ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

EECS: Energy Efficient Clustering Scheme in Wireless Sensor Networks

Sameer Muftah AmarAsayh
Department of Electronics and Communication,
J.K Institute of Applied Physics and Technology,
University of Allahabad.

Prof. Rajiv RanjanTewarii
Department of Electronics and Communication
J.K Institute of Applied Physics and Technology,
University of Allahabad.

Abstract:

Data gathering is a common but critical operationin many applications of wireless sensor networks. Innovative techniques that improve energy e±ciency toprolong the network lifetime are highly required. Clus-tering is an e®ective topology control approach in wire-less sensor networks, which can increase network scalability and lifetime. In this paper, we propose a novel clustering schema EECS for wireless sensor networks,which better suits the periodical data gathering applications. Our approach elects cluster heads withmore residual energy through local radio communication while achieving well cluster head distribution; further more, it introduces a novel method to balance theload among the cluster heads. Simulation results showthat EECS outperforms LEACH signi¯cantly with prolonging the network lifetime over 35%.

1 Introduction:

Continued advances of MEMS and wireless communication technologies have enabled the deploymentof large scale wireless sensor networks (WSNs) [1].The potential applications of WSNs are highly varied,such as environmental monitoring, target tracking andmilitary [2]. Sensors in such a network are equippedwith sensing, data processing and radio transmission units while the power is highly limited. Due to thesensors' limited power, innovative techniques that improve energy effiency to prolong the network lifetimeare highly required.Data gathering is a common but critical operationin many applications of WSNs, where data aggregation and hierarchical mechanism are commonly usedtechniques. Data aggregation can eliminate the data redundancy and reduce the communication load [3].Hierarchical (clustering) mechanisms are especially effective in increasing network scalability and reducingdata latency, which have been extensively exploited.

LEACH [4] which is the rst clustering protocol, proposes a two-phase mechanism based on single-hopcommunication. The plain node transmits the data tothe corresponding cluster head and the cluster headtransmits the aggregated data to the base station (BS). HEED [5] selects cluster heads through O(1)time iteration according to some metric and adoptsthe multi-hop communication to further reduce theenergy consumption. PEGASIS [6] improves the performance of LEACH and prolongs the network lifetimegreatly with a chain topology. But the delay is significant although the energy is saved. There are someother related work [7{9] which efficently use energythrough clustering.In this paper, we propose and evaluate an energyefficient clustering scheme (EECS) for periodical datagathering applications in WSNs. In the cluster headelection phase, a constant number of candidate nodesare elected and compete for cluster heads accordingto the node residual energy.

The competition process is localized and without iteration, thus it has muchlower message overhead. The method also producesa near uniform distribution of cluster heads. Furtherin the cluster formation phase, a novel approach isintroduced to balance the load among cluster heads.EECS is fully distributed and more energy efficientand the simulation results show that it prolongs thenetwork lifetime as much as 135% of LEACH. The remainder of this paper is organized as follows. Section 2 outlines the data gathering issues in WSNs.Section 3 exhibits the details of EECS and Section 4analyzes the properties of EECS. Section 5 evaluatesthe performance of EECS. Finally, Section 6 gives theconclusion and future work.2 Problem OutlineData gathering is a typical application in WSNs.Sensors periodical sense the environment and transmitthe data to the base station (BS), and the BS analyzesthe data to draw some conclusions about the activityin the area. We make a few assumptions about the network model and introduce the radio model beforethe problem statements.

2.1 Network Model:

To simplify the network model, we adopt a fewreasonable assumptions as follows:

1)Nsensors areuniformly dispersed within a square ¯eldA;

2)Allsensors and BS are stationary after deployment;

3) The communication is based on the single-hop;

4)Communication is symmetric and a sensor can compute the approximate distance based on the receivedsignal strength if the transmission power is given;

5)All sensors are location-unaware;

6)All sensors are of equal significance.

We use a simplied model shown in [4] for the radio hardware energy dissipation as follows. We refer readers to [4] for more details. To transmit an l –bitdata to a distance d, the radio expands:

$$E_{T_x}(l,d) = \begin{cases} l \times E_{elec} + l \times \epsilon_{fs} d^2, & d < d_{crossover} \\ l \times E_{elec} + l \times \epsilon_{mp} d^4, & d \geq d_{crossover} \end{cases} \quad (1)$$

The first item presents the energy consumption ofradio dissipation, while the second presents the energy consumption for amplifying radio. Dependingon the transmission distance both the free space ²fsand the multi-path fading ²mp channel models areused [11]. When receiving this data, the radio expends: ERx(l) = l£Eelec. Additionally, the operationof data aggregation consumes the energy as EDA.

2.2 Problem Statement:

Once a sensor node runs out its energy, we consider the network is dead because some area cannotbe monitored any more. Periodical data gathering applications in large scale sensor networks appeal the design of scalable, energy efficient clustering algorithms.

Thus our primal goals in EECS are as follows: 1) fullydistributed manner. Sensors interact with each otherthrough localized communication;

2) low control overhead. It is desirable to reduce control overhead to extend the time of data gathering;

3) load balancedclustering mechanism. Balance the load among thesensors, especially among the cluster heads. In thenext section, we will describe the EECS algorithm indetails.

3 EECS DetailsEECS is a LEACH-like clustering scheme, wherethe network is partitioned into a set of clusters withone cluster head in each cluster. Communication between cluster head and BS is direct (single-hop). Foreasy reference, we summarize the notations in Table 1.
In the network deployment phase, the BS broadcasts a \hello" message to all the nodes at a certain
Table 1: Meanings of the Notations

Notation	Meaning
T	a threshold between 0 and 1
$E_{residual}$	the residual energy of node
CH	the set of cluster heads
m_j	the sum of members in cluster j
P	the set of plain nodes
CH_i	the i^{th} node in CH
P_j	the j^{th} node in P
$d(x,y)$	the distance between node x and y
$EX(x)$	the expectation of x
$R_{compete}$	broadcast radius of candidate nodes

power level. By this way each node can computethe approximate distance to the BS based on the received signal strength. It helps nodes to select theproper power level to communicate with the BS. Aswill shown in Section 3.2, we will use this distanceto balance the load among cluster heads.

In clusterhead election phase, well distributed cluster heads areelected with a little control overhead. And In cluster formation phase, a novel weighted function is in-troduced to form load balanced clusters. Detailed descriptions of these two phases are in the following subsections.

3.1 Cluster head election:

In this phase, several cluster heads are elected.Nodes become CANDIDATE nodes with a probabilityT and then broadcast the COMPETE HEAD MSGs withinradio range Rcompeteto advertise their wills.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

EachCANDIDATE node checks whether there is a CANDIDATE node with more residual energy within the radiusRcompete. Once the CANDIDATE node finds a more powerful CANDIDATE node, it will give up the competitionwithout receiving subsequential COMPETE HEAD MSGs.Otherwise, it will be elected as HEAD in the end.

3.2 Cluster formation:

In this phase, each HEAD node broadcasts theHEAD AD MSG across the network, while the PLAINnodes receive all the HEAD AD MSGs and decide whichcluster to join. Most of existed metric for PLAIN nodesto make decisions is the distance metric. For example in [4] or [7], the PLAIN nodes choose the clusterhead that require minimum communication according to the received signal strength. However, pursuing efficient energy consumption of the PLAIN nodes onlymay lead HEAD nodes exhausted quickly during thedata transmission phase. In the data transmission phase, the consumed energy of cluster head i, E(CHi) is as follows, assuming d(CHi;BS) >dcrossover.

E(CHi) = milEelec+(mi+1)lEDA+l(Eelec+²mpd4) (2)

Observing formula 2, energy consumption ofE(CHi) is composed of three parts: data receiving,data aggregation and data transmission. In the field,several cluster heads may be near the BS, while someare far away. The energy expended during data transmission for far away cluster heads is signi¯cant, especially in large scale networks. Since d(CHi;BS) hasbeen ¯xed after cluster head election , we should justify the cluster size for each cluster head to balancetheir load across the network. The larger d(CHi;BS)is, the smaller member size mi the cluster head Chishould accommodate.Energy consumption of the PLAIN node Pjduringtransmitting the data to CHiobey the formula

1. LetE(Pj) be the energy consumed by Pj. If Pjalwayschooses the cluster head CHbest with min fE(Pj) g,CHbestmay be exhausted due to long distance datatransmission to the BS and immoderate cluster size,although the energy of Pjis saved. Thus, PLAIN node Pjin EECS chooses the cluster head by consideringnot only saving its own energy but also balancing theworkload of cluster heads,i.e. two distance factors: d(Pj ;CHi) and d(CHi;BS).We introduce a weighted function cost(j; i) for the PLAIN node Pjto make a decision, which iscost(j; i) = w£f(d(Pj ;CHi))+(1¡w)£g(d(CHi;BS));

(3)andPjchooses CHiwith min fcostgto join.In formula 3, f and g are two normalized functionsfor the distance d(Pj ;CHi) and d(CHi;BS) respectively:

f = d(Pj ;CHi)df max

g = d(CHi;BS) ¡ dg min

dg max ¡ dg min

(4)where df max = EX(max fd(Pj ;CHi)g), dg max =max fd(CHi;BS)g and dg min = min fd(CHi;BS)g.f subfunction in cost guarantees that memberschoose the closest cluster head in order to minimizeenergy consumption of the cluster members, While gsubfunction makes the nodes join the cluster head withsmall d(CHi;BS) to alleviate the workload of the cluster heads farther from the BS. wis the weighted factorfor the tradeo® between f and g. The experiments inSection 6 will show that the optimal value of w depends on the specific network scale.

3.3 Synchronization issues:

Synchronization between each phase should beguaranteed that each node has enough time to complete the procedure; while within each phase, synchronization among the nodes is not necessary and idlenodes will turn to sleep till the phase ends. In EECS,it is achieved by having the BS periodically broadcastsynchronization signals to all nodes. 4 EECS AnalysisIn this section, we analyze the performance of EECS in details and explain how to set the parameters T and Rcompete. Lemma 1.The control overhead complexity acrossthe network is O(N), where N is the number of nodes. Proof.Observing EECS, every node sends outsmall constant-length control messages each roundwithout iteration.

Each HEAD node sends threemessages which are COMPETE HEAD MSG, HEAD AD MSGand SCHEDULE MSG; each CANDIDATE node sends two messages which are COMPETE HEAD MSGand JOIN CLUSTER MSG; while the others sendJOIN CLUSTER MSGs only. Clearly, the total controloverhead is NT + N, whose asymptotic order isO(N). Good quality HEAD nodes should be guaranteed byenough competition of the CANDIDATE nodes. SinceT is the only crucial factor which aspects the sum ofCANDIDATE nodes, it must be large enough to guarantee enough CANDIDATE nodes. On the other hand, the largerT is, the more overhead is produced in the cluster head election phase. So, we must properly set Tto reduce the overhead with guaranteed HEAD quality.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

In LEACH, there is no interaction during the cluster head election. So the control overhead is near optimal, which is 2NP +N(1¡P) = NP +N, where P is similar to T in [4]. Thus the overhead of EECS is only(1+T)=(1+P) times of LEACH. In HEED, HEAD nodesare elected with iteration. Although the communication is localized and the algorithm terminates in O(1) iteration, HEED still produces much more overheadwith the upper bound Niter£N. Clearly, our approachis better than HEED. The above property shows thatthe control overhead of EECS is low signi¯cantly.Lemma 2.There is at most one cluster head in everyRcompete radio covered range.
Proof.Let S be the set of all sensor nodes. And for8x 2 CH, let Cx= fyjd(y; x) • Rcompete; y 2 Sg.For contradiction, we assume that there is a nodey 2 Cxwhich is also a cluster head.

Accordingto the competition metric in cluster head election,x:Eresidual> z:Eresidual; 8z 2 Cx. Since y 2Cx, thenx:Eresidual > y:Eresidual. The communication issymmetric in the network model of EECS. If y is thecluster head, y:Eresidual> x:Eresidual as x is withinthe distance Rcompete, which is a contradiction.So, for 8x 2 CH, 8y 2 Cx, there is y =2 CH.In [4], the author proves that there is an optimalnumber of cluster heads koptin a given scene. SinceEECS is a LEACH-like protocol, we want to elect koptcluster heads every round. According to Lemma 2, Rcompetea®ects the cluster heads directly. So we compute the optimized value of Rcompete, denoted by Roptin the following lemma.Lemma 3.There is an optimal range Ropt forRcompete, which is
q
A
¼kopt
, where kopt is the optimal
range of jCHj.
Proof.Let P(CANDIDATE) be the probability ofone node being CANDIDATE node, so the sum ofCANDIDATE nodes n is P(CANDIDATE) £ N. Inthe Rcompeteradius range, there are m nodes inCANDIDATE state(boundary cases are ignored), wherem = n £ ¼R2compete
A .
Since all nodes have the same capacity, these mnodes have equal probability to be HEAD, then theprobability of one node being HEAD node P(HEAD) =P(HEADjCANDIDATE) = A
¼R2competeN. So the expectation of the sum of cluster heads EX(jCHj) =N £ P(HEAD) = A
¼R2compete.

In order to optimize energy consumption, we wantto let EX(jCHj) equal to koptin [4]. Combining theinduction in [4] and the formula of EX(jCHj) , we can find that the optimal radius RoptisqA¼kopt.In LEACH, cluster heads are elected simply at random. As a result, the distribution of the cluster headsare not ensured and may be non-uniform. Some members have to expend much more energy to communicate with the corresponding cluster heads far away.The last two lemmas show that there is one and onlycluster head within any Rcompetewith high probability. Thus the cluster heads in EECS are distributedevenly.

5 Simulation :

In this section, we evaluate the performance ofEECS protocol implemented with MATLAB. For simplicity, we assume the probability of signal collisionand interference in the wireless channel is ignorable.And we adapt the same MAC protocols in EECS as in

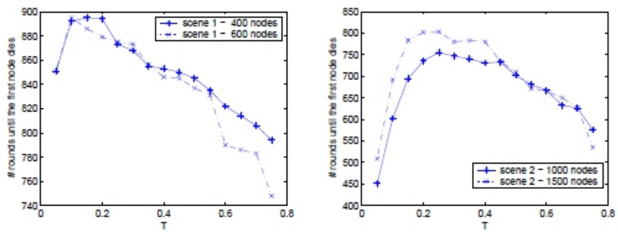


Figure 1: The impaction of T on the network lifetime:(a) normal scene, (b) large scene

LEACH. In order to explain the relations between thenetwork scale and the parameters in EECS, we runeach kind of simulation in two different scenes, whichare normal scale scene (scene 1) and large scale scene(scene 2) respectively.

The parameters of simulations are listed in TABEL2, and the parameters of the radiomodel are the same as LEACH [4]. Unless otherwisespeci¯ed, every simulation result shown below is theaverage of 100 independent experiments where each experiment uses a different randomly-generated uniform topology of sensor nodes.

Table 2: Parameters of Simulations

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Parameter	scene 1	scene 2
Area	100×100	200×200
Location of BS	$(50, 200)$	$(100, 350)$
N	400(600)	1000(1500)
Initial energy	$0.5J$	$1.0\ J$
E_{elec}	$50\ nJ/bit$	
ϵ_{fs}	$10\ pJ/bit/m^2$	
ϵ_{mp}	$0.0013\ pJ/bit/m^4$	
$d_{crossover}$	$87\ m$	
E_{DA}	$5\ nJ/bit/signal$	
Packet size	$4000\ bits$	

Lifetime is the criterion for evaluating the performance of sensor networks. In the simulation, we measure the lifetime in terms of round when the first nodedies. We use the energy utilization rate ´ to evaluatethe efficiency of energy consumption which is definedas the ratio of the total energy consumed when the first node dies to the initial total energy. A high implies that energy consumption is distributed wellacross the network. We first examine the impact of T on the networklifetime, as the scales are different. We have done two

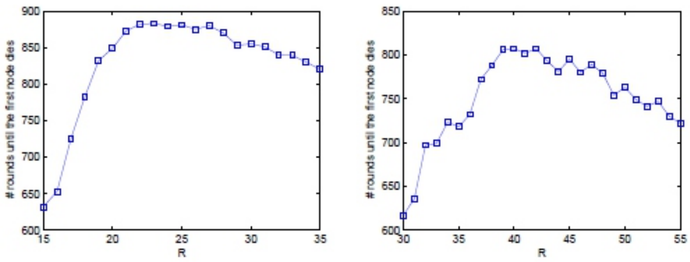
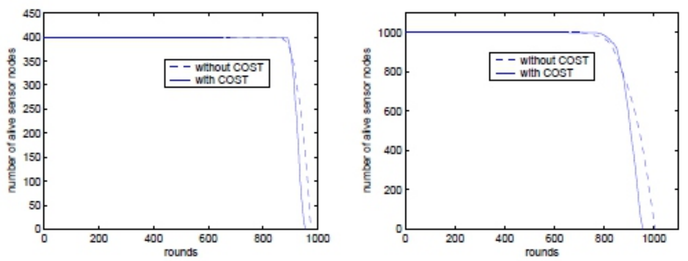


Figure 2: The impaction of Rcompeteon the networklifetime:(a) normal scene, (b) large scene.



(a) normal scene (b) large scene

Figure 3: The impaction of cost on the network lifetime:(a) normal scene, (b) large scene

independent experiments in di®erent scales. In normal scale, N = 400; 600, Rcompete= 26; 22, w = 0:8;in large scale, N = 1000; 1500, Rcompete= 40; 35,w = 0:6.

As T varies from 0.05 to 0.75, Figure 1shows the relation between T and the network lifetime.There is an optimal range for the value of T, which isabout 0:1 » 0:3 in the given scene. According to theexplanation about T in Section4, T must be properlyset with guaranteed HEAD quality and low overhead. Another point needed to be mentioned that the optimal value Toptdecreases when the network densityincreases. It can be explained that there is an optimalsum of CANDIDATE nodes in a given network coveragesize.In the experiment shown in Figure 2, we demonstrate Lemma 3 by observing the relation betweenRcompeteand the network lifetime. In scene 1, N =400 and kopt= 4 » 7, so the optimal value Roptis between 21 » 28; In scene 2, N = 1000 andkopt= 6 » 10, so Roptis between 36 » 46. Observingthe impact on network lifetime when Rcompetevaries, Figure 2 suggests that the optimal value of Rcompeteis about 25 in scene 1 and about 40 in scene 2. Bothresults fall into the optimal range computed prior.

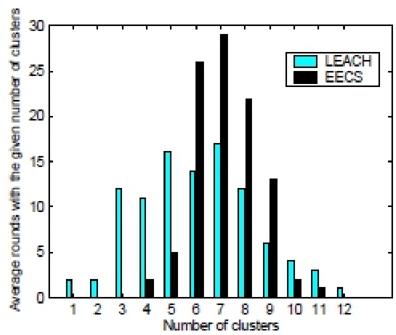


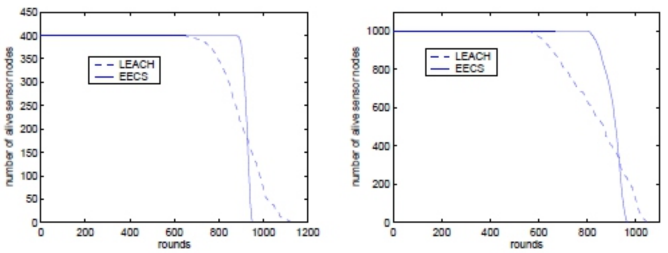
Figure 4: The number of clusters in each round inboth EECS and LEACH (scene 1)

In Figure 3, the experiment shows the efficiency ofcost introduced to balance the load among the cluster heads, where the dash line denoted as the methodwithout considering the cluster heads' load balance issue. We set w at 0.8 in scene 1 and 0.6 in scene 2respectively. Comparing the without ¡ cost method(w = 1) with the with¡costmethod, we find that thecost indeed extends the network lifetime.

The value ofw is determined by the specific scene. While the network grows larger, the di®erence among d(CHi;BS)simpacts the load balance among the cluster headsmore and more distinctly. So w should be decreasedand the PLANE node will consider more about the loadof cluster head when joining the cluster. That's whythe value of w is bigger in scene 1 than in scene 2.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

In this paper, the cost function is simple, and we will optimize the cost function in the next work.Finally, we compare the performance of EECS with the original-LEACH [4] based on the same assumptions in [4]. In scene1, kopt= 6, T = 0:2, Rcompete=26 and w = 0:8; in scene2, kopt= 9, T = 0:15,Rcompete= 40, w = 0:6. In Figure 4, it exhibits thedistribution of the number of clusters in random selected 100 rounds in both EECS and LEACH. Shownas the figure, the number of clusters varies widely ineach simulation run in LEACH; on the other hand,the cluster number varies narrowly at the koptrangein EECS. In LEACH, the clusters in each round is notcontrolled although the expectation is aware; while in EECS, we use the Roptradio radius to set up koptclusters in all probability in each round. Figure 5 showsthe variation of total number of sensors still alive whenthe simulation time lapses. In scene1, EECS prolongsthe lifetime over 35% against LEACH. The energy utilization rate is about 93% in EECS, while only 53%in LEACH. The reason is that EECS always achieves the well distributed cluster heads with considering theresidual energy; further, we consider to balance theload among the cluster heads with weighted function.



(a) normal scene (b) large scene

Figure 5: Performance comparison of EECS andLEACH:(a) normal scene, (b) large scene

In Figure 5-b, the efficiency of EECS is more distinct when the network scale grows. In [5], the author mentions that the original LEACH outperformsHEED When based on the same assumptions in [4]which is identical with EECS. In order to save energyfurther, HEED adopts the multi-hop communicationamong the cluster heads during the inter-cluster com- munications in the data transmission phase. Noticethat we focus on the cluster set-up algorithm but notthe data transmission approach in our current work.Future work will consider the multi-hop technique inthe inter-cluster communication. Readers should refer to [12] for details about the multi-hop routing inclustered networks.

6 Conclusion and Future Work:

In this paper, we present a novel distributed, energyefficient and load balanced clustering scheme appliedfor periodical data gathering. EECS produces a uniform distribution of cluster heads across the networkthrough localized communication with little overhead.What's more, a novel approach has been introducedto distribute the energy consumption among the sensors in the cluster formation phase.

Simulation results show that EECS prolongs the network lifetime as much as 135% of LEACH and the total energy isefficiently consumed.All of our contributions here are focused on thecluster set-up stage. There are still much space to improve the performance of data transmission. In thelarge scale sensor networks, multihop communicationis a mainstream technique for energy saving. We willremove the assumption of singlehop and design an energy efficient protocol for both intra-cluster and intercluster data transmission in the future work.

References:

[1] D. Estrin, et. al., \Instrumenting the world withwireless sensor networks," in the International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2001.

[2] Akyildiz I.F., et. al.,\A Survey on Sensor Networks," IEEE Communications Magazine, Au- gust, pp. 102-114, 2002.

[3] B. Krishnamachari, et. al., \The impact of dataaggregation in wireless sensor networks," in the22nd International Conference on DistributedComputing Systems Workshops (ICDCSW'02).

[4] W. Heinzelman,et. al., \An application-specificprotocol architecture for wireless icrosensor networks," IEEE Transactions on Wireless Communications, 1(4):660-669, 2002.

[5] O. Younis, et. al., \HEED: A Hybrid, EnergyEfficient, Distributed Clustering Approach for Ad Hoc Sensor Networks," IEEE Transactions onMobile Computing, 3(4):660-669, 2004.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

[6] S. Lindsey, et. al.,\PEGASIS: Power-EffcientGathering in Sensor Information Systems," IEEEAerospace Conference Proceedings, Vol. 3, 9-16pp. 1125-1130, 2002.

[7] S. Bandyopadhyay, et. al., \An Energy-EfficientHierarchical Clustering Algorithm for WirelessSensor Networks," in IEEE INFOCOM'03.

[8] H. Chan, et. al., \ACE: An Emergent Algorithmfor Highly Uniform Cluster Formation," in theFirst European Workshop on Sensor Networks(EWSN), January 2004.

[9] J. Kamimura, et. al., \Energy-E±cient Clustering Method for Data Gathering in Sensor Net-works," in the Annual International Conferenceon Broadband Networks, October 2004.

[10] J. Hill, et. al., \System Architecture Directionsfor Networked Sensors", Architectural support forProgramming Languages and Operating Systems,pp. 93-104,2000.

[11] T. Rappaport, Wireless Communication: Principles & Practice. Prentice-Hall, 1996.

[12] V. Mhatre, et. al., \Design guidelines for wirelesssensor networks: communication, clustering andaggregation," Ad Hoc Networks Journal, Elsevier Science, 2(1):45-63, 2004.