

Neighbor Node Discovery Algorithm for Energy-Efficient Clustering in Ubiquitous Sensor Networks

Ji Young Choi¹, Chung Gu Kang^{1,*}, Yong Suk Kim², and Kyeong Hur³

¹ Department of Radio Communications Engineering, Korea University
{jychoi2004, ccgkang}@korea.ac.kr

² SAMSUNG Advanced Institute of Technology (SAIT), Korea
yongsuk@samsung.com

³ Department of Computer Education, Gyeongin National University of Education
khur@ginue.ac.kr

Abstract. Clustering algorithm is an essential element to implement a hierarchical routing protocol, especially for a large-scale wireless sensor network. In this paper, we propose a new type of energy-efficient clustering algorithm, which maximizes the physical distance between cluster head and gateway by a neighbor node discovery mechanism. Furthermore, a slave/master patching scheme is introduced as a useful means of further improving the energy-efficiency. It has been shown that the number of cluster heads can be reduced by as many as 21% as compared with the existing clustering algorithms.

1 Introduction

Numerous routing protocols have been developed for discovering and maintaining routes in the large-scale wireless sensor networks (WSNs). A cluster-based hierarchical routing protocol is one of those, especially useful for prolonging the lifetime of WSNs. In a flat routing protocol, multi-hop connectivity is provided to communicate with a base station (BS). In a cluster-based hierarchical routing protocol, however, a sensor field region is divided into several clusters. Each cluster has its own cluster head, while all nodes in a cluster are within a direct radio range of the cluster head. Nodes located within the radio range of more than one cluster heads are designated as gateways, which is used for inter-cluster communication. The data gathered by a sensor node is transmitted to a cluster head, which subsequently relays it to a BS through a hierarchy of cluster heads and gateways. As only cluster heads and gateways participate in the routing process, hierarchical routing tends to be more energy-efficient than a flat routing structure.

Data aggregation is another energy-aware element in WSNs. Once a particular event is detected in a region, it is transmitted to the corresponding cluster head, which usually aggregates the received data before transmitted to the BS. Data aggregation is to combine the data coming from nearby region so as to eliminate the data redundancy for reducing the overall communication load [4]. Several clustering algorithms [1-3, 6-8] are essential elements to realize hierarchical routing and data aggregation.

* Member IEEE.

If more than two sensor nodes in the same cluster transmit at the same time, intra-cluster collision is incurred. Two broad categories of wireless channel access scheme are used for multiple access communication: contention-based and reservation-based. A contention-based scheme such as CSMA-CA protocol with RTS/CTS handshaking signals can be used for intra-cluster communication, i.e., communication between a cluster head and sensor nodes. However, the number of retransmissions is increased with a large number of sensor nodes, incurring unnecessary power consumption. Meanwhile, a reservation-based scheme such as TDMA protocol, allocates unique time slots to each sensor node, can significantly reduce the power consumption in multiple access communication. Furthermore, it allows sensor nodes to turn off their receiver and/or transmitter when they are not scheduled to communicate.

Although TDMA protocol can be more energy-efficient than CSMA-CA protocol, it is hard to coordinate the time slots for inter-cluster communication, especially when a single wireless channel (code) is employed. In other words, intra-cluster communication may be interfered with another inter-cluster communication unless there is some centralized means of coordinating two different levels of communication. Due to the distributed nature of WSN, the centralized TDMA protocol may be too expensive for handling multi-cluster-based communication in the course of hierarchical routing. The different approach is to combine TDMA and CSMA-CA protocols as in HiPERMAC [Grant No. KOR-2005-0046460(2005)], in which TDMA is used for intra-cluster communication while CSMA-CA is used for inter-cluster communication.

As mentioned earlier, cluster heads and gateways must be always turned on since they involve with data aggregation and/or hierarchical routing process while a gateway cannot use the energy-efficient TDMA protocol due to inter-cluster collision. Therefore, one of the most important design objectives is to reduce the number of cluster heads and gateways in the course of clustering process, which can increase the overall lifetime of WSNs. In this paper, we propose a new type of energy-efficient clustering algorithm, in which a simple protocol is designed for counting the number of neighbor nodes.

The remainder of this paper is organized as follows. In Section 2, we summarize operational characteristics of the previous related works. In Section 3, we present details of the proposed clustering algorithm. Simulation results are given to evaluate its performance in Section 4. Finally, Section 5 concludes the paper.

2 Related Works

Many clustering algorithms over a single wireless channel (code) have been proposed to choose cluster heads for operation of energy-efficient wireless media access control and hierarchical routing in WSNs. Among those, Lowest ID clustering algorithm [1], also known as identifier-based clustering, proceeds as follows. Each node is assigned a unique ID and it periodically broadcasts a list of its neighbors (including itself). A node which only hears nodes with ID higher than itself becomes a “cluster head.” A node which can hear only one cluster head becomes an “ordinary node.” A node which can hear two or more cluster heads becomes a “gateway” and it lies within the radio range of two or more cluster heads. The gateway nodes are generally used for routing between clusters. Although Lowest ID is one of the most simple and popular

clustering schemes, its drawback is that too many cluster heads and gateways may be generated. Topology Discovery algorithm [3] attempts to reduce the number of cluster heads and gateways by maximizing a physical distance between a cluster head and a gateway. It assumes that propagation delay is proportional to their physical distance between sending and receiving nodes. A cluster head broadcasts a topology discovery request packet and the node which has received it sends back a response to the cluster head. The cluster head measures this propagation delay from each node and elects one which has a longest propagation delay as a gateway. Meanwhile, the gateway elects a cluster head of neighbor cluster in the same way. This same election procedure is repeated until each node is assigned to at least one cluster. The assumption that “propagation delay is proportional to their physical distance between sending and receiving nodes” cannot be guaranteed when multi-path fading is considered in the general wireless channel environment. Furthermore, it may be difficult to distinguish a slight difference of each propagation delay since radio range of WSNs is generally within ten meters.

3 Neighbor Node Discovery (NND) Algorithm

We first present the underlying concept of our clustering algorithm and then, illustrate its detailed operation.

3.1 Design Objective and Constraint

We aim at forming the multiple clusters to cover all communicating nodes, while facilitating an energy-efficient operation for the large-scale WSNs. As opposed to the existing protocols with multiple frequency channels, e.g., in terms of orthogonal codes or frequency bands, we consider a system with a single frequency band, which allows for the cost-effective transceiver implementation and flexible network deployment. In general, it is a challenging objective to design the access schemes for both intra-cluster and inter-cluster communication only by using a single frequency band. In our design, a dynamic reservation TDMA scheme is considered for intra-cluster communication as a means of sharing the wireless link in an energy-efficient manner. Meanwhile, a contention-based CSMA-CA scheme is considered for inter-cluster communications as a means of avoiding inter-cluster collision. A superframe is divided into two different periods, one for TDMA and the other for CSMA-CA protocol. Fig. 1 shows the organization of sensor nodes in a hierarchical routing structure, illustrating four different types of nodes found in the course of clustering: cluster head (CH), initial node (IN), ordinary node (ON), and gateway (GW). A cluster is a set of nodes within one-hop range from a CH, which coordinates the TDMA time slots sharing among those nodes within the cluster while providing a link with the gateway nodes for inter-cluster communication. The INs are defined as a set of nodes that have not been covered by any cluster. INs are turned into either CHs, ONs, or GWs in the course of performing the clustering algorithm. CH and GW should be always turned on since they participate in data aggregation and/or routing process. Therefore, they consume more energy than an ON, depleting their battery faster. The overall network lifetime can be prolonged by reducing the number of CHs and GWs. Meanwhile, note that ON

is more energy-efficient than GW since it can turn off its receiver and/or transmitter when it is not scheduled for TDMA time slots. In this case, the network lifetime can be further prolonged by increasing the number of ONs. Our design objective for the multi-cluster routing architecture operating with a single frequency band is to control the numbers of the different types of nodes for increasing the overall energy efficiency, i.e., minimizing the number of CHs and GWs.

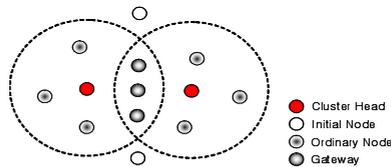


Fig. 1. Types of sensor nodes in a hierarchical routing architecture

3.2 Basic Operation

The essence of our approach is to discover neighbor nodes. One of the nodes at the cluster edge is designated as a GW, while a node which is farthest from GW is designated as a CH of neighbor cluster. Basic operation of the NND algorithm consists of 4 different procedures: initial cluster set-up, gateway election, cluster head election, and gateway re-election procedures. All these procedures are combined and repeated until all INs are turned into either CH, GW, or ON. A sequence of these procedures is shown with a flow chart in Fig. 2. The detailed operation is described in the following subsections.

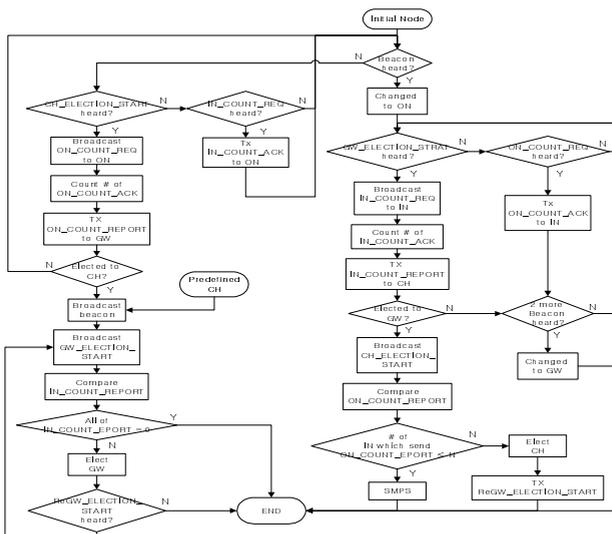


Fig. 2. Flow Chart of Neighbor Node Discovery Algorithm

3.2.1 Initial Cluster Set-Up Procedure

One of the CHs is designed as an originating CH, from which other clusters are progressively formed. It broadcasts an advertisement message within its radio range and an IN which has received it sends a joining request message to the CH (refer to Fig. 3). The CH stores the identity of the IN which has sent a joining request message and registers it as a cluster member. The originating CH broadcasts a beacon signal within its radio range and the IN which has heard the beacon becomes an ON.

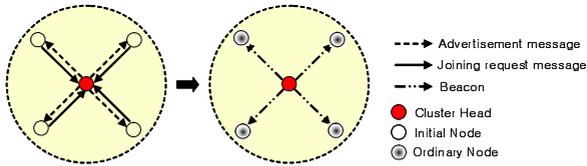


Fig. 3. Initial cluster set-up procedure

3.2.2 Gateway Election Procedure

In this procedure, the farthest ON from a CH is elected as a GW among one-hop neighbors of a CH, which allows for maximizing the physical distance between a CH and a GW. To realize this particular procedure, we assume that INs are uniformly distributed, which leads to an acceptable observation that “The farthest ON from a CH will discover the largest number of INs.” Each ON among one-hop neighbors of a CH counts the number of INs within its radio range and then, reports it to the CH. The ON which discovers the largest number of INs is elected as a GW by the CH. Fig. 4 illustrates a simple example, in which Node A includes three INs while node B includes two INs within its radio range. According to the above conclusion, it is desirable that node A is elected as a GW. A contention-based channel access scheme will be used during the gateway election procedure to avoid the collisions between nodes.

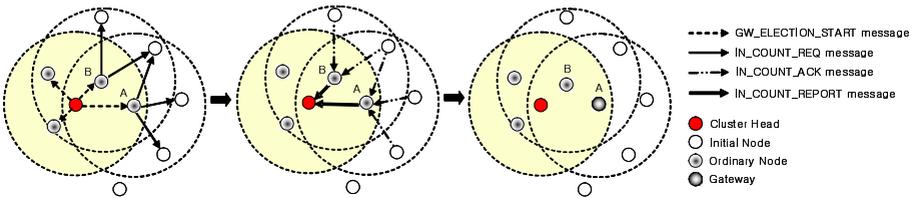


Fig. 4. Gateway election procedure

A CH broadcasts a GW_ELECTION_START message within its radio range. An ON which has received this message counts the number of INs within its radio range by the following steps: The ON broadcasts an IN_COUNT_REQ message to INs. The IN which has received this message transmits an IN_COUNT_ACK message to ON, limiting the transmitted number of this ACK messages to only once even if it has received the multiple IN_COUNT_REQ messages from a large number of ONs. The ON reports the received number of ACKs from INs through an IN_COUNT_REPORT

message to the CH. The received number of ACKs is equal to the number of INs within the radio range of ON. Analyzing all the received `IN_COUNT_REPORT` messages, CH elects the ON which includes the largest number of INs as a GW. In case of a tie, one with the lowest ID becomes GW. If the values of the received `IN_COUNT_REPORT` messages are all equal to zero, i.e., there is no IN around ONs, the CH does not elect a GW and the clustering algorithm is finished. In case that there is at least one `IN_COUNT_REPORT` message with its value higher than zero, the CH elects a GW and then, proceeds to the cluster head election procedure.

3.2.3 Cluster Head Election Procedure

In this procedure, the farthest IN from GW is elected as a CH of the neighbor cluster among one-hop neighbors of GW, which allows for maximizing the physical distance between GW and CH of a neighbor cluster. To realize this particular procedure, we resort to an acceptable observation that “the farthest IN from a GW will discover the least number of ONs.” Each IN among one-hop neighbors of the GW counts the number of ONs within its radio range and reports it to the GW. The IN which discovers the least number of ONs is elected as a CH of neighbor cluster by the GW. Fig. 5 illustrates a simple example, in which Node A includes three ONs while node B includes one ON within its radio range. Following our procedure, the node B is elected as a CH of the neighbor cluster. A contention-based channel access scheme will be used in the course the cluster head election.

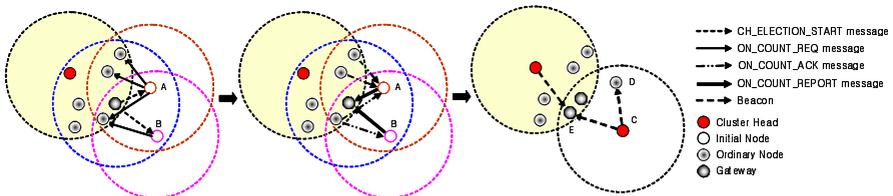


Fig. 5. Cluster head election procedure

A GW broadcasts a `CH_ELECTION_START` message within its radio range. An IN which has received this message counts the number of ONs within its radio range by the following steps: The IN broadcasts an `ON_COUNT_REQ` message to ONs and the ON which has received this message transmits an `ON_COUNT_ACK` message to the IN, limiting the transmitted number of this ACK messages to only once even if it has received the multiple `ON_COUNT_REQ` messages from a large number of INs. The IN reports the received number of ACKs from ONs through the `ON_COUNT_REPORT` message to the GW. The received number of ACKs is equal to the number of ONs within the radio range of IN. Analyzing all the `ON_COUNT_REPORT` messages, GW elects the IN which includes the least number of ONs as a CH of neighbor cluster. In case of a tie, one with the lowest ID becomes CH. If the number of IN which sends the `ON_COUNT_REPORT` message is zero, i.e., there is no IN around GW, the GW does not elect a CH of neighbor cluster and immediately, the clustering algorithm is finished. At least one IN transmits an `ON_COUNT_REPORT` message, the GW elects a CH of neighbor cluster and proceeds to the next gateway election procedure. A newly-elected

CH (Node C) broadcasts an advertisement message within its radio range and the IN which has received it sends a joining request message to a CH. The CH stores the identity of IN which has sent the joining request message and registers it as a cluster member. The CH broadcasts a beacon and an IN (node D) which has heard it becomes an ON. The ON (Node E) which has heard more than two beacons from the different CH becomes a GW.

3.2.4 Gateway Re-election Procedure

In order to reduce the overall clustering time throughout the network, it will be useful to perform the clustering process towards every direction in a parallel manner, as illustrated in Fig. 6. A GW which finishes the election of a CH for neighbor cluster (Node B) transmits a GW_ReELECTION_START message to a CH (Node A), previously elected itself as a GW. The CH (node A) which has received GW_ReELECTION_START message elects another GW out of ONs within its own cluster, following the previous gateway election procedure.

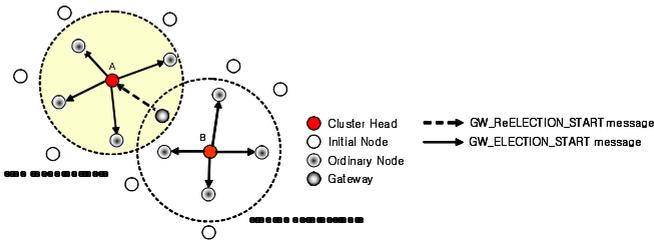


Fig. 6. Gateway re-election procedure

3.3 Slave/Master Patching (SMP) Scheme

Most of the clustering algorithms, including the proposed NND algorithm, inherit the problem that the number of GWs is rapidly increasing as a CH is elected although it includes no or a few ONs within its radio range. It is illustrated in Fig. 7. To solve this problem, we consider another feature based on a simple control scheme, named as a Slave/Master Patching (SMP) Scheme. A basic idea behind this scheme is that a CH

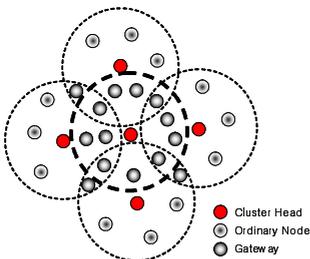


Fig. 7. Undesirable situation of too many GW formed: Illustrative Example

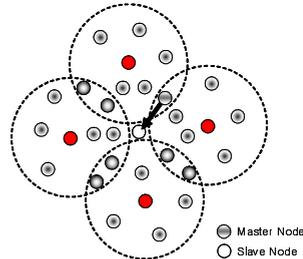


Fig. 8. Example of Slave/Master Patching Scheme

is not elected by GW if the number of INs within a radio range of GW (i.e., the number of received ON_COUNT_REPORT message) is lower than a predefined bound, termed as an election threshold N in the course of our cluster head election procedure ($N = 0, 1, 2, \dots$). To this end, the remaining INs become slave nodes (SNs) of a GW and then, the GW becomes a master node (MN) of the remaining INs, as shown in Fig. 8. The resource and traffic load of the SN can be controlled by MN.

4 Performance Analysis

Computer simulation has been performed to evaluate the performance improvement of the proposed NND algorithm over two existing clustering algorithms, Lowest ID algorithm [1] and Topology Discovery algorithm (TopDisc) [3].

The performance measures in our analysis include the numbers of CHs and GWs that have been reduced and the number of ONs that have been increased by the NND algorithm. We assume that sensor nodes are randomly deployed in a field of $50\text{m} \times 50\text{m}$ as varying the number of sensor nodes from 500 to 1,500. A radio range of each node is fixed to 5m. All results are obtained by taking an average of 500 different runs. Unless stated otherwise, we have chosen the predefined election threshold in the SMP scheme as $N = 2$, i.e., GW does not elect CH if the number of INs within its radio range is lower than two.

Fig. 9 through Fig. 11 show the average number of CHs, GWs and ONs formed by each clustering algorithm as the total number of nodes is varied. In Fig. 9, our NND algorithm shows about 6% improvement in the average number of CHs over Lowest ID algorithm. Not much difference from TopDisc algorithm can be found, even if multi-path fading is not taken into account in our simulation. Furthermore, we find the SMP scheme, as a part of our NND algorithm, achieving a significant improvement, as much as 21% over Lowest ID algorithm. Fig. 10 shows that the NND algorithm achieves as much as 13% and 8% decreases in the average number of GWs respectively, as compared with Lowest ID and TopDisc algorithms. Furthermore, the NND algorithm *with* the SMP scheme achieves the average 40 % decrease as compared with Lowest ID algorithm. Meanwhile, Fig. 11 shows that it achieves as much as 29% increase in the average number of ONs as compared with Lowest ID algorithm while achieving the average 11 % increase as compared with TopDisc

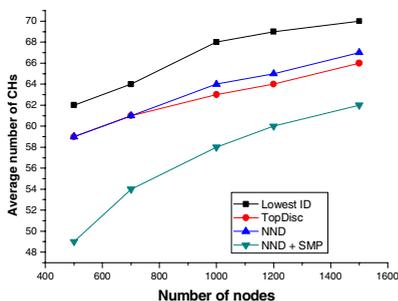


Fig. 9. Average number of CHs

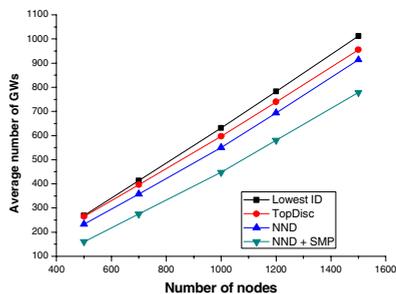


Fig. 10. Average number of GWs

algorithm. Finally, the NND algorithm *with* SMP scheme can achieve as much as 66% increase in the average number of formed ONs as compared with Lowest ID algorithm. In Fig. 12 through Fig. 14, we investigate an effect of the election threshold N on the performance of our SMP scheme. As N increases, it shows that the numbers of CHs and GWs are decreasing while the numbers of ONs and SNs are increasing. Important implication is that a mix ratio of all different types of nodes can be controlled by varying the election threshold N in our approach.

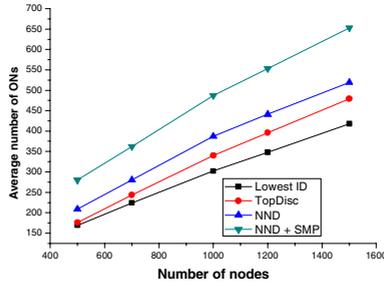


Fig. 11. Average number of ONs

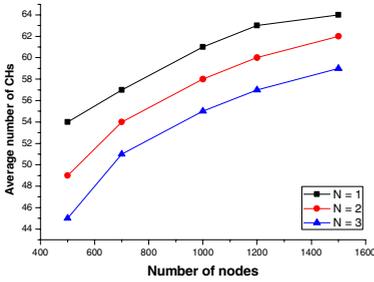


Fig. 12. Average number of CHs

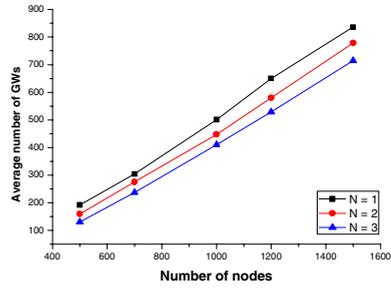


Fig. 13. Average number of GWs

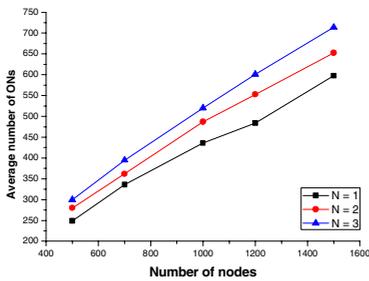


Fig. 14. Average number of ONs

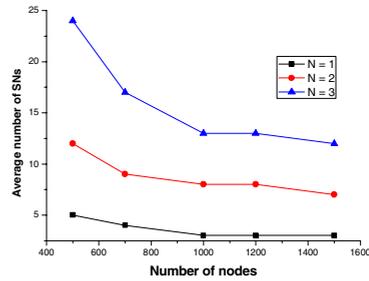


Fig. 15. Average number of SNs

5 Conclusion

In this paper, we have proposed a new type of clustering algorithm for a coverage extension, facilitating the energy-efficient medium access and routing controls in the large-scale wireless sensor networks (WSNs). It has been found that the proposed NND algorithm in conjunction with the slave/master patching scheme is effective for implementing the various routing architectures with a single frequency band channel operation. As a mix ratio of all different types of nodes can be controlled by the SMP scheme, it will be a useful and flexible means of deploying the WSNs subject to a limited network lifetime. In particular, the NND algorithm has been successfully employed in our recent development of MAC protocol for WSNs, HiPERMAC (Hierarchically-Paired Evolutionary Radio MAC Protocol) [9], which is designed for implement the hierarchical routing protocol subject to a single frequency channel.

Acknowledgment

This work has been supported in part by the SAMSUNG Advanced Institute of Technology (SAIT) of Korea and in part by grant No. R01-2003-000-10155-0(2004) from the Basic Research Program of the Korea Science & Engineering Foundation.

References

1. M, Gerla., J, Tsai.: Multiclustler mobile multimedia radio network. ACM Baltzer Journal of Wireless Networks (1995)
2. D.J, Baker., A, Ephremides.: The Architectural Organization of a Mobile Radio Network via a Distributed Algorithm. IEEE Transactions on Communications, Vol. 29, No. 11 (1981) 1694 - 1701
3. Budhaditya, Deb., Sudeept, Bhatnagar., Badri, Nath.: A Topology Discovery Algorithm for Sensor Networks with Applications to Network Management. Technical Report DCS-TR-441, Department of Computer Science, Rutgers University (2001)
4. B, Krishnamachari.: The impact of data aggregation in wireless sensor networks. in the 22nd International Conference on Distributed Computing Systems Workshops (2002)
5. Zhihui, Chen., Khokhar, A.: Self organization and energy efficient TDMA MAC protocol by wake up for wireless sensor networks. Sensor and Ad Hoc Communications and Networks,2004. IEEE SECON 2004. (2004) 335 - 341
6. T.J, Kwon., Mario, Gerla.: Efficient flooding with Passive Clustering (PC) in ad hoc networks. ACM SIGCOMM Computer Communication Review Volume 32 Issue 1 (2002)
7. Lin, C.R., Gerla, M.: Adaptive clustering for mobile wireless networks. Selected Areas in Communications, IEEE Journal on Volume 15, Issue 7 (1997) 1265 - 1275
8. Chatterjee, M., Sas, S.K., Turgut, D.: An on-demand weighted clustering algorithm (WCA) for ad hoc networks. Global Telecommunications Conference, GLOBECOM '00. IEEE Volume 3 (2000) 1697 - 1701
9. HiPERMAC: Hierarchically-Paired Evolutionary Radio MAC Protocol, Internal Report, Samsung Advanced Institute of Technology, June 2005 (also filed for patent).