the nodes is normally implemented using radio communication over the popular ISM bands. However, some sensor networks use optical or infrared communication, with the latter having the advantage of being robust and virtually interference free.

Power Consumption. As we have already seen, many of the challenges of sensor networks revolve around the limited power resources. The size of the nodes limits the size of the battery. The software and hardware design needs to carefully consider the issues of efficient energy use. For instance, data compression might reduce the amount of energy used for radio transmission, but uses additional energy for computation and/or filtering. The energy policy also depends on the application; in some applications, it might be acceptable to turn off a subset of nodes in order to conserve energy while other applications require all nodes operating simultaneously.

IV. PARAMETERS USED FOR CH SELECTION IN WIRELESS SENSOR NETWORKS

In this section, we discuss the optimal clustering parameters used to improve the network lifetime in WSN's. The notation used in this paper are listed in Table 1

Table: 1

Notation	Definition
N_{total}	Total number of network nodes
N	Number of nodes in each cluster
$E_{current}$	Remaining energy of the sensor nodes
E_{max}	Maximum energy of a sensor node when the battery is fully charge
d_{BS}	Distance between the node and BS
d_{far}	Distance of farthest node to BS
dis_i	Distance between the node and its i th neighbor
dis_{max}	Distance between the node and its furthest neighbor
$count_{nig}$	The number of neighbors of the node
K_{opt}	The optimal number of clusters in the network
Threshold	The optimal number of nodes in a cluster
$N^+(i)$	The neighbor set of node i that have higher score than it
$N^-(i)$	The neighbor set of node i that have lower score than it
$score(i)$	Score of node i
$RCR_i(v)$	Node i 's tendency to selects node v as its CH
$SRCR_i(i)$	Sum of $RCR_i(i)$

A. The Residual Energy of Node

In a large scale sensor network with nodes that have limited energy sources, we expect CHs to consume more energy to communicate with member nodes [8]. If required, they combine their own data together and send them to the BS or the next-step CH. Therefore, a CH node consumes more energy than a non-CH node and is very likely to run out of energy quicker than its member nodes [8]. Therefore, the clustering process should include the periodical assignment of the CH role among nodes with the highest residual energy. So, score parameter (SP) is defined as a criterion for a node with higher residual energy to obtain more score than other nodes to be qualified as the cluster head. SP1 is therefore, defined as follows

$$SP1 = \frac{E_{current}}{E_{max}} \quad (3)$$

where $E_{current}$ denotes the volume of residual energy and E_{max} is the maximum volume of the energy of a node when it is fully charged. Since residual energy appears in the numerator, nodes with higher residual energy will more likely be selected as CH

B. The Distance Between Node and BS

A node selected as a CH is tasked with collecting data from member nodes, deleting repeated data, conducting a limited processing operation on the data, and finally sending them to the BS [9]. Since data transmission, especially over long distances, leads to the highest energy consumption in a sensor node [10], the closer the CH candidate is to the BS the lower will be the amount of energy it consumes to send the data to the BS. As a result, CH lifetime and the duration of the steady-state phase increase in each round of clustering (the interval between two set-up phases of the CH selection increases), that prevents extra energy consumption for reselecting CHs., such increases the network lifetime consequently. Thus, nodes that are closer to the BS more likely become a CH. SP2 is defined by Eq. (4) as a criterion for the closer node to BS, in order to obtain more score than other nodes as a prerequisite to become a cluster head

$$SP2 = 1 - \frac{d_{bs}}{d_{far}} \quad (4)$$

where, d_{bs} the distance between each node and BS.

C. The Distance Between Node and Its Neighboring Nodes-

There is a direct relationship between distance and the energy consumed for transmitting data since each member node sends the data it has locally sensed to its CH. Therefore, the closer the member nodes are to the CH, the lesser will be the amount of energy they consume to transmit data they have sensed [7, 8]. SP3 is defined by Eq. (5) as a criterion for the node that is closer to its neighboring nodes to obtain more score than others as a pre-requisite to become qualifies as a cluster head.

$$SP3 = 1 - \frac{\sum_1^N dis_i}{N \times dis_{max}} \tag{5}$$

where, dis_i denotes the distance between the node and its i^{th} neighbor. dis_{max} denotes the distance between the node and its remotest neighboring node and N is the number of neighboring nodes. Since the total distance is the numerator and $N \times dis_{max}$ is the denominator in Eq. (5), and also considering that this statement is multiplied by a negative number, the smaller the total distance between a node and its neighbors is, the lower the SP3 will be. In other words, smaller total distance between a node and its neighbors, increases its chance to become a CH

D. The Number of Neighboring Nodes

There must be a good balance between the number and size of clusters [11]. Therefore, it is necessary to set a threshold for the number of cluster members such that nodes with number of neighbors close to the threshold will have the higher chance to be selected as CH. SP4 is defined by Eq. (6) as a criterion for attaining more scores by the nodes with number of neighbors close to the threshold, in order to become qualifies as a cluster head. where, $count_{nig}$ denotes the number of neighboring nodes of each node and $threshold$ stands for the optimal number of neighbors for each node. In Eq. (6), SP4 will be equal to 1 if the number of the node's neighbors $count_{nig}$ is equal to the optimal number of neighbors (threshold), and less than 1 if counting is either higher or lower than the threshold. The threshold is obtained as follows:

$$SP4 = 1 - \frac{(count_{nig} - threshold)^2}{threshold^2} \tag{6}$$

where, $count_{nig}$ denotes the number of neighboring nodes of each node and $threshold$ stands for the optimal number of neighbors for each node

In Eq. (6), SP4 will be equal to 1 if the number of the node's neighbors $count_{nig}$ is equal to the optimal number of neighbors (threshold), and less than 1 if $count_{nig}$ is either higher or lower than the threshold. The threshold is obtained as follows:

$$threshold = \frac{N_{total}}{K_{opt}} \tag{7}$$

where, N_{total} denotes the total number of the nodes and K_{opt} is the optimal number of clusters.

E. Network density

The network density is the number of nodes per square meter. It varies from one deployment to another and from one node to another within the same deployment depending on the node distribution. According to authors[1], this parameter does not have a fixed value to be used as a reference. The ideal value is application and environment dependent. In addition, this parameter has a network management importance as it helps to identify the dense zones of the network and the non-well covered zones. Hence, it may lead to redeployment of more nodes in some zones for a better coverage.

F. Node position within the network

Another parameter related to the network deployment has also to be studied carefully, due to its importance as we will explain in this section. This parameter is the position (P) of the agent node in the network. We define three types of node positions: (1) normal, (2) edge and (3) critical. The normal position is the position inside the network where the node has multiple neighbors. This kind of nodes may tend toward the cooperative behavior, to maximize the amount of the important information collected in the network. The edge node (E in figure 1) is a node in the border of the network, which has a restricted view of the network limited to only one neighbor.

A node is considered in a critical position (C in figure 1) if it connects two parts of the network. That means, if the node runs out of battery, it may divide the network and multiple nodes behind it will become unreachable and in the best case they will require a longer route to communicate their data to the sink. This longer route is expensive in term of energy as the number of hops is increased. For example, in figure 1 , if a C node runs out of battery, the network will be divided in two parts.

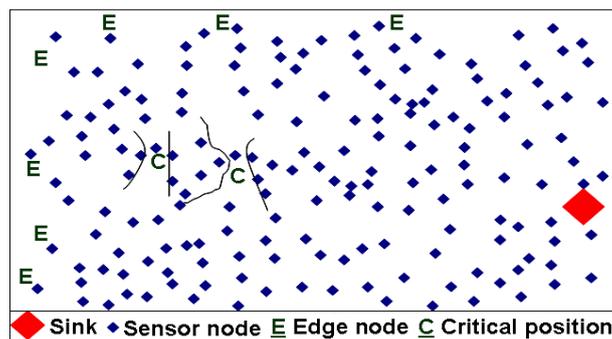


Fig. 1. Nodes' positions in the network

A good strategy should allow a sensor node in a critical position to decrease its power consumption to maintain the connection between the two parts of the network the longest possible time. Thus, the value of the importance factor of the node position should help the sensor node to apply a selfish behavior and hence, e.g, it should be greater than or equal to the energy or the information importance degree factors

To facilitate the computation of P, we propose a fixed value for each type of node position. These values are 10%, 50% and 100% for the normal, edge and critical, position respectively.

V. CONCLUSION

The energy consumption of the wireless sensor network directly determines the lifetime of the wireless sensor network. Reasonably deploying the wireless sensor nodes can improve the coverage effect of the wireless sensor network and reduce the movement distance of the wireless sensor nodes. Clustering is a good technique to reduce energy consumption and to provide stability in wireless sensor network our main focus is to maximize the network lifetime and to minimize the energy consumption of the sensor nodes. We assume that the network life time is the time from the deployment of the WSN till the first gateway dies. Therefore, network life can be maximized by using the parameter discussed in this paper. If we can minimize the energy consumption of the CH nodes then energy consumption of the sensor nodes can be minimized if we can minimize their distance from their corresponding CH's. Future work will involve the simulation of working mechanism using above mentioned parameters based on certain design issue.

REFERENCES

- [1] F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks:A survey. Computer Networks, 38(4):393–422, 2002
- [2] W. Heinzelman, A. Chandrakasan and H. Balakrishnan. 2000. Energy-Efficient Communication Protocol for Wireless Microsensor Networks. Proceedings of the 33rd Hawaii International Conference on System Sciences (HICSS '00).
- [3] S. Lindsey, C. Raghavendra. 2002. PEGASIS: Power- Efficient Gathering in Sensor Information Systems. IEEE Aerospace Conference Proceedings, Vol. 3, 9-16 pp. 1125-1130.
- [4] A. Manjeshwar and D. P. Agarwal. 2001. TEEN: a routing protocol for enhanced efficiency in wireless sensor networks. In 1st International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing.
- [5] A. Manjeshwar and D. P. Agarwal. 2002. APTEEN: A hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks. Parallel and Distributed Processing Symposium, Proceedings International, IPDPS, pp. 195-202.
- [6] Ossama Younis and Sonia Fahmy. 2004. Distributed Clustering in Ad-hoc Sensor Networks: A Hybrid, Energy-Efficient Approach. In Proceedings of IEEE INFOCOM, Hong Kong, an extended version appeared in IEEE Transactions on Mobile Computing, 3(4).
- [7] Cook, B. S., Vyas, R., Kim, S., Thai, T., Le, T., Traille, A., et al. (2014). RFID-based sensors for zero power autonomous wireless sensor networks. IEEE Sensors Journal, 14(8), 2419–2431.
- [8] Barati, H., Movaghar, A., Rahmani, A. M., & Sarmast, A. (2012).A distributed energy aware clustering approach for large scale wireless sensor networks. International Journal on Technical and Physical Problems of Engineering, 4(4), 125–132.
- [9] Zheng, J.,&Jamalipour,A. (2009). Wireless sensor networks: A networking perspective. Hoboken: Wiley.
- [10] Kang, S. H., & Nguyen, T. (2012). Distance based thresholds for cluster head selection in wireless sensor networks. IEEE Communications Letters, 16(9), 1396–1399.
- [11] Chen, H.,&Megerian, S. (2006). Cluster sizing and head selection for efficient data aggregation and routing in sensor networks. In IEEE wireless communications and networking conference, pp. 2318–2323