Energy Conserving Passive Clustering for Efficient Routing in Wireless Sensor Network

Md. Mamun-or-Rashid, Muhammad Mahbub Alam and Choong Seon Hong
Department of Computer Engineering, Kyung Hee University
1-Seochun-ri, Kiheung-eup, Youngin-si, Kyunggi-do, Korea, 449-701
{mamun, mahbub}@networking.khu.ac.kr and cshong@khu.ac.kr

Abstract*-In case of sensor network, hierarchical routing has received a great magnitude of interest than flat routing for achieving energy efficiency. While the overhead of creating hierarchical control structure, should be kept reasonably low. Directed Diffusion (DD) is one of the most prominent and widely accepted example of hierarchical data centric routing. Directed Diffusion suffers from huge flooding problem and further improved using passive clustering to reduce flooding for interest propagation. Passive clustering technique is not energy efficient as it does not consider residual energy of nodes. Also it does not have any parameter to restrict the number of clusterheads and gateways. To achieve energy efficiency and prolonged network lifetime we introduce residual energy and distance based cluster creation. Energy Conserving Passive Clustering technique will select better nodes as clusterhead and gateway in terms of residual energy and distance which will in turn reduce the overlapping region and number of gateway to conserve energy. At the same time we have applied periodic sleep and awake among cluster members to achieve prolonged lifetime. Our proposed algorithm will add some extra overhead for cluster creation and maintenance in compare to passive clustering but will perform better in terms of energy dissipation and enhanced lifetime of the network.

***Keywords* — Sensor Network, Clustering, Energy Efficient Routing**

1. Introduction

Sensor network can be envisage as a collection of large number of sensor nodes and each node is consisting of a sensing circuitry, a processor, a radio transceiver for short range communication and a battery-supplied power. The deployment of sensor nodes are random, in the inaccessible terrain and left as unattended. Some unique characteristics of sensor network has been presented in [1]. Mostly sensor nodes may fail due to power constraint or physical damage or environmental interference. This necessitates huge deployment of sensor nodes. Dense deployment of nodes will in turn increase the probability of collision, redundant data dissemination and huge flooding. The ultimate outcome of these mentioned problems is wasteful energy consumption which eventually reduces the overall lifetime of sensor network and non-uniform sensing load of a particular terrain may even partition the network and make the whole network non-operational. So sensor network routing protocol must be designed with regard to the trade off between two design issues robustness and energy efficiency. Hierarchical and data-centric routing protocols received much interest for sensor network due to the following motivational aspects:

- i. In sensor networks, hierarchical control structures contribute to improved efficiency of resource use by creating contexts for:
  - a) Managing the whole network with near optimal number of nodes for energy conservation
  - b) Managing wireless communications among multiple nodes to reduce channel contention
  - c) Forming routing backbones to reduce network diameter
  - d) Periodic sleep and awake technique can be applied to conserve energy
- ii. Data-centric routing can make efficient use of resources in the following way
  - a) To combine data from different sensors to eliminate redundant transmission
  - b) From multiple sources to one destination a data-centric routing approach allows in-network consolidation of data
  - c) Most of the sensor network application data is requested based on certain attribute so data-centric routing suits most for sensor network

Hierarchical and data-centric routing algorithms must execute based on local coordination and should not depend on global topology information as localized protocols are rewarded with good performance in terms of energy consumption and backbone size [11]

Our work is motivated by Directed Diffusion (DD) [2] and Passive Clustering Directed Diffusion (PCDD) [3]. DD [2] is a data-centric which is suitable for sensor network application and adding passive clustering to DD has enriched the routing quality in terms of energy conservation. Directed Diffusion (DD) is a prominent example of data-centric and application aware routing protocol where all data generated by sensor nodes is named by attribute value pairs. DD suffers from flooding problem for interest propagation and was further improved in [3] combining the idea of passive clustering [4] along with the DD to reduce redundant transmission and huge

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD)" (KRF-2006-521-D00394)

flooding for interest propagation. While passive clustering is popular for its on demand clustering technique which can avoid potential set up time and reduced message overhead. Passive Clustering Directed Diffusion (PCDD) generates huge number of gateways and does not consider nodes energy status as a cluster setup parameter. Hence the possibility of conserving energy is still in PCDD by reducing gateways and data redundancy by finding common regions in terms of sensing and transmission range.

In this paper, we have proposed Energy Conserving Passive Clustering (ECPC) algorithm which takes account of both residual energy and distance for becoming cluster head and gateway and also eliminate the problem of idle listening through periodic sleep and awake among the cluster members. Our algorithm outperforms DD and PCDD in terms of energy dissipation and network lifetime. It also generates much less gateway nodes than PCDD algorithm.

Next section of our paper organizes as follows: section 2 articulates brief literature survey about hierarchical routing protocols of sensor network, section 3 describes the parameter for clusterhead and gateway selection of our algorithm, section 4 summarizes our proposed protocol, section 5 describes the query propagation and data dissemination, section 6 show the improvement of our protocol through simulation and finally we conclude with future works in section 7.

2. Related Works

Hierarchical [2][3][5][6][7][8] and data-centric[2][9] routing algorithms are well suited for sensor networks due to its dense deployment and application nature. Hierarchical structure will allow a set of nodes to perform the routing operation and thus reducing the amount of redundant transmission and collisions. While the amount of overhead for the setup and maintenance cost of hierarchical structure should be kept reasonably low [10].

LEACH [5] is one of the most popular hierarchical algorithm 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. It uses single-hop routing where each node can transmit directly to the cluster-head and the sink. Therefore, it is not applicable to networks deployed in large regions. Furthermore, the idea of dynamic clustering brings extra overhead, e.g. head changes, advertisements etc., which may diminish the gain in energy consumption. PEGASIS [6] is an improvement of the LEACH protocol. Although the PEGASIS approaches avoid the clustering overhead of LEACH, they still require dynamic topology adjustment since sensor’s energy is not tracked. For example, every sensor needs to be aware of the status of its neighbor so that it knows where to route that data. Such topology adjustment can introduce significant overhead especially for dense networks. TEEN [7] and APTEEN [8] claimed that they outperform LEACH [5] in terms of energy dissipation and network lifetime. The main drawbacks of the two approaches are the overhead and complexity of forming clusters in multiple levels, implementing threshold-based functions and dealing with attribute-based naming of queries.

Directed diffusion [2] is a popular data dissemination paradigm for WSNs. DD is a data-centric and application-aware approach in the sense that all data generated by sensor nodes is named by attribute-value pairs. The main idea of the DD is to combine the incoming data from different sources to enroute (in-network aggregation) by eliminating redundancy, minimizing the number of transmissions and thus saving network energy and prolonging its lifetime. Unlike traditional end-to-end routing, DD routing finds routes from multiple sources to a single destination that allows in-network consolidation of redundant data. Unfortunately the idea of directed diffusion does not consider the dense deployment of sensor nodes that causes huge amount of unnecessary re-forwarding of messages.

The noble idea of Directed Diffusion (DD) is further improved in [3] combining the idea of Passive Clustering (PC) proposed by Kwon and Gerla [2]. Passive clustering is an on demand creation and maintenance of the cluster substrate which can avoid potential long set-up time and reduce the cost of initial interest propagation of Directed Diffusion. But passive clustering does not consider residual energy for becoming cluster head. As a result network life time is reduces if the low energy nodes become cluster head. Moreover this approach does not consider the idle listening time of the nodes that eventually consume huge energy of the total energy consumption.

As energy conservation is the primary and covetable target for sensor network, we intend to extend the idea of passive clustering by introducing residual energy and distance parameter for cluster creation. Our algorithm will choose the best node as clusterhead in terms of residual energy and distance and also restricts the number of gateway. Also periodic sleep and awake approach among cluster members will conserve energy for better lifetime of the network.

3. Energy Conserving Passive Clustering

Passive clustering [2] is clustering technique which does not use dedicated, protocol specific control packets or signals. One of the major reason is conventional clustering algorithms require advertising neighbor information. In wireless ad hoc network collecting this information requires O(k) where k is the possible minimum node degree. In case of sensor network as the deployment is huge, at a time an optimal set of sensor should be responsible for sensing and disseminating data. The amount of overhead for cluster setup and maintenance is much lower than the amount of redundant transmission and collision. As minimum as the set of sensors can be kept without destroying connectivity property, the amount of redundant transmission and collision will be reduced.

Clusterhead Selection: Unlike passive clustering our cluster head selection does not depend on “first declaration wins” rule. As sensor network is very dense, for a certain period of time optimal number of sensor will be responsible of communicating with the sink and dynamic routing capability based on power availability, position and reachability. So, our clusterehad selection algorithm will prioritize following parameters:

i. Clusterhead node must have higher residual energy with in the 1-hop neighbors.
ii. The distance is another criterion that should be considered to form better clusters.

To accomplish this each node n_i will calculate its waiting time before declaring itself as a clusterhead using following equation

$$T_{Cwait} = \frac{D_i}{R_i} \times T_e \qquad (1)$$

Where, D_i is the distance from which node it has received the query, R_i is the residual energy of node n_i and T_e is constant value defined by the ratio of nodes maximum energy transmission range.

Gateway Selection: Sensor nodes in the overlapping region of two or more clusters can claim themselves as a gateway. The number of gateway is directly proportional to the reliable data delivery and inversely proportional to energy efficiency. The more the number of gateway the amount of overhearing, idle listening and redundant transmission. So the maximum number of gateway between two or more than two clusterhead should be specified to save the energy as well as getting stable network performance. Our gateway selection procedure prioritizes the following parameters:

i. Residual energy of the node which wants declare itself as a gateway
ii. Average distance of clusterhead nodes with in 1-hop.

Accordingly, each node between two or more clusterheads, should calculate average distance from the nodes it has received clusterhead notification message and wait before declaring itself as a gateway node. Average distance can be calculated using the following equation:

$$D_{avg} = \frac{\sum_{i=1}^{n} Dc_i}{n} \qquad (2)$$

Waiting time before declaring gateway can be calculated using the following equation:

$$T_{Gwait} = \frac{D_{avg}}{R_i} \times T_e \qquad (3)$$

4. Energy Conserving Passive Clustering Algorithm

We summarize our proposed algorithm as follows:

i. There are six possible states: ordinary, cluster_member, clusterhead_ready, custerhead, gateway_ready and gateway
ii. At the beginning of each round of cluster creation, all nodes will be in ordinary status. As like directed diffusion sink will inject a query into the sensor network
iii. Nodes in ordinary status receiving the query will change there status to clusterhead ready status.
iv. All clusterhead_ready status nodes will calculate individuals T_{Cwait} using equation 1. Each node will wait T_{Cwait} time before transmitting ADV_CLUSTERHEAD message indicating its node ID and residual energy and also forward the query
v. Nodes in ordinary status receiving ADV_CLUSTERHEAD message will change their status to gateway_ready status and forward the query injected by the sink. At the same time each node in ordinary and gateway_ready status will store the clusterhead node ID and distance (calculated based on received signal strength) from the clusterhead
vi. Any node in gateway_ready status receiving two or more than two ADV_CLUSTERHEAD message will calculate average distance with the stored clusterhead's distance using equation (2) and after that it will calculate T_{Gwait} time using equation (3). The node will wait T_{Gwait} time and before transmitting ADV_GATEWAY message indicating own ID and node ID of the clusterheads for which it is responsible.
vii. Nodes in gateway_ready status receiving ADV_GATEWAY message will store the gateway node ID and the corresponding clusterhead's ID. Note that it is important to keep track of the mapping of gateway and clusterhead as gateway_ready status nodes having clusterheads in common will restrict itself from declaring itself as gateway. Node n_i in gateway_ready status will restrict itself from transmitting ADV_GATEWAY message if the following equation is true:

$$N(n_i(C_{id})) = \bigcup_{j=0}^{n} N(n_j(N(C_{id}))) \qquad (4)$$

Where $N(n_i(C_{id}))$ is the neighbor clusterhead IDs of node n_i, n_j's are the neighbor gateways of n_i and $N(n_j(N(C_{id}))$ is the clusterhead IDs of n_i's neighbor gateways. The equation will force once node to become gateway only when it is necessary.

viii. If nodes in ordinary status and gateway_ready status will change their status to cluster_member if they are not necessary for the communication.

5. Query Propagation and Data Dissemination

Our algorithm must choose the better nodes as clusterheads and gateways considering energy and topological position in the network through equation 1 and 3. Number of gateway is restricted using equation 4. Also the algorithm executes based on local coordination as each node competes for clusterhead

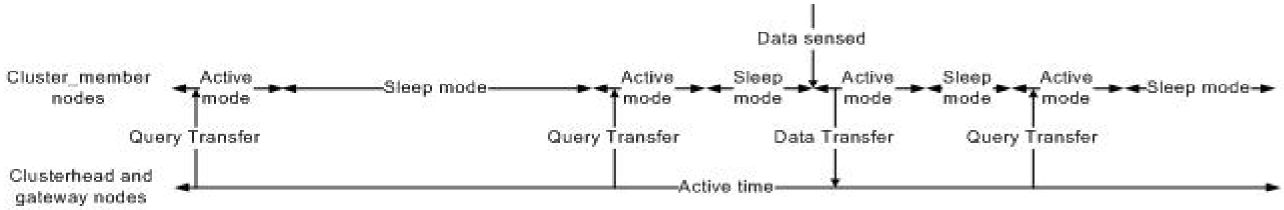

Figure.1. Query Propagation and Data Dissemination between Clusterhead, Gateway and Cluster Member Nodes

and gateway based on 1-hop residual energy and distance information and does not depend on global topological information.

Nodes is clusterhead and gateway status will always keep their radio on and cluster member will keep their radio on and off in a periodic fashion for query transfer and data dissemination. After the deployment the cluster setup will run within certain time interval and each time new set of nodes will be selected clusterheads and gateways

After cluster setup data and query need to be propagated through the clusterhead and gateways. As we have introduced periodic sleep and awake for the cluster members we define two types of communication for query propagation.

Intra-Cluster Communication: Within a cluster, clusterhead which is reachable from all cluster_member nodes by one hop. The cluster_member nodes and clusterhead should be synchronized in such a way, that all the nodes will be in active state when the clusterhead diffuses the query. During the sleep time if cluster head receives any new query, it waits up to the next active time and then broadcast the query. The active time should be small enough and sleep time should be multiples of the active time. The round time (if we consider one consecutive idle time and sleep time as a round) should be selected properly. If it is too small then the overhead of turning the radio on and off frequently will decrease the energy savings. On the other hand if it is too large, then the new query propagation will be delayed. These timings are based on the type of application and the type of requirements. Cluster_member nodes will continue to sense within the sensing range and if anything sensed within the vicinity, turn on the radio and disseminate sensed data to the respective clusterhead. It is to be observed that query propagation from clusterhead to cluster_member nodes is one to many communication. At the time of query propagation all nodes within the cluster will receive data while sensed data is purely random and the communication is one-to-one with clusterhead. As a consequence we have adopted periodic query propagation and on demand data dissemination between clusterhead and cluster_member nodes with in the cluster. This will reduce the chance of idle listening and collision which in turn will conserve energy.

Inter-Cluster Communication: Query and data will be propagated between source and sink via the clusterhead and gateway by inter-cluster communication. So the clusterhead and gateway will remain active all the time throughout their turn.

6. Simulation

Our simulation starts with 50 nodes uniformly distributed within an area of 100X100 and latter on we increase the number of nodes for the same area. We have developed our simulation using C++. According to specification of Mica2mote, transmit require 81mw, receive and idle require 30 and sleep status require 0.003 mw. As the ratio of transmit and receive is 3:1, in our simulation we consider 0.0003 unit/bit/m^2 unit energy for transmit, 0.0001 unit/bit/m^2 unit energy for receiving and 0 unit energy consumption for idle and sleep respectively. All nodes are initialized with a transmission of 20. Our simulation uses packet size of 2Kbits. We compare our proposed Energy Conserving Passive Clustering (EEPC) with Directed Diffusion (DD) and Passive Clustering Directed Diffusion (DDPC) in terms of following matrices:

i. Total backbone size: After the completion of cluster setup what is the total number of clusterhead and gateway nodes. Note that these nodes will act as a backbone for query propagation and data dissemination.

ii. Energy Dissipation: Amount of energy dissipation under various work load. In our case we increase traffic bye increasing the number of sources for a fixed size packet. Less energy dissipation leads to prolonged lifetime.

iii. Hop count: Number of hops required to reach from source to sink. optimal number of hops can be calculated using following equation:

$$H_c = \left\lceil \frac{d(\text{source,sink})}{R_t} \right\rceil \qquad (5)$$

Where $d(\text{source,sink})$ is the Euclidian distance between source and sink and R_t is the transmission range. This is very important for the performance of routing protocol. A good routing protocol should have the hop count value near to H_c for any source and sink pair.

Figure 2 shows the number of clusterhead and gateway generated by PCDD and our proposed ECPC algorithm. In case of ECPC, each node will check its necessity before declaring itself as gateway. Thus ECPC reduces the number of gateway significantly by reducing gateway for the same common clusterheads. We perform the experiment starting from 50 nodes up to 300 nodes for the same 100X100 deployment area.

In Fig. 3. we have plotted the optimal hop count according to transmission range and the distance between source and sink and the average hop count for increasing amount of sources for DD, DDPC and ECPC. Proposed ECPC requires near optimal number of hops for propagating query and data.

Energy dissipation (In Fig. 4.) of ECPC is much lower than DD and DDPC as ECPC will reduce the amount of redundant transmission by restricting number of forwarding nodes (clusterhead and gateway). Periodic sleep and awake among the cluster member will significantly conserve energy by reducing idle listening and overhearing. Our algorithm dissipates less energy in compare to DDPC and DD algorithm.

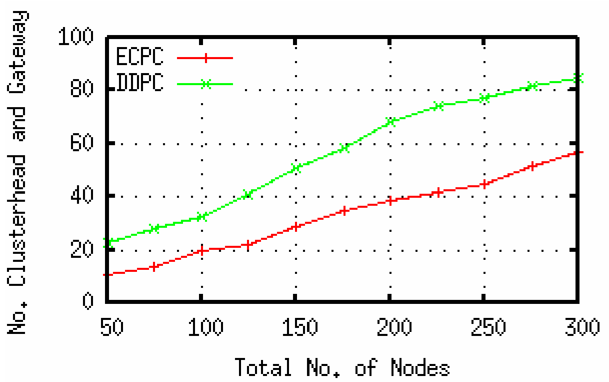

Figure 2. Average Number of Clusterhead and Gateway Generated by ECPC and DDPC Algorithm

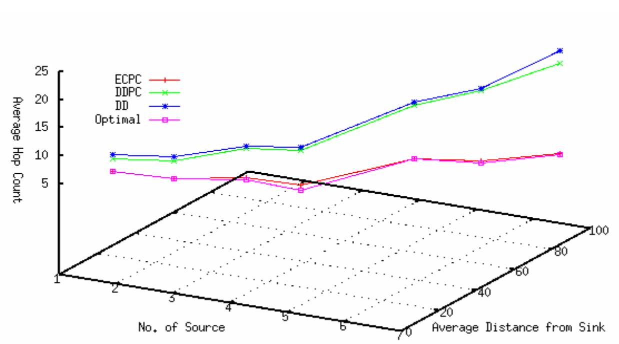

Figure 3. Average Hop Count under Increasing Number of Sources and Average Distance of Sources from the Sink for 100 Nodes Deployed Randomly in A 100X100 Area

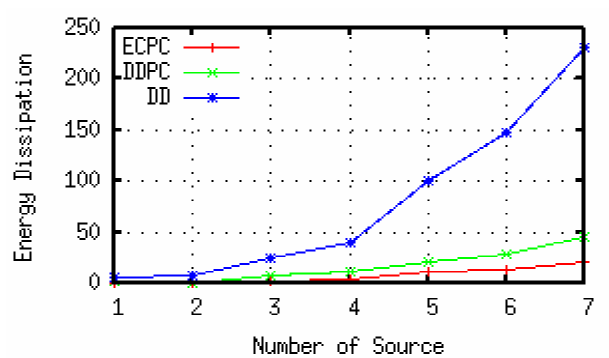

Figure 4. Energy Dissipation for 100 Nodes Placed Randomly into 100X100 Area with Varying Number of Sources from 1 to 7

7. Conclusion

Our proposed energy conserving passive clustering algorithm introduces some amount of overhead for cluster creation but through simulation it shows better performance in terms of energy dissipation. According to the selection parameter our algorithm chooses better nodes as clusterheads and gateways considering energy and topological position. We also reduce the number of redundant gateways for reduced collision. Our algorithm executes based on 1-hop residual energy and distance information and does not depend on global topological information. One important consideration is the time interval of cluster formation. It is to be noted that the interval time is fully depended on the amount of traffic generated by the network and hence application dependent. Interval parameter will control the balance energy consumption among all nodes and thus can restrict nodes early death. Finding the optimal interval time will result more energy efficiency. We intend to work on this in future. We intend to evaluate our algorithm under multiple sink and also introduce sink mobility as a future work.

References

[1] lan F. Akyildiz, WellJan Su, Yogesh Sankarasubramaniam, Erdal Cayirci, “A Survey on Sensor Network”, Communications Magazine, IEEE, August, (2002), Vol. 40 102 - 114

[2] C. Intanagonwiwat, R. Govindan, and D. Estrin, Directed diffusion: a scalable and robust communication paradigm for sensor networks, Proceedings of ACM MobiCom '00, Boston, MA, (2000) 56-67

[3] Handziski V., Köpke A., Karl H., Frank C. and Drytkiewicz W. “Improving the Energy Efficiency of Directed Diffusion Using Passive Clustering”, in Proc of the 1st European Workshop on Wireless Sensor Networks (EWSN), LNCS 2920, Berlin, Germany, (2004) 172-187

[4] Gerla, M., Taek Jin Kwon; Pei, G., On-demand routing in large ad hoc wireless networks with passive clustering, Wireless Communications and Networking Conference, WCNC. 2000 IEEE, Vol. 1, (2000) 100 - 105

[5] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, Energy-Efficient Communication Protocol for Wireless Micro-sensor Networks, Proceedings of the 33rd Hawaii International Conference on System Sciences (HICSS '00), (2000)

[6] S. Lindsey, C. Raghavendra, PEGASIS: Power-Efficient Gathering in Sensor Information Systems, IEEE Aerospace Conference Proceedings, Vol. 3, (2002) 1125-1130

[7] A. Manjeshwar and D. P. Agarwal, TEEN: a routing protocol for enhanced efficiency in wireless sensor networks, In1st International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing, (2001)

[8] A. Manjeshwar and D. P. Agarwal, APTEEN: A hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks, Parallel and Distributed Processing Symposium., Proceedings International IPDPS, (2002) 195-202

[9] Azzedine Boukerche,Xiuzhen Cheng,Joseph Linus,"Energy-aware data-centric routing in microsensor networks", Proceedings of the 8th international workshop on Modeling analysis and simulation of wireless and mobile systems,MSWiM 03, San Diego, California, USA, (2003) 42 - 49

[10] O. Younis, and S. Fahmy, **“**Distributed clustering in ad-hoc sensor networks: a hybrid, energy-efficient approach”, INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies, Vol. 1 (2004)

[11] S. Basagni, M. Mastrogiovanni, A. Panconesi and C. Petrioli, “Localized Protocols for Ad Hoc Clustering and Backbone Formation: A performance comparison”, IEEE Transaction on Parallel and Distributed Systems, Vol. 17, No.4, (2006) 292-306