Journal of Theoretical and Applied Information Technology
30th April 2014. Vol. 62 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A SURVEY OF STATE OF THE ART: HIERARCHICAL ROUTING ALGORITHMS FOR WIRELESS SENSOR NETWORKS

[*1]HUSAM KAREEM, [2]S.J. HASHIM, [3]A. SALI, [4]SHAMALA SUBRAMANIAM
[1, 2, 3]Faculty of Engineering, Universiti Putra Malaysia
[4]Faculty of Computer Science and Information Technology, Universiti Putra Malaysia
E-Mails: [*1]husam_kf@yahoo.com , [2]sjh@upm.edu.my ,
[3]aduwati@upm.edu.my , [4]shamala@fsktm.upm.edu.my

ABSTRACT

Generally, any device that has the ability to sense the surrounding environment can be considered as a sensor node. A wireless sensor network is a group of sensor nodes that cooperate with each other, it may contain a few numbers of sensors or it can consist of thousands of sensors. Wireless sensor networks can be used in wide range of applications such as measuring temperature, humidity, pressure, noise level, monitoring the vehicular movement, military applications and health applications. One of the main crucial issues can face the operation and the surveillance of wireless sensor networks is energy consumption. Energy in wireless sensor networks mostly powered by battery and the greatest share of this energy is consumed during data transmission. Many researches have been done to solve this problem or at least find a solution to decrease the energy consumption. One of those solutions is using efficient routing algorithm. The most efficient type of algorithms that can be used for WSNs in large areas is hierarchical routing algorithms. In this article, we will present a review for the state of the art for recent hierarchical algorithms. Moreover, we will use criteria to classify the hierarchical algorithms that never used before in any article, which is the mechanism that implemented in the hierarchical algorithm such as clustering, chaining, or hybrid between cluster and chain.

Keywords: *Wireless Sensor Networks, Hierarchical Routing Algorithms, Clustering Routing Algorithms, Chain Routing Algorithms, PEGASIS, LEACH.*

1. INTRODUCTION

Since there are numerous applications for WSNs, this leads to propose many new algorithms to be compatible with different applications and different network structure. These routing algorithms have considered in their design WSNs characteristics such as the energy limitation, communication range, and memory size. Our concern in this paper is to introduce a valuable review about the recent work in hierarchical routing algorithms, and introduce an explanation about the difficulties and challenges might face the design of routing algorithms. In addition, compared their solutions that they utilized in their mechanism in order to overcome those issues. Moreover, we will try to clarify what are the main advantages and disadvantages in each category of hierarchical algorithms, whether if it is cluster approach or chain approach. The main objective of hierarchical routing algorithms is to maintain the energy consumption in wireless sensor networks by using multi hope routing technique also using data aggregation and fusion in order to minimize the number of transmitted messages to the sink [1]. Low-Energy Adaptive Clustering Hierarchy (LEACH) was the first hierarchical routing algorithm [2]. After that, many algorithms were proposed using the main idea of LEACH such as Power-Efficient GAthering in Sensor Information Systems (PEGASIS) [3]. Mostly, in such algorithms, they use clustering technique or chaining technique and some of them try to take the advantage of both techniques and come up with a hybrid one that apply both clustering and chaining. We will classify the recent hierarchical algorithms based on the hierarchical style used in each algorithm into three categories, clustering, chaining, and Hybrid routing algorithms.

2. ROUTING DIFFICULTIES AND DESIGN CHALLENGES IN WSNS

Although there are numerous applications of wireless sensor networks, these networks have many limitations, such as limited energy, limited

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

computing abilities and limited bandwidth [4]. One of the main important goals to design a routing algorithm for WSNs is to perform data communication while trying to extend the lifetime of the network and prevent connectivity degradation by employing aggressive energy management techniques. Many factors can act as challenges in design a routing protocol for wireless sensor network and in order to achieve an efficient communication in WSNs, these challenges must be overcome. We summarize some of the challenges that can affect the routing protocol design in WSNs.

2.1 Nodes Deployment

Deployment of nodes in WSNs depends on the application of the network under consideration, it can be randomized or deterministic. In random deployment, the nodes are randomly scattered in the target area. Random deployment, in practice, is faster than deterministic method and more feasible. On the other hand it costs more sensor nodes to get more data accuracy [5]. In deterministic deployment, the nodes are manually installed according to predetermined distance. Although it is time consuming to deploy the sensor nodes manually but it can use fewer nodes in compare with random deployment with same accuracy level. An example of deterministic deployment is grid deployment [5].

2.2 Energy Consumption

During creating the nodes' infrastructure, the process of designing the routing protocol is greatly affected by energy constraint. The energy consumption in wireless radio transmission is proportional to the square of the distance of transmission and sometimes higher than this in the presence of obstacles. To decrease the consequences of this issue, we have to choose one of two options, multi-hop routing or single-hop routing. Multi-hop routing is more energy conservation but it can cause more overhead for topology management [5]. On the other hand, single-hop routing has less overhead but it consumes much more energy than multi-hop because of the direct transmission to the sink especially, when the sensor nodes are relatively far from the sink. So that, using multi-hop routing, in most applications, become unavoidable.

2.3 Data Aggregation

Since sensors might generate redundant data, they implement a data aggregation to eliminate the duplicated data and consequently decrease the number of transmissions [1]. The aggregated data is a combination of data collected from different sources and combined together in order to decrease the data packet size. There are some functions can be used to aggregate the data such as minima, maxima, duplicate suppression and average. This technique is very efficient to improve the wireless sensor network lifetime, especially when multi-hop routing is applied in that network, because data aggregation can be very applicable technique when it is used along with multi-hop routing technique at the same algorithm.

2.4 Scalability

In any sensing area, the number of nodes deployed, can reach hundreds or even thousands of sensors. Therefore, the routing scheme for such areas should consider this amount of nodes and must be able to handle all nodes in the sensing field. In addition, routing protocol should have the ability to handle any probable scaling for the network [4].

3. RELATED WORK

The increasing interest in wireless sensor networks and the ongoing development of new techniques and applications in such networks inspired some researchers' efforts for exploring the characteristics, applications, and communication techniques for this kind of networks. In this section, we demonstrate the features that differentiate our work from the previous work and clarify the difference in scop.

The objective of [1] is to introduce a comprehensive survey that makes a classification for different routing algorithms based on several architectural factors. It considered as a good introductory for readers who concern about routing in wireless sensor networks in general. They did not focus deeply on one type of routing algorithms. Authors of [6] introduced a survey on cluster-based routing protocol in wireless sensor networks, they highlighted clustering algorithms in WSNs and introduced a basic classification for clustering algorithms. The classification considered three characteristics of clustering technique, which are the factors they used to decide upon which nodes will function as cluster heads, the clustering mechanism is done by the base station (centralized) or by the sensor nodes (localized), and the hops inside the cluster, between cluster members and the cluster head. This work was focusing only on clustering technique and discuss some of the design challenges might face clustering algorithms. Therefore, it cannot be considered as a comprehensive survey on hierarchical routing algorithms. Our survey is more focused on hierarchical routing algorithms and classifies these

Journal of Theoretical and Applied Information Technology
30th April 2014. Vol. 62 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

algorithms into three categories based on the mechanism that they used in order to route the data from sensor nodes to the base station, which are chain, cluster, and hybrid routing algorithms. To the best of our knowledge, our survey is the first work that makes classification within the hierarchical routing algorithms based on the mechanism that they used for routing the data packets. Moreover, our work presents the current state of the art in routing algorithms by including the recently introduced hierarchical routing algorithms.

4. MAIN HIERARCHICAL ROUTING ALGORITHMS

In this part, we will highlight the two main hierarchical routing algorithms that all other algorithms presented based on their techniques. The two main hierarchical algorithms are LEACH and PEGASIS [2] [3].

4.1 LEACH

Low-Energy Adaptive Clustering Hierarchy (LEACH) is one of the most famous hierarchical routing protocols for wireless sensor networks. LEACH is a self-organizing, adaptive clustering algorithm that utilizes randomization to distribute the energy load equally among the sensor nodes in the network [2]. In this algorithm, the sensor nodes organize themselves in groups and each group act as a dependent cluster, then, one of those nodes functions as a local base station or cluster head for each cluster. The cluster head will be responsible for data collection from sensor nodes, data aggregation and finally sending the aggregated data to the base station. LEACH utilizes a randomized rotation to select the cluster head in order to rotate the job of being cluster head among other nodes. If the selection of cluster head were fixed during the network lifetime, then the unlucky nodes that chosen to be cluster head would die very fast.

Moreover, LEACH performs local aggregation for the data to minimize the number of the messages and the amount of the data that would be sent from the clusters to the base station. Furthermore, this data aggregation will decrease energy consumption and consequently increase network lifetime. The number of cluster heads is about 5% of the total number of nodes and nodes will join the closest cluster to them [2]. The main advantage of LEACH is more energy conservation about 7-8 factors in compare with direct transmission. The main disadvantage of LEACH is the nodes use single-hop to transfer the data to the cluster head then again by single-hop the cluster head will transfer the data to base station so that, it is still consuming more energy.

4.2 PEGASIS

Power-efficient GAthering in Sensor Information Systems (PEGASIS) is another hierarchical routing algorithm and considered as an improvement of LEACH. Instead of forming multiple clusters, PEGASIS forms chain that includes all nodes in the play field and select one node among them to be the chain head. In PEGASIS, the main idea, each node transmits to and receives from the closest neighbor and nodes take turn to function as chain head. The formation of the chain is done by: either the sensor nodes use greedy algorithm or the base station computes the chain and broadcast it to all sensor nodes within the network. The method of gathering data in PEGASIS, each sensor node receives data from the closest neighbor and fuses the received data with its own data then send it to other closest neighbor on the chain [3]. The way of selecting chain head is using the number of the current sensing round modulus the number of sensor nodes in the field. Therefore, each node will transmit and receive one data packet in each round and each node will function as chain head once every (N) rounds, where (N) is the number of sensor nodes in the field. PEGASIS achieves high amount of energy conservation but the main drawback in this algorithm is the excessive delay during data transmission.

5. RECENT WORKS BASED ON HIERARCHICAL ROUTING ALGORITHMS

In this topic, we will explore the works that have been done based on the idea of hierarchical routing algorithms and the strategies used to select the head nodes because it is the area of interest in our research. We divide the work based on the hierarchical approach that implemented in the routing algorithms whether they are based on i) cluster or ii) chain approach and in some cases, the routing algorithm is iii) hybrid because it is combining cluster and chain to introduce another type of hierarchical algorithms.

5.1 Chain Based Routing Algorithms

This In this section, we will explore some of the hierarchical routing algorithms for WSNs, which based on chain approach.

Authors in [7] tried to utilize chain approach to come up with routing algorithm that can prolong the network lifetime. They propose EBSCR (Energy-Based selective Cluster-Head Rotation)

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

protocol [7]. EBSCR consider as an improvement on PEGASIS protocol because they use the same way of constructing the chain, either done by using greedy route selection or centralized, when the base station computes the chain and broadcast the result to the sensor nodes. The only improvement is the way of selecting the chain head amongst other nodes, at the beginning of sensing rounds a node is selected as a chain head and a threshold for its energy is set. When the energy of the chain head become less than the threshold, another node with maximum remaining energy will be the chain head and so on. They use the following equation to calculate the threshold and this threshold.

$$\text{Threshold} = \mu\text{Ec} \qquad (1)$$

Where μ is a constant parameter and *Ec* is the energy left in the sensor node when it becomes the chain head. The advantage of this protocol is that it could achieve longer lifetime for the network. On the other hand the long chain, that connects all nodes and they should transfer their data from the first node in the chain to the chain head. Therefore, it will make great delay when the chain head and the first node are within the both ends of the chain.

Another improvement on PEGASIS algorithm has introduced in 2008. The authors introduced a new routing algorithm based on chain approach which is called SPER (Safe and Power –Efficient Routing Algorithm in Wireless Sensor Networks) [8]. This protocol divides the sensor network into multiple chains instead of one complete chain contains all nodes. It starts when each node on the edge of the sensor network try to connect to the base station using the shortest path by utilizing the greedy algorithm. If two chains, through the way to base station, meeting each other at a certain point, then they will form one chain to the base station. There is a threshold of the weight for each node on the chain, the node should compute the value of the threshold to decide if it will accept when a nearby node asks to join the chain, and it will unconditionally accept if that node has compelling sign. In case, the node was rejected to join a nearby node, it will search for other nearby nodes and ask for connection. After the node is rejected to be connected from all nearby nodes, it will generate a compelling signal and connect to the node that has maximum energy. This algorithm can consider as more load balancing than PEGASIS because when the node generate a compelling signal it will connect to the node that has maximum energy and those nodes with minimum energy will only serve as end nodes in order to maximize their lifetime [8].

The disadvantage of this algorithm is the great overhead that happened during threshold calculations in each round.

In 2010 the authors of [9] introduced routing algorithm based on chain technique that called EECB (Energy Efficient Chain Based Routing Protocol). They come up with an idea to solve the problem that happen when some nodes are relatively distant from their neighbours then when they are elected as a chain head their neighbours will consume more energy to connect to them. This algorithm works within three stages, chain construction, leader election, and data transmission. In chain construction stage, the nodes will use the greedy algorithm to form the chain starting with the furthest node to the base station. The link between node *v* and node *v*+1 is denoted as the distance threshold [9]. In case, the distance between neighbour nodes is equal to or more than $d_{threshold}$, then the link between those neighbours called *LL*.

$$d_{\text{threshold}} = \propto \max(d_v), v = 1,2, \ldots i-1 \qquad (2)$$

Where the distance of the link *v* is *dv*, α denotes the constant that defined by the user, $d_{threshold}$ is the distance threshold for *LL*. If $di <= d_{threshold}$, it means that the link between the node *i* and node *i*+1 is not *LL* and they can connect to each other and join the chain simply by following the greedy algorithm. If *di* > *d* threshold it means that the link is *LL* [9]. If the link *i* is equal to *LL*, node *i*+1 is prohibited to connect directly to node *i*. So that, node *i*+1 will search for the nearest node to connect with. If the node *i* not the nearest node, node i will be one of the ends for the chain and node *i*+1 will connect to the nearest node. If node I is the nearest node to node *i*+1, in this case *LL* link cannot be avoided and node *i*+1 will connect to node *i*. After constructing the chain, a head for the chain will be selected based on the residual energy in each node and the distance between the node and the base station.

In 2011, authors introduced an algorithm called IEEPB (Improved Energy Efficient Based Protocol) [10]. This algorithm works within three stages; chain construction, leader selection stage and data transmission stage. The base station will broadcast a hello message to all nodes to get some main information about the sensor nodes within the network like node *ID* and the distance between the node and the base station. The node that has the furthest distance to the base station will be the first node in the chain and then it will search for the nearest node to it to start forming the chain. The leader selection is based on weighting method that

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

includes both the remaining energy in the node and the distance between the node and the base station. The nodes will estimate the distance to the base station based on the signal strength from the base station to the sensor nodes, as we mentioned earlier, the base station will broadcast hello message at the initial work so each node can estimate the distance to base station by using that hello message. Then each node will calculate the combined weight that includes the distance to base station and the residual energy for each node. The node that has the minimum combined weight will be the chain head. The chain in IEEPB has more than two ends. Each node will receive data from its closest neighbour and then fuses this received data with its own data and send it to the closest neighbour. This algorithm achieved an improvement on extending the lifetime of the network, on the other hand, it has great overhead caused by weight calculation for each node in each round.

Another algorithm has presented in 2011, Chain-Chain Based Routing Protocol (CCBRP) introduced based on chain mechanism [11]. CCBRP divides sensors network into horizontal chains and works within two stages. First stage starts when one node in each vertical chain randomly selected to act as a chain head for other chain members. Every node will send its data to its closest neighbour, when node receives the data from its neighbour, fuses the received data with its own data and forward it to the next neighbour and so on [11]. This process will continue in each chain until all data are delivered to the chain head. The second stage starts when one of chain heads randomly selected to act as a main head for the network. All chain head nodes will form a vertical chain, one of the node in the vertical chain will be selected randomly to function as a main head for the network [11]. first node in the vertical chain will forward its data to its closest neighbour and same as first stage, when node receives the data, fuses the data with its own data and forward it to its closest neighbour and so on until all data be gathered in the main head node. When the main head receives this data, it will fuse it with its own data and forward it to the base station.

The authors of [12] introduced energy efficient chain based network (ECBSN) that consists of multiple chains and operates in two stages: chain formation stage and data transmission stage. The way of forming the chains is as follow, a source node *S* broadcast a route request message for one hop *TTL* = 1, this message containing threshold energy value. When a node receives the route message, compare its energy with the threshold energy. If the energy of the node is more than the threshold, an active reply message will be sent to the source node [12]. All active reply messages scanned by the source and the neighbour that has the shortest path will be selected to route the data. The number of chains in the network that has *N* number of nodes will be *N*/*M*, where *M* is the length of each chain. The chain leader selection based on the residual energy in each node, and the main leader will be selected based on the distance to the base station, the leader that has the minimum distance to base station will be the main leader of the network. The main leader will pass the aggregated data to the base station, and the chain construction process will happen again when 20% of the nodes die. Therefore, the selection of the leaders and main leader will also have delay, this delay gain some advantages in compare with the selection of leaders in each round, these advantages are: i) Less communication overhead. ii) Reducing the time wasted for selecting the leaders.

Authors in [13] introduced new routing algorithm that focus on solving the problem of the chain head fault. They introduced EBRT (Energy-Based Re-transmission Algorithm of the Leader node's neighbour node for Reliable Transmission in the PEGASIS) to function as a reliable routing algorithm. In this approach, if the base station dose not receives the sensed data within the calculated time, it will broadcast an error message to all nodes in the network. When the two nearest neighbours of the leader node receive the error message, they will make direct connection between them and compare their remaining energy. The neighbour that has more remaining energy will be the leader of the chain. The disadvantage of this algorithm is the overhead that caused by the calculations of transmission required time and energy comparisons between neighbours.

In 2013 a new technique is introduced based on sink mobility and chain routing approach. The authors in [14] proposed MIEEPB (Mobile sink improved energy-efficient PEGASIS-based routing protocol). This algorithm divide the sensing field to 4 regions and nodes in each region will form their own chain. Sinks sends hello message to all nodes and the node that has the furthest distance to the base station will represent the first node in the chain in each region. Then each node will search for the nearest node to connect with until forming the complete chain. The selection of the chain head is based on the weight of each node *Qi*. Where *Qi* is represented by dividing the remaining energy for the node *Ei* with its distance to the base station *Di*.

Journal of Theoretical and Applied Information Technology
30th April 2014. Vol. 62 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$Qi = Ei/Di \quad (3)$$

In MIEEPB, they assumed that the sink has unlimited amount of energy and it has the ability to move in a fixed path from one region to another, stay in each region for a specific period of time and collect the sensed data from each location [14].

In 2013, an improvement on ECBSN is introduced [15]. The new routing algorithm IECBSN (Improved Energy Efficient Chain Based Sensor Network) adopts new factor for the functionality of selecting the chain head, this factor called selection value *SV*. IECBSN consists of four stages; network construction, chains formation, leader selection, and data transmission. A source node *S* will broadcast a hello message to other nodes, when the S receives the reply from other nodes it can compare the distance for each node based on the signal strength from the sensor nodes. Then it will start forming the chains by connecting to the closest neighbour. Each network with *N* number of nodes, if the length of each chain is equal to *M*, the number of chains in each network *P* is equal to *N* divided by *M*. After forming the chains, the next step is to select the leader node for each chain. They select the leader based on the selection value *SV* [15].

$$SV_i = E_{r(i)} \frac{1}{adist\,(n_i, n_{BS})} \quad (4)$$

Where *Er (i)* is the remaining energy of each node in the chain, adist (n_i, n_{BS}) is the distance from the node *i* to the base station. In each chain, the node that has the highest *SV* will be the chain leader. The node that has minimum distance to the base station will be the main leader (leader of the chain leaders). The main advantage of IECBSN algorithm is gaining more energy balancing by adopting new factor for selecting the head node, which called *(SV)*.

In 2014, a new routing algorithm is presented based on chain technique and its applicable for grid topology. The protocol called (TSCP) Two Stages Chain Protocol for wireless sensor networks [16]. TSCP algorithm works within two stages, forming the horizontal chains, forming the vertical chain. The first stage starts with dividing the sensors into horizontal chain and selecting one of the nodes in each chain to work as a head for other sensor nodes in the chain. The selection way is done periodically (e.g. in first sensing round, first node in each horizontal chain works as the chain head and so on) [16]. The second stage starts when all chain head nodes form a vertical chain and one of these nodes will be selected to function as a main head for the entire network. The way of selecting the main head node is based on the residual energy in each node in the vertical chain, the node with highest remaining energy will be selected to function as a main head [16].

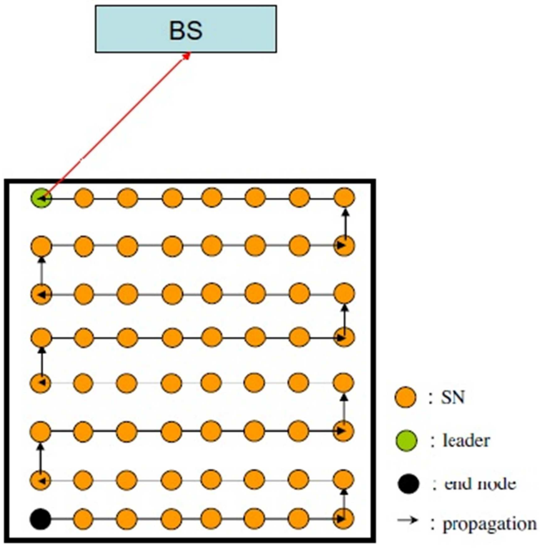

Fig.1. General Scenario for chain routing technique.

5.2 Cluster based routing algorithms

This In this part, we will review some of the previous work that has been done in order to improve routing algorithms in WSNs, using clustering approach. In [17] the authors proposed new technique to select the cluster head, which is using Decision Trees for Wireless Sensor Networks. The proposed algorithm is based on four different factors; the distance between the node and the centre of the cluster, residual energy of the node, mobility of the node, and finally Vulnerability index. They categorized the nodes in two classes, nodes in class I can function as a cluster head while nodes in class II cannot be a cluster head. The classification is conducted by using the factors that we mentioned earlier. This algorithm operates in two phases, start-up phase and steady – state phase. In the start-up phase, the base station sends hello message to all sensor nodes and then sensor nodes will reply the hello message and send their control information. The base station will use the control information and run the decision tree to decide the proper sets of the cluster heads amongst all sensor nodes and then send the list that contain the cluster heads. Each cluster head will send a notification to its cluster members. After that, each sensor node attaches itself to one cluster head. The steady state phase including the transmission of data from normal sensor nodes to

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the cluster heads and then from the cluster heads to the base station.

In 2009 a routing algorithm based on LEACH was introduced, Re-cluster-LEACH, Re-cluster Control Algorithm Based on Density for Wireless Sensor Network [18]. In this algorithm, the authors tried to enhance the operation of leach algorithm by solving the problem of unbalanced clusters. They consider each cluster has a certain circular area and just those nodes that located within that area can join the cluster. The factor that controls the areas construction is the density of the nodes. So each spot that has high density of sensor nodes will be considered as an independent area. Each node has a timer to be a cluster head and this timer is randomly chosen. When a node receives the message from the cluster head, it will compare the distance to the cluster head and the diameter of its area. If that distance is less or equal to the diameter, it will send an ACK message to cluster head and join the cluster. If the distance to cluster head is more than the diameter of its area, it will ignore the message and wait for a message from another cluster head. In case the cluster head did not receive any ACK message, it will degrade to an ordinary node [18].

In 2009, another improvement on LEACH protocol was proposed, Leach-H is an improvement over LEACH protocol that uses simulated annealing algorithm as an optimized method to select the cluster heads set in the first round [19]. After the first round, new cluster heads will be selected among the nodes that belong to the same cluster. They make two collections, the first one is collection *C*, which represents the node that are currently cluster heads, the second collection is *C'*, which represents the nodes that are new cluster heads. The energy consumed by collection *C* is represented by f (*C*), while f *(C')* represents the energy consumed by collection (C'). If *f(C') > f(C),* leach-H will set collection (C') to be the next cluster head set. Otherwise, leach-H will determine *C* based on the probability *p* [19].

Authors in [20] have done new improvement over LEACH algorithm based on energy. They developed new method to determine the number of nodes that operate as a cluster heads in each sensing round. The improved protocol, BEC (Based Energy Clustering) calculates the number of clusters based on the following function:

$$k = \frac{\sqrt{27} * M}{2 * r^2 * \pi^2} \qquad (5)$$

Where *k* is the number of cluster heads, *M* is the area of monitoring region, and *r* is the perceived radius of the nodes. The electing strategy of the cluster head is based on the remaining energy in the nodes and it is done by applying the following equation.

$$T(n) = \frac{E_i * k}{E_{total}} \qquad (6)$$

Where *T (n)* is energy threshold of a node to be cluster head, the node with highest threshold would have highest probability to be cluster head. E_i is the residual energy of the node, E_{total} is the total residual energy in the network, and *k* is the number of cluster heads. The way of forming the cluster is the same as the way of selecting cluster heads in LEACH.

Authors in [21] have introduced a routing algorithm consider as an extension of LEACH, that developed a strategy for selecting the cluster head and the nodes that allowed sending to the cluster head in a certain sensing round. The algorithm called W-LEACH, it sets a weight Wi for each sensor node. The weight *Wi* depends on two factors which are the density *di* and the remaining energy. The density *di* represents the ratio of all live sensor nodes that allocated in the range *r* of sensor node to the total number of live sensor nodes in the network. The area with high density will have a short distance to communicate with the cluster head. Hence, there will be less energy consumption. On the other hand there would be a redundant of sensed data because nodes are close to each other so that, their data would be highly correlated [21]. To determine the weight for each node, they developed the following function.

$$w_i = \begin{cases} e_i * d_i, & \text{if } d_i > d_{threshold} \\ d_i\,, & \text{otherwise} \end{cases} \qquad (8)$$

Where *di* = (1+alive sensor in the range *r*) / *n*, and $d_{threshold}$ is a density threshold to define the set of sensor nodes in a low density area [21]. The nodes in low density would be highly probable to be chosen to send their data and this will shorten their life. In order to extend their life as long as possible, W-LEACH does not require nodes with *di* <= $d_{threshold}$ to transmit their data in each round. Even though, W-LEACH can achieve more energy conservation, it is not suitable for many applications especially for real time applications because some nodes in low-density areas will not send their data in every sensing round and this will lead to non-negligible data loss.

In 2011, a multiple cluster-heads routing protocol (MCHRP) was proposed for load

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

balancing. A vice cluster heads used in this algorithm for data acquisition, fusion, and transmission. MCHRP uses the strategy of cluster heads rotation. The rotation operation is done within two stages, clustering stage and transmission stage [22]. The selection of main cluster heads is done by the base station using the function *F(i)*. After this, the cluster heads broadcast a message to all sensor nodes in the network to join the cluster. Then, the main cluster heads select vice cluster heads and alternative cluster heads among the nodes that belong to its cluster based on the function *G (i, N)*. The jobs of vice cluster head are data collection, fusion, and then transfer the data to the main cluster head as we mentioned earlier while the job of the main cluster head is to transfer the data to the base station.

$$f(i) = p_{remain}(i) + \frac{1}{p_{frequency}(i)} \quad (8)$$

$$G(i, N) = p_{remain}(i) + \frac{1}{p_{distance}(i, N)} + \frac{1}{p_{frequency}(i)} \quad (9)$$

Where $P_{remain}(i)$ is, the current energy of *node(i)* divided by the initial energy of the node. $P_{frequency}(i)$ indicates how many times the *node(i)* have become main cluster head, alternative cluster head or vice cluster head. $P_{distance}(i)$ is the information about *node(i)* to the main cluster head [22]. This algorithm is successful to gain more load balance than basic LEACH but it has more calculations to select the main cluster head, vice cluster head and alternative cluster head.

Authors in [23] proposed an improvement on LEACH to optimize the selection of cluster head nodes, they named the improved algorithm LEACH-R. They improved the threshold equation that decides which node can be a cluster head.

$$T_R(n) = \begin{cases} \frac{p}{1 - p * \left(r \bmod \frac{1}{p}\right)} [\delta p + 1 - \delta p] * \frac{E_{residual}}{E_o} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

Where E_o is the initial energy of each node, $E_{residual}$ is the remaining energy in each node, δ is the number of consecutive rounds that the node has selected as a cluster head. Once node selected as a cluster head, the value of δ will reset to be (0). After forming the clusters, they will select a relay node *R* among the cluster heads based on the value of λ, which depends on the residual energy of the cluster head and its distance to the base station.

$$\lambda = \frac{E_{residual}}{d_{to\ Bs}} \quad (11)$$

The cluster head with highest λ will function as *R*. Cluster heads, those are not selected as *R*, will send their data to *R* instead of sending their data to the base station and *node(R)* will send the aggregated data to the base station.

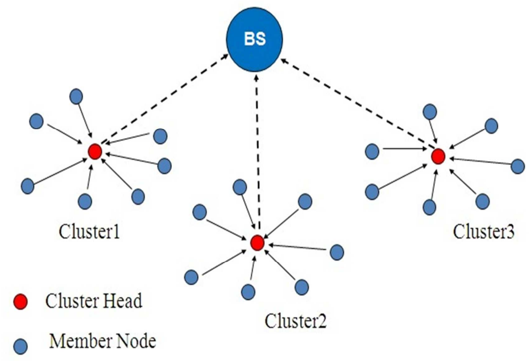

Figure 2: General Scenario for cluster routing technique

5.3 Hybrid hierarchical routing algorithms

In this part, we will highlight the recent works that has been introduced in WSNs routing algorithms based on mixing the cluster approach together with chain approach. In [24] the authors introduced a new routing algorithm that combine the idea of clustering with chain approach to reduce energy consumption in clustering algorithms. This algorithm works within two phases, initialization phase and transmission phase. The initialization phase involving chains and clusters formation, which is done by the base station in order to get better distribution for the nodes between the clusters [24]. The selection of cluster head is based on the residual energy of the node and its distance to the base station. The formation of the chain started with the farthest node in the cluster then it will connect to the closest node until all nodes are in the chain. Although this algorithm achieves an enhancement on energy saving, they did not introduce a method or a factor that control the number of nodes in each cluster.

In 2009, the authors of [25] proposed an algorithm that divides the sensor nodes into suppositional cells, the algorithm termed CFDASC (Chain-based Fast Data Aggregation Algorithm Based on Suppositional Cells). Nods within each cell will form a cluster and operate as a cluster head in turn. Cluster head nodes will connect to each other in a vertical direction, creating vertical chains, nodes within one chain function as a chain head,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

that responsible for data aggregation and data transmission to the base station, in turn. The main achievement of this algorithm is an improvement over PEGASIS in term of latency in about 50.03%.

In 2010, the authors proposed a hybrid algorithm in order to prolong the network lifetime called as CBERP (Cluster Based Energy Efficient Routing Protocol for Wireless Sensor Network)[26]. CBERP protocol divides the network into clusters and selects a cluster head node for each cluster, the node with highest *E* value will be selected by the base station as a cluster head.

$$E = \frac{E_{resi}}{E_{init}} \times CH_{pnt} \quad (12)$$

Where E_{resi} is the remaining energy of the node, E_{init} is the initial energy of the node, and CH_{pnt} is the percentage of the nodes to function as a cluster head. All nodes send their data to the cluster head, the cluster head aggregate the received data with its own data. After all the data packets are gathered in the cluster head nodes, cluster head nodes form a chain based on greedy algorithm. Each node in the chain received the data from the closest neighbour, aggregates the received data with its own data and sends it to the next closest neighbour until all data are gathered in the closest node to the base station. The closest node will aggregate the received data with its own data and send it to the base station.

Another routing algorithm was presented in 2010, which is called CCM (Chain-Cluster based Mixed routing) [27]. This algorithm is applicable for grid topology and works within two stages. In the first stage, which is chain based routing, the network will be divided into horizontal chains and one node will be selected as a chain head periodically in each sensing round. Each node will send its data to its closest neighbour, the next node will aggregate the received data with its own data and send it to the closest neighbour until all data be delivered in the chain head. The second stage, which is cluster based routing, all chain head nodes will form a cluster and one node will be selected to be the cluster head. The selection of cluster head is based on the remaining energy in each node in the cluster, the node with maximum remaining energy will be selected as cluster head [27].

Authors in [28] introduced an improvement on LEACH using chain approach among cluster heads. The idea of this algorithm is simple and straightforward. At the beginning, the nodes will be organized in a clusters and the node with maximum remaining energy will function as a cluster head. After forming the clusters, each node will send its data to the cluster head, each cluster head will aggregate the received data with its own data. The cluster heads will form a chain started from the furthest node to base station, the node that has closest distance to base station will act as a chain leader. Cluster heads will send their data to the nearest neighbour on the chain until reach the chain leader. The chain leader will aggregate the received data with its own data and then send it to the base station.

The improvement in this algorithm is in power saving, the reason of this improvement is instead of all cluster heads send their data directly to the base station and definitely this process will consume too much energy, only one node will transmit the data to base station, which is the chain leader. This technique achieves more energy conservation and decrease the number of transmissions to the base station [28].

In 2012, authors of [29] introduced routing algorithm termed ECCP (Energy Efficient Cluster-Chain based Protocol). The main objectives of this protocol is to reduce the energy consumption of the sensor nodes, extend the network lifetime, and achieve an even distribution for energy load among all sensor nodes. The ECCP algorithms works according to three phases, clustering phase, chain formation phase, and data transmission phase. In clustering phase, each node broadcast a message that contains its location and its remaining energy. The selection of cluster head is based on the remaining energy in each node, nodes will join the closest cluster based on the signal strength from the cluster head message.

Clusters in ECCP are fixed, it will not repeat the formation of cluster in each round, just when a node die, the cluster head will notify the base station that the cluster need to be reconstructed in the next round. Nodes in each cluster will form a chain and send their data to the cluster head. After aggregating the data in the cluster head, the base station will form a chain that contains all cluster heads based on greedy algorithm and started from the furthest node to the base station the second farthest until reach the nearest node to the base station. The node that has closest distance to the base station will act as the main leader and send the aggregated data to the base station. This algorithm achieves advantages over LEACH protocol in terms of load balancing, network lifetime, energy consumption and network stability.

In [30], the authors proposed an algorithm termed REC+ (A Reliable and Energy-efficient Chain-cluster Based Routing Protocol). REC+ is operates within three phases, cluster formation

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

phase, cluster head selection and chain cluster forming phase, and steady state phase. This algorithm divides the network into some clusters based on the value of Y-coordination for each node, the base station deal with each cluster independently without loss the generality [30]. After that each cluster will be divided into sub-clusters and inside each cluster, nodes transmit their data to the closest neighbour until the data reach the cluster head based on chain approach. This algorithm achieves more energy conservation and end-to-end reliability.

The authors of [31] introduced a routing algorithm named as CCCP (Chain-Clustered Communication). CCCP is an improvement over PEGASIS and LEACH based on chaining and clustering concepts. They divided the sensor network into 5 clusters and the sensor nodes will form a chain inside each cluster starting from the furthest node to the base station. The cluster heads will select a leader node among them based on the formula (r mod n), r is the number of sensing rounds and n is the number of cluster heads [31]. The selection of the leader node is done every sensing round while the selection of new cluster head is done when the current cluster heads energy becomes less than $k*E$, where k is a constant and E is the initial energy for the node. The new cluster head will be selected depends on the remaining energy, the node that has the highest remaining energy will be the new cluster head.

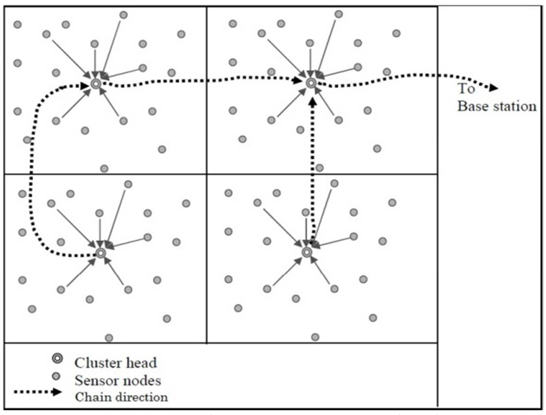

Figure 3: General Scenario For Hybrid Routing Technique

6. OPEN RESEARCH ISSUES

Although many researches have been done to increase the lifetime of WSNs by improving the routing algorithms, there are still no optimum routing algorithm can achieve energy efficiency and extend network lifetime without more overhead on the network. On the other hand the factors that they used to improve the network lifetime by selecting cluster head or chain head such as remaining energy, distance, and density are not very reliable when they implemented them. For instance, some nodes have a little more energy but very far from the base station or they have less nodes density and same things for other factors. The reason of these limitations and trade-offs in the performance versus the lifetime is that they used fixed logic to decide upon these factors. Therefore, more research should be done on using artificial intelligence with routing algorithms. This can lead to the emergence of new smart routing algorithms that can be almost optimal algorithms. However using such technique will come up with more overheads to the sensor nodes, but this overheads can be avoided by making the clustering or chaining technique centrally done (by the base station).

7. CONCLUSION

Decreasing energy consumption without compromising the performance of the routing algorithm (i.e. load balancing and network stability), considered as an important factor to enhance the robustness of WSNs routing algorithms. Increasing the applications of wireless sensor networks inspired many researchers to explore the routing algorithms that developed for such networks and study their architectures and techniques. On the other hand, the previous work did not focus on deep details in hierarchical technique and did not investigate the types of hierarchical techniques. The significant aspect in our work is that we classified the hierarchical routing algorithms based on the way of routing the data packets for three categories, chain, cluster, and hybrid routing algorithms. Broadly speaking, hierarchical algorithms utilize two main approaches to route the data from sensor nodes to the base station, which are clustering approach and chain approach. Generally, each approach has one main drawback. The main drawback of clustering algorithms is the high-energy consumption due to the long distance from the each node to the cluster head and the long distance from the cluster head nodes to the base station. Whereas, the main drawback in chain approach is the delay that happens because of the long chain that includes all sensor nodes. In order to overcome these drawbacks, some researchers introduced hybrid algorithms that combine cluster and chain approaches within one algorithm.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Regardless of all factors we mentioned, the selection of routing algorithm mainly depends on the network application. In this section, we described the hierarchical algorithms and we classified them into three categories, chain, cluster, hybrid routing algorithms.

As part of our future work, we will try to present a hierarchical routing algorithm by implementing fuzzy logic and genetic algorithm in the mechanism of the algorithm.

REFRENCES:

[1] Akkaya, Kemal, and Mohamed Younis. "A survey on routing protocols for wireless sensor networks." *Ad hoc networks,* vol. 3.p.p.325-349, (2005)

[2] Heinzelman, Wendi Rabiner, Anantha Chandrakasan, and Hari Balakrishnan. "Energy-efficient communication protocol for wireless micro sensor networks." In *System Sciences, Proceedings of the 33rd Annual Hawaii International Conference, 2000*, pp. 10-pp.

[3] Lindsey, Stephanie, and Cauligi S. Raghavendra. "PEGASIS: Power-efficient gathering in sensor information systems." In *Aerospace conference proceedings, 2002. IEEE*, vol. 3, pp. 3-1125.

[4] Al-Karaki, Jamal N., and Ahmed E. Kamal. "Routing techniques in wireless sensor networks: a survey." *Wireless Communications, IEEE*, vol.11. p.p. 6-28, (2004).

[5] Bellavista, Paolo, Jiannong Cao, Falko Dressler, Domenico Ferrari, Mario Gerla, Hisashi Kobayashi, Sergio Palazzo et al. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*. Springer, 2010.

[6] Wei, Chunjuan, Junjie Yang, Yanjie Gao, and Zhimei Zhang. "Cluster-based routing protocols in wireless sensor networks: a survey." *In Computer Science and Network Technology (ICCSNT), 2011 International Conference on, vol. 3, pp. 1659-1663. IEEE, 2011.*

[7] Hu, Feng Ming, Yun Ho Kim, Kyung Tae Kim, Hee Yong Youn, and Chang Won Park. "Energy-based selective cluster-head rotation in wireless sensor networks." In *Proceedings of the 2008 International Conference on Advanced Infocomm Technology*, p. 121. ACM, 2008.

[8] Zhao, Yan, and GaocaoXu. "SPER: A Safe and Power-Efficient Routing Algorithm in Wireless Sensor Networks." In *High Performance Computing and Communications. HPCC'08. 10th IEEE International Conference*, 2008, pp. 1025-1029.

[9] Yu, Yongchang, and Yichang Song. "An Energy-Efficient Chain-Based Routing Protocol in Wireless Sensor Network." In *Computer Application and System Modelling (ICCASM), International Conference on*, 2010, vol. 11, pp. V11-486.

[10] Sen, Feng, Qi Bing, and Tang Liangrui. "An improved energy-efficient pegasis-based protocol in wireless sensor networks." In *Fuzzy Systems and Knowledge Discovery (FSKD), Eighth International Conference on*, 2011, vol. 4, pp. 2230-2233.

[11] Ali, Samia A., and Shreen K. Refaay. "Chain-Chain Based Routing Protocol." *IJCSI International Journal of Computer Science Issues*, vol. 8, pp. 1694-0814, (2011).

[12] S. Mahajan and J. Mahotra, "A novel chain-based wireless data sensor network (ECBSN) technique*," International Journal of Computer Science and Telecommunications*, vol. 2, pp. 83-97, November 2011.

[13] Lim, Se-Jung, A. K. Bashir, So-Yeon Rhee, and Myong-Soon Park. "Energy-Based Re-transmission Algorithm of the Leader Node's Neighbour Node for Reliable Transmission in the PEGASIS." In *Ubiquitous Computing and Multimedia Applications*, vol. 150, pp. 120-128. Springer Berlin Heidelberg, 2011.

[14] Mohsin RazaJafri, Nadeem Javaid, Akmal Javaid, Zahoor Ali Khan, "Maximizing the Lifetime of Multi chain PEGASIS using Sink Mobility*", World Applied Sciences Journal*, vol. 21, pp. 1283-1289, 2013.

[15] Shilpa Mahajan, Jyoteesh Malhotra, Sandeep Sharma, "Improved Enhanced Chain Based Energy Efficient Wireless Sensor Network", *Wireless Sensor Network*, vol. 5, pp. 84-89, 2013.

[16] Husam kareem, S.J.Hashim, Shamala Subramaniam, Aduwati Sali. "Energy Efficient Two Stage Routing Protocol (TSCP) for Wireless Sensor Networks." *Journal of Theoretical and Applied Information Technology* 59, no. 2, PP. 442-450, (2014).

[17] Ahmed, Ghufran, Noor M. Khan, Zubair Khalid, and Rodica Ramer. "Cluster head selection using decision trees for wireless sensor networks." In *Intelligent Sensors, Sensor Networks and Information Processing. ISSNIP. International Conference*, pp. 173-178. IEEE, 2008.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[18] Yi, Gao, Sun Guiling, Li Weixiang, and Pan Yong. "Recluster-LEACH: A recluster control algorithm based on density for wireless sensor network." In *Power Electronics and Intelligent Transportation System (PEITS), 2nd International Conference on*, vol. 3, pp. 198-202. IEEE, 2009.

[19] Wang, Wei, Qian ping Wang, Wei Luo, Mengmeng Sheng, Wanrong Wu, and Li Hao. "Leach-H: An improved routing protocol for collaborative sensing networks." In *Wireless Communications & Signal Processing. WCSP. International Conference*, pp. 1-5. IEEE, 2009.

[20] Wang, Linlin, Jie Liu, and Wei Wang. "An improvement and simulation of leach protocol for wireless sensor network." In *Pervasive Computing Signal Processing and Applications (PCSPA), First International Conference*, pp. 444-447. IEEE, 2010.

[21] Abdul Salam, Hanady M., and Layla K. Kamel. "W-LEACH: Weighted Low Energy Adaptive Clustering Hierarchy aggregation algorithm for data streams in wireless sensor networks." In *Data Mining Workshops (ICDMW), IEEE International Conference*, pp. 1-8. IEEE, 2010.

[22] Tang, Da, Xiang Liu, Yuqian Jiao, and Qianjin Yue. "A load balanced multiple Cluster-heads routing protocol for wireless sensor networks." In *Communication Technology (ICCT), IEEE 13th International Conference*, pp. 656-660, 2011.

[23] Kamath, H. Srikanth. "Energy Efficient Routing Protocol for Wireless Sensor Networks." *International Journal of Advanced Computer Research*, vol. 3, pp. 95-100, 2013.

[24] Bilami, Azeddine, and Djallel Eddine Boubiche. "A hybrid energy aware routing algorithm for wireless sensor networks." In *Computers and Communications, ISCC. IEEE Symposium*, pp. 975-980. IEEE, 2008.

[25] Wu, Hao, Yi-ming Ding, and Zhangdui Zhong. "A chain-based fast data aggregation algorithm based on suppositional cells for wireless sensor networks." In *Power Electronics and Intelligent Transportation System (PEITS), 2nd International Conference*, vol. 1, pp. 106-109. IEEE, 2009.

[26] Lee, H. Y., Kyoung Oh Lee, H. L. Lee, and Aries Kusdaryono. "CBERP: Cluster based energy efficient routing protocol for wireless sensor network." In *Proc. 12th Int. l. Conf. Networking, VLSI and Signal Processing. University of Cambridge UK*, pp. 24-28. 2010.

[27] Tang, Feilong, Ilsun You, Song Guo, Minyi Guo, and Yonggong Ma. "A chain-cluster based routing algorithm for wireless sensor networks." *journal of intelligent manufacturing* , vol. 23, pp. 1305-1313, 2012.

[28] Pal, Subhajit, Debnath Bhattacharyya, and Tai-hoon Kim. "Energy Efficient, Chain Based Clustering Routing Protocol for Wireless Sensor Networks." In *Security-Enriched Urban Computing and Smart Grid*, pp. 472-481. Springer Berlin Heidelberg, 2010.

[29] Sheikhpour, Razieh, and Sam Jabbehdari. "A Cluster-Chain based Routing Protocol for Balancing Energy Consumption in Wireless Sensor Networks." *International Journal of Multimedia & Ubiquitous Engineering*, vol.7, pp. 1-16, 2012.

[30] Taghikhaki, Z.; Meratnia, N.; Havinga, P.J.M., "A reliable and energy-efficient chain-cluster based routing protocol for Wireless Sensor Networks," *Intelligent Sensors, Sensor Networks and Information Processing, IEEE Eighth International Conference*, pp.248,253, 2-5 April 2013

[31] Sharma, Swati, and Pradeep Mittal. *"Chain* Clustered *Communication Protocol for WSN*." International Journal of Computer Trends and Technology (IJCTT), vol. 4, pp. 1-4, May, 2013.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: Simulation software, performance metric(s) and routing technique used in each algorithm

Algorithm name	Simulation tool	Performance metric(s)	Routing technique
LEACH	MATLAB	Network lifetime	Cluster
PEGASIS [3]	C language	Network lifetime	Chain
EBSCR [7], SPER[8], EECB[9]	Not mentioned	Network lifetime	Chain
IEEPB[10]	MATLAB	Network lifetime	Chain
CCBRP [11]	JAVA Language	Energy consumption, transmission delay, Energy x Delay	Chain
ECBSN[12], IECBSN[15]	MATLAB 7.0	Network lifetime, Energy consumption	Chain
EBRT[13]	Not mentioned	Network lifetime, Reliability	Chain
MIEPPB[14]	MATLAB	Energy consumption, Network lifetime	Chain
TSCP [16]	MATLAB	Load Balancing, Network Lifetime and stability interval, energy consumption	Chain
Ahmed, Ghufran, et al [17]	MATLAB 7.0	Network lifetime	Cluster
Recluster – Leach [18]	OMNet++	Network lifetime	Cluster
Leach-H [19]	NS-2	Energy consumption, Network lifetime	Cluster
BEC[20], Leach-R[23]	MATLAB	Energy consumption, Network lifetime	Cluster
W-Leach[21]	C-Language	First node and Last node died, Number of alive sensors, Remaining energy	Cluster
MCHRP[22]	OMNet++ 4.0	Network lifetime, Energy consumption	Cluster
Bilami, Azeddine, and DjallelEddineBoubiche [24]	NS-simulator	Network lifetime	Hybrid
CFDASC[25]	OMNet++	Transmission delay, Network lifetime	Hybrid
CBERP[26]	NS-2	First node died	Hybrid
CCM [27]	SWANS	Energy consumption, transmission delay, Energy x Delay	Hybrid
Pal, Subhajit, et al [28]	Not mentioned	Energy dissipation, Network lifetime	Hybrid
ECCP[29]	MATLAB	Network lifetime, Stability period, Instability period, Load balancing, Energy consumption, Network throughput, Communication overhead	Hybrid
REC+ [30]	Not mentioned	Network lifetime, Energy consumption, Delay, Energy x Delay, End-to-End reliability	Hybrid
CCCP[31]	MATLAB	Network lifetime	Hybrid