A Hybrid Clustering Approach for Prolonging Lifetime in Wireless Sensor Networks

Peyman Neamatollahi, Hoda Taheri
Dept. Computer Engineering
Islamic Azad University, Mashhad branch, Young Researchers Club
Mashhad, Iran
{neamatollahi.peyman, h.taheri.mshd}@gmail.com

Mahmoud Naghibzadeh
Dept. Computer Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
naghibzadeh@um.ac.ir

Mohammad-Hossein Yaghmaee
Dept. Computer Engineering
Islamic Azad University, Mashhad branch
Mashhad, Iran
yaghmaee@ieee.org

***Abstract*—Clustering is an effective approach for organizing the network into a connected hierarchy, load balancing, and prolonging network lifetime. Clustering protocols in wireless sensor networks are classified into static and dynamic. In static clustering, clusters are formed once, forever and role of the cluster head is scheduled among the nodes in a cluster. However, in dynamic clustering the time is divided into rounds and clustering is performed in the beginning of each round. This paper presents a Hybrid Clustering Approach (HCA). Whenever a cluster head consumes a prespecified part of its energy, it indirectly informs all other nodes so, clustering will be done in the beginning of the upcoming round. Therefore, clustering is performed on demand. To evaluate the efficiency of proposal, the well known distributed clustering protocol, HEED, is used as baseline example. By means of simulation results, we demonstrate that significant energy saving can be achieved using HCA.**

Keywords-sensor networks; clustering; network lifetime; energy efficiency; distributed algorithms; hybrid approach.

I. INTRODUCTION

Wireless sensor networks (WSNs) provide reliable monitoring from far away distances. These networks are basically data gathering networks where data are highly correlated and the end user needs a high level description of the environment the nodes are sensing [1]. The requirements of these networks are ease of deployment, long system lifetime, and low-latency data transfers. The main task of a sensor node in a sensor field is to detect events, perform quick local data processing, and then transmit the data [2]. As it is mentioned in [3, 4], the nodes have typically low mobility and they are limited in capabilities, energy supply and bandwidth. The sensor network should perform as long as possible. On the other hand, battery recharging may be inconvenient or impossible. Therefore, all aspects of the sensor node, from the hardware to the protocols, must be designed to be extremely energy efficient [5]. In a sensor node, energy consumption can be “useful” or “wasteful”. Useful energy consumption can be either due to the following items:

- transmitting/receiving data
- processing query requests
- forwarding queries/data to neighboring nodes.

Wasteful energy consumption can be due to the following items:

- idle listening to the media
- retransmitting due to packet collisions
- overhearing
- generating/handling control packets [6].

In direct communication WSNs, the sensor nodes directly transmit their sensing data to the Base Station (BS) without any coordination between them. However, in Cluster-based WSNs, the network is divided into clusters. Each sensor node exchanges its information only with its cluster head (CH), which transmits the aggregated information to the BS. Aggregation and fusion of sensor node’s data at the CHs motivate significant reduction in the amount of data sent to the BS; therefore cause saving energy and bandwidth resources. Once the clusters are constructed, each sensor node will be given an exclusive time slot; therefore, each sensor node knows when to transmit. Consequently, a node does not require being awake during the complete time-division multiple access (TDMA) frame but only at its specific time slot [7]. To sum up, clustering coordinates the transmissions of the sensor nodes with a common schedule in the steady state phase, which removes collisions, idle listening, and overhearing. Therefore, clustering achieves an important improvement in terms of energy consumption. Besides, it is particularly crucial for scaling the network to hundreds or thousands of nodes [8]. On the other hand, the clustering is one of the basic approaches to design energy-efficient distributed WSNs. In many applications, cluster organization is a natural way to group spatially close sensor nodes, so that exploit the correlation and eliminate redundancy that often exists among the sensor readings [9]. However, in many cases these benefits, compared to the direct communication WSN, result in extra overhead due to cluster formation’s message exchanges.

Generally, clustering sensor nodes is performed in two ways: centralized and distributed. In centralized clustering protocols (e.g. [15]), the BS is responsible for gathering the clustering information from the network. So, it has global knowledge of the network. Because the BS is unlimited in terms of processing, memory and energy supply, it can execute the best clustering algorithms to form the network. Therefore, the BS determines CHs and also respective clusters. These protocols are not applicable for large scale networks because collecting the clustering information at the BS is both time and energy consuming. However, in distributed clustering protocols (e.g. [5, 6, 8, 13, and 14]) every node decides to whether become CH or not only based on local information rather than global information.

In view of cluster formation, there are two types of approaches in the literature: static and dynamic. In static clustering protocols, the clusters are constructed only once and permanently. Some of these protocols (e.g. [15]) predefine CHs forever. These protocols are often useful for heterogeneous networks, in which the nodes with higher energy and more capabilities become CHs during the network lifetime. In many cases, static clustering protocols apply centralized approach to form ideal clusters one time. In dynamic protocols (e.g. [5, 6, 13, and 14]), the clustering is done in the beginning of each round. Most of them are consist of two phases: setup phase and steady state phase. In setup phase, the CH election is often performed at first, and then CHs announce some other nodes to join them. After that, each CH schedules the members to transmit data and inform them of their respective time slot. In steady state phase, each regular (non CH) node sends the sensing data to its CH during the respective time slot. Then the CHs after aggregating the received data, forward them to the BS through single hop or multi hop fashion. After a certain period of time that is elapsed, the network goes into the clustering phase again and enters the next round. The disadvantage of dynamic clustering approach compared to the static clustering approach is the extra overhead, due to performing setup phase in each round, imposed on the network.

Considering the disadvantages of two mentioned approaches, static and dynamic, we present a hybrid method for the clustering protocols in WSNs named HCA. In our approach the clustering is not performed in each round, because it is triggered when at least one CH has consumed a prespecified part of its residual energy. Using this method, we propose a novel mechanism to improve the energy efficiency in cluster-based WSNs. To evaluate the efficiency of HCA, we implement it on a well known distributed clustering protocol, HEED [6], then compare the output with HEED and LEACH [5] protocols. Simulation results (in MATLAB) show that the method conserves energy and consequently extends the network lifetime.

The rest of the paper is organized as follows: Section 2 introduces the reference protocols, Section 3 describes the hybrid clustering approach, Section 4 shows the simulation results in MATLAB software, and conclusion is at the end.

II. Reference Protocols

In the following, we review two famous distributed dynamic clustering protocols used in our simulation.

A. Low Energy Adaptive Clustering Hierarchy

LEACH [12] minimizes energy dissipation in sensor networks due to constructing clusters. This protocol does not consider node's residual energy in the clustering process. LEACH operation is done in two phases, setup phase and steady state phase. In the setup phase, a sensor node selects a random number between 0 and 1. If this number is less than the threshold *T(n)*, the node becomes a CH. *T(n)* is computed as:

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

where r is the current round; *p,* the desired percentage for becoming CH; and *G*, is the collection of nodes that in the last $1/p$ rounds have not been elected as a CH. After electing CHs, every CH announces all sensor nodes in the network that it is the new CH. When each node receives the announcement, it chooses its desired cluster to join based on the signal strength of the announcement from the CHs to it. So, the sensor nodes inform their appropriate CH to join it. Afterwards, the CHs based on a TDMA approach assign the time slot to each node so that a member can send its data to its CH in this period. The sensor nodes can initiate sensing and transmitting data to the CHs during the steady state phase. The CHs also aggregate data received from the nodes in their cluster before sending these data to the BS via a single hop fashion.

B. Hybrid Energy Efficient Distributed

Younis and Fahmy [6] proposed an iterative clustering protocol, named HEED. HEED is different from LEACH in the way CHs are elected. Both, electing the CHs and joining to the clusters, are done based on the combination of two parameters. The primary parameter depends on the node's residual energy. The alternative parameter is the intra cluster "communication cost". Each node computes a communication cost depending on whether variable power levels, applied for intra cluster communication, are permissible or not. If the power level is fixed for all of the nodes, then the communication cost can be proportional to (i) node degree, if load distribution between CHs is required, or (ii) 1/node degree, if producing dense clusters is required. The authors define *AMRP* the average of the minimum power levels needed by all *M* nodes within the cluster range to access the CH u, i.e. $AMRP(u) = \frac{\sum_{i=1}^{M} MinPwr(i)}{M}$. If variable power levels are admissible, *AMRP* is used as the cost function. In this approach, every regular node elects the least communication cost CH in order to join it. On the other hand, the CHs send the aggregated data to the BS in a multi hop fashion.

III. The Hybrid Clustering Approach

Suppose *n* nodes are distributed in a field. Let L_i denote the lifetime of node *i*. Let the network lifetime, which is defined as *L* be the time elapsed until the first node in the network depletes its energy. In other words, $L = min(L_1, L_2, ..., L_n)$. The major purpose is to maximize *L*, which requires using the energy of all nodes uniformly.

In order to overcome the disadvantages of static and dynamic clustering approaches, we design the hybrid method in a way that clustering is not performed in every round. To do so, CHs after that clusters formed, save their residual energy in their memory (say into E_{CH} variable) at the end of each setup phase. Whenever a CH finds that its $E_{residual}$ becomes less than αE_{CH} (α is a constant number and $0 < \alpha < 1$), it sets a specific bit in a data packet which is ready to be sent to the BS in the current TDMA frame. The CH through forwarding this packet to the BS (in a multi hop fashion) informs the BS that the sensors should hold the setup phase at the beginning of the upcoming round. After that, the BS sends specific synchronization pulses in a multi hop fashion to all nodes.

E_{CH}: Residual energy of the CH node in the end of each setup phase.
α: A floating point number which is constant and $0 \leq \alpha \leq 1$.
$S = \{s \in S_{CH} | E_{residual}(s) > 0\}$,
$\forall x \in S$: IF $E_{residual}(x) < \alpha E_{CH}(x)$ THEN
The node x sets a specific bit in its data packet and informs the BS in a multi hop fashion.
IF *the synch pulse has received* THEN
The node x becomes ready to hold the clustering for the upcoming round.
The BS receives a data packet:
IF $specific\_bit = true$ THEN
The BS broadcasts the synch pulse one time per round in the network via multi hop fashion.

Figure 1. The pseudo code of HCA.

These pulses are quickly dispersed in the network according to the approach presented in [17]. When each node receives a pulse, it prepares itself to perform clustering. So, CH election and consequently the cluster formation are done on demand. For details see Fig. 1. Using this method, an overhead due to consecutive setup phases is tremendously reduced. Besides, we have decrease in energy consumption of nodes and increase in network lifetime totally.

IV. SIMULATION RESULTS

In this section, we compare the implementation of HCA on HEED (referred as HCA in figures) with HEED and LEACH protocols from different aspects. This comparison is performed via MATLAB software using the following assumptions and system parameters:

- All nodes have the equal amount of initial energy (2 J)
- The total number of nodes in the system is $N = 100$
- Sensor nodes are randomly dispersed in a square field (between (0, 0) and (100, 100) m)
- The BS is located outside the supervised area at the coordinate (50, 175)
- The energy required for data aggregation is set as E_{DA}=5nJ/bit/signal and CHs perform ideal data aggregation
- We assume a simple model for energy dissipation of the radio hardware in which the receiver dissipates energy to run the radio electronics and the transmitter dissipates energy to run the power amplifier and the radio electronics, as shown in Fig. 2 which is redrawn from [5]. Thus, for transmitting a k-bit message a distance, the radio expends

$$E_{Tx}(k,d) = E_{Tx-elec}(k) + E_{Tx-amp}(k,d)$$

$$E_{Tx}(k,d) = \begin{cases} kE_{elec} + k\epsilon_{fs}d^2, & d < d_0 \\ kE_{elec} + k\epsilon_{mp}d^4, & d \geq d_0 \end{cases}$$

and for receiving this message, the radio expends:

$$E_{Rx}(k) = E_{Rx-elec}(k) = kE_{elec}.$$

- Nodes always have data to send to the end user and nodes situated in each other's close proximity have correlated data
- The other parameters are listed in TABLE 1.

In HEED protocol, T_{CP} and T_{NO} are defined as follows:

- T_{CP} (the period of clustering process) is the time interval used by the clustering protocol to cluster the network

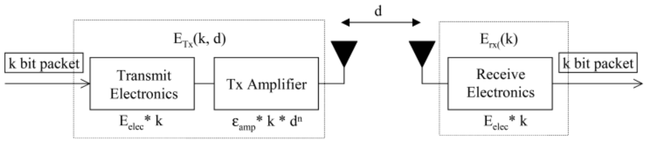

Figure 2. Radio energy dissipation model.

TABLE I. PARAMETER SETTINGS

Parameter	Value
ϵ_{fs}	10 pJ/bit/m^2
ϵ_{mp}	0.0013 pJ/bit/m^4
E_{elec}	50 nJ/bit
E_{DA}	5 nJ/bit/signal
Idle power	13.5 mW
Sleep power	15 µW
Initial energy per node	2 J
Round time	20 sec
Network grid	From (0,0) to (100,100)
Base Station	At (50,175)
Threshold distance (d_0)	75 m
Initial cluster radius (R_C)	45 m
Data packet size	100 byte
Control packet size	25 byte
Round (T_{NO})	5 TDMA frames
α	0.8

- T_{NO} (the network operation interval) is the time between the end of a T_{CP} interval and the start of the subsequent T_{CP} interval.

In order to reduce overhead we must ensure that $T_{NO} >> T_{CP}$. Note that in contrast with other dynamic clustering protocols that perform cluster formation in each round, our method accomplishes it on demand rather than each round. Therefore, it is possible that some rounds do not include T_{CP}; instead T_{NO} extends during these rounds. As a result, the length of T_{CP} interval is fixed but the length of T_{NO} interval is variable in the network lifetime.

We show α-diagram in Fig. 3 and then investigate the number of live nodes, the network lifetime, total residual energy of the network, and the number of the CH election during the network lifetime in Fig. 4 to 7.

As it is mentioned before, the CH election is done on demand; When a CH consumes a specific part of its energy (i.e. $E_{residual} \leq \alpha E_{CH}$) it informs the other nodes indirectly that the CH election must be performed for the upcoming round. In order to obtain the best α, we ran HCA on HEED with different communication costs and different values of α. In Fig. 3, α differs from 0 to 1 and each plot demonstrates the average of 5 times execution for the specified cost when the first node dies (FND). When α is equal to zero it means that no CH election is done during the network lifetime (i.e. static clustering). In homogenous networks which nodes have similar capabilities and the same amount of energy, CHs depletes respective energies quickly. Therefore, when a CH dies, the respective cluster becomes dysfunctional. When α equals to one, it means that the CH election is performed each round, similar to LEACH and HEED protocols. The rest of the figures in this simulation are achieved from $\alpha = 0.8$ because it gains better network lifetime for different communication costs totally.

Fig. 4 shows the total number of nodes that remain alive over the simulation round with node degree, 1/node degree and AMRP communication costs. As it is shown in this figure, HCA considerably increases network lifetime more than HEED and LEACH protocols. The main reason is that the number of performed clustering are reduced considerably because this protocol does not hold the clustering in all rounds. Therefore, the extra overhead due to consecutive setup phase decreases

To better compare network lifetime for the mentioned communication costs, we can summarize the time the first node dies (FND), only one half of the nodes which are alive (HNA) and the last node dies (LND) in the Fig. 5. It is obvious that HCA acts better in any definition of network lifetime than HEED and LEACH protocols.

Fig. 6 shows total residual energy of all nodes in the network with the node degree communication cost. This figure also shows that HCA conserves more energy than others. The plots of other two costs are similar, so they are removed.

By not performing the setup phase in all rounds in HCA, we can compare the number of clustering after each round for the protocols in Fig. 7. The plots of LEACH and HEED protocols are the same because both of them perform the setup phase in each round. As it is shown in this figure, the number of CH elections in HCA is much less than this value in other two protocols. During the network lifetime in HCA protocol, by increasing the number of rounds the number of CH elections increases progressively. The reason is that by decreasing energy of nodes, the number of clustering that should be held becomes more.

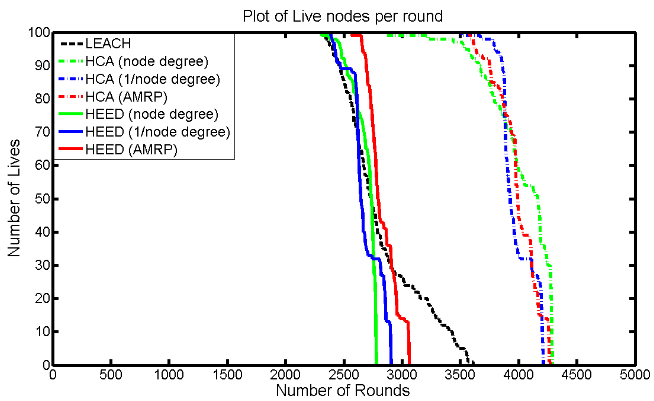

Figure 4. The number of live nodes in HCA, LEACH and HEED protocols.

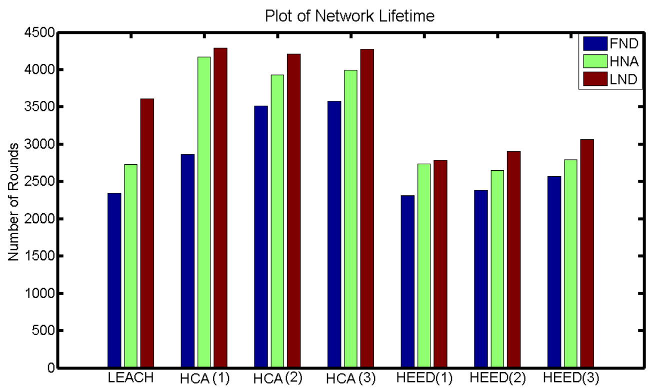

Figure 5. The comparison of different definitions of network lifetime in HCA, LEACH and HEED protocols with different communication costs: (1) node degree, (2) 1/node degree, and (3) AMRP.

V. Conclusion

In this paper, we proposed a hybrid approach for clustering WSNs. Our approach can be useful for applications that require scalability and prolonged network lifetime. After the first setup phase, the clustering will not be performed until at least one of the CHs consume a predefine part of its energy, because doing the clustering at the beginning of each round imposes lots of overhead on the network. We evaluated HCA on HEED by comparing it with HEED and LEACH protocols using the MATLAB. The simulation results show that HCA is approximately 30% more efficient in terms of network lifetime than the two protocols. The main reason is that the clustering is executed on demand. We will attempt to expand the on demand setup phase approach to the other well known protocols.

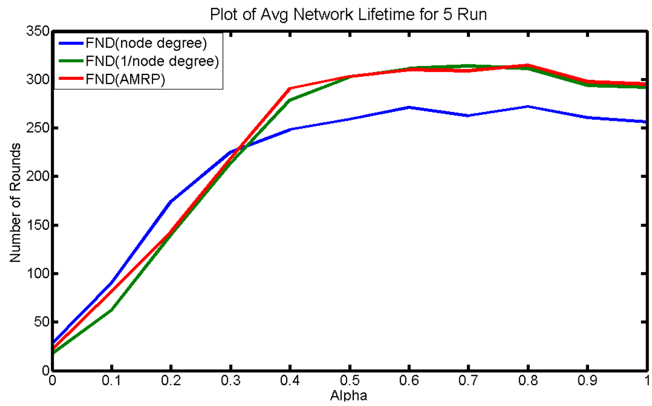

Figure 3. The α-diagram in HCA with different communication costs.

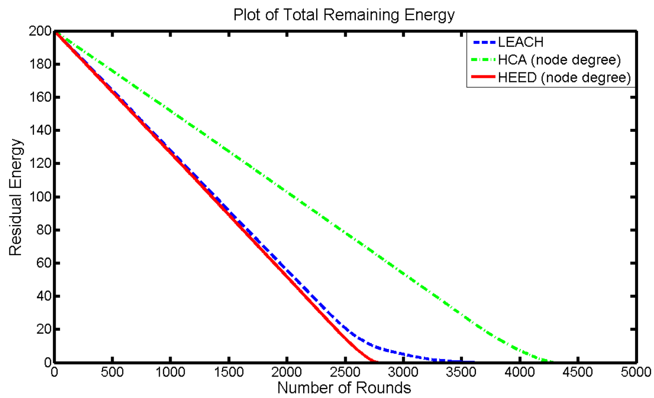

Figure 6. The comparison of total remaining energy in HCA, LEACH and HEED protocols with node degree communication cost.

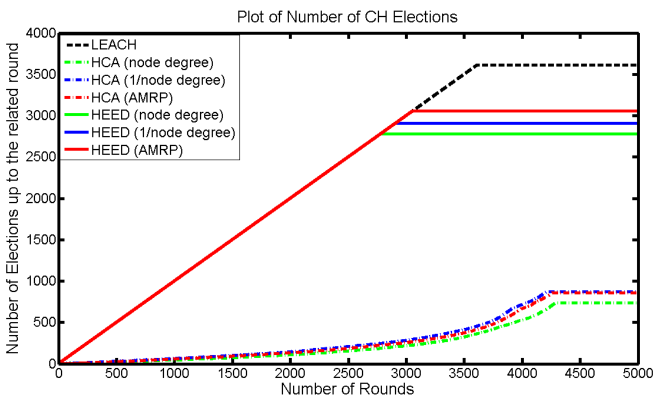

Figure 7. The comparison of the number of the CH election up to each round in HCA, LEACH and HEED protocols with different communication costs.

REFERENCES

[1] W. B. Heinzelman, "Application-specific protocol architectures for wireless networks," *Ph.D. Thesis, Massachusetts Institute of Technology*, June 2000.

[2] I. Akiyldiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.

[3] O. Younis and S. Fahamy, "Distributed clustering in ad-hoc sensor networks: A hybrid, energy-efficient approach," in *Proc. IEEE INFOCOM*, Mar. 2004, vol. 1, pp. 629–640.

[4] K. Akkaya, M. Younis, "A survey on routing protocols for wireless sensor networks," *Elsevier Journal of Ad Hoc Networks*, vol. 3, no. 3, 325–349, 2005.

[5] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 660–670, Oct. 2002.

[6] O. Younis, S. Fahmy, "HEED: A Hybrid, Energy-Efficient, Distributed clustering approach for Ad Hoc sensor networks," *IEEE Transactions on Mobile Computing,* vol. 3, no. 4, pp. 366–379, 2004.

[7] N. Bouabdallah, M. E. Rivero–Angeles, and B. Sericola, "Continuous Monitoring Using Event-Driven Reporting for Cluster-Based Wireless Sensor Networks," *IEEE Trans. Vehicular Technology.*, vol. 58, no. 7, pp. 3460–3479, Sept. 2009.

[8] X. Zhu, L. Shen, and T. P. Yum, "Hausdorff Clustering and Minimum Energy Routing for Wireless Sensor Networks," *IEEE Trans. Vehicular Technology*, vol. 58, no. 2, 2009.

[9] S. Soro, W. Heinzelman, "Cluster head election techniques for coverage preservation in wireless sensor networks", Ad Hoc Networks, vol. 7, no. 5, pp. 955–972, July 2009.

[10] A. A. Abbasi, M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Computer Communications*, vol. 30, pp. 2826–2841, June 2007.

[11] O. Younis, M. Krunz, and S. Ramasubramanian, "Node Clustering in Wireless Sensor Networks: Recent Developments and Deployment Challenges*," IEEE Network,* May/June 2006.

[12] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," *Proc. of the Hawaii International Conference on System Sciences*, pp. 1–10, January 2000.

[13] M. C. M. Thein, T. Thein, "An Energy Efficient Cluster-Head Selection for Wireless Sensor Networks," *International Conference on Intelligent Systems, Modeling and Simulation*, 287-291, 2010.

[14] S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in *Proc. IEEE INFOCOM*, Apr. 2003, vol. 3, pp. 1713–1723.

[15] T. Kaur, J. Baek," A Strategic Deployment and Cluster-Header Selection for Wireless Sensor Networks," *IEEE Trans. on Consumer Electronics*, Vol. 55, No. 4, 2009.

[16] S. Singh, M.Woo, and C. Raghavendra, "Power-aware routing in mobile ad hoc networks," in *Proc. 4th Annual ACM/IEEE Int. Conf. Mobile Computing Networking (MobiCom)*, Oct. 1998.

[17] J. Greunen, J. Rabaey, "Lightweight Time Synchronization for Sensor Networks," *WSNA'03*, Sept. 2003.