

# HEER- A Delay-aware and Energy-Efficient Routing Protocol for Wireless Sensor Networks

D. Yi, and H. Yang, *Member, IEEE*

**Abstract**— Minimizing energy consumption to maximize network lifetime is one of the crucial concerns in designing wireless sensor network routing protocols. Cluster-based protocols have shown promising energy-efficiency performance, where sensor nodes take turns to act as cluster heads (CHs), which carry out higher-level data routing and relaying. In such case the energy consumption is more evenly distributed for all the nodes. However, most cluster-based protocols improve energy-efficiency at the cost of transmission delay. In this paper, we propose an improved delay-aware and energy-efficient clustered protocol called Hamilton Energy-Efficient Routing Protocol (HEER). HEER forms clusters in the network initialization phase and links members in each cluster on a Hamilton Path, constructed using a greedy algorithm, for data transmission purpose. No cluster reformation is required and the members on the path will take turns to become cluster head. The design allows HEER to save on network administration energy and also balance the load comparing to traditional cluster-based protocols. The algorithms designed in HEER also means that it does not suffer long delay and does not require each node to have global location information comparing with classic chain-based protocols such as PEGASIS and its variations. We implemented the HEER protocol in MATLAB simulation and compared it with several cluster-based and chain-based protocols. We found that HEER is able to achieve an improved network lifetime over the current protocols while maintaining the average data transmission delay. In the simulation, HEER achieved 66.5% and 40.6% more rounds than LEACH and LEACH-EE, which are cluster-based protocols. When compared with chain-based protocols (PEGASIS and Intra-grid-PEGASIS), HEER managed 21.2 times and 16.7 times more rounds than PEGASIS and Intra-grid-PEGASIS respectively. In addition, HEER can eliminated 90% of transmission delay comparing to LEACH and LEACH-EE and 99% comparing with PEGASIS and Intra-grid-PEGASIS.

**Keywords**— routing, WSNs, Hamilton Path, clustering

## 1 INTRODUCTION

WIRELESS Sensor Networks (WSNs) are widely used for various applications such as target tracking, habitat monitoring, military operation, surveillance system, vehicle motion, earthquake detection, patient monitoring systems and pollution control system etc. WSNs usually consist of Sensor Nodes (SNs) with small form factor, low cost and low energy

consumption. In most cases SNs are battery powered and are randomly deployed in the Field of Interest (FoI) in large numbers. It is therefore difficult to receive recharge or replacement batteries frequently, if possible at all [1]. As a result, performance of WSNs is often constrained by energy, of which the consumption efficiency is one of the most important concerns for designing WSNs [2].

Wireless communication is one of the major power consuming activities on SNs and its usage is controlled mainly by the routing protocol in action. An appropriate protocol design is therefore crucial for minimizing the energy consumption. There are three main types of routing protocols in WSNs, namely flat, hierarchical and location based protocols.

In Flat routing protocol, all nodes communicate in a peer-to-peer manner without any hierarchy in the structure. The advantage is having a simple, expandable and low maintenance structure. However, it lacks local and central management of communication resource and usually suffers from long and varied delay.

Location-based routing protocols, on the other hand, take in consideration of the location of SNs. Relative coordinates of neighbouring nodes can be extracted from the RF communication between them or, in some cases, by incorporating on-board GPS [3]. However, location tracking devices are huge energy consumers and their performance drops in harsh environment such as coal mines, undersea and are therefore strictly limited by workspace.

In hierarchical routing protocol, SNs form local structure called clusters, which consist of a cluster head (CH) and several non-cluster heads (non-CHs). CH and non-CHs in a cluster are called Cluster Members (CMs). Non-CHs transfer data through its CH, which manages the non-CHs in its cluster by allocating transmission slots for each of them. All the CHs then form the communication backbone of the network at a higher hierarchical level. Comparing with flat and local-based routing protocols, hierarchical protocol has better energy consumption efficiency and would be what we focus on in this paper.

In hierarchical protocol, CHs undertake various tasks such as node association, authentication, data aggregation, data fusion and task assignments [4]. It is therefore understandable that CHs would usually have much higher energy consumption compared to non-CHs. To prevent CHs from early dying and increase network lifetime the energy consumption within the network needs to be balanced by distributing the high load of CHs among all nodes. In this paper we propose a load distribution mechanism based on Hamilton path that relief

D. Yi is a M.Sc. candidate in Department of Computer Science, Loughborough University, LE11 3TU, UK. (E-mail: D.Yi-13@student.lboro.ac.uk).

H. Yang is Lecturer with the Computer Science Department, Loughborough University, LE11 3TU, UK. (E-mail: H.Yang@lboro.ac.uk).

much of the energy consumption happened in CHs data receiving and cluster reformation.

In additional, when the size of clusters grows, the pressure on CHs in aggregating and receiving data also increases. Controlling size of cluster in appropriate size is also important for balancing the load. In HEER mechanism to optimize the size of clusters is also designed based on the length of data for each SN per round. (The operation of hierarchical protocol can be divided into round and non-CHs send data to CHs and CHs transfer aggregated data to Base Station (BS) in per round.)

In this paper we aim to reduce and balance the network energy consumption by using Hamilton Path in graph theory to construct data transmission path. Definition of Hamilton Path is a path that traverses a graph and that accesses each node only once. A graph would be the abstract network structure in the context of this paper. In some hierarchical protocols, CH receives from each non-CHs' messages several times in a round. HEER travels all nodes once time in cluster for each round so that nodes are not repeatedly accessed. This feature reduces frequent access to CH. Moreover, HEER only forms clusters in first round the energy for cluster reformation can be saved. The lifetime of WSN can therefore be prolonged. Last but not least, in HEER the relationship between maximum packet length and each SN's detected data length dictates the average cluster size configuration to make sure that all the data in a cluster can be aggregated and transmitted to CH by making the most of a single packet so that the best possible energy-efficiency can be achieved during routing and data aggregation.

The contribution in this paper is threefold: firstly, we adopt the concept of Hamilton Path for linking members in each cluster formed without the need for the individual node's global position information, this reduces transmission distance for each CM as well as power/traffic pressure at CHs. Secondly, we design HEER in a way that clusters are formed only once so that the network administrative overhead is significantly reduced. Last but not least, we control our cluster size and link it to the maximum frame length in 6LowPan and the data payload of each CM so that the usage of network transmission capability is maximized.

The remainder of this paper is organized as follows: In Section 2, the methodologies which are used in the research are introduced. Then the related works in hierarchical WSNs are mentioned in Section 3. After that, the conception of design in HEER is indicated in Section 4. Next implement details which include phases and operation of HEER are shown in Section 5. The mathematical analysis of proposed protocol would be illuminated in Section 6 and simulation results and comparisons are discussed in Section 7. Eventually, Section 8 concludes the paper.

## 2 METHODOLOGY

### 2.1 Basic idea and approaches

The proposed protocol is based on Hamilton Path (HP) to save energy and balance load. Hamilton Path index of a given topology is NP-complete problem [5]. NP stands for

Non-deterministic Polynomial time and this means that the problem can be solved in Polynomial time using a Non-deterministic Turing Machine. Basically, a solution has to be testable in poly time. A NP problem can be resolved by the given problem with modified input then the problem is NP-complete [6]. Thus, NP-complete problem cannot be solved in polynomial time.

A Hamilton Path is a graph path between two vertices of a graph that visits each vertex exactly once [7] and Hamilton Path index is a NP-complete problem so finding Hamilton Circuits or Paths in arbitrary graphs is proved to be among the hardest problems of computer science [8]. In order to prove existence of Hamilton paths in the WSNs, two solutions can be considered to demonstrate the existence of Hamilton Path. One is to make arbitrary two nodes connecting freely in the cluster and the other is that to build a universal path which can make arbitrary two SNs can connect with each other. In this paper, former solution is applied to find Hamilton Path and relevant work (simulation, mathematical justification) are carried out following the first solution. Thus, the proposed protocol illustrates that nodes connection situation is fully connection in the cluster. In this situation, Ore's Theorem and Rahman's Theorem, which are degree based, are sufficient conditions for the existence of Hamilton Path in the WSN.

Congestion Control is another important goal in wireless sensor network. In the wireless sensor network, multiple sensor nodes sense the same event and are active for transmitting the information when the event occurred. Transfer rate could be varied due to multiple events occurred simultaneously [9] so the information would occur crash when non-CHs send to message to CH simultaneously in a round. Our proposed protocol is based on Hamilton Path and each sensor node just receives message from its neighbour once time in a round therefore information cannot crash or congestion in the cluster.

### 2.2 Graph Theory for resolving Routing Problem

A WSN can be abstracted as a graph with each SN as a node. Therefore, Graph Theory has great potential using in WSNs' routing technology. The node's connection degree in the Graph Theory can stand for the amount of neighbours that a sensor node have in the WSNs. Our presented protocol needs to find Hamilton Path in the cluster in order to convince that there is least one Hamilton Path existing in the cluster. Analogy method uses to prove the existence of Hamilton Path in the WSN's cluster. So WSN can analogy as a graph, we proves the existence of data collection chain by proving that WSN application can satisfy sufficient condition for the existence of Hamilton Path in Graph Theory more provable details mentions in Section 4.

### 2.3 Evaluation Platform

The proposed protocol is evaluated by simulations conducted in MATLAB. Selected reference protocols from the literature are also evaluated in the same simulation environment for comparison purpose. LEACH is the first clustered routing protocol in hierarchical wireless sensor network. LEACH-EE is an improved LEACH protocol with better strategy in electing

cluster heads, where the sensor nodes with more residual energy have a higher probability to be selected as cluster heads. LEACH-CHT is another CH selection based protocol which is not only considering residual energy of sensor nodes but also considering their times of being cluster head and the distance of sink node. The proposed protocol will form chain in clusters, hence typical chain based (PEGASIS) protocol's CH selection strategy needs to be discussed as well.

In this paper, we choose the occurrence of the first dead node as the end of network lifetime. Two sets of simulation are completed in the paper. One is about optimal CH selection, we compared three different optimal CH selection based protocols including LEACH, LEACH-EE, LEACH-CHT and a chain based protocol's CH selection strategy which is PEGASIS, in order to select the best CH selection strategy for special application. The other is to test the lifetime in specific threshold, LEACH, PEGASIS, Intra-grid PEGASIS, HEER are compared to obtain results.

### 3 RELATED WORK

A fundamental challenge in WSNs is to expand their lifetime. SNs are usually supported only by batteries, so the power resource is limited. Among the operations of a SN, wireless communication is one of most energy hungry event. Routing protocol needs to be carefully designed to make the best use of the resource available. The operation of WSNs in most applications relies on the SNs in the FoI to stay alive for as long as possible so it is also important to consume the power of each SN as evenly as possible [10].

Grid-based Routing and Aggregator Selection Scheme (GRASS) is to maximize the network lifetime by utilizing data aggregation and in-network processing techniques [11]. Data aggregation aims to combine and summarize data packets from several sensor nodes so that communication bandwidth and energy consumption are reduced [12]. In the hierarchical protocols, CHs undertake data aggregation task and then transmitted entire cluster's sensing information to BS.

#### 3.1 Flat and Hierarchical Routing Protocol

Flat protocols [13] are excellent in terms of their capability of using power-aware metrics to choose minimum power consuming paths. However flat protocols are not energy-efficient enough and are not ideal for load balance. They are because the network very often ends up using some paths more frequently than others and there's usually no central control. Hence, the SNs in those preferred paths run out power faster, while there's still precious power left on many other nodes. Comparing with flat protocols, hierarchal protocols are better in the way that they distribute energy dissipation among all SNs in the network. In hierarchical protocols, SNs are assigned different roles, which are able to switch after a pre-set time. The roles in hierarchal WSN include cluster head (CH) and non-cluster heads (non-CHs). Non-CHs cannot communicate with Base Station (BS) directly. They need to transmit data to BS via CH. On the other hand, CHs undertake data aggregation function so the overhead for data transmission is reduced. Tyagi and Kumar summarized hierarchical routing

protocols as being based upon parameters of CH selection, multi-hop transmission, structure of deployment (homogeneous and heterogeneous structure), mobility, security, spare time management, specific conditions in application and cluster size (including fixation and variable) [14]. There are a number of different hierarchical routing protocols proposed for different application scenarios. Various parameters are considered based on the categories of clustering protocols as shown in Figure 1.

Clustering has many benefits. It reduces unnecessary and redundant data to transmit because CH will transmit data to BS after aggregating the data of cluster. The core of clustering is to improve performance of WSNs based upon lifetime enhancement and load balancing, power efficiency, energy efficiency, optimal CH calculation and prolonging of lifetime of network. In this paper we focus on battery life enhancement by improving load balancing and CHs calculation as well as optimizing cluster size.

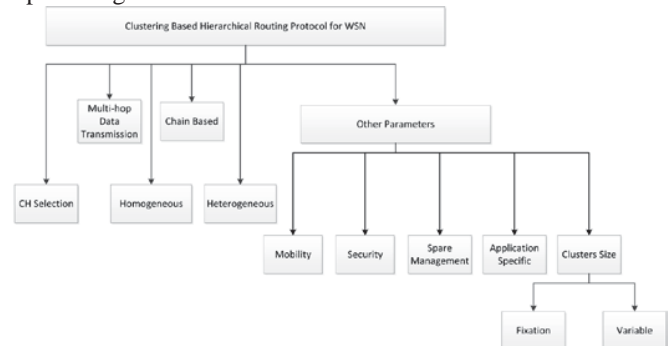


Figure 1 Categorization of Clustering Routing Protocols for WSNs

#### 3.2 Considerations and Techniques in Hierarchical Routing Protocol

LEACH protocol is the first hierarchical protocol in wireless sensor network. There are two main phases in clustering algorithms, which are node clustering and maintenance [15]. In LEACH, the operation of the network is broken into rounds that last for a pre-set time [16]. A number of SNs are selected as CHs in per round and the other SNs join them to form clusters depending on signal strength. At the beginning of each round, clusters will be re-formed. If a SNs' distance to the nearest node is farther than distance to BS, it will choose to directly communicate with BS. Generally speaking, cluster based protocol includes algorithms for distributed and adaptive cluster forming and cluster header position changing. The technique of distributed cluster forming ensures self-organization of most target nodes. The adaptive cluster forming and cluster header position changing algorithms ensure the energy dissipation is shared fairly among all nodes to prolong the lifetime of the whole system. In cluster based protocol, each round contains two phases: setup phase and steady phase in Figure 2. In setup phase, it forms cluster in self-adaptive mode and setup phase can be further divided into two stages which are advertisement stage and Cluster Set-up stage [17]. In the advertisement stage, CHs informs SNs in WSNs with advertisement messages to indicate their role. Non-CHs then join cluster according to the signal strength of advertisement message. Advertisement is a broadcast process that consumes fair amount energy. Furthermore, advertisement

stage happens in each round so its energy consumption will add up to a significant amount while the time goes by. On the other hand, in steady Phase, the main activity is data transmission from SNs to CHs and then to BS. The period of steady phase is usually much longer than the time of set-up phase to save the protocol overhead [16]. Increasing the length of steady phase in a round can save energy consumed by network management but will also reduce the fairness of energy consumption among the SNs.



Figure 2 Cluster-based Protocol Phases

A lot of variations of hierarchical protocols then followed after it came out. In this section we briefly review the mechanisms for CH selection, cluster size, delay, chain-based and congestion control in hierarchical protocols.

### 3.2.1 Selection of Cluster Head (CH)

Several protocols are proposed to optimize CH selection process so that the most suitable SNs are selected as CHs so that lifetime of WSNs can be prolonged.

In LEACH-EE [18] protocol, optimal CHs are considered to be the SNs who have more energy remained. The idea of LEACH-EE is therefore to let those SNs with more residual energy have a high probability of becoming CHs [15]. As an improvement, LEACH-CHT is proposed and considers not only each node's residual energy but also its distance to the BS. In Addition, to avoid repeated selection of a specific node as CH, the number of rounds that a node has been selected as CH is also considered [19].

Besides, forming clusters needs CHs to do broadcast and it would consume a significant number of energy. To address this, our proposed protocol just forms clusters in the initial round. Therefore, the proposed protocol can reduce energy dissipation by avoiding forming clusters repeatedly.

### 3.2.2 Cluster Size

Many existing protocols have provided methods to find optimal size of cluster in WSNs where nodes are deployed uniformly and randomly. 'Arranging clusters' sizes and transmission ranges for wireless sensor network (ACT)' [20] argues that the size of different clusters should be different. The parameter that affects the size of each cluster is the distance between its CH and the BS. The closer the distance, the smaller the cluster size as the CH closer to BS will relay more data. ACT tries to consume each CH's energy equally by adjusting the size of cluster, which reduces gradually from the farthest clusters to the closest ones.

ACT controls the size of a cluster by adjusting its radius. Another method for cluster size control is by managing the number of CMs in a cluster. LEACH and LEACH Based protocol (LEACH-C [16], BDBCP [21]) achieve this by controlling the number CHs in the network and assuming a uniform cluster size. They set a constant  $p$  which is the probability for a node to be selected as CH. The number of CMs in each cluster is therefore expected to be  $1/p$ .

Optimal Energy Consumption Model (OECM) [22] is able to obtain different optimal number of clusters according to different network performance factors such as radius of the

network, number of round, packet length and circuit energy dissipation and so on. Supposing that the number of SNs is uniform in LEACH and OECM WSNs, the number of CHs in OECM is not constant comparing with LEACH (the CHs number is constant in LEACH). OECM derived a relationship between optimal CHs number and various parameters. OECM forms clusters per round and it controls the number of cluster by adjusting the number of CHs in each round. However, OECM needs to reform clusters at the beginning of each round and data cannot be transmitted when forming clusters so delay will be increased. For overcoming the problem, we will form clusters at the first round and time for forming cluster can be saved when we design improved protocol. To avoid the high cost of long-range RF transmissions, HEED [23] operates in multi-hop network using a flexible power transmission for the inter-clustering communication and selects its CHs based on the node residual energy levels and cluster size for clusters will increase with distance from BS. However, in HEED, hot spot issue appears in areas that are close to BS as nodes that are closer to BS will relayed more data. In order to obtain a well-balanced network payload, the energy consumption of data communication and control overhead caused by route discovery should be considered. For this purpose, Energy-efficient Clustering (EC) [24] algorithm is proposed. EC determines suitable cluster sizes considering their hop distance to BS, EC effectively controls cluster sizes (which allows an approximately uniform use of the overall energy resources of a WSN) by adjusting the probability of a node being a CH.

Our proposed protocol uses a method similar to LEACH to manage the cluster size as ACT's method requires the CHs to detect BS position first find its appropriate next relay node. In worse situation, some CHs can be far away from the BS and end up having a large number of CMs. Those CHs will endure heavy traffic pressure and exhaust their energy very quickly. However, LEACH's method fixes the cluster size while building WSNs. In order to be more adaptive, our proposed protocol would adjust cluster size based on the relationship between CMs' data length and underlying protocols' maximum packet capacity. As a result, our proposed protocol will be able to adjust the number of CMs in a cluster dynamically. More details will be discussed in Section 6.2.

### 3.2.3 Chain Based Protocol

Network lifetime can be prolonged by linking CMs in chain(s). There are many mechanisms proposed for this in protocols such as PEGASIS, EBPEGASIS, EECBIG, IEEEPB, Hop PEGASIS and Intra-grid-PEGASIS.

PEGASIS [3] is a very typical chain based protocol, which links all SNs in a chain so that each node transmits and receives only from a neighbour. Just one SN is elected from the chain to relay all data to the BS. The data will be collected, aggregated and fused through the chain. In a randomly deployed WSN, each SN performs data detection, wireless communication and data fusion function. The position of each SN is supposed to be obtained by global positioning system (GPS). However, data fusion ability is always limited. Moreover, secure data aggregation [25] is a challenging task in WSNs. Thus, data cannot aggregate appropriately in some situations. When the number of SNs extends the maximum fusion ability, the data cannot fuse into a single packet length. The energy-efficiency

would be reduced. In addition, PEGASIS requires knowledge of the position of all the sensor nodes by GPS, which itself consumes a significant amount of power. Intra-grid-PEGASIS [26] is based on PEGASIS but divides sensor area into several network grids before connecting the SNs in each of them in a chain. The head of all network grids are then connected. The intra-grid-PEGASIS is an improvement over PEGASIS as energy dissipation can be reduced by avoiding formation of long distance chains. EBPEGASIS [3] is an energy efficient chaining protocol where nodes consider the average distance of formed chain to avoid forming long chains. It not only saves total energy consumption but also balances it better among all sensor nodes. IEEEPB [27] considers both nodes energy and distance between node and BS by normalizing the two factors and distributing different weight coefficients. Hop PEGASIS [28] makes sensor network more energy-efficient by dividing inter clusters into five different levels based on their distance to BS. The CH in low level, which means further away from BS, will forward data to higher-level cluster's CH. The CHs in five different level clusters will form a chain. PEGASIS and most of the improved PEGASIS protocols (intra-grid-PEGASIS [26], IEEEPB an EPEGASIS [29]) require that all the SNs are linked in a chain with the leader sensor node (CH) transmitting the data of the entire WSN to BS. With the size of the network growing, the number of SNs transmits data through the chain to leader sensor node by multi-hop which introduces a growing amount of delay that will eventually become unacceptable. Controlling length of chain is therefore important to control the average delay.

Generally speaking, chain based protocols have better performance in terms of network lifetime. In this paper, we inherit energy efficient design from chain-based protocols. Our HEER protocol will form clusters before linking SNs inside each cluster with a Hamilton Path. The details will be discussed in Section 5.1.

#### 3.2.4 Delay of WSN

In this paper, we focus mainly on network lifetime of WSNs. However, another important factor that needs to be considered is the delay. The energy consumption of the entire WSNs should be reduced while the delay should be kept acceptable. PEGASIS performs well in energy efficiency but its application is seriously limited as the detected data may need to travel the whole network before reaching BS. Therefore, PEGASIS's delay of transmission can be very high even though it is very energy efficient. PEDAP (Power Efficient Data gathering and Aggregation Protocol) [30] is a protocol, which uses near optimal minimum spanning tree based routing scheme to transmit data. The same as PEGASIS, main problem of PEDAP is high delay. In order to overcome the limitation of PEGASIS and PEDAP, a multi-layer energy-efficient and delay-reducing chain-based data gathering protocol (MEDC) was proposed in [31], where it puts forward the idea of multi-layer chain and uses the minimum total energy algorithm to construct the chain. As a result, multi-layer chain helps MEGC to reduce delay. Shi [32] proposed 'TDMA Scheduling with Optimized Energy Efficiency and Minimum Delay in Clustered Wireless Sensor Networks', which achieves high power efficiency, zero conflict and reduced end-to-end delay with a two-step approach that derives TDMA schedules. The first step is to formulate the problem via cross-layer optimization in order to get the most

energy-efficient flows on every link. The second step obtains a TDMA schedule with least frame length so that the least frame length guarantees minimum delay for the derived TDMA schedules. However, cross-layer optimization model is a nonlinear model so complexity of finding optimization result will increase with WSN grows. Moreover, network design should avoid cross-layer design. Our proposed protocol reduces delay by controlling cluster size and avoiding forming clusters repeatedly. Data cannot be transmitted or received during cluster setup phase. The wait for reforming cluster causes delay in the network. In our proposed protocol, the cluster formation process is executed only once at during network initialization in order to avoid waited time. PEGASIS and several improved PEGASIS (EECB [33], DS-PEGASIS [34]) protocols link all the SNs in the field in a chain. Those protocols do not need to form clusters repeatedly and then those protocols can avoid waited time problem. However, those protocols cause delay in another aspect. All the SNs in the WSN connect in a chain after leader sensor node receives all SNs' detected data and then leader sensor node would send all SNs' detected data to BS. PEGASIS and improved PEGASIS defines that there is only one leader sensor node in each round. Leader sensor node needs to wait all SNs' detected data arriving and then it is able to relay the all detected data to BS. Assuming that in the WSN, most of SNs are very close to leader sensor node and only several SNs in a far position from leader sensor node. According to the condition, PEGASIS and improved PEGASIS need to wait for detected data from several far SNs although it has received most of SNs' detected data in the WSN. The farthest SN from leader sensor node decides delay for transmitting. Our protocol avoids this kind of delay by controlling the size of cluster. Supposing that just several SNs connect in a chain rather than all SNs in the WSN connect in a chain, this problem is able to avoid. The farthest SN just affects SNs in its chain and the other chains would not affect by the farthest SN. Hence, average delay time of entire WSN is able to decrease much.

#### 3.2.5 Congestion Control

In WSNs, the data packets from SNs can be classified into three models, which are clock-driven, event-driven and query-driven [35]. Clock-driven data packets are produced by SNs and sent to BS in a fixed period. Event-driven data packets are generated when sensing data is over the highest threshold value and SNs identify current event as emergency transmitting to BS immediately. In query-driven model, BS queries SNs' data actively when a SN receives querying request, the SN will send back data as soon as possible. The event-driven packets and query-driven packets have higher priorities to be delivered than the clock-driven data packets [36]. In most hierarchical protocols, CHs allocate TDMA slots to its non-CHs. CHs adhere the TDMA schedule to send sensing data. In the FoI, emergency events are unpredictable. If an emergency event happens in SNs waking time, it does not have priority to transmit data because it has to follow allocated TDMA schedule to transmit data. We consider congestion in CHs when design new protocols which is mentioned in Section 6.4.

#### 4 DESIGN CONCEPT

We design our protocol based on a hierarchical WSN architecture. The assumption is that the SNs in a cluster form a single-hop fully connected graph and their data transmission is to be performed along a Hamilton Path that traverses the graph and ends at the CH rather than all transmitting directly to the CH. Such design reduces the average communication distance for each SN as well as the CH's receiving load so that energy consumption of each SN can be reduced and more evenly distributed among all SNs. The proposed protocol uses a procedure similar to LEACH to form clusters. It then establishes Hamilton Path among CMs (includes CH) in each cluster using a greedy algorithm. The data transmission in a cluster is then carried out along such path and the useful data will be aggregated on the way until it reaches the CH. Aggregating data along a Hamilton Path can reduce repetitive network overhead and avoid heavy bottleneck traffic at CHs. We argue that this reduces the overall network energy consumption as well as distributing it more evenly. We also link the data size and the selection of average cluster size, which is a parameter in our protocol to get the most out of the design. In this section, we describe some of those design concepts and the models we will be using in detail before presenting the protocol and its performance analysis in the following sections.

##### 4.1 Definition of Hamilton Path and Hamilton Circuit

A sample path in a graph  $G$  that passes through every vertex exactly once is called a Hamilton Path, and a circuit in a graph  $G$  that passes through every vertex exactly once is called a Hamilton Circuit.

From the definition of Hamilton Path/Circuit, conditions for the existence of a Hamilton Circuit in a graph are harsher than that for a Hamilton Path. Hamilton Circuit is sufficient condition for the existence of Hamilton Path and Hamilton Path is a necessary condition for existence of Hamilton Circuit.

In graph theory, Hamilton path does not always exist in arbitrarily graph. Before we can use it we need to prove its existence in the first place. The condition for Hamilton Path to exist can be presented by conditions set out in *Theorem 1* and *Theorem 2*, which both demonstrate that Hamilton Path exists in LEACH based protocol in WSNs.

##### 4.2 Theorem 1

In 1960, Ore proposed a lower bound for degree sum of nonadjacent pairs of vertices result for forcing the existence of a Hamiltonian cycle. Especially, Ore proved the *Theorem 1* as following where  $d_u$  denotes the degree of the vertex  $u$ .

According to Ore's theorem, let  $G$  be a (finite and simple) graph with  $n \geq 3$  vertices. We denote by  $deg v$  the degree of a vertex  $v$  in  $G$ , i.e. the number of incident edges in  $G$  to  $v$ . Then, Ore's theorem states that if  $deg v + deg w \geq n$  for every pair of non-adjacent vertices  $v, w \in G^*$ , then graph  $G$  has a Hamilton Circuit (includes Hamilton Path) [37].

Rahaman and Kaykobad presented [38] a sufficient condition for the existence of Hamilton Path in a graph basing upon a shortest distance so the parameter  $\delta(u, v)$  which denotes the length of the shortest path between  $u$  and  $v$ .

##### 4.3 Theorem 2

Let  $G = (V, E)$  be a connected graph with  $n$  vertices such that for all pairs of distinct nonadjacent vertices  $u, v \in V$  one has  $d_u + d_v + \delta(u, v) \geq n + 1$ . Then,  $G$  has a Hamilton path.

Let  $G = (V, E)$  be an undirected connection graph of a WSN with  $N$  nodes indexed by  $n \in \omega = \{1, 2, \dots, N\}$ , with the sink node having index  $N + 1$  and where each edge  $(m, n) \in E$  denotes a communication link from node  $m$  to node  $n$ .

In homogenous WSN, clusters always form in the homogeneity structure. In this case, assuming that SNs' communication range can cover the whole cluster, which is a prerequisite condition for performing the cluster forming phase in the first place, each cluster can be considered a single-hop fully connected network/graph as illustrated in Figure 3.

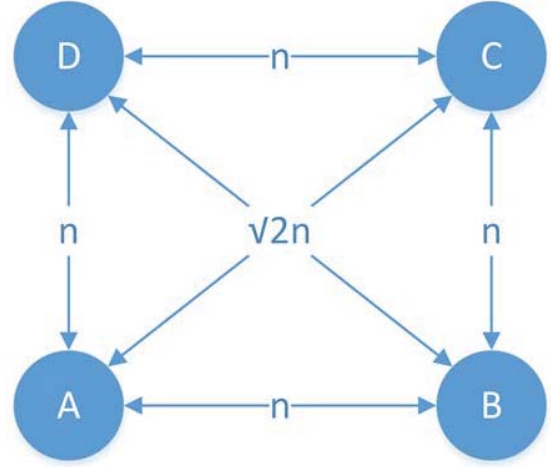


Figure 3 Single-hop Fully Connected WSN Model

A single-hop fully connected network can always satisfy the condition in *Theorem 2* so a Hamilton Path is assured to present in the cluster. Actually, there are usually more than one type of Hamilton Path in a single-hop fully connected network. In the case of Figure 3, it includes two types of different Hamilton Paths, as shown in Figure 4 (1), (2). Both of them can be used as the route for traversing the CMs (includes CH) and gathering data. The Hamilton Path in Figure 4-(1) is A->B->C->D and the Hamilton Path in Figure 4-(2) is A->C->B->D. Hence, we can find that the connected structure which is in Figure 4-(1) is more energy efficient than the other one in Figure 4-(2) according to the total transmission distance of Hamilton Path in Figure 4-(1) is shorter than in Hamilton Path in Figure 4-(2). It can therefore be concluded that the Hamilton Path in Figure 4-(1) is better than Hamilton Path in Figure 4-(2) in terms of energy-efficiency.

In order to find the optimal Hamilton Path in the cluster such as Figure 4 (1), the Greedy Algorithm [39] is proposed because Greedy Algorithm will select current optimal route to transmit data. For example, if node A is start node to form Hamilton Path, node A will select shortest path to form Hamilton Path (node A has two options to form Hamilton Path which are node B or node C). The distance between node A and node B is farther than the distance between node B and node C. Thus, node A will select node B as next member in the Hamilton Path. Then, node B can form Hamilton Path by selecting node C or

node D. According to current optimum in Greedy Algorithm, node B will select node C as next member in the Hamilton Path. When comes to node C, node C only have one option to form Hamilton Path. As a result, A->B->C->D rather than A->C->B->D will be formed Hamilton Path as show in Figure 4-(1). The process of the Greedy Algorithm is shown in Figure 5. The Step3 and Step 4 in Greedy Algorithm find all movable points and pick the point which is closest to home (current point) then move to the point. Therefore, Greedy Algorithm is able to choose the closest neighbour (in current set) from current node to be assigned as the next node in the path. In this case, it always generates the shortest Hamilton Path like Figure 4-(1). This is able to avoid forming non-energy-efficient Hamilton Path like Figure 4-(2) and prolonging lifetime of WSN.

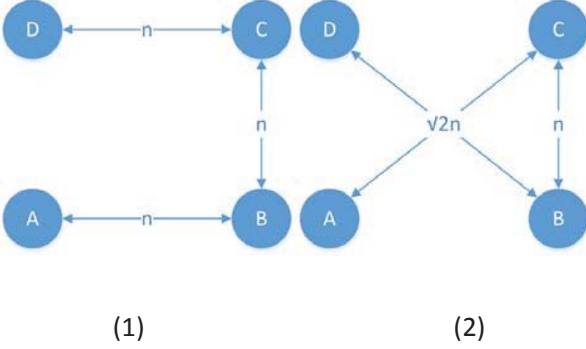


Figure 4 Data Gathering Chain

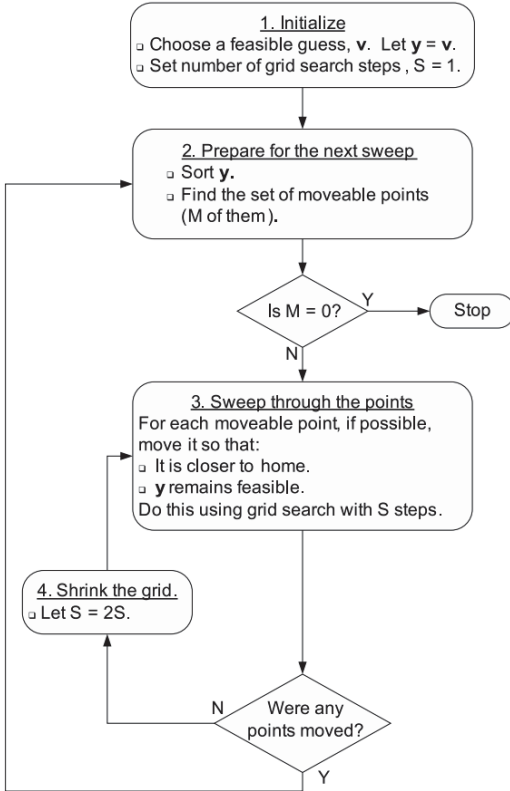


Figure 5 Greedy Algorithm for finding Hamilton path [40]

#### 4.4 Radio Energy Dissipation Model

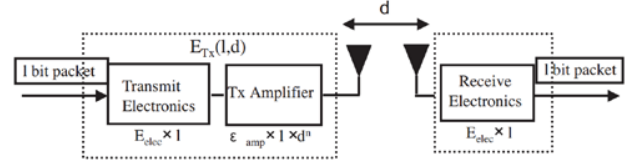


Figure 6 Data Transmission Model [26]

The proposed protocol uses the data transmission model in Figure 6 to estimate the total energy consumption in the WSN. As shown in Figure 6, for transmission of  $l$  bits data from the transmitter to the receiver at a distance  $d$ ,  $E_{Tx}(l, d)$  is the sum of energy consumption in transmitting electron and in power amplifier and is as shown in equation (1). The energy consumption for transmission can be given as  $E_{Tx-elec}(l, d) = E_{elec} \times l$  and the energy consumption in the power amplifier is  $E_{Tx-amp}(l, d) = l \times \epsilon_{amp} \times d^\alpha$ .  $E_{elec}$  is the energy dissipation for 1-bit data transmission in the transmitter electronic circuit so  $E_{Tx-elec}(l, d) = E_{elec} \times l$  denotes the energy consumption for transmitting  $l$  bits. The values of  $\epsilon_{amp}$  and  $\alpha$  vary depending on distance  $d$  and data decay rate. With the longer the communication range (distance) the more power a SN need to transmit the data. Threshold value  $d_0 = \sqrt{\epsilon_{fs} / \epsilon_{mp}}$  is used to distinguish whether data transmission happens in free space model or multi-path model. If current value of  $d < d_0$ , it proves that the transmission takes place in free space model and  $\epsilon_{fs}$  is the energy constant for amplifier to transmit 1 bit in free space model ( $d^2$  power loss). When  $d \geq d_0$ , it indicates that the signal transmission happens in multi-path decay model and  $\epsilon_{mp}$  is the energy constant for amplifier to transmit 1 bit in multi-path decay model ( $d^4$  power loss).  $\alpha$  is the attenuation exponent of wireless electromagnetic wave [26]. In different models, the relationship of energy dissipation and distance would be different. In general,  $\alpha$  is usually 2 in free space model and 4 in multi-path decay model.  $E_{Rx}(l)$  is the energy consumed for receiving  $l$  bits in receiver and can be given as  $l \times E_{elec}$  as shown in equation (2), where  $E_{elec}$  is the energy dissipation for 1-bit data transmission in the transmitter electronic circuit without amplifier and for 1-bit receiving in the receive electronics.

$$E_{Tx}(l, d) = E_{Tx-elec}(l) + E_{Tx-amp}(l, d) \\ = l \times E_{elec} + l \times \epsilon_{amp} \times d^\alpha \quad (1)$$

$$= \begin{cases} l \times E_{elec} + l \times \epsilon_{fs} \times d^2, & d < d_0 \\ l \times E_{elec} + l \times \epsilon_{mp} \times d^4, & d \geq d_0 \end{cases} \\ E_{Rx}(l) = l \times E_{elec} \quad (2)$$

#### 4.5 The mathematical model of delay

Assumes there are  $N$  (nodes) in the entire wireless sensor network. The probability of each sensor node being cluster

head is  $P$ . Thus, expected number of cluster heads in the whole WSN is  $P_i \sum_{i=1}^N n_i$ . When each SN has the same probability of becoming CH, the expected number of CHs is  $(NP)$ . In this paper, we assume that the SNs are identical and therefore have the same probability of becoming CH.

Assuming that there are  $N$  nodes in the WSN and that the area covered by the WSN is  $M \times M$ . The probability of a SN becoming cluster head is  $P$ . Therefore,  $(1/P)$  is the ideal number of cluster members in a cluster. The ideal number of CHs (or clusters) is  $(NP)$  in the WSN. In ideal network architecture the distance between sensor nodes is similar and the sensing region of each sensor node can be treated as circle, which is known by some as the pane model. Thus,  $d_{neighbour}$  is the distance between neighbour sensor node which is equal with diameter of each SN sensing region and can be given as  $2\sqrt{M^2/(N\pi)}$ . Let  $T_{forward}$  denote the delay time for one hop transmission.

SNs being CH can be divided into two cases. One case is that CH is in the body of chain. The other case is that CH is in the end of chain. If CH is in the body of chain, data collection will start with two ends simultaneously. Leader node (CH) will relay entire chain data until all data reach lead node. In the case, chain can be treated cut into two sections and delay is determined by longer section. As chain is symmetric, we only need to calculate sum of half situations delay and then double it how when the number of SNs on a chain is odd should add a minimum delay as well. We will get the total delay of different chain member being CH. After that we divide it by the number of chain member because every chain member can be CH, average delay on a Hamilton Path can be obtained. If CH is at the end of chain, maximum delay takes place in the case. All non-CHs data needs number of chain members minus one hops to reach CH in maximum delay case. When CH is exact at the middle of chain, it achieves minimum delay for collecting data.

Chain based protocols always form all SNs in a chain and one SN will be elected as leader (CH) to transmit all data in FoI, therefore average delay, maximum delay and minimum delay are given by

$$T_{AVE} = \left\{ \begin{array}{l} 2 \times \sum_{i=1}^{N/2} (N-i)(T_{forward}) / N, N \in \text{even} \\ [2 \times \sum_{i=1}^{N/2} (N-i)(T_{forward}) + (N/2)(T_{forward})] / N, N \in \text{odd}. \end{array} \right\} \quad (3)$$

Leader is end node (maximum delay):

$$T_{\max} = (N-1)(T_{forward}) \quad (4)$$

Leader is middle node (optimum delay):

$$T_{\min} = (N/2)(T_{forward}) \quad (5)$$

Forming all SNs in a chain causes high delay to collect data so we propose that CMs in a cluster forms chain rather than forming all SNs in a chain in order to reduce delay of collecting data. A CH will transmit data to BS until all data of its CMs arrive at CH. Thus, delay in a cluster is the time that all CMs' data arriving at CH. When CH is the end SN on the chain, the maximum delay occurs in HEER and is represented by the

delay of transmission through the full chain, as shown in Figure 7. When CH is at the exact middle of the chain, it achieves minimum delay as shown in Figure 8. When the other nodes on the chain are selected as CH, the delay is between the maximum and minimum delay. The average delay is therefore the mathematical expectation of the delay of each of the nodes as the CH.

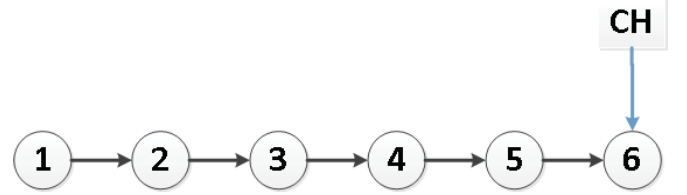


Figure 7 HEER Maximum Delay

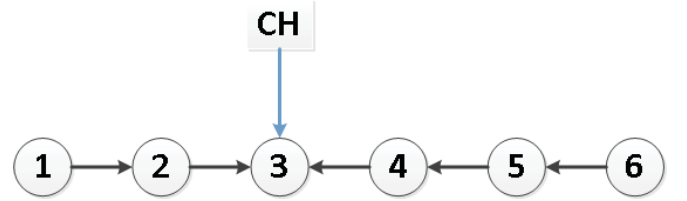


Figure 8 HEER Minimum Delay

The expected number of CMs in a cluster is  $(1/P)$  so average, maximum and minimum delay for collecting all the data on a Hamilton Path in HEER protocol can be given as:

$$T_{HEER\_AVE} = \left\{ \begin{array}{l} 2 \times \sum_{i=1}^{(1/P)/2} (1/P-i)(T_{forward}) / (1/P), (1/P) \in \text{even} \\ [2 \times \sum_{i=1}^{(1/P)/2} (1/P-i)(T_{forward}) + (1/P)(T_{forward})] / (1/P), (1/P) \in \text{odd}. \end{array} \right\} \quad (6)$$

CH is end node (maximum delay):

$$T_{HEER\_max} = ((1/P)-1)(T_{forward}) \quad (7)$$

CH is middle node (optimum delay):

$$T_{HEER\_min} = ((1/P)/2)(T_{forward}) \quad (8)$$

The average delay, maximum delay and minimum delay for most chain based protocols are show in equation(3), (4), (5). In HEER, they denote as equation(6), (7), (8). Comparing with average, maximum and minimum delay in HEER and traditional chain based protocols, HEER can effectively decrease the average, maximum, minimum delay for collecting data by reducing the length of chain.

## 5 DESCRIPTION OF HEER PROTOCOL

### 5.1 HEER Protocol Phases

LEACH Protocol contains two main phases in each round including Set-up Phase and Steady Phase. Set-up Phase can be further divided into two stages, which are Advertisement stage and Cluster Set-up stage, while the Steady Phase can also be broken down into Schedule Creation Stage and Data Transmission Stage in order to improve power efficiency. HEER protocol runs Set-up Phase and Steady Phase in first round. After the first round, HEER will no longer carry out the Advertisement stage and Cluster Set-up stage again. Instead, it maintains the clusters and chooses the next available CM on the Hamilton chain formed in the first round as the new CH in each round. This reduces energy consumption by avoiding the intense radio activities during re-clustering. In addition, it also reduces the delay incurred during such procedure as



advertisement and response time can be saved. In this case, each cluster's membership is fixed and finding Hamilton Path in each cluster happens only in the first round.

We illustrate the various phases in different rounds of HEER protocol in Figure 9, where the phases of LEACH are also shown as comparison. As shown in the figure HEER protocol contains two main phases, which are Set-up Phase and Steady Phase. In the first round, Set-up Phase is further divided into Advertisement stage, Cluster Set-up stage and Hamilton Path Discovery stage. Steady Phase that follows is further divided into Schedule Creation stage and Data Transmission via Hamilton Path stage. In the other rounds, HEER separates into three stages including CHs Selection stage, which belongs to Set-up Phase, Schedule Creation stage and Data Transmission via Hamilton Path stage which belongs to Steady Phase illustrating in Figure 10. Advertisement stage and Cluster Set-up stage can avoid instead HEER protocol needs to select optimal CHs which means SNs whom has more residual energy at Set-up Phase in each round.

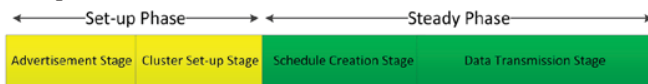


Figure 9 LEACH Protocol Phases



Figure 10 HEER Protocol First Round Phase



Figure 11 HEER Protocol the Rest of Round Phase

## 5.2 Operations of HEER Protocol

Step 1: Forming clusters like LEACH in the first round. Randomly selecting some nodes as cluster heads and the other nodes would join the cluster which is closest to its location. Each round selects CHs according to node's residual energy. Therefore, node who has more residual energy has more probability to become CHs. Because HEER forms clusters like LEACH, the SNs who are closer than all CHs in WSNs will directly communicate with BS. It saves energy maximally. The details are shown in Section 7.2.

Step 2: After forming clusters, the next task is to find a Hamilton Path in the clusters. Finding Hamilton Path is a NP-complete problem. It means only an approximate algorithm can be used to calculate Hamilton Path. The cluster members' relative location information can be gathered by CHs when forming clusters and CHs use the information to find Hamilton Path by Greedy Algorithm.

Step 3: The end nodes start to transmit their own data to their neighbor who is closer to CH and the neighbor would aggregate its detected data and received data until it reaches the CH. After CH receives all non-CH nodes' data, CH sends entire detected data in cluster to sink base. Next round would select a CH in Hamilton Path again and execute Step 3 again.

As shown in Figure 12, for our protocol only needs to form clusters once time in the whole lifetime so that the repeatedly forming clusters' energy consumption can be saved compared with LEACH and LEACH-based protocol. The chain that connects all cluster members will then be established with a

greedy algorithm. According to only forming clusters for once time, clusters and each cluster member would be fixed. This is good for data collection through chain because each cluster member is fixed therefore the chain can be repeatedly used each round. This saves energy to find chain each round and reduces delay of the WSNs.

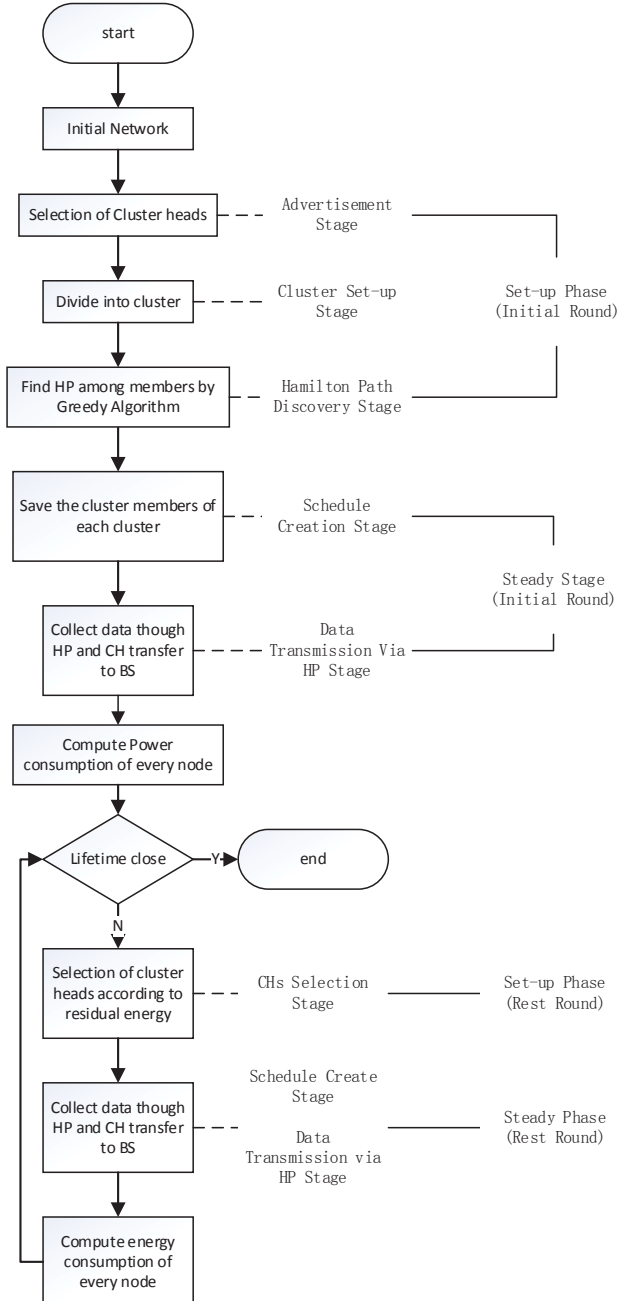


Figure 12 Flowchart of HEER Protocol

## 6 PROTOCOL ANALYSIS

In the proposed protocol, the initial process is separating nodes to form clusters. In this phase, some SNs are selected as CHs, of which the number is based on a predefined percentage. The CHs, once selected, will broadcast a message to inform the other nodes about their selection. Those not selected will then decide on which cluster to join depending on the received

signal strength to the CHs. This means that CHs are more likely to be located near the centre of their own clusters.

Cluster forming phase consumes a significant amount of energy, HEER performs such phase for just once in order to save power. After forming the clusters, the members in each cluster will form a Hamilton Path for data relaying. The Greedy Algorithm is a proven performer for building the path with minimum total length [41]. Data collection would then be carried out from end node (nodes) to CH, where the entire data payload is aggregated in order to be passed on to the sink node in the layer containing all the CHs. As HEER will keep the clusters and their paths formed in the initial round, CMs will then take turns to become the CH based on their energy level after a pre-set round time. The more residual energy a CM has, the higher probability it will be selected as the next CH. In this section, we discuss the performance of each phases of HEER as an evaluation as well as the justification for some of the designs.

### 6.1 Cluster head selection

Cluster Head selection mechanism is an important step as optimal CH selection can save more power in the long term. In our design, all nodes in a cluster are chained on a Hamilton Path, which leads us to think that those nodes can simply act as the cluster head in turns in order to be more energy efficient, as CH re-election process in traditional protocols is proved to be very power consuming. The total weight (distance) on Hamilton Path does not change and therefore no matter which node on the path is acting as the CH, the weight (distance) for collecting all data inside the cluster remains the same. Moreover, CH changing will occur after operating for a pre-set round time. In proposed protocol, the sensor node has more residual power more likely to be CH. In this case, the nodes in a cluster will have similar power consumption, meaning a fairer and more even power consumption distribution, which will prolong the lifetime of the cluster. Please note that the lifetime of network/cluster is bound by the occurrence of the first dead node.

### 6.2 Cluster Size Control

Commonly in WSN applications, SNs generate sensing data periodically with predefined time interval. In evenly distributed WSN, each SN would usually generate the same size of data during one interval. In our proposed protocol, this data would be transmitted and aggregated along a Hamilton Path that traverses a cluster and eventually reaches the CH. The CH would then transfer the whole data payload of the cluster together to the network sink node. The cluster size can therefore be optimized based on the maximum payload available in the underlying standard and the data generated by each SN to get the most out of the payload part available in standard packets. In this paper we take IEEE802.15.4 as example but the result can be easily generalized to different standards. According to IEEE 802.15.4 specification, maximum length of a frame is 127 bytes (1016 bits). The head of each frame is 25 bytes (200 bits) so it leaves a maximum of 102 bytes (816 bits) as payload at the media access control layer. When all the cluster members are sending messages to cluster head directly as in the traditional protocols, there is 25 bytes (200 bits) header information received repeatedly by the cluster head. This puts on additional

high pressure on the CHs in data receiving. In HEER, accumulating data among the cluster members before finally arriving at the cluster head is an efficient method for distributing the receiving load. The Hamilton path would be the most efficient if the aggregated data from all CMs can fit into one frame. As each frame has a limited payload size, the optimized size of cluster should be decided by the maximum frame payload size and the size of data generated by each SN during one interval. The LEACH protocol uses a predefined probability of a node being selected as cluster head to control the size of cluster. HEER applies the same mechanism but uses the frame payload size and SN data size to decide the value for such probability in order to achieve the optimized cluster size.

$$P = (l_{message} / F_{Max}) \quad (9)$$

Equation (9) shows how the probability being CH is calculated, where

$P$  : is the probability of a node being selected as CH

$F_{Max}$  : is maximum size of a Frame

$l_{message}$  : is the size of data from each node

### 6.3 Energy consumption model

HEER protocol energy consumption occurs in three parts. The first part is the energy consumption of clusters formation. The second part is the Hamilton Path discovery in clusters and the last part is the energy consumption for data transmission, which can perform in two different cases depending on whether the cluster head's (CH) position on the Hamilton path is at either end of it or in the middle.

*Part I: Energy consumption for cluster formation:*

The cluster formation algorithm ensures that the expected number of clusters in initial round is  $(NP)$  and expected number of CMs in a cluster is  $(1/P)$ . There are  $(1/P - 1)$  non-CHs and one CH in a cluster. Thus, the expected number of receiving advertisement for CH in a cluster is  $(1/P - 1)$  and CH receives  $(1/P - 1)$  responding message for joining cluster. We assume that advertisement message and responding message are the same length. The energy dissipation during cluster formation is twofold: the dissipation on CH and dissipation on non-CH members. For the CH it broadcasts advertisement and receives responding (REP) messages, while for the non-CH members they receive the advertisement from CH and transmit request (REQ) messages to CH. Let  $l_{adv}$  be the advertisement packet length, the energy dissipation of CH and each non-CH during cluster formation phase can be given by

$$E_{CH} = E_{broadcast} + E_{RX-REP} \\ = (E_{elec} \times l_{adv} + l_{adv} \times \epsilon_{fs} \times d_{toCH}^2) + E_{elec} \times l_{adv} \times (1/P - 1) \quad (10)$$

$$E_{non-CH} = E_{RX-ADV} + E_{TX-REQ} \\ = (E_{elec} \times l_{adv}) + (E_{elec} \times l_{adv} + l_{adv} \times \epsilon_{fs} \times d_{toCH}^2) \quad (11)$$

Where  $d_{toCH}$  is the distance from the node to the CH. The area of FoI is  $M^2$  so each cluster is occupying roughly an area of  $(M^2 / NP)$ . The region of cluster can be an arbitrary-shaped area with a node distribution  $\rho$ . Assuming CH at the centre of

cluster and the area is a circle with radius  $(M / \sqrt{(NP)\pi})$ , the expected squared distance from non-CHs to CH is

$$d_{toCH}^2 = \rho \iint_D x^2 + y^2 d\delta \quad (12)$$

$$D: x^2 + y^2 \leq \frac{M^2}{\pi(NP)}$$

In order to reduce the complexity of calculation, the equation (12) can be converted to Polar Coordinate as follows:

$$d_{toCH}^2 = \rho \int_{\theta=0}^{2\pi} d\theta \int_{r=0}^{M/\sqrt{\pi(NP)}} r^2 \cdot r dr \quad (13)$$

$$= \rho \frac{M^4}{2\pi N^2 P^2}$$

Assuming that the density of SNs is uniform throughout the cluster area then  $\rho$  will be a constant, which should have a value of  $\rho = (1/(M^2/NP))$ . Hence, the expected squared distance between CMs and CH can be presented as:

$$d_{toCH}^2 = \frac{(NP)}{M^2} \frac{M^4}{2\pi N^2 P^2} = \frac{M^2}{2\pi NP} \quad (14)$$

The energy dissipation for forming cluster is calculated by

$$E_{for \min g\_cluster} = E_{CH} + ((1/P) - 1) \times E_{non-CH}$$

$$= (E_{elec} \times l_{adv} + l_{adv} \times \epsilon_{fs} \times \frac{M^2}{2\pi NP}) + E_{elec} \times l_{adv} \times ((1/P) - 1)$$

$$+ ((1/P) - 1) \times \left[ (E_{elec} \times l_{adv}) + (E_{elec} \times l_{adv} + l_{adv} \times \epsilon_{fs} \times \frac{M^2}{2\pi NP}) \right] \quad (15)$$

$$= (3(1/P) - 2) \times E_{elec} \times l_{adv} + (1/P) \times l_{adv} \times \epsilon_{fs} \times \frac{M^2}{2\pi NP}$$

### Part II: Energy consumption for Hamilton path discovery

In this part, energy is consumed by actions performed for discovering the chain (Hamilton path). A greedy algorithm is used for forming such chain. When forming cluster, CH will receive response messages from the other CMs, CH can know the number CMs in the cluster and is able to use RSS measurement to identify the furthest CM in the cluster from CH and then send it Hamilton Path Declaration Message (includes the number of CMs in the cluster). The furthest CM in the cluster from CH is used as starting point of the chain and becomes the first active node in the algorithm. It will start by trying to find the closest neighbour node by broadcasting a chain forming advertisement message (with length of  $l_{ADV\_C}$ ) to the other CMs in the cluster with the value for broadcast range set to the diameter of its cluster  $2(M / \sqrt{(NP)\pi})$ . It then receives respond from the other CMs and chooses the nearest node as the next node based on the RSS measurement. It will then send Hamilton Path Declaration Message (with length of  $l_{HP}$ ) to the identified nearest SN (Hamilton Path Declaration Message will include the member of CMs in the cluster when current SN finds a new SN to add in Hamilton Path the number of CMs will minus one then duplicate to next SN. The next SN will repeat the procedure until the SN receive Hamilton Path Declaration Message with value of CMs equalling one that means the SN is the last node in the Hamilton Path), which will record MAC address and channel used by the Hamilton Path

Declaration Message. This nearest node then becomes the active node and will repeat the same procedure to find its closest neighbour node as the next node on the path. To avoid duplication the nodes that are already on the path will no longer participate in the practice. The procedure repeats until all nodes in the cluster are picked into the chain and the Hamilton path is therefore established.

Let  $d_{neighbour}$  be the distance from a node to its closest neighbour. Assuming that sensor node are evenly deployed in the FoI and by using ideal plate model we have  $d_{neighbour} = 2\sqrt{M^2/(N\pi)}$ . As the expected CM number in a cluster is  $(1/P)$ , there are  $(1/P - 1)$  CMs that need to perform chain forming advertisement and Hamilton Path Declaration Message according to the last node is the end of Hamilton Path no needing to finding next neighbour. Each current node will receive responding message from CMs which are not already in the Hamilton Path for responding chain forming advertisement. Assume that length of respond message is the same as chain forming advertisement. Let  $d_{CH\_Diameter}$  be the diameter of a cluster area,  $E_{broadcast}$  as the energy dissipation for broadcasting chain forming advertisement can therefore be presented as:

$$E_{broadcast} = E_{elec} \times l_{ADV\_C} + l_{ADV\_C} \times \epsilon_{fs} \times d_{CH\_Diameter}^2 \quad (16)$$

Let  $E_{RX-REP}$  denote the energy dissipation for receiving the responding message of chain forming advertisement. Hence,  $E_{RX-REP}$  can be presented as:

$$E_{RX-REP} = E_{elec} \times l_{ADV\_C} \quad (17)$$

The energy dissipation of announcing SN to become start point of Hamilton Path can be given by

$$E_{HP-ANN} = E_{elec} \times l_{ADV\_C} + l_{ADV\_C} \times \epsilon_{fs} \times d_{CH\_Diameter}^2 \quad (18)$$

Furthermore, the energy dissipation of sending Hamilton Path Declaration Message, denoted as  $E_{TX-HP}$ , can be presented as:

$$E_{TX-HP} = E_{elec} \times l_{HP} + l_{HP} \times \epsilon_{fs} \times d_{neighbour}^2 \quad (19)$$

The number  $i$  represent the sequence number of SN in the Hamilton Path. The overall energy dissipation of forming Hamilton Path can be given as:

$$E_{for \min g\_chain} = ((1/P) - 1) \times (E_{broadcast} + E_{TX-HP}) +$$

$$\sum_{i=1}^{((1/P)-1)} [(1/P - i) \times E_{RX-REP}] + E_{HP-ANN}$$

$$= ((1/P) - 1) \times (E_{elec} \times l_{ADV\_C} + l_{ADV\_C} \times \epsilon_{fs} \times d_{CH\_Diameter}^2$$

$$+ E_{elec} \times l_{HP} + l_{HP} \times \epsilon_{fs} \times d_{neighbour}^2) + \sum_{i=1}^{((1/P)-1)} [(1/P - i) \times E_{elec} \times l_{ADV\_C}]$$

$$+ E_{elec} \times l_{ADV\_C} + l_{ADV\_C} \times \epsilon_{fs} \times d_{CH\_Diameter}^2$$

$$= ((1/P) - 1) \times (E_{elec} \times l_{ADV\_C} + l_{ADV\_C} \times \epsilon_{fs} \times 4(M^2 / (NP\pi)))$$

$$+ E_{elec} \times l_{HP} + l_{HP} \times \epsilon_{fs} \times 4(M^2 / N\pi) + \sum_{i=1}^{((1/P)-1)} [(1/P - i) \times E_{elec} \times l_{ADV\_C}]$$

$$+ E_{elec} \times l_{ADV\_C} + l_{ADV\_C} \times \epsilon_{fs} \times 4(M^2 / (NP\pi)) \quad (20)$$

Part III: the energy consumption of data collection

Let  $d$  be the distance between sensor nodes and  $d_{toBS}$  be the distance from BS to CH. The energy dissipation of collecting data on the Hamilton path chain can be analysed in two cases depending on the CH's position.

Case 1: CH in the middle of the path

This means that the CH is one of the nodes that are not at either end of the path. In this case, we consider three types of nodes: the CH, NP-3 other mid-path nodes and two end nodes. The end nodes will initiate data transmission, which will occur on the path all the way until it reaches the CH. Each mid sensor node receives data from its neighbour, aggregate the data with its own before transmitting to the next neighbour on the path. The CH receives data from two neighbour nodes (neighbour1 and neighbour2) at last and aggregates data, including its own, to be relayed to BS. Hence, the energy dissipation of those three types of sensor nodes can be defined as following respectively.

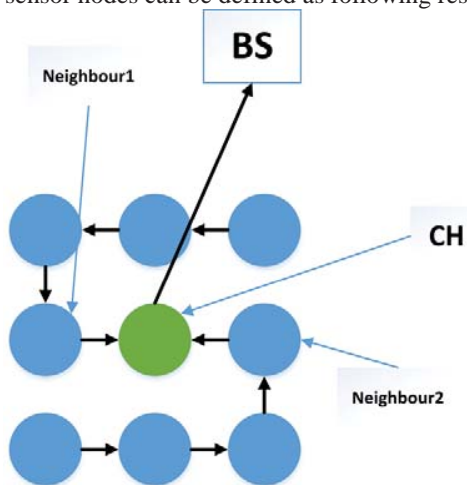


Figure 13 CH in the middle of the path

$$E_{CH} = E_{elec} \times (l_{neighbour1} + l_{neighbour2}) + E_{DA} \times (l_{own} + l_{neighbour1} + l_{neighbour2}) + l_{data\_to\_BS} \times \epsilon_{mp} \times d_{toBS}^4 \quad (21)$$

$$E_{mid} = E_{elec} \times (l_{neighbour}) + E_{DA} \times (l_{own} + l_{neighbour}) + l_{data\_to\_next\_neighbour} \times \epsilon_{fs} \times d^2 \quad (22)$$

$$E_{end} = E_{elec} \times (l_{own}) + l_{data\_to\_next\_neighbour} \times \epsilon_{fs} \times d^2 \quad (23)$$

Where  $E_i$  denotes the energy consumption of mid sensor node  $i$ . As different mid-path node receives and transmits data with different length, the energy consumption will be slightly different. The energy dissipation in one round for case 1, denoted by  $E_1$ , is given by

$$E_1 = E_{CH} + \sum_{i=1}^{(NP-3)} E_{(i)mid} + 2 \times E_{end} \quad (24)$$

Case 2: CH as an end node

In this case, we again consider three types of nodes: CH, NP-2 mid-path nodes and the non-CH end node. The only non-CH end node will initiate data transmission, which will go through the path until it reaches the CH at the other end. The mid-path node performs similar actions to what they perform in case 1. The difference is that the CH receives data from only one neighbour node (the only neighbour on the path) before aggregating data and relaying to BS. The energy dissipation of

those three types of nodes can be derived as follows respectively:

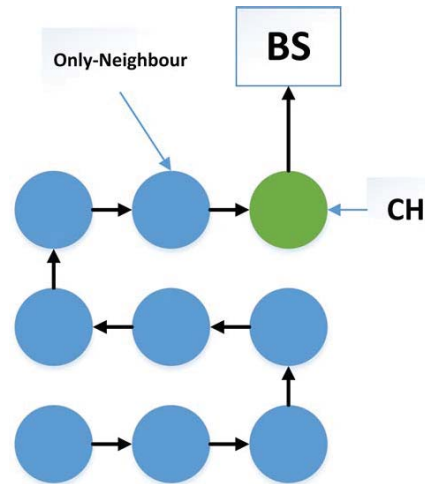


Figure 14 CH is end node

$$E_{CH} = E_{elec} \times (l_{only\_neighbour}) + E_{DA} \times (l_{own} + l_{only\_neighbour}) + l_{data\_to\_BS} \times \epsilon_{mp} \times d_{toBS}^4 \quad (25)$$

$$E_{mid} = E_{elec} \times (l_{neighbour}) + E_{DA} \times (l_{own} + l_{neighbour}) + l_{data\_to\_next\_neighbour} \times \epsilon_{fs} \times d^2 \quad (26)$$

$$E_{end} = E_{elec} \times (l_{own}) + l_{data\_to\_next\_neighbour} \times \epsilon_{fs} \times d^2 \quad (27)$$

Where  $E_i$  is the energy consumption of mid-path node  $i$ . Therefore, the energy dissipation in one round for case 2, denoted by  $E_2$ , is given as:

$$E_2 = E_{CH} + \sum_{i=1}^{(NP-2)} E_{(i)mid} + E_{end} \quad (28)$$

As the probability of having the situation in case 1 is  $(NP-2)/NP$  while that of having case 2 is  $2/NP$ . The overall mathematical expectation of the energy dissipation in one round per cluster, denoted by  $E_{rnd}$ , can be derived as:

$$E_{rnd} = (NP-2)/NP \times E_1 + (2/NP) \times E_2 \quad (29)$$

#### 6.4 Congestion Control in HEER

In most hierarchical protocols, CHs allocate TDMA slots to its non-CHs. CMs adhere the TDMA schedule to send sensing data. In the FoI, emergency events are unpredictable. If an emergency event happens in SNs waking time, it does not have priority to transmit data because it has to follow allocated TDMA schedule to transmit data. HEER overcomes this weakness by making CHs only communication with at most two nodes (SNs in Hamilton Path only communicate with neighbour SN or SNs) in SNs waking time. Thus, slots of CHs' TDMA schedule can be reduced and this decreases the probability that clock-driven data packet crashes with event data packet. In addition, CHs need to gather all CMs' sensing data before sends to BS in lots of hierarchical protocols. When some sensing packets arrive early in CHs, they have to wait in buffer until all CMs' data have arrived. In HEER, at most two

SNs to communicate with CHs (CHs need to wait for two neighbour nodes at most) in the cluster so average occupied time of CHs buffer should be decreased and less congestion may happen in CHs. HEER can reduce congestion in CH by forming all CMs in a cluster in Hamilton Path but the main propose of the paper is focused on energy-efficiency and load balancing so we will not discuss a lot in congestion control aspect.

## 7 PROTOCOL IMPLEMENTATION AND PERFORMANCE EVALUATION

In this paper, we used MATLAB to simulate the proposed protocol as well as selected reference protocols from the current research. The following configurations and assumptions hold in our simulations:

- We simulate 100 sensors deployed in an area with size  $300 \times 300 m^2$ ;
- Two different coordinates of sink sensor is set at (150, 450) or (150, 150) for comparison purpose;
- All the nodes are assumed identical in hardware and initial energy;
- Detected data is not compressed;
- The energy consumption parameters during the operational state of the network are shown in Table 1.

Table 1. Simulation Parameters.

Variables	Value
Initial energy	$E_{init} = 0.1J$
Based station location	(150, 450) / (150, 150)
Number of bits	$l = 90bits$ (Message Header is 200bits)
Radio attenuation exponent	free space $\alpha = 2$ multipath fading $\alpha = 4$
Amplifier dissipation	$\epsilon_{fs} = 10 pJ / bit / m^2$ $\epsilon_{mp} = 0.0013 pJ / bit / m^4$
Data compression ability	Detected Data cannot compress

### 7.1 Global CH selection strategy

Different CH selection strategies have different effect in various clustering protocols. In order to find appropriate cluster head selection strategy for our proposed protocol, we simulated four different CH selection strategies that can be found in PEGASIS, LEACH, LEACH-CHT and LEACH-CHT, which use different parameters to decide which SNs are chosen as CHs. The detailed selection criteria are described below.

In the CH selection process of LEACH protocol, every node has to select a random number between 0 and 1. When the number is less than a threshold  $T(n)$ , the node becomes a CH for the current round.  $T(n)$  is decided by equation (30), where  $p$  is the expected percentage of CHs,  $r$  is the number of current round and  $G$  is the set of sensor nodes that were not

selected as CHs in the last  $(1/p)$  rounds. Those parameters are also used with the same meaning in LEACH-EE, LEACH-CHT and PEGASIS.

$$T(n) = \begin{cases} \frac{p}{1 - p[r \bullet \text{mod}(1/p)]}, n \in G \\ 0, \text{otherwise} \end{cases} \quad (30)$$

Comparing with LEACH, LEACH-EE uses similar procedure but considers the nodes residual energy while making CH selection in an attempt to consume energy more evenly. The equation used for calculating  $T(n)$  in LEACH-EE is shown in equation (31), where  $E_{residual}$  is node residual energy and  $E_o$  is the initial energy.

$$T(n) = \begin{cases} \frac{p}{1 - p[r \bullet \text{mod}(1/p)]} \frac{E_{residual}}{E_o}, n \in G \\ 0, \text{otherwise.} \end{cases} \quad (31)$$

For further improvement, LEACH-CHT is proposed and considers not only each node's residual energy but also its distance to the BS. In Addition, to avoid repeatedly selecting a specific node as CH, the number of rounds that a node has been a CH is also considered. The formula for calculating  $T(n)$  in LEACH-CHT is given by equation (32), where  $d_{toBS}$  is the distance between current node and BS.  $CHT$  is the number of rounds when the current node has been CH. The formula is improved as

$$T(n) = \begin{cases} \frac{p}{1 - p[r \bullet \text{mod}(1/p)]} \frac{E_{residual}}{E_o} \frac{1}{d_{toBS}} \frac{1}{CHT}, n \in G \\ 0, \text{otherwise.} \end{cases} \quad (32)$$

Comparing to LEACH and its other variations, PEGASIS uses a rather different concept where there is only one CH in every round and all nodes chain together to pass data to the CH. Each node is given a node ID. When a node dies, the network re-forms the chain and node will get new IDs. The node with a node ID number that equals to current round number ( $r$ ) mod total number of nodes alive ( $N_{alive}$ ) is selected as the CH, which relays data to BS. Let  $CH(n)$  denote the ID number of the node selected as CH in the current round, it is calculated as follows:

$$CH(n) = \begin{cases} r \text{ mod}(N_{alive}), n \in N_{alive} \\ 0, \text{otherwise.} \end{cases} \quad (33)$$

In our simulation we use the occurrence of first dead SN as the end of network life and compare those different strategies. Figure 15 and Figure 16 show the result of how the minimum residual energy (the energy level of the node with the least energy in the network) in the network changes over rounds when the base station is located at (150,450) and (150,150) respectively. The four CH selection strategies have shown different performance. LEACH-CHT protocol does not demonstrate its energy efficient performance in the simulation, mostly because the simulation scenario focuses on load balance by watching for the first node death. The result indicates that LEACH-CHT has shorter life time than LEACH in this case. On the other hand, round number achieved by PEGASIS is very

poor (20 rounds for BS at 150,150 and 1 round for BS at 150,450) especially when no data compression is assumed in our setup. Furthermore, round number reached 266 for LEACH, 153 for LEACH-CHT and 315 for LEACH-EE for BS at (150,150). While BS is located at (150, 450), the first node dies at 10th round in LEACH-EE, 8th round in LEACH and 7th in LEACH-CHT. From the result, we can conclude that the CH selection mechanism in LEACH-EE performs best in our scenario.

Table 2. Simulation Parameter for CH Selection Strategy.

Variables	Value
Initial energy	$E_{init} = 0.1J$
Based station location	(150,150) / (150,450)
Number of bit	$l = 90bits$ (Message Header is 200bits)
Radio attenuation exponent	free space $\alpha = 2$ multipath fading $\alpha = 4$
Amplifier dissipation	$\epsilon_{fs} = 10 pJ / bit / m^2$ $\epsilon_{mp} = 0.0013 pJ / bit / m^4$
Data compression ability	Detected Data cannot compress

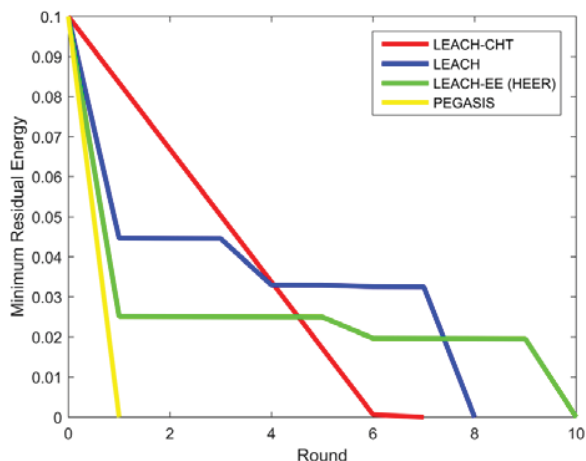


Figure 15 Selecting Cluster Head with BS (150,450)

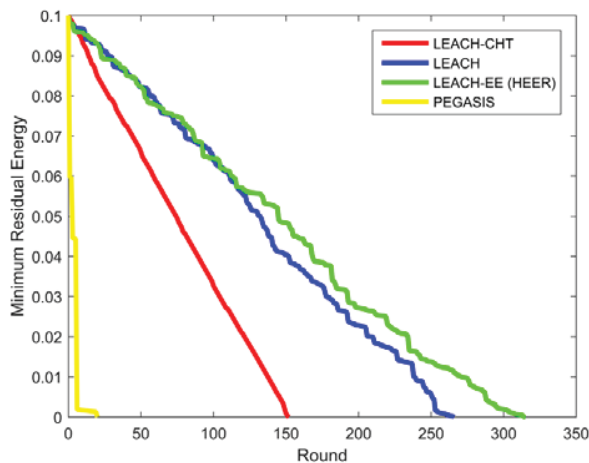


Figure 16 Selecting Cluster Head with BS (150,150)

As a result, LEACH-EE's cluster head selection strategy (as shown in equation (31)) is selected to be applied in the HEER protocol. As HEER performs global CH selection and cluster formation only once, this selected strategy applies only in the first round. The clusters will then be fixed and the CHs will be selected based on node residual energy thereafter to promote evenly distributed energy consumption.

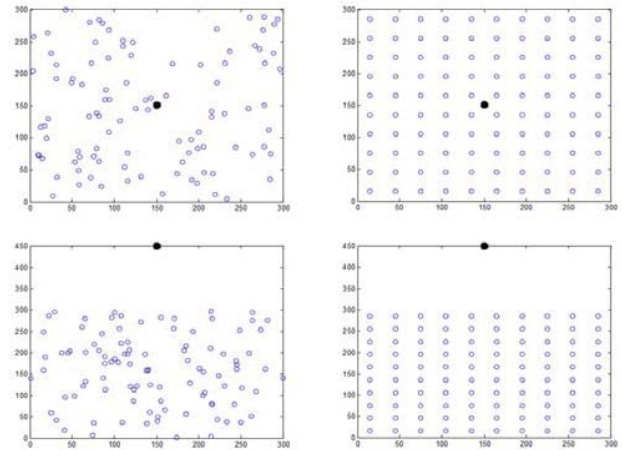


Figure 17 Four deployment scenarios in the Simulation

## 7.2 Topology Architecture in Cluster-based Network

We test HEER and LEACH in various network deployment scenarios, which are shown in Figure 17, to make sure that the result is generic. We do not show the topology structure in LEACH-EE, LEACH-CHT because all of LEACH, LEACH-EE and LEACH-CHT are cluster head selection related protocols and their topology architecture are very similar. Due to HEER forms clusters like LEACH in initial round, if SNs' distance to the nearest CH is farther than distance to BS, it will choose to directly communicate with BS. LEACH forms clusters in Figure 18 and Figure 20. If the SNs are closer to BS than all CHs, directly transmit data to BS is the most energy saving method. However, PEGASIS and Intra-grid-PEGASIS do not consider this kind of situation. It forms all SNs in a single chain in the WSN so this will bring extra energy dissipation. HEER overcomes the problem by forming clusters in first round and then all CMs in the same cluster forms a chain in Figure 19 and Figure 21. During cluster formation period, SNs that are closer to BS than any CHs will not join a cluster. Instead it will directly communicate with BS and this is the same in HEER protocol.

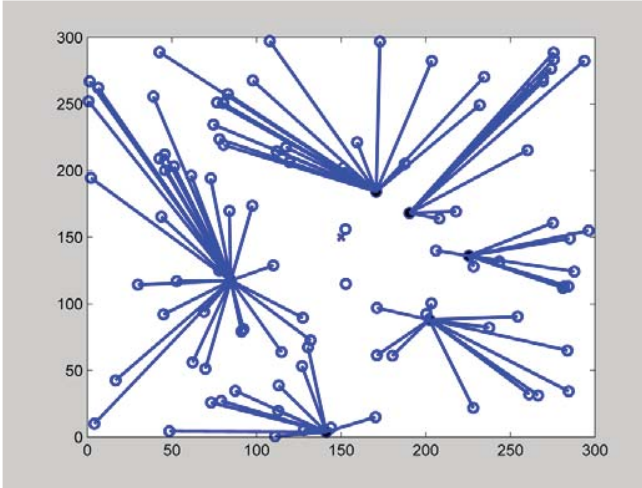


Figure 18 LEACH Topology Structure with BS (150, 150)

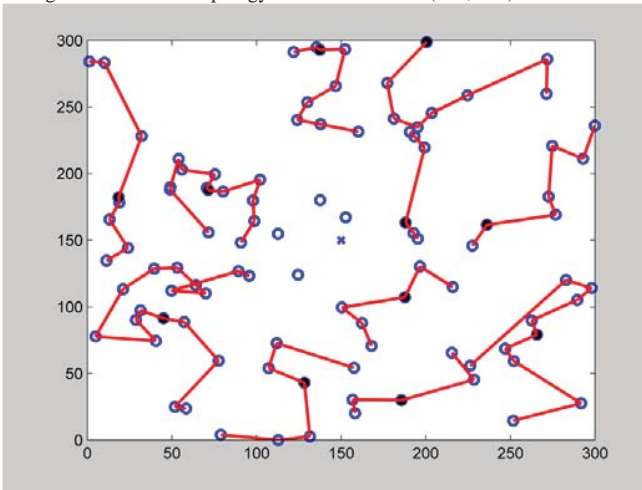


Figure 19 HEER Topology Structure with BS (150, 150)

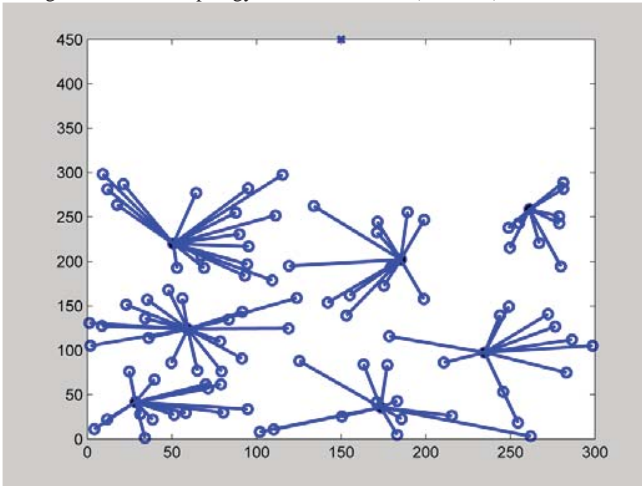


Figure 20 LEACH Topology Structure with BS (150, 450)

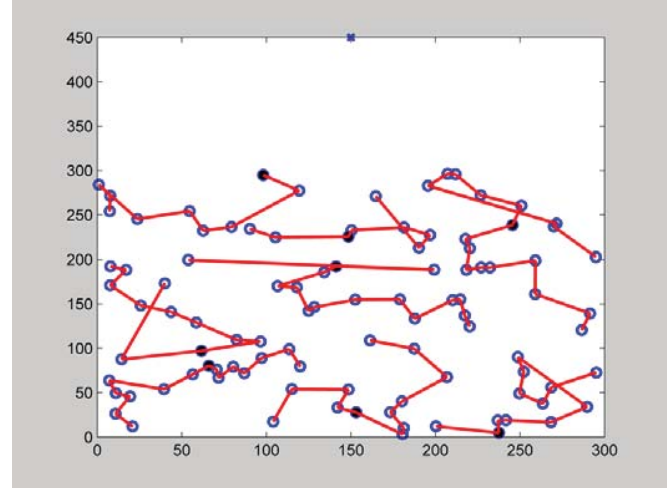


Figure 21 HEER Topology Structure with BS (150, 450)

### 7.3 Cluster-based network operation performance

In our simulations we compare HEER, PEGASIS, intra-grid-PEGASIS, LEACH and LEACH-EE in the four different deployment scenarios with BS positioned at (150, 450) and (150, 150). The simulation result indicates that HEER achieves the maximum number of rounds before the occurrence of first node death in the WSN regardless of the location of BS.

When BS is positioned at (150,450), HEER is able to achieve 13 rounds in grid deployment and 10 rounds in evenly distributed deployment until the first node dies while PEGASIS and intra-grid-PEGASIS both run 1 round in grid deployment and 1 round in evenly distributed deployment. LEACH can only run 8 rounds in evenly distributed deployment and 11 rounds in grid deployment until the first node dies. Moreover, LEACH-EE can run 10 rounds in evenly distributed deployment and 11 rounds in grid deployment until the first node is dead. It indicates that LEACH-EE is less affected by deployment of SNs. Comparing with different protocols in Figure 22 and Figure 23, HEER has the best performance in both evenly distributed deployment and grid deployment. HEER managed a lifetime 10 times longer than PEGASIS and intra-grid-PEGASIS in the evenly distributed deployment and 12 times longer in the grid deployment. HEER's lifetime achieved is also 37.5% longer in the evenly distributed deployment and 18.2% longer in grid deployment comparing to LEACH. When comparing with LEACH-EE, HEER outperformed it by 10% in evenly distributed deployment and 8.2% in grid deployment.

On the other hand, when BS is positioned at (150, 150) and deployment of SNs is evenly distributed, the first node dies in HEER at 443th round, while in PEGASIS and intra-grid-PEGASIS the first node death occurred in the 20th round and 25th round respectively. LEACH's first node death appears in the 266th round, with LEACH-EE managed to run until the 315th rounds. Therefore, HEER outperforms PEGASIS, intra-grid-PEGASIS, LEACH and LEACH-EE by at least 41% in evenly distributed deployment as shown in Figure 22.

In addition, the lifetime of HEER is nearly 11 times longer than PEGASIS and intra-grid-PEGASIS in grid deployment, The reason that the latter two protocol has the same performance is because in grid deployment

intra-grid-PEGASIS basically degraded to PEGASIS. Moreover, HEER's lifetime achieved is also 55% longer than LEACH and 51% longer than LEACH-EE in the same grid deployment as shown in Figure 23.

Furthermore, PEGASIS and intra-grid-PEGASIS consume more power at initial stages they need to form a chain that connects all SNs in the network. In the simulation, there are 100 SNs in the WSN it clearly showed that finding a chain that traverses the entire network consumes a great deal of energy. When the SNs are not evenly deployed throughout the area, non-ideal chains with long transmitting distance are usually formed. As a result, significant amount of energy would be spent on data transmission stages as well as forming the chain itself. Figure 22 and Figure 23 shows that the minimum residual energy among all SNs in the network of PEGASIS and intra-grid-PEGASIS decrease rapidly in initial period. This is because CH consumes a lot of power to receive and transmit data.

Furthermore, from the result showed in Figure 22 and Figure 25, we can clearly see that the location of BS has a significant impact on the lifetime of WSN. When the position of BS is at the centre of WSNs, lifetime of PEGASIS, intra-grid-PEGASIS can achieve 46 rounds in the grid deployment but only 1 round when the position of BS is on the edge of the area. Comparing with PEGASIS and intra-grid-PEGASIS, HEER is less sensitive to the BS location, though the lifetime still increased 36.7 times when BS moves from the edge to the centre, while LEACH and LEACH-EE had an increase of 28 times. This is because HEER is more like a hybrid protocol of LEACH and PEGASIS, which also share a similar operation pattern HEER's operation at CHs level.

In addition, we also simulate the scenarios that 50% nodes are alive in evenly distributed deployment and grid deployment. As shown in Figure 24, in evenly distributed deployment with BS at (150, 450), HEER is able to run 20.3% and 12.7% longer than LEACH and LEACH-EE respectively. When compared with PEGASIS and Intra-grid-PEGASIS, HEER can run 31.5% longer than PEGASIS and 29.1% longer than Intra-grid-PEGASIS. In grid deployment with BS at (150, 450), HEER is able to run 86 rounds which is 32.3% longer than LEACH and 28.3% longer than LEACH-EE. When compare with PEGASIS and Intra-grid-PEGASIS, HEER can achieve 50.9% longer for both of them, assuming that PEGASIS and Intra-grid-PEGASIS form the same chain in the grid deployment. In the evenly distributed deployment with BS (150, 150), we find that LEACH, LEACH-EE and HEER have very similar performance. Lifetime of HEER is slightly shorter than LEACH and LEACH-EE because HEER is more focused on the load balance among all SNs while LEACH and LEACH-EE are to achieve less total energy consumption. Comparing with HEER, LEACH and LEACH-EE, PEGASIS and Intra-grid-PEGASIS are far behind. From Table 5 and Table 6, we can conclude that HEER has the best performance with BS far away from FoI, while the performances of HEER, LEACH and LEACH-EE are similar when BS is at the centre of FoI. This is because SNs join clusters when its distance from BS is farther than its distance from closest CH. Hence, when BS is far away from FoI, all SNs will join various clusters to transmit data. With SNs dies in FoI, the distance between SN and CH will increase to a point when the energy dissipation

model will change from free space to multi-path. Therefore, the energy dissipation increases rapidly and kills SNs faster. In HEER, the SNs in a cluster will form a chain so that distance to transmit data will remain at a low value to keep the SNs alive. With BS at the centre of FoI, more SNs chose to transmit data directly to BS and the mechanism of HEER becomes less significant. Hence, performance of HEER, LEACH and LEACH-EE will be similar in this case.

In summary, HEER can consume energy more evenly comparing with LEACH, LEACH-EE, PEGASIS, intra-grid-PEGASIS, as it can operate more rounds until the first dead node appears in the network. Simulation also indicates that HEER has optimal performance regardless of the BS location. In addition, HEER can operate more rounds until half of the SNs die in the network when BS is far away from FoI. With more SNs die, chains become harder to achieve multi-path decay model in HEER comparing with LEACH, LEACH-EE. With BS at the centre of FoI, more SNs chose to transmit to BS directly when more SNs die and the performance of HEER, LEACH and LEACH-EE become similar.

Table 3 Simulation Results in Evenly Distributed deployment

BS Position \ Protocol	(150, 150)	(150, 450)
HEER	443	11
PEGASIS	20	1
Intra-grid-PEGASIS	25	1
LEACH	266	8
LEACH-EE	315	10

Table 4 Simulation Results in Grid deployment

BS Position \ Protocol	(150, 150)	(150, 450)
HEER	478	13
PEGASIS	46	1
Intra-grid-PEGASIS	46	1
LEACH	308	11
LEACH-EE	316	11

Table 5 Simulation Results when 50% nodes alive in Evenly Distributed Deployment

BS Position \ Protocol	(150, 150)	(150, 450)
------------------------	------------	------------



HEER	635	71
PEGASIS	210	54
Intra-grid-PEGASIS	222	55
LEACH	678	59
LEACH-EE	698	63

Table 6 Simulation Results when 50% nodes alive in Grid Deployment

BS Position \ Protocol	(150, 150)	(150, 450)
HEER	673	86
PEGASIS	230	57
Intra-grid-PEGASIS	230	57
LEACH	708	65
LEACH-EE	732	67

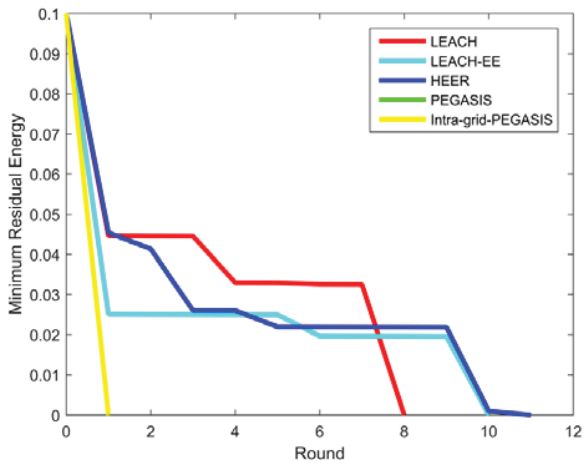


Figure 22 Comparing with BS (150,450) in Evenly Distributed deployment

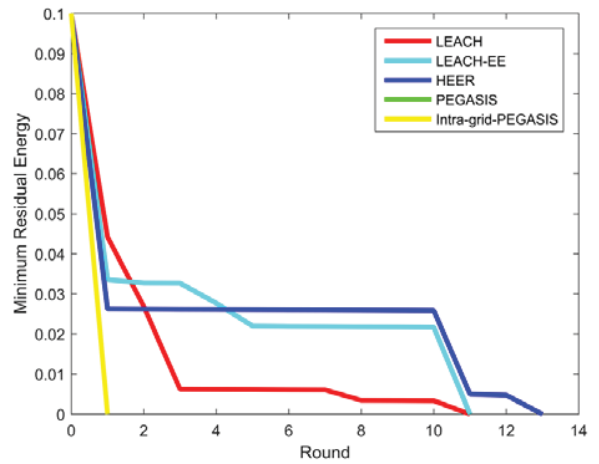


Figure 23 Comparing with BS (150,450) in Grid deployment

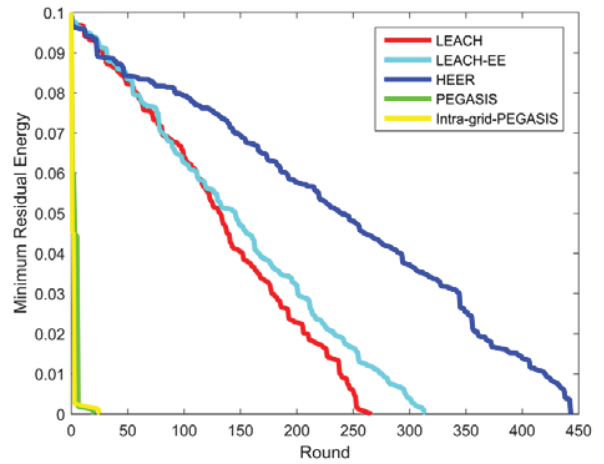


Figure 24 Comparing with BS (150, 150) in Evenly Distributed deployment

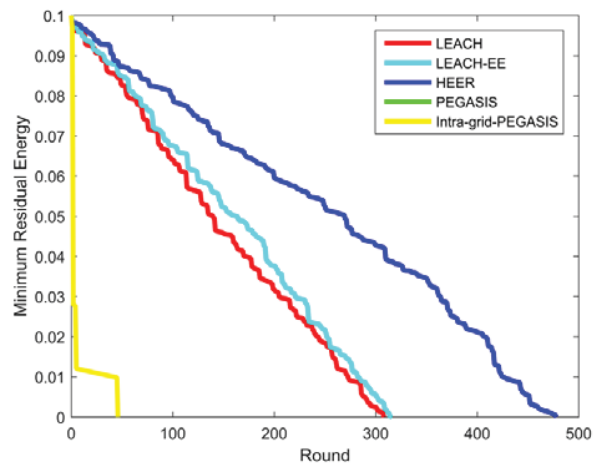


Figure 25 Comparing with BS (150, 150) in Grid Deployment

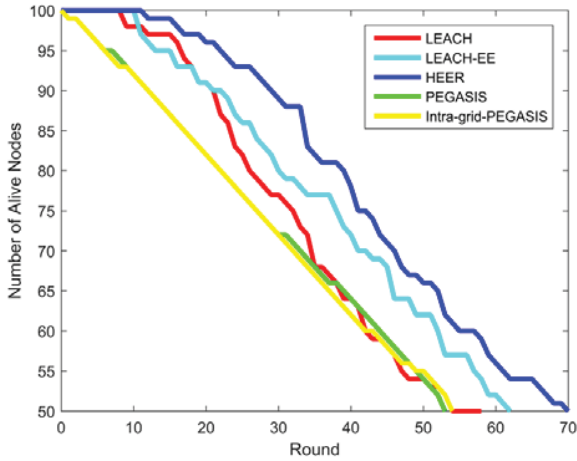


Figure 26 Comparing with BS (150, 450) in Evenly Distributed Deployment (when 50% nodes alive)

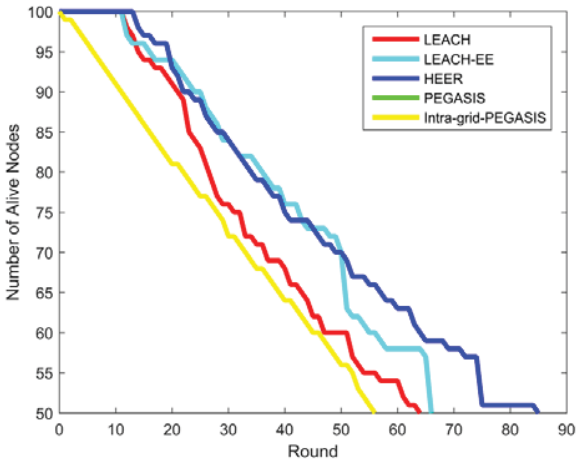


Figure 27 Comparing with BS (150, 450) in Grid Deployment (when 50% nodes alive)

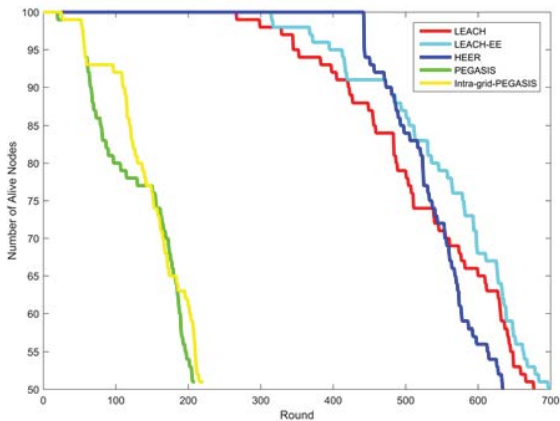


Figure 28 Comparing with BS (150, 150) in Evenly Distributed Deployment (when 50% nodes alive)

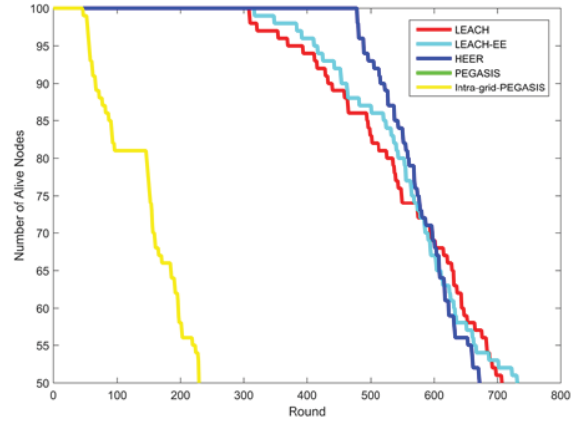


Figure 29 Comparing with BS (150, 150) in Grid Deployment (when 50% nodes alive)

#### 7.4 Transmission Delay in Cluster-based Network

In addition, we also simulate the transmission delay in HEER, PEGASIS, Intra-grid-PEGASIS, LEACH, LEACH-EE. In HEER, transmission delay is calculated from stable phase as HEER needs to form Hamilton Path in a cluster at initial phase but keeps the path until a new dead node appears. The results indicate that HEER has the lowest transmission delay. LEACH and LEACH-EE have very similar performance in transmission delay because they use the same mechanism for data transmission. As we would expect, PEGASIS and Intra-grid-PEGASIS have very poor performance due to data going through long chain of nodes before reaching BS. PEGASIS and Intra-grid-PEGASIS have very similar transmission delay as they use the same mechanism to transmit data. HEER has the best performance because HEER collects data on the path from both ends simultaneously. In addition, HEER also considers the utilization of each frame when forming Hamilton Path so less network traffic is generated. In LEACH and LEACH-EE, CHs need to broadcast a message in each round and then wait for non-CHs to join. After CHs finish receiving joining message from all non-CHs, non-CH can transmit detected data to CHs. This means the transmission delay in LEACH and LEACH-EE consists of two parts: in cluster forming phase and in actual data transmission. This produces a transmission delay higher than HEER. In Figure 30, x-axis is the number of Round and y-axis is the transmission delay in current round. Comparing with LEACH and LEACH-EE, HEER is able to reduce transmission delay by 90%. While comparing with PEGASIS and Intra-grid-PEGASIS, HEER can reduce 99% transmission delay showing in Figure 30 as HEER does not form the long chain used in them.

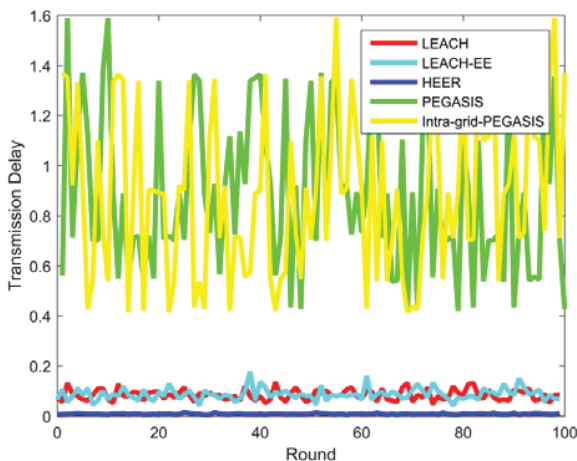


Figure 30 Transmission Delay in Different Protocols

## 8 CONCLUSION

In this paper we proposed an energy and delay-aware routing protocol based on clustering and the concept of Hamilton path. We transmit and aggregate data payload through a Hamilton path formed with all cluster members in order to reduce total network energy consumption by saving transmission range. The design also achieves more balanced power consumption for CHs. When the CM number grows the cost of forming Hamilton path and data aggregation both increase. Therefore we introduced cluster size control in our protocol design. This assures that the utilization of each packet is maximized while the average delay for data transmission to BS can be controlled. Furthermore, with the path formed, HEER protocol forms its clusters only once at the initial round and the CMs on the path takes turns to become CH. Hence, HEER saves more network management overhead comparing with other hierarchical protocols that need to form clusters again in each round.

To evaluate our design in the HEER protocol, we simulated LEACH, LEACH-EE, PEGASIS, intra-grid-PEGASIS as well as HEER itself to compare the performance. The evaluation considered various typical WSN deployment scenarios, such as grid deployment and random even deployment. The results indicate that HEER protocol has optimal performance in all scenarios, in terms of network lifetime when first dead node appears.

## REFERENCES

- [1] K. Maraiya, K. Kant and N. Gupta, "Wireless Sensor Network: A Review on Data Aggregation," *International Journal of Scientific & Engineering Research*, vol. 2, no. 4, pp. 1-6, April 2011.
- [2] T. Du, S. Qu, F. Liu and Q. Wang, "An energy efficiency semi-static routing algorithm for WSNs based on HAC clustering method," *Information Fusion*, vol. 21, pp. 18-29, January 2015.
- [3] V. Choudhary and M. K. Mahajan, "Energy-Efficient Protocols in Wireless Sensor Networks: A Survey," *IOSR Journal of Engineering (IOSRJEN)*, vol. 4, no. 6, pp. 39-43, 2014.
- [4] H. Yang and B. Sikdar, "Optimal Cluster Head Selection in the LEACH Architecture," in *Performance, Computing, and Communications Conference*, New Orleans, LA, 2007.
- [5] Z. Ryjáček, G. J. Woeginger and L. Xiong, "Hamiltonian index is NP-complete," *Discrete Applied Mathematics*, vol. 159, no. 4, pp. 246-250, February 2011.
- [6] M. R. Garey, D. S. Johnson and L. Stockmeyer, "Some simplified NP-complete graph problems," *Theoretical Computer Science*, vol. 1, no. 3, pp. 237-267, February 1976.
- [7] T. Gallai, "On finding simple paths and circuits in a graph," *Theory of graphs*, vol. 15, pp. 115-118, 1968.
- [8] M. R. Garey and D. S. Johnson, "Computers and intractability: a guide to NP-completeness," 1979.
- [9] S. A. Kahurke and B. N. Mahajan, "Implementation of Priority Based Scheduling and Congestion Control Protocol in Multipath Multi-hop WSN," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 1, no. 2, pp. 103-108, June 2012.
- [10] J. Yick, B. Mukherjee and D. Ghosal, "Wireless sensor network survey," *Computer Networks, Elsevier*, vol. 52, no. 12, pp. 2292-2330, 2008.
- [11] J. N. Al-Karaki, R. Ul-Mustafa and A. E. Kamal, "Data aggregation and routing in Wireless Sensor Networks: Optimal and heuristic algorithms," *Computer Networks*, vol. 53, no. 7, pp. 945-960, 2009.
- [12] S. Ozdemir and Y. Xiao, "Secure data aggregation in wireless sensor networks: A comprehensive overview," *Computer Networks*, vol. 53, no. 12, pp. 2022-2037, 2009.
- [13] A. E. Abdulla, H. Nishiyama and N. Kato, "Extending the lifetime of wireless sensor networks: A hybrid routing algorithm," *Computer Communications*, vol. 35, no. 9, p. 1056-1063, 2012.
- [14] S. Tyagi and N. Kumar, "A system atic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks," *Journal of Network and Computer Applications*, vol. 36, no. 2, p. 623-645, 2013.
- [15] D. Dembla and S. H. Mehta, "Energy Efficient LEACH protocol for Wireless Sensor Network," *International Journal of IT & Knowledge Management*, vol. 6, no. 2, pp. 165-169, 2013.
- [16] W. B. Heinzelman, A. P. Chandrakasan and H. Balakrishnan, "An applicationspecific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, p. 660-670, 2002.
- [17] K. V. Shukla, "Research On Energy Efficient Routing Protocol LEACH For Wireless Sensor Networks," *International Journal of Engineering Research & Technology (IJERT)*, vol. 2, no. 3, pp. 1-5, March 2013.
- [18] Y. Qin, Q. Tang, Y. Liang, X. Yue and X. Li, "An Energy-Efficient Cooperative MIMO Scheme for Wireless Sensor Networks Based on Clustering," in

- Computational Science and Engineering (CSE), 2011 IEEE 14th International Conference*, Dalian, Liaoning, 2011.
- [19] N. ZHANG and X.-a. TONG, "An Improving Method of Cluster Head Option of LEACH in Wireless Sensor Networks," *Science Technology and Engineering*, vol. 12, no. 13, pp. 3139-3141, 2012.
- [20] W. K. Lai, C. S. Fan and L. Y. Lin, "Arranging cluster sizes and transmission ranges for wireless sensor networks," *Information Sciences*, vol. 183, no. 1, pp. 117-131, 15 January 2012.
- [21] S. D. MURUGANATHAN, D. C. F. MA and R. I. BHASIN, "A Centralized Energy-Efficient Routing Protocol for Wireless Sensor Networks," *Communications Magazine, IEEE*, vol. 43, no. 3, pp. 8 - 13, 2005.
- [22] K. Yang, Y.-M. Wu and H.-B. Zhou, "Research of Optimal Energy Consumption Model in Wireless Sensor Network," in *Second international conference on*, Chengdu, 2010.
- [23] O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *Mobile Computing, IEEE Transactions on*, vol. 3, no. 4, pp. 366-379, 2004.
- [24] D. Wei, Y. Jin, S. a. M. K. Vural and R. Tafazolli, "An Energy-Efficient Clustering Solution for Wireless Sensor Networks," *IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS*, vol. 10, no. 11, pp. 3973-3983, 2011.
- [25] S. Kannadhasan, G. KarthiKeyan and V. Sethupathi, "A Graph Theory based Energy Efficient Clustering Techniques in Wireless Sensor Networks," in *2013 IEEE Conference on Information and Communication Technologies (ICT 2013)*, JeJu Island, 2013.
- [26] Y.-L. Chen and J.-S. Lin, "Energy efficiency analysis of a chain-based scheme via intra-grid for wireless sensor networks," *Computer Communications*, vol. 35, no. 4, p. pp. 507-516, 2012.
- [27] F. a. B. Q. Sen and T. Liangrui, "An Improved Energy-Efficient PEGASIS-Based Protocol in Wireless Sensor Networks," in *2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, Shanghai, 2011.
- [28] M. A. Ali and A. K. Rajpoot, "Development of energy efficient routing protocol using Hop PEGASIS in Wireless Sensor Networks," *International Journal of Computer Science & Engineering Technology (IJCSET)*, vol. 5, no. 02, pp. 124-131, 2014.
- [29] S.-M. Jung, Y.-J. Han and T.-M. Chung, "The Concentric Clustering Scheme for Efficient Energy Consumption in the PEGASIS," in *Advanced Communication Technology, The 9th International Conference*, Gangwon-Do, 2007.
- [30] Korpeoglu, Tan, H. Ozgur and Ibrahim, "Power Efficient Data Gathering and Aggregation in Wireless Sensor Networks," *ACM SIGMOD Record*, vol. 32, no. 4, pp. 66 - 71, December 2003.
- [31] L. Yuan, Y. Zhu and T. Xu, "A Multi-Layered Energy-Efficient and Delay-Reduced Chain-Based Data Gathering Protocol for Wireless Sensor Network," in *Mechronic and Embedded Systems and Applications, 2008. MESA 2008. IEEE/ASME International Conference on*, Beijing, 2008.
- [32] L. Shi and A. O. Fapojuwo, "TDMA Scheduling with Optimized Energy Efficiency and Minimum Delay in Clustered Wireless Sensor Networks," *IEEE TRANSACTIONS ON MOBILE COMPUTING*, vol. 9, no. 7, pp. 927 - 940, July 2010.
- [33] Y. Yu and Y. Song, "An Energy-Efficient Chain-Based Routing Protocol in Wireless Sensor Network," in *2010 International Conference on Computer Application and System Modeling (ICCA SM 2010)*, Taiyuan, 2010.
- [34] J.-E. Lee and K. Kim, "Diamond-Shaped Routing Method for Reliable Data Transmission in Wireless Sensor Networks," in *2008 International Symposium on Parallel and Distributed Processing with Applications*, Sydney, NSW, 2008.
- [35] C. Wang, K. Sohrawy, V. Lawrence, B. Li and Y. Hu, "Priority-based Congestion Control in Wireless Sensor Networks," in *Proceedings of the IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing*, 2006.
- [36] H.-J. Liu, M.-W. Huang, W.-S. Hsieh and C. J. Jan, "Priority-based Hybrid Protocol in Wireless Sensor Networks," in *11th IEEE International Conference on High Performance Computing and Communications*, Seoul, 2009.
- [37] S. Ronghua, "2-Neighborhoods and Hamiltonian Conditions," *Journal of Graph Theory*, vol. 16, no. 3, pp. 267-271, 1992.
- [38] M. S. Rahman and M. Kaykobad, "On Hamiltonian cycles and Hamiltonian paths," in *Information Processing Letters*, 2005.
- [39] P. J. Looges and S. Olariu, "Optimal greedy algorithms for indifference graphs," Birmingham, AL, 1992.
- [40] M. A. Wolters, "A Greedy Algorithm for Unimodal Kernel Density Estimation by Data Sharpening," *Journal of Statistical Software*, vol. 47, no. 6, pp. 1-26, April 2012.
- [41] S. Lindsey and C. S. Raghavendra, "PEGASIS: Power-Efficient GATHERing in Sensor Information Systems," in *Aerospace Conference Proceedings, 2002. IEEE (Volume:3 )*, Los Angeles, CA, USA, 2002.