

# **An Energy based Data Gathering Approach for Wireless Sensor Networks**

**N. Kaleeswari**

Research Scholar, Anna University of Technology,  
Coimbatore, India

**K. Baskaran, PhD.**

Associate Professor Department of CSE,  
Government College of Technology, Coimbatore,  
India

## **ABSTRACT**

In Wireless Sensor Networks (WSNs), sensor nodes are organized in a random manner. These sensor nodes are mainly deployed in physical environment to gather data and deliver the data to the destination wherever requires. Sensor nodes have limited computational capability, limited bandwidth, energy source and memory. The main challenge in data gathering is to maximize the network lifetime and the given energy constraints of sensor nodes. If energy consumption per node can be balanced as well, a near optimal data gathering and routing scheme can be achieved in terms of network lifetime. We have proposed the algorithm called Energy based Data Gathering Approach (EDGA) to make the correct balance between the three factors like residual energy, distance and data gathering ratio. By simulation results we have shown that the EDGA achieves good throughput, high network lifetime, high residual energy while attaining low delay than the existing schemes LEACH and EBRP.

## **General Terms**

Data Gathering, network life time.

## **Keywords**

WSN, EDGA, data gathering, network lifetime, delay, residual energy and throughput

## **1. INTRODUCTION**

Wireless sensor network (WSN) is a group of sensor nodes (SNs) working in uncontrolled areas and organized into cooperative network [1]. It is composed of huge number of sensor nodes which can monitor the environment by collecting, processing as well as transmitting collected data to the remote sink node through direct or multi hop transmission. WSNs have attracted lots of attention in recent years due to their wide applications such as battlefield surveillance, inventory and wildlife monitoring, smart home and healthcare etc.

Each node has processing capability, a radio, sensors, memory and a battery [2]. Since the sensor nodes are usually operated by a limited battery power which may not be replaceable once deployed, it is therefore, vital that the sensor network is energy balanced in order to ensure an extended network lifetime and efficient data gathering.

## **2. DESIGN GOALS OF WIRELESS SENSOR NETWORKS WSNs**

Based on the application, different architecture, goals and constraints have been considered for WSNs. As discussed in [3], the following design goals are given.

## **2.1 Network Dynamics**

There are three main components in a sensor network. These are the sensor nodes, sink and monitored events. Aside from the very few setups that utilize mobile sensors most of the network architectures assume that sensor nodes are stationary. On the other hand, supporting the mobility of sinks or cluster-heads (gateways) is sometimes deemed necessary. For instance, in a target detection/tracking application, the event (phenomenon) is dynamic whereas forest monitoring for early fire prevention is an example of static events. Monitoring static events allows the network to work in a reactive mode, simply generating traffic when reporting. Dynamic events in most applications require periodic reporting and consequently generate significant traffic to be routed to the sink.

## **2.2 Node Deployment**

This is application dependent and affects the performance of the routing protocol. The deployment is either deterministic or self-organizing. In deterministic situations, the sensors are manually placed and data is routed through pre-determined paths. However in self-organizing systems, the sensor nodes are scattered randomly creating an infrastructure in an ad hoc manner. In that infrastructure, the position of the sink or the cluster-head is also crucial in terms of energy efficiency and performance. When the distribution of nodes is not uniform, optimal clustering becomes a pressing issue to enable energy efficient network operation.

## **2.3 Energy Considerations**

During the creation of an infrastructure, the process of setting up the routes is greatly influenced by energy considerations. Since the transmission power of a wireless radio is proportional to distance squared or even higher order in the presence of obstacles, multihop routing will consume less energy than direct communication. However, multi-hop routing introduces significant overhead for topology management and medium access control. Direct routing would perform well enough if all the nodes were very close to the sink. Most of the time sensors are scattered randomly over an area of interest and multi-hop routing becomes unavoidable.

## **2.4 Data Delivery Models**

Depending on the application of the sensor network, the data delivery model to the sink can be continuous, event-driven, query-driven and hybrid. In the continuous delivery model, each sensor sends data periodically.

## **2.5 Data Aggregation/Fusion**

Since sensor nodes might generate significant redundant data, similar packets from multiple nodes can be aggregated so that the number of transmissions would be reduced. Data

aggregation is the combination of data from different sources by using functions such as suppression (eliminating duplicates), min, max and average. In some network architectures, all aggregation functions are assigned to more powerful and specialized nodes. Data aggregation is also feasible through signal processing techniques. In that case, it is referred as data fusion where a node is capable of producing a more accurate signal by reducing the noise.

## 2.6 Data Gathering

Data gathering is a common function of sensor networks, where information sampled at sensor nodes needs to be transported to central base stations for further processing and analysis. In view of the severe energy constraints of sensor nodes and the limited transport capacity of multihop wireless networks, an important topic addressed by wireless sensor networks community has been in-network data aggregation. The idea is to pre-process sensor data in the network by sensor nodes endowed with computational power, so as to reduce expensive data transmission. Wireless sensor network where  $n$  sensors are randomly deployed in a finite geographical region. Each sensor measures independent field values at regular time intervals and sends these values to sinks. The union of all sensing values from  $n$  sensors at particular time is called a snapshot. The task of data collection is to deliver these snapshots to the sinks. Due to spatial separation, several sensors can successfully transmit at the same time if these transmissions do not cause any destructive wireless interferences.

## 2.7. Problem Statement

Some of the existing algorithm on data gathering implicitly assumes routing techniques similar to those in wireline networks, neglecting the characteristics of wireless transmission. On the one hand, wireless transmission is error-prone. Sequential forwarding of packets along a fixed path may incur many retransmissions, and thus exhaust scarce network resources such as energy and capacity. On the other hand, wireless transmission is broadcast in nature. The chance that all the neighboring nodes fail to receive the packet is small (multiuser diversity in packet reception). Moreover, multiple receptions of a packet by different nodes can also be exploited for opportunistic data compression. By leveraging the wireless broadcast advantage and multiuser diversity, we can reduce the number of wireless transmissions needed for data gathering.

## 3. PREVIOUS WORK

Jim chou et al [4] have demonstrated a new approach to decrease energy consumption in sensor networks using a distributed adaptive signal processing framework. Here the sensor nodes compress the readings with respect to one another without the requirement of explicit and energy expensive to achieve the compression. But the communication from the destination node to the source nodes representing the number of bits required turn into very high if the correlation in sensor data is varying in nature.

Honguang et.al [5] proposed rate assignment method to extend the lifetime of WSN. The research's development has been progressed in the area of data aggregation. But many of them concentrated in lossless algorithms for wireless sensor networks. The main objective is to design simple prediction algorithm which predicts the sensor data and whenever the predicted value is more than the threshold value the data is transmitted using adaptive compression algorithm. The

compression algorithm adapts to changing correlation in sensor data.

W. R. Heinzelman et.al [6] proposed LEACH algorithm which exploited the randomized rotation of cluster heads for distributing evenly the energy load. From the results LEACH can achieve as much as a factor of 8 reductions in energy dissipation compared with conventional routing protocols. In order to achieve a better balance of energy consumption, the cluster-head selection algorithm must be performed often, which will include the excessive processing and communication overheads to the network.

S. Lindsey [7] demonstrated the method called PEGASIS to reduce the number of sensor nodes communicating directly with the base station to one by forming a chain. This chain is formed by passing through all sensor nodes where each sensor node receives from and transmits to the closest possible neighbor node. The data is gathered starting from each endpoint of the chain until the randomized head-node is reached. The data is fused each time it moves from node to node. The designated head-node is responsible for transmitting the final data to the base station. From the results they have shown that PEGASIS achieves a better performance than LEACH by between 100% and 300% in terms of network lifetime.

Efthymiou et al.[8] mentioned the difficulty of energy balanced data propagation in WSNs. Here, the sensor nodes are assumed to be uniformly deployed in a fan-shaped or circular area and all the sensor nodes are assumed to have the similar packet generation rate. The energy balance is totally achieved by controlling the fraction of packets for direct transmission and hop-by hop transmission.

D.C. Hoang, and R. Kumar, [9] proposed the Harmony Search Algorithm which is the Clustering technique. It is one of the techniques utilized to extend lifetime of the network by applying data aggregation and balancing energy consumption in the middle of sensor nodes of the network. It is used for minimizing the intra-cluster distance and optimizing the energy consumption of the network. The algorithm is music based meta heuristic optimization method which is analogous with the music improvisation process. The proposed algorithm is used to reduce energy consumption and improve the network lifetime.

Lingyun Yuan et al [10] explored the Multi-Layer Energy-Efficient And Delay-Reducing Chain-Based Data Gathering Protocol. This protocol mainly puts forward the idea of multi-layer chain and also uses the minimum total energy algorithm to construct the chain. Including this, the maximum residual energy of sensor nodes is the standard for selection of leaders. But this method can not only prolong the network lifetime, but also reduce the network delay remarkably.

Yi-hua Zhu et.al [11] demonstrated the Energy-efficient Routing Algorithm to Prolong Lifetime (ERAPL) which is able to dramatically prolong network lifetime while efficiently expends energy. A data gathering sequence (DGS) is used to avoid mutual transmission and loop transmission among sensor nodes. Each sensor node proportionally transmits traffic to the links confined in the DGS. A mathematical programming model is also used here to achieve minimal remaining energy of nodes and to optimize network lifetime. Furthermore some of the genetic algorithms are also used to find the optimal solution of the proposed programming

problem. ERAPL outperforms them in terms of network lifetime.

Ying Liang [12] proposed Novel clustering algorithm (OCABTR) which fully consider the characteristic of target occurrence in monitor area. The energy is adopted the strategy of forming cluster first and selecting cluster-head afterward. The cluster is shaped by genetic algorithm to optimize and partition the adjacent sensor nodes which will sense similar target into one cluster. This approach can effectively reduce redundant data transmission by improving the rate of data aggregation in clusters. Here the main operation of OCABTR is divided into rounds. Each round starts with a set-up phase when the clusters are organized and followed by a steady-state phase when data are transferred from the sensor nodes to the cluster head and on to the Base Station (BS). This algorithm significantly outperforms previous methods, in terms of system lifetime.

Jun Yang, et al.[13] proposed the Energy-Efficient Data Gathering Protocol (EEDGP). This protocol consists of a clustering method of balancing energy consumption, a data prediction transmission strategy and an energy-aware multi hop routing algorithm. In clustering method, the initial probability of node for cluster head election is derived from mathematical relation between application's seamless coverage fraction and numbers of required cluster heads. In data aggregation phase, the spatial correlation of data within a cluster is utilized by cluster head to aggregate sampling data. In order to alleviate the hot spot problem among cluster heads, a greedy geographic and energy-aware multi hop routing algorithm is presented for inter-cluster communication. EEDGP outperforms in terms of network lifetime by balancing energy consumption and decrease of transmission while meeting desired application-specific requirements.

S.Madden et.al [14] proposed TAG (Tiny Aggregation) approach where each epoch separates time slots and these time slots are used to specify the different levels of routing tree in reversal form. In this method, each node depends on its situation in the tree, and in its related time slot will send its data. The node synchronization of this approach is used for sending and receiving data could effectively reduce the average energy consumption.

The paper is organized as follows. The Section 1 describes introduction about WSNs, data gathering and issues. Section 2 deals with the previous work which is related to the data gathering algorithms. Section 3 is devoted for the implementation of proposed algorithm. Section 4 describes the performance analysis and the last section concludes the work.

## 4. IMPLEMENTATION OF PROPOSED ALGORITHM

Approach (EDGA) is used. Here, we consider whole network consists of several clusters with the same size. Cluster means group of sensor node. Here each node is connected to the sub tree and each sub tree contains together to form a cluster. Each root node of the sub tree is considered as the head node of the cluster. Each cluster has a Cluster-Head (CH). The clustering operation and the selection of cluster-head (CH) can be done by using any existing protocol LEACH. Fig 1 illustrates the cluster formation.

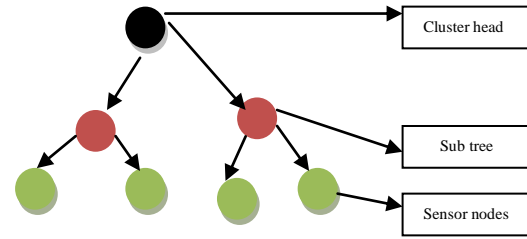


Fig.1. Cluster Formation and Selection of Cluster Head (CH)

### 4.1 Data Compression

Here the data compression approach is used with conditional entropy. The goal of data compression is removing the data over-head of sensors and reducing the correlations of data for achieving beneficial information for the base station.

A network consists of  $N$  nodes ( $X_1, X_2, \dots, X_N$ ) in which the data production of each node is equal to  $H(X_n) = H_1 \{n = 1, 2, \dots, N\}$ . The minimum number of bits is compressed by means of conditional entropy,

$$H(X_m / X_n) = \left[ 1 - \frac{1}{\frac{k}{p} + 1} \right] H_1 \{n = 1, 2, \dots, N\}$$

In this equation  $p$  is a constant which presents the amount of data correlation and  $k$  is the distance between nodes  $X_m$  and  $X_n$ .

### 4.1.2 Steps involved in the proposed algorithm are as follows

#### Step 1:

The source node establishing the route by flooding the network with a broadcast message. The message consists of neighbor id, node's position, current energy level and hop count. The selection probability of neighbor nodes are computed and inserted into the Neighbor List.

#### Step 2:

The root node of cluster broadcasts the information packet to its neighbor nodes. When a sensor node receives the information packet, it considers the sender as one of its possible parents and stores its information. After that the node's position, current energy and data label fields of the packets are updated which increments the hop count and transmit the packet to its neighbors. This process will be done until all the sensor nodes in the cluster receive the information packet.

#### Step3:

If the entire nodes received the proposed information packet, each sensor node selects its parent which should send its data to it. This selection process will be done based on the following conditions given below:

- Choosing the sensor node among the possible parent which has the least hop distance from the cluster head. This node should be the closest node to cluster head.
- If there is more than one sensor node with the least hop distance, the sensor node which has the most current energy and the least data correlation will be selected as the parent.

From the above conditions

- The shortest path from a node to cluster head can be selected.
- The network lifetime can be decreased by participating most durable nodes.
- The networks overhead can be reduced by checking the data correlation of the nodes.

**Step 4:**

When a node observes new data it initiates the process of routing. If multiple neighbors have smaller hop counts, the neighbor with the highest remaining energy is selected as the next hop. If a node does not have a neighbor with a smaller hop count, it selects the neighbor with the highest remaining energy from neighbors with an equal hop count to it. If the node does not have a neighbor with an equal hop count to it, the message is discarded. But selecting next hop in proposed algorithm, a sensor node searches into its neighbor table for the neighbors with highest probability. The energy level of data sender node is attached to original data packet. Then the data packet is forwarded to neighbor with highest probability.

**Step 5:**

All neighbors of data sender node receive the forwarded data packet, by overhearing technique. They only update the remaining energy value in the neighbor table for the neighbor that sent the data packet.

The operation of the proposed algorithm can be shortened as follows:

- The source node establishing the route by flooding the network with a broadcast message.
- All sensor nodes append their entire neighbor’s information to their neighbor tables.
- The sensor node is selected based on the highest probability and data packet is forwarded to it.
- Neighbor nodes update the energy level of data sender node is updated by its.
- Neighbor List of previous data sender node in each hop contains the probability of data sender node.

**4.1.3 Energy Model**

The energy model of proposed algorithm is given below. In this model energy consumption for transmitting M bit is equal to:

$$E_{tr}(M, d) = E_{elec} \times M + \delta_{amp} \times M \times d^2 - E_{wast}(P_{drop})$$

M = bit contain some information like current energy level of the node, data label, node’s location and hop count.

$E_{elec}$  = Energy to be Transmitted and Received electronic device module (75 nJ/bit).

$\delta_{amp}$  = Transmitter Amplifier (150 pJ/bit/m<sup>2</sup>)

d = distance between the two nodes.

$E_{wast}(P_{drop})$  = Energy wasted on packet dropping.

And the energy for receiving K bit is equal to:

$$E_{rr} = E_{elec} \times M$$

**4.1.4 Proposed Packet Format**

The proposed packet format of EDGA is given below;

Node’s Position	Present Energy	Hop Count	Neighbor Id	Data Label	CRC	MAC
2	1	1	2	1	1	4

**Fig. 2. Proposed packet format**

In fig 2, the packet format of proposed algorithm is shown. Here the first field occupies 2 bytes where each sensor node

should determine now it location in prior. The second field occupies 1 byte field where the remaining energy of a node. In next field, hop count fills 1 byte where it calculated the number of hops from cluster head. Neighbor id occupies 2 byte field. Data label occupies 1 byte field where the data value which is sensed by a node. Cyclic Redundancy Check (CRC) occupies 1 byte field for error correction and detection. Finally the MAC (Message Authentication Code) occupies 4 byte field for authenticating message

**5. PERFORMANCE ANALYSIS**

We use Network Simulator (NS2) to simulate our proposed algorithm. In our simulation, 200 mobile nodes move in a 1200 meter x 1200 meter square region for 60 seconds simulation time. All nodes have the same transmission range of 250 meters. Our simulation settings and parameters are summarized in table 1.

**Table.1.Simulation Settings and Parameters**

No. of Nodes	200
Area Size	1200 X 1200
Mac	802.11
Radio Range	250m
Simulation Time	60 sec
Traffic Source	CBR
Packet Size	512 bytes
Mobility Model	Random Way Point
Transmitter Amplifier	150 pJ/bit/m <sup>2</sup>
Package rate	5 pkt/s
Protocol	DSR

**5.1 Performance Metrics**

We evaluate mainly the performance according to the following metrics.

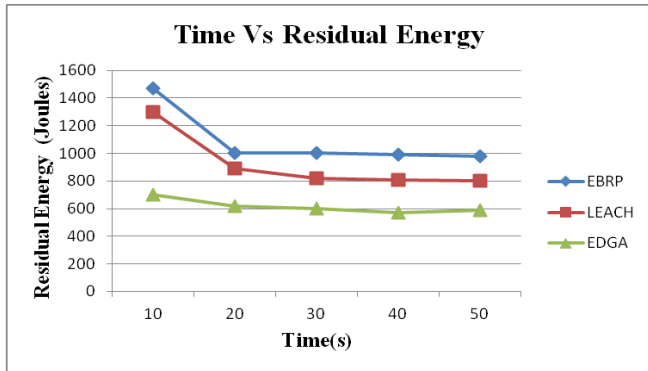
**End-to-end delay:** The end-to-end-delay is averaged over all surviving data packets from the sources to the destinations.

**Data Availability Ratio:** It is defined as the making the copies of data items which shared by several users in a particular point of time.

**Throughput:** It is defined as the number of packets received at a particular point of time

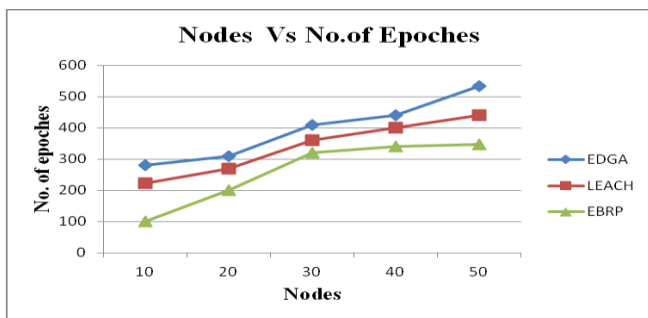
The simulation results are presented in the next part. We compare our proposed algorithm (EDGA) with EBRP [15] and LEACH in presence of clustering environment.

Figure 3 shows the results of average residual energy for varying the time from 10 to 50. From the results, we can see that EDGA scheme has lower residual energy than the LEACH and EBRP scheme



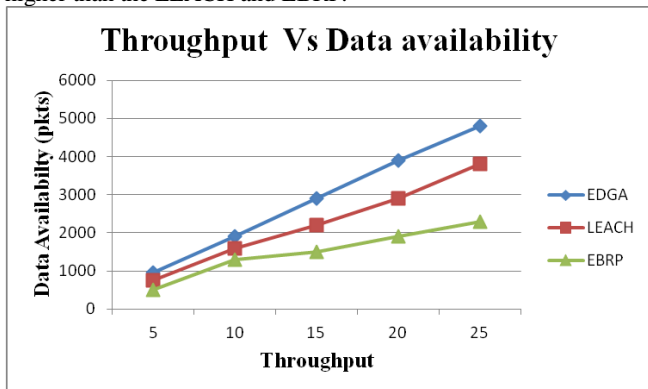
**Fig. 3. Time Vs Residual Energy**

Fig. 4, presents the network lifetime comparison for EDGA, LEACH, EBRP. It is clearly seen that number of epochs consumed by EDGA is high compared to LEACH and EBRP.



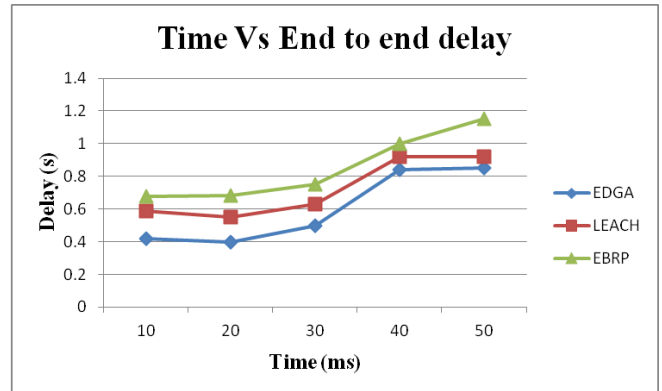
**Fig. 4. Increasing the network lifetime**

Fig. 5, presents the comparison of data availability. It is clearly shown that the data availability ratio of EDGA is higher than the LEACH and EBRP.



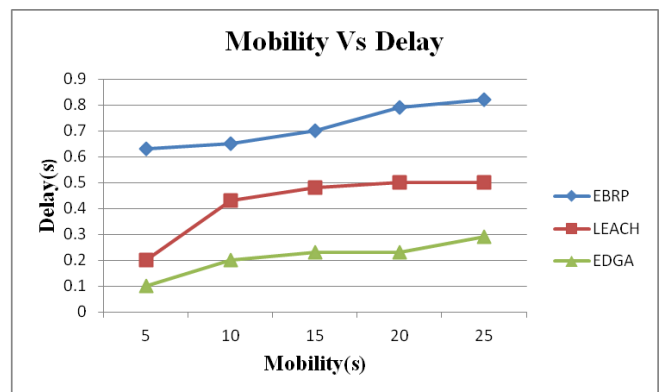
**Fig. 5. Throughput Vs Data Availability**

Figure 6 shows the results of Time Vs End to end delay. From the results, we can see that EDGA scheme has slightly lower delay than the EBRP and LEACH scheme because of authentication routines.



**Fig. 6. Time Vs End to end delay**

Fig. 7, presents the comparison of delay while varying the mobility from 5 to 25. It is clearly shown that the delay of EDGA is lower than the EBRP and LEACH.



**Fig. 7. Mobility Vs Delay**

## 6. CONCLUSION

In WSNs, the best route is being determined by choosing efficient strategy to forward the data to the base station. Due to that, the node consumes more energy unnecessarily. In this paper, we have developed a Energy based Data Gathering Approach which attains energy model which attains high network life time and throughput to the sensor nodes. In the first phase of the scheme, network lifetime is increased. In second phase, residual energy consumption is increased using energy model. It uses following factors called distance , residual energy, mobility factor, mobility factor and data correlation to favor packet forwarding by maintaining high residual energy consumption for each node. We have demonstrated the energy estimation of each node. By simulation results we have shown that the EDGA achieves good throughput, high network lifetime, high residual energy while attaining low delay than the existing schemes LEACH and EBRP while varying the number of nodes, time, node throughput and mobility.

## 7. REFERENCES

- [1] Akyildiz IF, Su W, Sankarasubramaniam Y, Cayirci E, “ Wireless sensor networks: a survey”. Computer Networks, 38(4): 393-422., 2002.
- [2] Baronti P, Pillai P, Chook VWC, Chessa S, Gotta A, Hu YF , “Wireless sensor networks: A survey on the state of the art and the 802.15.4 and ZigBee standards”, Computer Communication, 30(7): pp. 1655- 1695, 2007.

- [3] Kemal Akkaya and Mohamed Younis , “A survey on routing protocols for wireless sensor networks” Elsevier, Ad hoc Networks, 2005, pp.325-349.
- [4] Jim Chou, D Petrovic and K Ramachandran,2003. “A distributed and Adaptive Signal processing approach to reducing energy consumption in sensor network”, Proceedings, IEEE Infocom’03, march 2003.
- [5] HonguangWang, Dangming Rend, Weiwang and Hamid Sharif, 2007. “A Rate-Oriented Routing Scheme to support Distributed Source Coding in Wireless Sensor Network”, Information Technology Journal 6(5) pp156 - 760.
- [6] W.R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energyefficient communication protocol for wireless microsensor networks,” in 33rd Annual Hawaii International Conference on System Sciences,2000, pp. 3005 – 3014.
- [7] S. Lindsey and C. S. Raghavendra, “Pegasis: Power-efficient gathering in sensor information systems,” in IEEE Aerospace Conference, March 2002.
- [8] Charilaos Efthymiou et.al , “Energy Balanced Data Propagation inWireless Sensor Networks “IST Programme of the European Union, pp.1-8, 1999.
- [9] Hoang, D.C. Yadav, P. Kumar, R. Panda, S.K, ”A Robust Harmony Search Algorithm Based Clustering Protocol for Wireless Sensor Networks”, Communications Workshops (ICC), 2010 IEEE International Conference on ,Issue Date : 23-27 May 2010, page(s): 1 – 5.
- [10] Lingyun Yuan Yunlong Zhu Tianwei Xu, “A Multi-Layered Energy-Efficient and Delay-Reduced Chain-Based Data Gathering Protocol for Wireless Sensor Network”, Mechtronic and Embedded Systems and Applications, 2008. MESA 2008. IEEE/ASME International Conference on Issue Date : 12- 15 Oct. 2008 On page(s): 13 – 18.
- [11] Yi-hua Zhu , Wan-deng Wu, Jian Pan , Yi- ping Tang, “An energy-efficient data gathering algorithm to prolong lifetime of wireless sensor networks”, Original Research Article,Computer Communications, Volume 33, Issue 5, 15March2010,Pages639-647.
- [12] Ying Liang, “An energy-efficient clustering algorithm for data gathering and aggregation in sensor networks”, Industrial Electronics and Applications, 2009. ICIEA 2009. 4th IEEE Conference on 25-27 May2009, pp.3935-3939.
- [13] Soyoung Hwang Gwang-Ja Jin Changsub Shin Bongsoo Kim,” Energy-Aware Data Gathering in Wireless Sensor Networks”,Consumer Communications and Networking Conference, 2009. CCNC 2009. 6th IEEE Issue Date : 10-13Jan.2009 On page(s): 1 – 4
- [14] S. Madden, M. J. Franklin, J. M. Hellerstein and W. Hong, “TAG: A Tiny AGregation Service for Ad-Hoc

## 8. AUTHORS PROFILE

**Dr. K.Baskaran** received his Bachelor of Engineering degree in Electrical and Electronics Engineering from Annamalai University, India in 1989, Master of Engineering degree in Computer Science Engineering from Bharathiar University, India in 2002 and Ph.D degree from Anna University-Chennai, India in 2006.He is a member of IEEE and member of ISTE. He has about 22 years of teaching experience, since 1990.He is now Associate Professor in Computer Science and Engineering, Government College of Technology, Coimbatore, India.

**N.Kaleeswari** received her B.Sc degree in Electronics from Bharathiar University, India in 2000, M.Sc degree in Applied Electronics from Bharathiar University, India in 2002, M.E. degree in Electronics and Communication Engineering (Applied Electronics) from Anna University of Technology – Coimbatore, India in 2010 and she is currently pursuing full time Ph.D. degree in Electronics and Communication Engineering from Anna University of Technology – Coimbatore, India. Her research interests include Wireless Sensor Networks, System Networking, Digital Communication, Adhoc Networks and Pervasive Computing.