2015 IEEE Power, Communication and Information Technology Conference (PCITC)
Siksha 'O' Anusandhan University, Bhubaneswar, India.

An Enhancement of PEGASIS Protocol with Improved Network Lifetime for Wireless Sensor Networks

Alekha Kumar Mishra
Department of MCA
Silicon Institute of Technology
Bhubaneswar, India
alekha@silicon.ac.in

Rukshan Ur Rahman, Rahul Bharadwaj* and Rohit Sharma[Ψ]
Department of Computer Science and Engineering
Silicon Institute of Technology
Bhubaneswar, India
rukshanurrahman@gmail.com,
*rahulbharadwaj855@gmail.com,
[Ψ]rohit.sharma06596@outlook.com

Abstract— **Major Wireless Sensor Network designs are focused on improving the network lifetime. The data aggregation process is the simplest way to achieve this. The PEGASIS routing protocol is a promising approach for efficient data aggregation in Wireless Sensor Networks. It is also reported in the literature that the PEGASIS performs better than LEACH and HEED protocol towards improving the network lifetime. However, in PEGASIS, the selection of a leader located far away from the base station still remains exposed to higher energy consumption for sending aggregated data. In this paper, we propose an enhancement to the chain formation process of PEGASIS to further improve its lifetime, and called it PEGASIS with Improved Network Lifetime (PEGASIS-INL). In contrast to original PEGASIS protocol, the PEGASIS-INL selects a node as a leader only if it is within a strong communication range of base station. A multiple overlapped chain formation and the aggregation technique is used for transmission of sensed information to the base station. Experimental results demonstrate that the PEGASIS-INL algorithm outperforms PEGASIS in terms of the product of consumed energy and delay, weighting the overall performance of both the energy consumption and the network lifetime.**

Keywords- Wireless Sensor Networks, Data Aggregation, PEGASIS, Chain Formation, Leader Selection, Network Lifetime

I. Introduction

Wireless Sensor Networks (WSN) [1] consist of spatially distributed sensor nodes that are used to monitor physical phenomena, such as temperature, sound, pressure, vibration etc. The sensor nodes cooperatively forward their data to the sink often called as the base station (BS). Recent technological developments are able to produce inexpensive sensors that are relatively smaller in size and have advanced sensing, processing and communication capabilities. Despite of advances in all units of a sensor node, the energy still remains a limiting factor. Since, it has been observed that the energy consumption of communication unit is significantly higher than rest of the units of a sensor node. Therefore, most of the communication mechanisms in WSNs aim to minimize power consumption. The WSN supports a wide range of applications. Some of these applications include enemy troop movement tracking, traffic monitoring, wildlife monitoring, structure monitoring, and soil moisture and fertility monitoring in agricultural applications [2].

In WSN, the foremost responsibility of a sensor is to gather data from the environment, and collaborate with other sensors to route them to the base station. To deliver accurate data it is required that all the data that are collected by sensor nodes reaches the BS. To improve data delivery, fault tolerance and sensing quality, one solution is to increase data redundancy by deploying sensors with high density. However, the redundancy increases the energy consumption in the network. Researchers found that energy consumption can be reduced to a significant level by aggregating sensed data using various mathematical functions without losing its characteristics. This mechanism helps to provide only relevant information to the BS instead of flooding raw sensed data.

A number of data aggregation techniques are available in literature [3]. The most popular routing algorithms that support data aggregation are Low Energy Adaptive Clustering Hierarchy Protocol (LEACH) [4], Hybrid Energy Efficient Distributed Clustering Protocol (HEED) [5], and Power-Efficient GAthering in Sensor Information Systems (PEGASIS) [6]. The mentioned algorithms aim to provide energy efficient data aggregation mechanism with higher performance. The key idea of PEGASIS is to form a chain among the sensor nodes so that each node will communicate to a close neighbor. In PEGASIS, gathered data moves from node to node, get fused, and eventually a node designated as leader transmits it to the BS. The PEGASIS protocol is energy efficient compared to LEACH. A number of variations of PEGASIS are proposed to improve its efficiency. However, some issues corresponding to data aggregation still remains the same. We draw our attention to the following issues in PEGASIS protocol:

i. It is assumed that all nodes have global knowledge of the network. However, this assumption may not work in a real world scenario, where sensors are deployed in unknown environment.
ii. There is no specific mechanism for selection of farthest node in the network.
iii. Leaders that are located at far away from the BS will dissipate a huge amount of energy to send fused data to the BS in compared to closer nodes.
iv. The random position of a leader in the network would significantly increase the end-to-end delay.

In this paper, we propose a variation of the PEGASIS protocol called PEGASIS with Improved Network Lifetime (PEGASIS-INL). The proposed mechanism aims to address the above issues in WSN. The contributions are enlisted as follows: First, an energy-efficient mechanism to determine the nodes that are eligible to become a leader in the network; secondly, an improved greedy concurrent chain formation mechanism; and finally, the data aggregation process to transmit data to BS. The proposed scheme is energy efficient and has improved network lifetime. The performance of the proposed mechanism is compared with PEGASIS using parameters such as energy consumption, and network lifetime.

Rest of the paper is organized as follows: Section II briefly explains existing protocols for data aggregation that are reported in the literature. Assumptions about the network are provided in Section III. Section IV provides the energy model for energy-efficiency analysis. Section V elaborates the proposed variation of PEGASIS protocol. Analysis of proposed mechanism is made in Section VI. Section VII shows the result obtained through simulation and a few conclusions are made in Section VIII.

II. RELATED WORKS

There have been many efforts made toward efficient data aggregation in WSNs. In this section, we will briefly discuss the existing schemes for efficient data aggregation and routing that are reported in the literature.

Heinzelman *et al.* Proposed a Low Energy Adaptive Clustering Hierarchy Protocol, well known as LEACH protocol [4]. LEACH is a self-adaptive and self-organized hierarchical routing protocol. LEACH protocol uses round as unit, each round is made up of two stages: i) Cluster set-up stage, and ii) Steady-state stage. The nodes transmit data to cluster heads, and the cluster heads aggregate and compress the data and forward it to the BS. LEACH assumes that each node has a radio powerful enough to directly reach the base station or the nearest cluster head. In LEACH, a cluster-head is elected periodically to balance energy consumption. However, LEACH has a disadvantage of voting process that spends additional energy and incurs network traffic overhead.

Younis *et al.* proposed a Hybrid Energy Efficient Distributed Clustering (HEED) protocol [5]. HEED uses residual energy as primary parameter and node degree and distance to neighbors as secondary parameters for cluster-head selection. It extends LEACH protocol to support multi-power level sensor nodes in the network. The energy dissipated in cluster formation is lesser in HEED compared to LEACH.

Lindsey *et al.* proposed a chain-based protocol called Power-Efficient GAthering in Sensor Information Systems (PEGASIS). The key idea of PEGASIS is to form a chain among the sensor nodes so that each node will receive from and transmit to a closer neighbor. Gathered data moves from node to node, get fused, and eventually a designated node transmits to the BS. Nodes take turns transmitting to the BS so that the average energy spent by each node per round is reduced. It is found that PEGASIS performs better than LEACH.

A number of variations of PEGASIS have been proposed to improve the performance of PEGASIS. We briefly discuss a few of these schemes as follows:

Jung *et al.* proposed an enhanced PEGASIS protocol based on the concentric clustering scheme [7]. The concentric clustering scheme is used to consider the location of the base station to provide elongated lifetime of the WSN. This scheme helps to avoid the data transmission attended with the redundant energy consumption. However, the concentric cluster scheme is feasible only for specific network topologies.

Guo *et al.* used an improved ant colony algorithm for chain formation in PEGASIS [8]. The proposed scheme forms a chain that makes the path more even-distributed and the total square of transmission distance much lesser. Sensor's energy is taken into consideration while selecting leader in each round of transmission.

Linping *et al.* used the concept of using double cluster heads for PEGASIS in their proposed scheme [9]. The scheme divides the network into different data levels. Two cluster-heads are named as primary and secondary cluster-heads. The primary cluster-heads are responsible for a particular data level that receives data from the chain, and transmits the fused data to secondary cluster-heads. The secondary cluster-heads are in-charge of transmitting the lower-level data from primary to upper-level cluster-head.

Sen *et al.* proposed an Improved Energy-Efficient PEGASIS-Based protocol (IEEPB) [10]. In their scheme, chain-building method avoids the formation of long links between neighboring nodes by comparing the distance twice and finding the shortest path to link two adjacent nodes. The leader of each round is selected by considering the normalized value of nodes energy and distance between node and BS. The scheme integrates PEGASIS with intersection-based coverage algorithm to reduce the energy consumption in the network.

Aliouat *et al.* proposed a multi-hops variation of PEGASIS [11], which is intended to minimize the energy consumption and extend the network life time. In this variation, each round consists of two phases: i) Initialization phase, and ii) Transmission Phase. The algorithm emphasizes on deciding the number of cluster heads. The improvement over performance is achieved by the use of the inter-clusters multi-hops routing leading to the BS.

Most of the above proposed schemes adopt clustering techniques. It has been observed that the selection process of cluster-head incurs additional communication overhead and computation task for each round. Moreover, adopting techniques such as ant colony algorithm is also not suitable for energy constraint network like WSN. In the proposed scheme, we intend to propose a simple leader selection mechanism and reduce the distance of leader from BS by allowing only a set of eligible nodes to become leader. In subsequent sections, we discuss the proposed scheme in detail.

III. NETWORK ASSUMPTIONS

The following assumptions are made about the network: i) Nodes are static and deployed at random, ii) The location of a node is unknown prior to deployment, iii) The sensor nodes do not have global knowledge about the network such as location of all nodes, iv) BS is considered to be more powerful than sensor nodes, and v) The communication among the nodes and with the BS is bidirectional.

IV. RADIO ENERGY MODEL

The major energy consuming task involves in a data aggregation technique is to transmit and receive the data. Therefore, we use the following formal radio energy consumption model as mentioned in [6], [7] to define the energy model for radio unit. The amount of energy a radio expends to transmit *k*-bit message to a distance *d* is given by

$$\begin{aligned} E_T(k, d) &= E_{Tx\_elec}(k) + E_{Tx\_amp}(k) \\ &= E_{elec} \times k + \varepsilon_{amp} \times k \times d^2 \end{aligned} \quad (1)$$

where, E_{elec} and ε_{amp} is the energy consumed by transmitter electronics and amplifier respectively to transmit 1 bit data. The amount of energy a radio expends to receive the same message is given by

$$\begin{aligned} E_T(k, d) &= E_{Rx\_elec}(k) \\ &= E_{elec} \times k \end{aligned} \quad (2)$$

V. PROPOSED WORK

In this section, we describe the proposed PEGASIS-INL in details. The proposed mechanism is divided into three phases: *A)* Candidate Leader Selection Phase, *B)* Chain Formation Phase, and *C)* Data aggregation and transmission phase. All the above phases are discussed in the subsequent sub-sections

A. Candidate Leader Selection Phase

In PEGASIS, a leader node is responsible for transmitting the aggregated data to the BS in each round of data acquisition. According to PEGASIS, each node in the network will become a leader in its turn. This concept equally distributes the data transmission responsibility among all the nodes in the network. However, a leader that is located at a far distance from the BS will dissipate a huge amount of energy to transmit data to the BS. A node located nearer to the BS may require significantly lesser amount of energy to perform the same task. In PEGASIS-INL, not all nodes are eligible to become a leader in the network. The sensor node that can become a leader node is selected using the following steps. Initially, BS sends an echo broadcast to all the nodes in the network. Since, it is assumed that the BS has high range transmission capability it is assumed that it can broadcast message to all the nodes in the network. The sensor nodes can measure the signal strength of the broadcast using RSSI. If a node finds the measured RSSI value of the broadcast message is higher than a predefined threshold power level, then it becomes eligible to be chosen as a leader in the data aggregation process. In other words, the nodes under this category are the candidates for becoming leader in each round.

B. Chain Formation Phase

The chain formation process of PEGASIS-INL is different from the chain formation of PEGASIS in the following way. In PEGASIS, the chain formation starts from the farthest node from BS that chooses the closer neighbor using greedy approach. However, in PEGASIS-INL, the multiple-overlapped chains are formed with candidate leaders as root. The multiple chain formation for each candidate leader helps to achieve faster aggregation and forwarding of data. The detail process of chain formation is discussed below. Every candidate leaders broadcast a candidature message in the network bearing its ID, Time-stamp, a hop-count, and time to leave field. On receiving a candidature message of the leader l_i, a node creates a route entry for the candidate in a route vector if not present. Otherwise, it updates route entry if the current route has lower number of hops to the candidate node. The node increments the hop-count field and decrement the time to leave field before forwarding the candidature message in the network. As a result, each node has a shortest-route to every candidate nodes in the network in its route-vector. In other words, a set of overlapped chain is formed for each

C. Data aggregation and transmission phase

A leader node is selected for each round of data transmission. The selection of leader in each round is done in the following way: each node sort the route-vector using leader IDs as a key. The candidate with *index* = R_i *% route_vector.size()* becomes leader for round R_i. Here, *route_vector.size()* is the size of route-vector. Since, every node can find the current leader and the route to it using their route vector. Therefore, PEGASIS-INL skips the process of token passing prior to data aggregation. Each node then uses route from its route-vector to send the data to leader node ID_i. Each intermediate node fuse the data received from the nodes of lower layer with its own before sending to the leader node. Intermediate nodes that common to more than one chain are called the merging points. A merging point waits for a fixed period of time to collect and aggregate data from the nodes of lower layer before sending to the leader nodes. When, the aggregated data reaches the leader node, it transmits to the BS.

D. Reconstruction of Chain on Node Failure

In PEGASIS-INL, whenever a node dies, all candidate leaders reconstruct chain using the process mentioned in Section V-B. However, when a candidate leader node dies all remaining nodes in the network simply discard the route-entry to this node from their route-vector

VI. ANALYSIS

A. Energy Consumption

In this section, we compare the energy consumption between PEGASIS and proposed scheme. We consider the energy model discussed in Section IV for comparison.

1) Energy consumption during data transmission:

In PEGASIS, the energy consumed by N nodes to transmit k-bit data to the leader node using chain [6], [7] is given by

$$E_{TX} = N \times E_{elec} \times k + \varepsilon_{amp} \times k \times \sum_{m=1}^{N} d^2_{(m,m-1)} \quad (3)$$

where, $d_{(i,j)}$ is the distance from node i to j. In PEGASIS-INL, this value remains same since each node transmit data to its neighbor close to the current leader.

2) Energy consumption during control token passing:

The control token used by leader in PEGASIS is a simply a control message without data. Therefore, the energy consumed by N nodes to transmit token to its neighboring node is

$$E_{token} = N \times E_{elec} \times k_{token} + \varepsilon_{amp} \times k_{token} \times \sum_{m=1}^{N} d^2_{(m,m-1)} \quad (4)$$

Where, k_{token} is the size of the token in number of bits. In proposed scheme, there is no cost associated with PEGASIS-INL to transmit token during data aggregation because each node maintains a route-vector for each candidate leader in the network.

3) Energy consumption during data reception:

For both PEGASIS and PEGASIS-INL, the energy consumed by N node to receive k-bit of data from its neighbor is

$$E_{Rx} = N \times E_{elec} \times k \quad (5)$$

4) Energy consumption during data aggregation:

In PEGASIS and proposed scheme every node aggregate the received data from its neighbor. Therefore, the energy consumed for data aggregation in the network is given by

$$E_A = N \times E_{agg} \quad (6)$$

where, E_{agg} be the energy spent by a node for data aggregation.

5) Energy consumption during data transmission by a leader node:

$$E_{TX\_BS} = E_{elec} \times k + \varepsilon_{amp} \times k \times d^2_{(li,BS)} \quad (7)$$

where, l_i is the leader node of round i. In PEGASIS, $l_i \in \{0, 1, 2, . . . , N\}$, whereas in PEGASIS-INL, $l_i \in S_{cand}$, such that $S_{cand} \subset \{0, 1, 2, . . . , N\}$. S_{cand} contains only nodes that are very close to the BS in comparison to rest of the node. Therefore, the average distance of leaders from the BS in PEGASIS-INL is quite less than the average distance of leaders in PEGASIS. This implies that E_{TX_BS} of PEGASIS will be significantly higher than the E_{TX_BS} of PEGASIS-INL. This shows that PEGASIS-INL is energy-efficient compared to PEGASIS.

B. End-to-End Delay of aggregated data transmission

In PEGASIS, the fused data flows from a node to the leader across the chain. The length of the chain is N. Therefore, average path length that an aggregated data packet travels to reach the leader node is $\frac{N}{2}$. However, in PEGASIS-INL an aggregated packet of a node follows the path with minimum number of hops to the leader nodes. In a distributed network of size N, the average length of a chain or path is $\sqrt{N}$ [12]. Therefore, the average path length travelled by an aggregated data in PEGASIS-INL is lesser compared to PEGASIS.

C. Memory Requirements

In proposed scheme each node maintains a route-vector in its cache. The size of the route-vector is equal to number of candidate leaders in the network. Higher the number of candidate leaders, higher is the size of route-vector. This is only possible when majority of the sensors nodes are deployed very close to the BS. The size of route-vector is the only memory overhead of a node in PEGASIS-INL. However, the PEGASIS does not incur additional memory overhead.

VII. SIMULATION AND RESULTS

TABLE I. SIMULATION PARAMETERS

Parameter	Value
Simulation Areas	500 x 500 m2
Network Size	500
Deployment type	Uniformly random
BS location	(0,250)
Packet Length	256 Bytes
Number of rounds	2500
Simulation time	3200 sec
Initial Node's energy	1 J
Communication range	25 meter
Eelec	50 nJ
Amp	100 pJ
Eagg	5 nJ
RSSI Threshold	-60dbm

We simulated PEGASIS-INL using Castalia-3.2 [13] simulator in the Omnet++ simulation environment [14]. Table I summarize the parameters considered for simulation environment. The metrics used for evaluating the performance of proposed scheme during the simulation are: *i)* Average energy consumed per node, *ii)* Average remaining energy per node *iii)* Number of nodes alive, and *iv)* Number of nodes died per round.

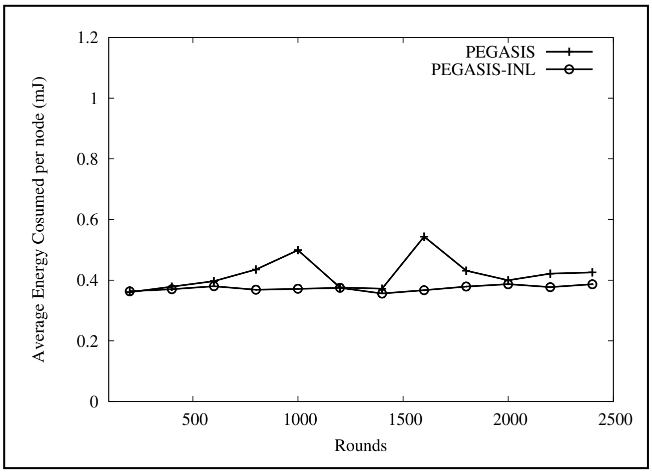

Fig. 1. Comparison of average energy consumption of a node per round *vs.* Number of Rounds

The comparison of average energy consumption of a node *vs.* number of rounds is shown in Fig. 1. It is observed from the figure that the average energy consumption of a node in PEGASIS-INL is marginally lower than PEGASIS. This is because of shorter distance communication between a leader node and the BS compared to PEGASIS. The fluctuations in PEGASIS graph is due to selection of leaders located relatively higher distance from BS.

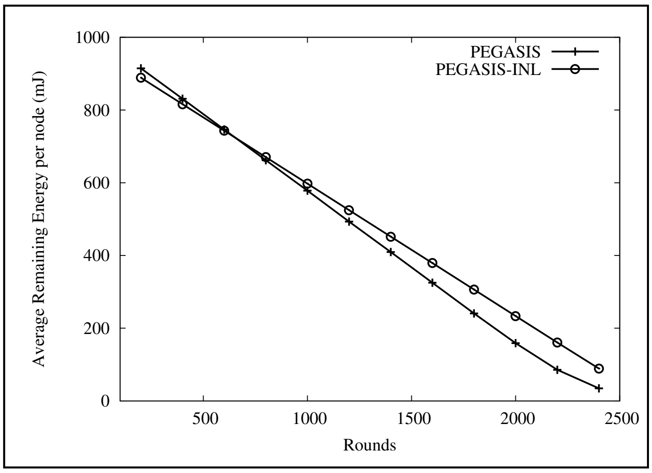

Fig. 2. Comparison of average remaining energy per node vs. Number of rounds.

Fig. 2 shows the comparison of remaining energy of nodes at the end of each round. It is observed from the figure that with the increase in number of rounds the remaining energy of PEGASIS decreases at higher rate compared to proposed PEGASIS-INL. This is mostly due to sending aggregated data to BS over large distance. T he comparison for number of alive nodes *vs.* number of rounds is shown in Fig. 3. It is observed from the figure that number of alive nodes is higher in PEGASIS-INL compared to PEGASIS. This is due to lesser energy consumption in each round of proposed scheme compared to PEGASIS. The comparison of number of nodes died per round is shown in Fig. 4.

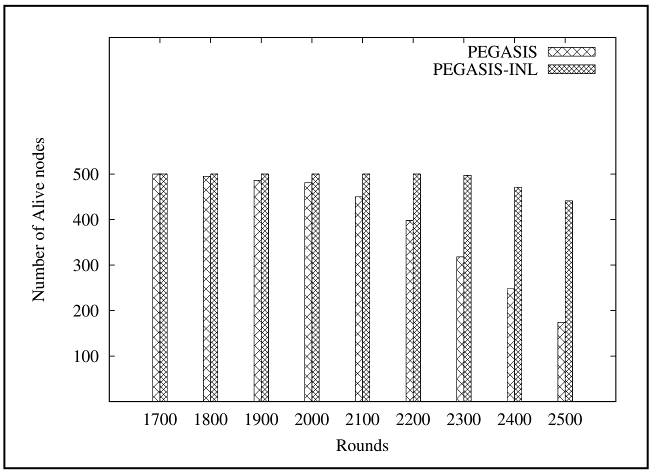

Fig. 3. Comparison of number of alive nodes vs. Number of rounds.

The number of nodes died in each round is higher in PEGASIS. This is attributed to the reason provided above. Therefore, from the all aboves result it is found that the proposed PEGASIS-INL has improved network lifetime compared to the PEGASIS protocol.

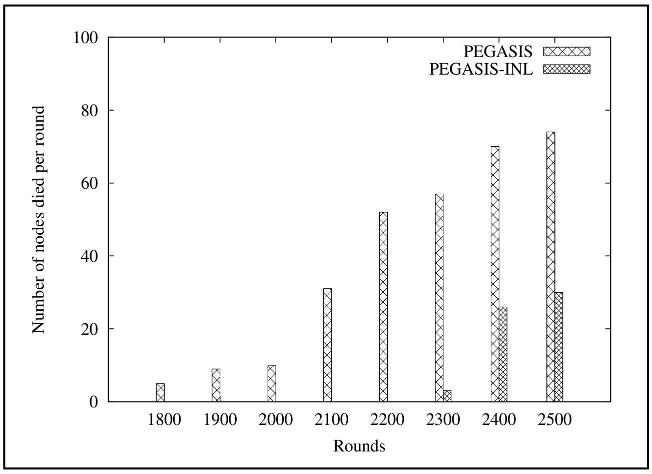

Fig. 4. Comparison of numberof nodes died per round vs. Number of rounds.

VIII. CONCLUSION

In this paper, we have proposed an efficient chain formation scheme for PEGASIS that improves the network lifetime. The scheme selects a subset of nodes in the proximity of BS as candidate leaders. The candidate leaders information is sent to all the nodes through a multi-hop broadcast process. In each round, nodes send the aggregated data to the leader node, which is responsible to send it to the BS. It is observed from the simulation results that the proposed PEGASIS-INL has higher network lifetime than PEGASIS protocol. The lower energy consumption in each round by each node attributed to enhance the lifetime of nodes in the network.

REFERENCES

[1] C. Buratti, A. Conti, D. Dardari, and R. Verdone, "An Overview on Wireless Sensor Networks Technology and Evolution," *Sensors*, vol. 2009, no. 9, pp. 6869 – 6896, August 2006.

[2] K. Sohraby, D. Minoli, and T. Znati, *Wireless Sensor Networks Technology, Protocols, and Applications*. John Wiley and Sons, 2007.

[3] R. Rajgopalan and P. K. Varshney, "Data Aggregation Techniques in Sensor Networks : A Survey," *IEEE Communications Surveys and Tutorials*, vol. 8, no. 4, pp. 48 – 63, 4th Quarter 2006.

[4] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy efficient Communication Protocol for Wireless Microsensor Networks," in *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, 2000*. IEEE, January 2000, pp. 01 – 10.

[5] O. Younis and 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, October- December 2004.

[6] S. Lindsey, C. S. Raghavendra, and K. M. Sivalingam, "PEGASIS: Power-Efficient Gathering in Sensor Information Systems," *IEEE Transaction on Parallel and Distributed Systems*, vol. 13, no. 9, pp. 924 –935, 2002.

[7] S.-M. Jung, Y.-J. Han, and T.-M. Chung, "The Concentric Clustering Scheme for Efficient Energy Consumption in the PEGASIS," in *Proceedings of The 9th International Conference on Advanced Communication Technology*, vol. 1, February 2007, pp. 260 – 265.

[8] W. Guo, W. Zhang, and G. Lu, "PEGASIS Protocol in Wireless Sensor Network Based on an Improved Ant Colony Algorithm," in *Proceedings of Second International Workshop on Education Technology and Computer Science*, vol. 3, March 2010, pp. 64 – 67.

[9] W. Linping, C. Zhen, B. Wu, and W. ZuJeng, "Improved algorithm of PEGASIS protocol introducing double cluster heads in wireless sensor network," in *Proceedings of International Conference on Computer, Mechatronics, Control and Electronic Engineering (CMCE)*, vol. 1, August 2010, pp. 148 – 151.

[10] F. Sen, Q. Bing, and T. Liangrui, "An Improved Energy-Efficient PEGASIS-Based Protocol in Wireless Sensor Networks," in *Proceedings of 8th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, vol. 4, July 2011, pp. 2330 – 2333.

[11] Z. Aliouat and M. Aliouat, "Efficient Management of Energy Budget for PEGASIS Routing Protocol," in *Proceedings of 6th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT)*, March 2012, pp. 516 – 521.

[12] B. Parno, A. Perrig, and V. Gligor, "Distributed Detection of Node Replication Attacks in Sensor Networks," in *Proceedings of IEEE Symposium on Security and Privacy*. IEEE, May 2005, pp. 49–63.

[13] A. Boulis, *Castalia 3.2, User's Manual*, National ICT Australia Ltd., Australia, 2011.

[14] A. Varga and R. Hornig, "An Overview of The Omnet++ Simulation Environment," in *Proceedings of 1st International Conference on Simulation Tools and Techniques for Communications, Networks and Systems and Workshops, SimuTools'08*, 2008, pp. 01–10.