



Journal of Applied Sciences

ISSN 1812-5654

science
alert

ANSI*net*
an open access publisher
<http://ansinet.com>

A Survey on Cluster-based Routing Protocols in Wireless Sensor Networks

Nooshin Nokhanji and Zurina Mohd Hanapi

Department of Communication Technology and Network, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Serdang, 43400, Selangor, Malaysia

Abstract: In recent years, Wireless Sensor Networks (WSNs) have become interesting for a wide range of applications and a hot research area. Cluster-based routing protocols are more adequate for WSN among different types of routing protocols due to the higher energy utilization rate and being more scalable. In this study, cluster-based routing protocols have been reviewed. In addition, the advantages and objectives of this group of routing protocols are sketched out. Furthermore, several cluster-based routing protocols are analyzed in detail and compared according to the several important metrics. In conclusion, the study is summarized and finalized with some directions for future cluster-based routing protocols.

Key words: Wireless sensor networks, cluster-based routing protocols, cluster head, cluster formation, intra-cluster communication, inter-cluster communication

INTRODUCTION

The combination of innovations in Micro-Electro-Mechanical systems and the development of low power radio technologies, in addition to the advances in low-power embedded micro-controllers, result in the appearance of a specific type of *ad hoc* network called a Wireless Sensor Network (WSN) (Heinzelman *et al.*, 2000; Lindsey *et al.*, 2002; Younis and Fahmy, 2004). WSNs have received a significant amount of attraction in recent years, because of their benefits in sensing and gathering different kinds of data in various environments. WSNs consist of hundred to thousand sensor nodes which measure a property from the environment as well as processing and transmitting the collected data to the Base Station (BS). They are being used in a varied range of applications, such as disaster relief operations, biodiversity mapping, agriculture, medicine and health care and facility management (Younis and Fahmy, 2004; Yick *et al.*, 2008; Liu, 2012). Due to the task-oriented feature of WSNs, they are more suitable for environmental monitoring applications (Feng *et al.*, 2012). Some of these applications, such as checking the water flow, gas flow and home security, make the “home automation” possible (Feng *et al.*, 2012).

However, WSNs have many constraints due to the limited capabilities of the sensor nodes. Some of these restraints are limited computational power; low memory;

limited transmission range, which leads to multi-hop communications; restricted battery power, as their batteries cannot be replaced after deployment and the low band width of wireless links. As the sensor nodes have limited and non-rechargeable energy resources, energy efficiency is a very critical issue in WSNs that affects the network life time significantly. Hence, the main concern in designing protocols is how to minimize the energy consumption and maximize the network lifetime (Gu *et al.*, 2013).

Routing protocols in WSNs are responsible for discovering and maintaining energy efficient routes, in order to make the communications reliable and efficient (Al-Karaki and Kamal, 2004). However, they have some features that make them challenging and different from the traditional routing protocols in other wireless networks. Firstly, using IP-based routing protocols is infeasible in WSN, since the IP address cannot be assigned to the sensor nodes. Secondly, the data flow characteristics are different in such networks, where the data from multiple sensor nodes are going to the BS. Thirdly, as the multiple sensors in the same area can generate similar data, the sources can create data redundancy in data traffic. Finally, routing protocols should meet network obstacles, such as limited energy, communication band width, etc. (Liu, 2012; Akyildiz *et al.*, 2002; Roseline and Sumath, 2011).

Various numbers of routing protocols are proposed for WSN with the common objective of getting the ideal

trade between energy consumption, latency and bit data rate (Haneef and Zhongliang, 2012). Among the different categories of routing protocols based on the network architecture, cluster-based (hierarchical) routing protocols are more energy efficient and increase the scalability as well as the lifetime of the network (Heinzelman *et al.*, 2000; Liu, 2012; Naeimi *et al.*, 2012).

Recently, numerous cluster-based routing protocols have been developed for WSNs. In this study, we discuss the challenges related to the design of new routing protocols in detail. In addition, we attempt to comprehensively review and discuss the prominent cluster-based routing protocols which have been proposed for WSNs. Then, these protocols are surveyed based on the important design factors. Our goal is to provide a deeper understanding of the approaches and methods used in the clustering process.

DESIGN CHALLENGES OF ROUTING PROTOCOLS IN WSNs

As mentioned previously, routing protocols in WSN are responsible for discovering and maintaining energy efficient routes in the network, in order to make communications reliable and efficient (Al-Karaki and Kamal, 2004). Due to the limitations in this kind of network, the main and definitive aim of routing protocol design is extending the network lifetime by keeping the sensors alive as long as possible. This issue results in keeping the network connected, for a long period of time (Roseline and Sumath, 2011; Haneef and Zhongliang, 2012).

There are some challenging factors which are important in designing routing protocols. These include energy consumption, node deployment, scalability, coverage, location awareness, the nature of the nodes, the complexity of the control overhead messages, Quality of Service (QoS) and application (Haneef and Zhongliang, 2012; Al-Karaki and Kamal, 2004). A short discussion is provided for these metrics as follows.

Energy consumption: The main goal of routing protocols is to deliver data among the sensors and sink in an efficient way. Each sensor node consumes energy in sensing, processing, receiving and transmitting information. Among these, data transmission is the most energy consuming task (Haneef and Zhongliang, 2012). Since, the sensor nodes have limited energy resources, energy depletion of some nodes results in great topology

and network connectivity changes, reorganization of the network and finding new routes. Therefore, there is a need to design routing protocols that can accommodate the trade off between energy consumption optimization and accuracy (Haneef and Zhongliang, 2012; Akyildiz and Vuran, 2010; Cordeiro and Agrawal, 2011).

Node deployment: Node deployment is very application dependent. Moreover, it influences the energy consumption, coverage, delay, lifetime and the throughput of the network. Node deployment can be manual (deterministic), randomized or non-uniform. In the first strategy, the nodes are located deterministically in order to meet the required performance objectives. In manual node deployment, the coverage of the area is satisfied with careful choice of node density. Although, this is a good solution when the nodes are expensive or their operations are influenced by their position (e.g. in underwater applications), it is not appropriate for harsh environments. Random deployment is a good solution for this kind of environment in which the sensor nodes are scattered in the environment randomly. Furthermore, the node density can be controlled. As the sensor nodes are very cheap and small, a large population of nodes can be utilized which results in a uniform node distribution (Younis and Akkaya, 2008). If the application is related to event detection, then it is efficient to have a uniform node deployment to get effective results (Haneef and Zhongliang, 2012). However, non-uniform node distribution is another method to utilize the influence of non-uniform energy expenditure. The main idea is that the nodes with different densities are scattered in the various regions of the field which results in balancing the communication load and energy expenditure of these regions (Yuan *et al.*, 2011).

Scalability: The number of the nodes that are deployed in the field may vary from tens to thousands. The routing protocol should be scalable to work efficiently with this enormous number of nodes (Haneef and Zhongliang, 2012; Al-Karaki and Kamal, 2004). When the number of nodes is extensive, it is infeasible that each node maintains global knowledge of network topology. Therefore, scalable routing protocols should be fully distributed in order to work efficiently with a limited knowledge about the network topology (Akyildiz and Vuran, 2010).

Coverage: In some cases depending on the application, the sensor networks have to support continuous

coverage of the area; the network should be able to provide the data from anywhere at any time in order to fulfill the requirement of application (Soro and Heinzelman, 2009). In addition, in some applications, high node density is used in order to fulfill the required coverage (Haneef and Zhongliang, 2012; Soro and Heinzelman, 2009). This factor not only affects the coverage and connectivity of the network but also avoids the nodes becoming isolated from each other. However, high node density increases the data redundancy.

Location awareness: Some routing protocols require the location information of sensor nodes in order to perform their tasks. The location information of nodes can be achieved via a Global Positioning System (GPS) or positioning algorithm. GPS increases the hardware cost of nodes and positioning algorithms need the sensor nodes to exchange many messages. This results in increasing the overhead messages and energy consumption. Therefore, it will be reasonably better if the routing protocols do not require location information (Liu *et al.*, 2007).

Nature of nodes: In WSN, the nodes that are scattered over the environment can be either homogeneous or heterogeneous which is very application dependent. When the nodes are homogeneous, they have the same capabilities, such as transmission range, battery life, processing power or speed of movement. The heterogeneous nodes have different capabilities, such as different initial energy. The existence of heterogeneous nodes introduces many technical issues related to data routing (Al-Karaki and Kamal, 2004).

Control messages: The overhead of the control messages exchanged for finding new routes and maintaining existing routes should be kept to the minimum. The control packets can reduce the throughput of the network due to collision with data packets and consume bandwidth (Murthy and Manoj, 2004).

QoS: The routing protocols should also be able to provide a certain level of QoS that is required by the application. The QoS parameters can be bandwidth, delivery delay, jitter, throughput etc. (Murthy and Manoj, 2004). For instance, target detection and tracking applications require low transmission delay for the time-sensitive data. While, multimedia networks require high throughput (Dargie and Poellabauer, 2010). However, in many applications, energy consumption

which is directly related to the network lifetime, is more important than the quality of the transmitted data (Al-Karaki and Kamal, 2004).

Application: The routing protocols are very application specific. The design of new routing protocols is affected by the environmental conditions, network topology and the requirements of the application (Kumarawadu *et al.*, 2008). In other words, different scenarios or network environments need different routing protocols. From the application's viewpoint, data can be collected from the environment using various methods. In time-driven methods, such as environmental monitoring, the sensor nodes send their data periodically to BS or gateways. In event-driven methods, the sensor nodes report the collected data when the event occurs, such as fire detection applications. Eventually, in query-driven methods, the BS requests data from the nodes and sends a query (Al-Karaki and Kamal, 2004; Dargie and Poellabauer, 2010).

The aforementioned factors are the design challenges of routing protocols which affect the performance of WSN. It should be noted once again that the routing protocols in WSN are very application specific and, hence, all of the factors mentioned previously affect them significantly.

CUSTER-BASED ROUTING PROTOCOLS IN WSNs

It has been proven by many researchers that cluster-based routing protocols are more energy efficient than non-cluster based routing protocols, in that they increase the scalability and lifetime of the network (Heinzelman *et al.*, 2000; Liu, 2012; Naeimi *et al.*, 2012; Yu *et al.*, 2011).

Cluster-based routing protocols are composed of rounds. Each round consists of four stages: Cluster head selection, cluster formation, intra-cluster communication and inter-cluster communication. In the first stage, the Cluster Head (CH) nodes are selected with the purpose of gathering data from the Cluster Members (CMs) or ordinary nodes. This stage is followed by constructing the clusters. In other words, the ordinary nodes select a CH to join and form the clusters. In intra-cluster communication stage, the CMs gather the data and send it to their CHs. After receiving the data from the CMs, the CHs perform data aggregation to omit the redundant data. In inter-cluster communication stage, the CHs send the data to the BS (Liu, 2012; Haneef and Zhongliang, 2012). Both intra and inter cluster communication phase might be performed in single hop or multi-hop communication.

Clustering protocols have numerous advantages which make them the most suitable and compatible protocols for WSN (Haneef and Zhongliang, 2012; Naeimi *et al.*, 2012). These main advantages are described as below. The total energy expenditure of data transmission is minimized. In addition, data collision is reduced and redundant data are eliminated in data aggregation process. Furthermore, the delay is reduced, since only CHs perform the task of data transmission. The bandwidth demand is lessened and the limited bandwidth is used effectively. The overhead of routing and topology maintenance is decreased. Moreover, the network manageability and scalability are enhanced (Liu and Shi, 2012).

PROMINENT ENERGY EFFICIENT CLUSTER-BASED ROUTING PROTOCOLS

In this section, a comprehensive survey of prominent cluster-based routing protocols is presented. These routing protocols are discussed in detail and their advantages and disadvantages are highlighted.

Low Energy Adaptive Clustering Hierarchy (LEACH): LEACH proposed by Heinzelman *et al.* (2000) is one of the well-known cluster-based routing algorithms. LEACH is a self-organized and adaptive clustering protocol that uses randomization in order to distribute energy equally among the network. The operation of LEACH is divided into rounds. Each round consists of a setup and a steady state phase. It is assumed that the sensor nodes have the

same initial energy. In the setup phase, each node decides to be a CH based on a threshold, $T(n)$. This decision is independent from the other nodes. The nodes that have elected themselves as the CH for the current round, broadcast an advertisement message. The non-CH nodes select a CH based on the signal strength they have received (i.e. they join the closest CH). After the clusters are formed, the CHs define a Time Division Multiple Access (TDMA) schedule for their CMs. In the steady state phase, the CMs send the collected data to their CH. The CHs accomplish data aggregation to omit the redundant data and compress the data into one single signal. Then, the CHs send this signal to the BS directly. After a predetermined period, the next round will be started. The topology of LEACH is shown in Fig. 1. The advantages of LEACH are described below.

The role of the CH is rotated among the nodes in the network in order to balance the energy consumption (Zungeru *et al.*, 2012). By performing data aggregation, LEACH reduces the amount of data that must be transmitted to the BS. No extra negation is required for CH selection which results in the reduction of control messages (Heinzelman *et al.*, 2000). The nodes are dead in a random fashion, thus no part of the network is not being sensed due to nodes death (Heinzelman *et al.*, 2000). LEACH does not require any control information from the BS or global knowledge of the network to operate, thus it is a completely distributed routing protocol (Heinzelman *et al.*, 2000; Liu, 2012). This routing protocol uses TDMA and Code Division Multiple Access (CDMA) in order to avoid collisions and interference (Liu, 2012).

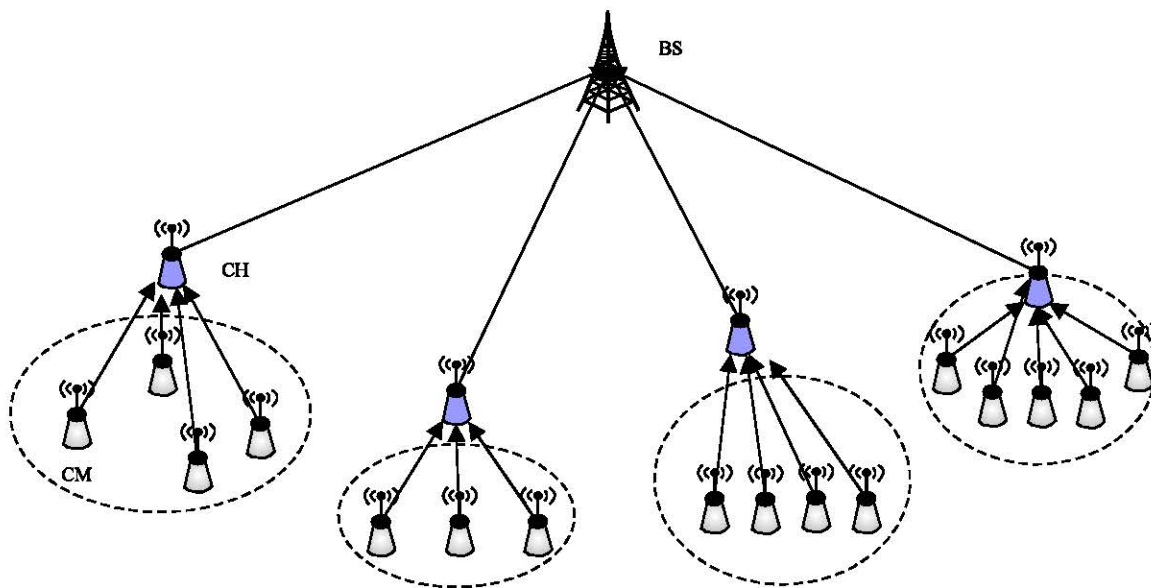


Fig. 1: Topology of LEACH

Although, LEACH is the simplest hierarchical protocol and can lead to saving energy in WSN, some problems exist. Firstly, the residual energy of the node is not considered in CH selection. Hence, the nodes with lower initial energy can be selected as CHs which results in premature death, coverage and energy hole problems (Liu, 2012; Abbasi and Younis, 2007). Furthermore, as the clusters are reformed in each phase, a significant amount of energy is wasted in order to construct the clusters (Liu and Shi, 2012). Moreover, CHs can be dense or sparse in different areas, as LEACH performs the CH-selection phase in terms of probabilities (Liu, 2012). Since, CHs send the aggregated data to the BS directly, the CHs far away from the BS will die earlier (Chen, 2012). Finally, due to the single hop transmission in the inter-cluster and intra-cluster communications, LEACH is not appropriate for large size networks. Thus, it is not a scalable routing protocol (Liu and Shi, 2012).

LEACH-centralized (LEACH-C): As LEACH does not guarantee the placement or number of CHs, another protocol LEACH-C was proposed by Heinzelman *et al.* (2000). It consists of two phases: The first phase uses a centralized clustering algorithm and the second phase is identical to the steady state phase in LEACH. The sensor nodes have the same initial energy at the beginning. During the setup phase, all the nodes send a message to the BS which contains their location and energy level. The BS calculates the average energy of the nodes to make sure that the energy is distributed uniformly among the nodes (Heinzelman *et al.*, 2000). Those sensor nodes whose energy level is less than the average cannot be selected as the CH for the current round. The BS tries to find the CHs among the remaining nodes using a simulated annealing algorithm. The algorithm tries to minimize the amount of energy for non-CH nodes by minimizing the total sum of squared distance between the non-CH nodes and the closest CH (Heinzelman *et al.*, 2002). As the CHs and their clusters are formed, the BS broadcasts a message which contains the ID of the determined CHs for each node. If the node ID is equal to the ID of CH, it is selected as the CH, otherwise, the node determines its TDMA slot for data transmission and goes to the sleep mode. The steady state phase is the same as LEACH. The advantages of LEACH-C are described below.

LEACH-C considers the energy level of nodes in CH selection. Thus, the nodes with more energy remaining have more chance of being chosen as the CH (Heinzelman *et al.*, 2002). In addition, the number of CHs in each round equals a prescheduled optimal value which is one of the reasons that LEACH-C outperforms LEACH

(Heinzelman *et al.*, 2002; Muruganathan *et al.*, 2005). Furthermore, it transmits 40% more data per unit of time, since the BS uses the global knowledge of the network to create better clusters. Therefore, less energy is required for data transmission with the creation of these clusters (Muruganathan *et al.*, 2005).

Although, LEACH-C outperforms LEACH, it has some limitations. Firstly, since the CHs have to transfer data to the BS in a single hop manner, they deplete a noticeable amount of energy (Muruganathan *et al.*, 2005). Moreover, a significant amount of control messages should be sent at the beginning of each round which consumes considerable energy. Finally, as the cluster setup phase is accomplished each round, energy expenditure and latency is increased (Bajaber and Awan, 2011).

Energy Efficient Unequal Clustering (EEUC): EEUC was proposed by Li *et al.* (2005) in order to solve the problem of hot spots. The hot spot problem is related to the early death of CH nodes which are closed to the BS. The CHs near to the BS have to relay more load in addition to the data of their own cluster. Consequently, their power is exhausted quickly. EEUC is a distributed CH competitive algorithm. It is assumed that the sensor nodes have the same initial energy. The tentative CHs are selected with the same probability which is a predefined threshold. Each tentative CH has a competition range which decreases as its distance to the BS decreases. The clusters near the BS have a smaller size in order to save more energy for their inter-cluster communications. Additionally, the final CHs are not in each other's competition diameter. The CH selection is primarily based on the residual energy of each node. Then, the tentative CHs compete to become the final CH in local regions. Once the CHs are defined, they start advertising. The nodes join the CHs that are closest to them. The intra-cluster communication is the same as LEACH in which the CHs perform data aggregation after gathering the data. Li *et al.* (2005) introduced a threshold, for multi-hop communications. If the distance between the CH and the BS is less than the threshold, it will communicate with the BS directly. Otherwise, the CHs should choose a relay node. Each CH broadcasts a message which contains information about the node ID, residual energy and the distance to the BS. CHs will choose an adjacent node with less distance to the BS and more residual energy in order to extend the network lifetime and reduce wireless channel interference. The advantages of EEUC are described below.

EEUC tries to resolve the hot spot problem by forming smaller clusters near the BS (Li *et al.*, 2005). Moreover, it considers the nodes with more residual

energy in CH selection and CH rotation to balance the energy expenditure across the network (Li *et al.*, 2005). In addition, it introduces a method for relay node selection that reduces wireless channel interference.

However, EEUC has drawbacks. Firstly, the CH selection (i.e. selection of tentative CH until final stage) requires broadcasting a significant amount of control messages which increases energy consumption. Secondly, EEUC increases the completion of the setup phase as it needs to determine the neighbor nodes and gather resource information (Naeimi *et al.*, 2012). Finally, it increases the overhead and declines the performance of the network due to extra global data aggregation (Boyinbode *et al.*, 2010).

Energy Efficient Clustering Scheme (EECS): The EECS was proposed by Ye *et al.* (2005). It is assumed that the sensor nodes have the same initial energy in EECS. In CH election, some nodes with a probability are selected as the candidate CHs. The candidate CHs broadcast an advertisement message in the range of R competition. If there is another candidate node with higher residual energy in this range, it will give up the competition. Otherwise, it will be selected as the final CH. Since the inter-cluster communications in the EECS are based on the single hop, the CHs far from the BS will lose a significant amount of energy and will die earlier. Thus, the cluster formation phase in the EECS is completely different from other approaches. In the EECS, a weighted cost function is introduced. The nodes choose their CHs based on two factors: Their distance to the CH and the CH distance to the BS. Thus, the clusters far from the BS become smaller than the ones near to the BS. In other words, each node joins the CH with minimum cost. This issue will result in an energy saving of the nodes and balance the load of the CH that they want to join (Naeimi *et al.*, 2012). It should be noted that the EECS is appropriate for periodic data gathering applications. The benefits of the routing protocol are described below.

The EECS develops the distribution of energy over the network which results in a better utilization of the resources and prolongs the network lifetime (Boyinbode *et al.*, 2010). Moreover, it solves the problem that the CHs that are far from the BS require more energy for inter-cluster communication.

However, it has some disadvantages. Firstly, clusters near the BS may be burdened which may result in the quick death of the CHs (Boyinbode *et al.*, 2010). Secondly, as communication with the BS in a single hop manner consumes a considerable amount of energy, EECS is not appropriate for large size networks (Liu, 2012). Also,

non-CH nodes need global knowledge relating to the location in order to compute their distance to the BS and CHs (Liu, 2012). Finally, since all the nodes must compete in CH election, this routing protocol increases the control overhead complexity (Liu, 2012).

Base Station Controlled Dynamic Clustering Protocol

(BCDCP): The BCDCP is a centralized clustering routing protocol proposed by Muruganathan *et al.* (2005). The BS, as the key component of BCDCP, is capable of complicated calculations. Cluster formation is the main idea of this protocol where the CHs are distributed uniformly over the network and serve equal number of the CMs. The sensor nodes have equal initial energy. The BCDCP consists of two phases: Setup phase and data communication phase. The setup phase consists of the following sub-phases: Cluster setup, cluster selection, CH-to-CH routing path formation and schedule creation for each CH. At the beginning of each round, the nodes send their current energy level to the BS. Then, the BS calculates the average energy of the network. Clustering is performed based on an iterative cluster-splitting algorithm. In other words, a group of nodes is chosen whose energy level is above the average energy of the network. The selected CHs for this round are chosen from this group. First, two nodes far away from each other are chosen. The remaining nodes are grouped in these two nodes based on their distance (i.e. the closest one). Next, the two groups become balanced in order to have approximately the same number of nodes. After balancing, the two sub-clusters are formed. In this way, minimum energy is required for the communication phase. This process will be repeated until the desired number of clusters is achieved. The BCDCP uses a CH-to-CH multi-hop transmission for inter-cluster communications. Once the clusters and CHs are formed, the BS chooses the lowest routing path and forwards the information about the selected CH, clusters grouping details and routing path to all the nodes. Routing paths are formed based on the MST (Shen, 1999). The CH which has to communicate with the BS directly, is chosen randomly. This randomization helps prevent constantly choosing the CHs near the BS as a relay node. The BCDCP utilizes a TDMA scheduling scheme to minimize collision among the nodes. The next phase is data communication which consists of data gathering, data fusion and data routing. Each sensor node sends data to its CH according to the defined time slot. CHs aggregate data and send it to the BS using the CH-to-CH routing path. It should be noted that the BCDCP also uses CDMA in order to avoid radio interference. The advantages of BCDCP are described below.

The BS performs all the severe tasks, such as CH selection, cluster setup, CH-to-CH path formation and randomized rotation of CH, in order to save a considerable amount of energy. Furthermore, clusters are distributed uniformly and each CH consumes the same amount of energy (Liu, 2012). In addition, since the BCDCP uses the idea of randomized rotation of CH, it prevents the quick death of the CH near the BS (Muruganathan *et al.*, 2005). It also prevents collision and radio interference by utilizing the TDMA and CDMA schemes.

However, there are a few limitations with the BCDCP. Firstly, since the BCDCP is a centralized algorithm, it is not appropriate for large scale networks (Liu, 2012). Secondly, sensor nodes must send their energy level and location to the BS at the beginning of each round which consumes a noticeable amount of energy. Furthermore, the complexity in the design of the network increases (Liu, 2012). Thirdly, if the CHs far away from the BS are chosen in CH randomization, their energy will be depleted. Finally, the BCDCP assumes that the sensor nodes always have data to send, thus it is not appropriate for event driven applications, such as the tracking of enemy activities, fault detection and diagnosis in machinery (Liu and Shi, 2012).

Energy Efficient Heterogeneous Clustered (EEHC): The EEHC protocol that was proposed by Kumar *et al.* (2009) considers a network with heterogeneous nodes in terms of energy. There are three heterogeneity levels in EEHC: Super nodes, advanced nodes and normal nodes. To be more specific, this protocol extends the network lifetime by using heterogeneous sensor nodes. It elects the CH in a distributed fashion (Katiyar *et al.*, 2011a). The CH selection is based on the weighted election probabilities of each node in accordance with the remaining energy. Although, the EEHC considers the residual energy of nodes, it does not consider the initial energy of nodes. The phases in EEHC are identical to LEACH. However, some differences exist. In LEACH, all the nodes are homogeneous. In addition, EEHC uses the weighted probabilities to obtain the threshold used to elect the CH in each round. The advantages of EEHC are described below.

EEHC considers heterogeneous nodes in terms of energy in order to improve the network lifetime (Kumar *et al.*, 2009). Furthermore, it considers the remaining energy of the nodes in the CH selection. Therefore, the nodes with more residual energy have more chance of becoming the CH for the current round.

However, single hop communications with the BS results in quick energy depletion of the CHs.

LEACH-Fuzzy Logic (LEACH-FL): LEACH-FL which is an improvement of the LEACH protocol based on the fuzzy logic, was proposed by Ran *et al.* (2010). The only difference between LEACH and LEACH-FL concerns the method of CH selection. LEACH-FL considers three attributes of nodes in the selection of the CH (i.e. distance, node density and battery level). These three attributes are used as three input functions to transform the system input into fuzzy sets. The BS selects nodes with a higher chance as the CH using if-then rules in fuzzy interface. The remaining operations of LEACH-FL are the same as LEACH. In addition, it is assumed that the sensor nodes are equal in terms of the initial energy.

LEACH-FL outperforms LEACH due to its consideration of the three attributes in CH selection. However, it has limitations. Since the BS should have knowledge about the location and other information of the nodes, LEACH-FL is not applicable for large scale networks (Taheri *et al.*, 2012). In addition, sending the local information at the beginning of each round is energy consuming. Furthermore, CHs should communicate with the BS directly. Therefore, the CHs far from the BS spend more energy and die quickly.

Far Zone-LEACH (FZ-LEACH): Far Zone-LEACH (FZ-LEACH) was proposed by Katiyar *et al.* (2011a) in order to solve the problem of the varying cluster sizes in LEACH. FZ-LEACH contains two phases: Cluster setup phase and steady state phase. The CH selection and cluster formation are identical to LEACH. The nodes have equal initial energy. After the clusters are formed, each CH creates a TDMA schedule list to assign nodes a time slot for data transmission. The CMs send their power level to their CH. The idea of power level is borrowed from HEED (Younis and Fahmy, 2004; Katiyar *et al.*, 2011b). The FZ-LEACH tries to improve intra-cluster communication. Each node sends the minimum power level that is required to communicate with its CH (MRP). Then, the CH calculates an average minimum reachability power (AMRP) for the communication cost. The nodes with a lower power level than the AMRP are considered to be in the "Far Zone". As far zones are formed, one node is randomly selected as the "Zone Head" (ZH) for the current round. The ZHs create time slots for zone members in order to transmit data to them. Then, the ZHs transmit data to the CHs. The far zone in a cluster can be considered as a "Level-2" cluster. The important point of FZ-LEACH is that the far zone is only formed when the nodes have a MRP less than the AMRP. FZ-LEACH is appropriate for large real life deployment of WSN. The simple topology of FZ-LEACH is indicated in Fig. 2. It is

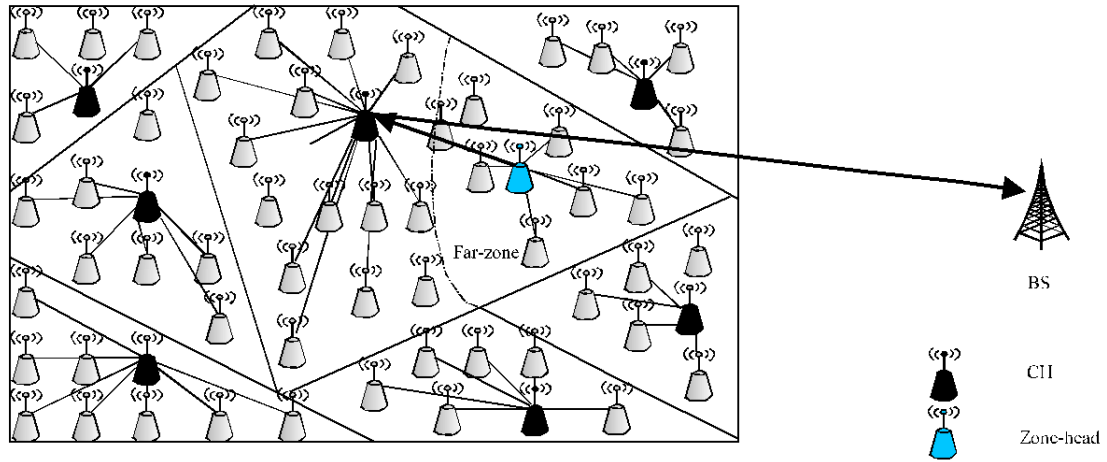


Fig. 2: Topology of FZ-LEACH

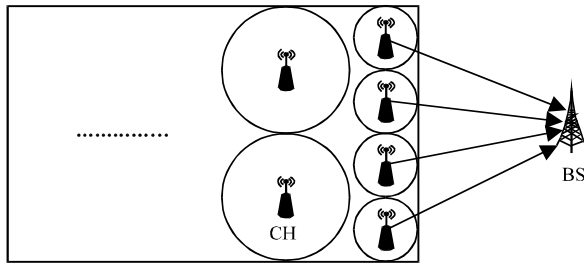


Fig. 3: Topology of ACT

suited for large networks, since it improves long distance transmission efficiently. However, it does not consider the distance between the ZH and the Far-Zone members in the ZH election. Moreover, it does not consider that ZH may not have any Far-Zone members (Yoo, 2012).

Arranging Cluster sizes and Transmission ranges (ACT):

ACT was proposed by Lai *et al.* (2012) with a topology in which the network is separated into multiple levels. The cluster sizes in each level are the same but different from other levels. It is assumed that the sensor nodes have the same initial energy in the ACT. ACT consists of three phases: cluster formation phase, data forwarding and cluster maintenance. In the cluster formation phase, the BS divides the network topology into K levels. Clusters near the BS are the smallest. The largest clusters are the ones farthest from the BS. Then, the BS calculates the radius of the clusters for each level, number of levels, cluster radius and determines the CHs in the ideal locations. The sensor node closest to the ideal location will be selected as the CH. A number is assigned to each CH. The transmission path of each CH can be assigned as a two-dimensional array (X) (Y); where X is

the level of CH and Y means the order of cluster in that level. After the cluster setup is done, the CHs collect data transmitted by the Cms using the MST concept (Shen, 1999) and perform data aggregation. As the CH performs data aggregation in “i” level, it forwards its data to the next CH in “i-1” level until the transmitted data reach the BS. In the cluster maintenance phase, as ACT avoids re-clustering in each round, CH rotation and cross-level data transmission are then performed. In CH rotation, if the CH in each cluster reaches a threshold (i.e. 15% of initial energy), another CH (i.e. closest to the ideal point) will be selected. During CH rotation, a change message is broadcasted. In cross-level data transmission, whenever the BS determines that each node in the first level cannot serve as CH anymore, it broadcasts a message to allow the CHs in the second level to transmit the data directly (the same for 3th, 4th... kth level). The topology of ACT is presented in Fig. 3. The advantages of ACT are described below.

Since, the BS determines the size of clusters according to their distance, the load of the CHs near the BS reduces. Therefore, the clusters near the BS can save more energy for inter-cluster communications. Furthermore, ACT allows each CH to expend almost the same amount of energy. Thus, clusters near the BS do not exhaust their energy so fast. In addition, it avoids cluster re-configuration every round to extend the network lifetime. Moreover, ACT utilizes cross-level transmission to prolong the network lifetime.

However, there are some limitations with ACT. It does not address the coverage problem (Lai *et al.*, 2012). In addition, although cluster sizes are arranged according to energy consumption, the location of the newly selected CH strays from the ideal ones. This causes the distribution of energy to fluctuate (Fu *et al.*, 2013).

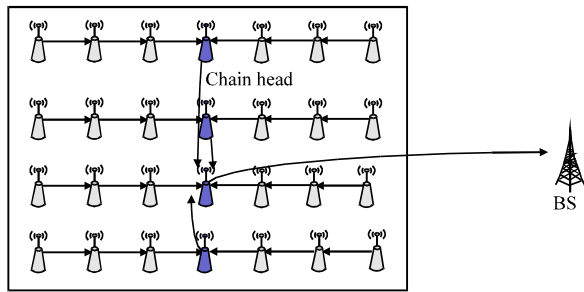


Fig. 4: Topology of CCM

Chain cluster-based mixed (CCM): The CCM routing protocol, proposed by Tang *et al.* (2012), tends to use and integrate the advantages of two algorithms: PEGASIS in terms of low energy consumption and LEACH in terms of short transmission delay. CCM consists of two phases: Chain based routing and cluster based routing. The sensor nodes have equal initial energy in CCM. In the first phase, each node can transmit data with the two closest node neighbors in the same chain. Generally, in this step, the nodes in each chain send their data to the chain head. In CCM, each node plays the role of the chain head in turn. The chain head generates two tokens and passes them to the two ends of the chain to transmit their data to their neighbors in parallel. Each node that receives a token will fuse the received data with its own and sends it to the next neighbor node. This process will be continued until all the data are transmitted to the chain head. Next, the chain heads aggregate data and destroy two tokens. The chain-based routing is performed using the same method in all chains. Then, the second phase of CCM is started. The chain heads form a cluster and the rest of the nodes go into sleep mode. Cluster-based routing includes two stages: Voting CH and data transmission. In the first stage, all the members broadcast their residual energy. The node with the highest residual energy will be chosen as the CH and advertise itself. Then, it assigns a time slot for each member by creating a TDMA schedule. After collecting all the data from the members, the CH performs data aggregation and sends the data directly to the BS. The applications of CCM are gathering the utility of water, electricity and gas in a remote way. Figure 4 indicates the topology of CCM.

CCM solves the weaknesses of the extra energy overhead for periodic CH selection in LEACH and the long transmission delay of PEGASIS. However, it has some drawbacks.

If the chains are not far enough or does not have CDMA functionality, the transmissions in the chains should happen sequentially which results in a delay (Tang *et al.*, 2012). In addition, if the selected CH is far

away from the BS, it will consume much energy to send data in a single hop manner. Finally, since in the cluster-based routing just one cluster is formed, CCM is not appropriate for large-scale networks.

Energy Aware Distributed Clustering (EADC): The EADC routing protocol was proposed by Yu *et al.* (2012) for networks with non-uniform node distribution. This cluster-based routing protocol consists of an energy aware clustering algorithm and a cluster-based routing algorithm. In EADC, it is assumed that the sensor nodes are heterogeneous in terms of the initial energy. EADC starts to make clusters in equal sizes. In the routing algorithm, EADC chooses the CHs with more remaining energy and less CMs as the relay node for the next hops. In the cluster-based routing protocol, a routing tree based on the CHs will be constructed. The energy consumption of the CHs in multi-hop communication is divided into intra-cluster energy consumption and inter-cluster energy consumption. As the CHs in EADC are distributed uniformly in the network with even cluster size, the intra-cluster energy consumption of the nodes are the same. However, due to non-uniform node distribution, the load is still imbalanced among the CHs. Therefore, EADC adjusts the intra-cluster and inter-cluster energy consumption of the CHs to prolong the network lifetime. Several CHs should be chosen to communicate with BS directly. Yu *et al.* (2012) introduced a distance threshold that is identical to the distance threshold in EEUC (Li *et al.*, 2005). The CHs select a relay node with more residual energy, less the CMs and nearer to the BS as their next hop. EADC has several advantages which are described below.

EADC considers the heterogeneity of nodes in terms of energy in the CH selection phase. In addition, the CHs are distributed uniformly among the network and they can cover all the nodes in the monitored area. Moreover, the cluster-based routing algorithm in EADC efficiently balances the energy expenditure between the CHs by choosing the CH with more remaining energy and less CMs as the next hop.

However, due to the different density of the nodes in some regions, in dense areas, there will be redundancy and duplication in data. This redundancy in data transmission is energy consuming and not considered in EADC.

COMPARISON OF PROMINENT CLUSTER-BASED ROUTING PROTOCOLS IN WSNS

In this section, the energy efficient cluster-based routing protocols in the previous section are compared in Table 1. We compare these routing protocols based on

Table 1: Comparison of prominent clustering algorithms in WSN

Protocol	Energy efficiency	Scalability	Delivery delay	Mobility	Control message complexity	Node distribution	Location awareness	Sensor type	Coverage	Application
LEACH	Very low	Very low	Very small	Stationary	low	Random	Not required	Homogeneous	Not considered	Time-driven
LEACH-C	Very low	Very low	Small	Stationary	Very high	Random	Required	Homogeneous	Not considered	Time-driven
EEUC	High	High	Moderate	Stationary	High	Uniform	Required	Homogeneous	Not considered	Time-driven
EECS	Moderate	Low	Small	Stationary	Very high	Uniform	Required	Homogeneous	Not considered	Time-driven
BCDCP	Very low	Very low	Small	Stationary	Very high	Random	Required	Homogeneous	Not considered	Time-driven
EEHC	Low	Very low	Very small	Stationary	Low	Random	Required	Heterogeneous	Not considered	Time-driven
LEACH-FL	Very low	Very low	Small	Stationary	Very high	Random	Required	Homogeneous	Not considered	Time-driven
FZ-LEACH	Low	Low	Very small	Stationary	Low	Random	Not required	Homogeneous	Not considered	Time-driven
ACT	Moderate	Very low	Small	Stationary	Very high	Uniform	Required	Homogeneous	Not considered	Event-driven
CCM	Low	Low	Moderate	Stationary	High	Grid	Required	Homogeneous	Not considered	Time-driven
EADC	High	High	Moderate	Stationary	High	Nonuniform	Not required	Heterogeneous	Not considered	Time-driven

the important metrics which are discussed in detail in section 2, as shown in Table 1. These important metrics have an influence on the design and performance of routing protocols.

CONCLUSION AND OPEN ISSUES

WSN has received much attraction recently due to its various usages in diverse applications. Routing has a key role in WSN, since it is the most energy-conserving factor. Among the different classifications of routing protocols based on the network architecture, cluster-based routing protocols are more suitable for WSN. In the past few years, there has been an increasing effort in designing efficient and adequate cluster-based routing protocols for WSN.

In this study, a comprehensive survey on cluster-based routing protocols in WSNs is presented. We have discussed the design challenges of the routing protocols in detail. Additionally, a few classical cluster-based routing protocols are deeply analyzed and compared based on the important metrics. Although the presented clustering routing protocols tried to improve the performance of WSN, still come challenges are remained. First, QoS problem is needed to be addressed in future studies which is the main problem of real-time applications. Secondly, coverage requirements need to be considered in the design of future cluster-based routing protocols, since it is an outstanding issue to preserve the complete coverage of area in a long period. Finally, future works need to investigate the methods to handle the challenges related to the mobility of the nodes in the network such as the topology changes in the node-mobility applications.

In conclusion, once again it should be noted that the routing protocols are very application specific in WSN. Moreover, each application has its own requirements which should be considered in the design.

REFERENCES

- Abbasi, A.A. and M. Younis, 2007. A survey on clustering algorithms for wireless sensor networks. *Comput. Commun.*, 30: 2826-2841.
- Akyildiz, I.F., W. Su, Y. Sankarasubramaniam and E. Cayirci, 2002. A survey on sensor networks. *IEEE Commun. Mag.*, 40: 102-114.
- Akyildiz, I.F. and M.C. Vuran, 2010. *Wireless Sensor Networks*. John Wiley and Sons, USA.
- Al-Karaki, J.N. and A.E. Kamal, 2004. Routing techniques in wireless sensor networks: A survey. *IEEE Wireless Commun.*, 11: 6-28.
- Bajaber, F. and I. Awan, 2011. Adaptive decentralized re-clustering protocol for wireless sensor networks. *J. Comput. Syst. Sci.*, 77: 282-292.
- Boyinbode, O., H. Le, A. Mbogho, M. Takizawa and R. Poliah, 2010. A survey on clustering algorithms for wireless sensor networks. *Proceedings of the 13th International Conference on Network-Based Information Systems*, September 14-16, 2010, Takayama, pp: 358-364.
- Chen, J., 2012. Improvement of leach routing algorithm based on use of balanced energy in wireless sensor networks. *Proceedings of the 7th International Conference on Advanced Intelligent Computing*, August 11-14, 2011, Zhengzhou, China, pp: 71-76.
- Cordeiro, C.M. and D.P. Agrawal, 2011. *Ad hoc and Sensor Networks: Theory and Applications*. World Scientific, Singapore, ISBN-13: 9789814338882, Pages: 636.
- Dargie, W. and C. Poellabauer, 2010. *Fundamentals of Wireless Sensor Networks: Theory and Practice*. John Wiley and Sons, New York, ISBN-13: 9780470975688, Pages: 336.
- Ferng, H.W., R. Tendeau and A. Kurniawan, 2012. Energy-efficient routing protocol for wireless sensor networks with static clustering and dynamic structure. *Wireless Personal Commun.*, 65: 347-367.

- Fu, S., J. Ma, H. Li and C. Wang, 2013. Energy-balanced separating algorithm for Cluster-based data aggregation in wireless sensor networks. *Int. J. Distributed Sensor Networks*, 10.1155/2013/570805
- Gu, X., J. Yu, D. Yu, G. Wang and Y. Lv, 2013. ECDC: An energy and Coverage-aware distributed clustering protocol for wireless sensor networks. *Comput. Electrical Eng.*, 10.1016/j.compeleceng.2013.08.003
- Haneef, M. and D. Zhongliang, 2012. Design challenges and comparative analysis of cluster based routing protocols used in wireless sensor networks for improving network lifetime. *Adv. Inform. Sci. Service Sci.*, 4: 450-459.
- Heinzelman, W.R., A. Chandrakasan and H. Balakrishnan, 2000. Energy-efficient communication protocol for wireless microsensor networks. *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, January 4-7, 2000, Maui, Hawaii, pp: 10.
- Heinzelman, W.B., A.P. Chandrakasan and H. Balakrishnan, 2002. An application-specific protocol architecture for wireless microsensor networks. *IEEE Trans. Wireless Commun.*, 1: 660-670.
- Katiyar, V., N. Chand and S. Soni, 2011a. A survey on clustering algorithms for heterogeneous wireless sensor networks. *Int. J. Adv. Network. Appl.*, 2: 745-754.
- Katiyar, V., N. Chand, G.C. Gautam and A. Kumar, 2011b. Improvement in LEACH protocol for Large-scale wireless sensor networks. *Proceedings of the International Conference on Emerging Trends in Electrical and Computer Technology*, March 23-24, 2011, Tamil Nadu, pp: 1070-1075.
- Kumar, D., T.C. Aseri and R.B. Patel, 2009. EEHC: Energy efficient heterogeneous clustered scheme for wireless sensor networks. *Comput. Commun.*, 32: 662-667.
- Kumarawadu, P., D.J. Dechene, M. Luccini and A. Sauer, 2008. Algorithms for node clustering in wireless sensor networks: A survey. *Proceedings of the 4th International Conference on Information and Automation for Sustainability*, December 12-14, 2008, Colombo, pp: 295-300.
- Lai, W.K., C.S. Fan and L.Y. Lin, 2012. Arranging cluster sizes and transmission ranges for wireless sensor networks. *Inform. Sci.*, 183: 117-131.
- Li, C., M. Ye, G. Chen and J. Wu, 2005. An energy-efficient unequal clustering mechanism for wireless sensor networks. *Proceedings of the International Conference on Mobile Adhoc and Sensor Systems Conference*, November 7-10, 2005, IEEE Press, Washington, USA., pp: 1-8.
- Lindsey, S., C. Raghavendra and K.M. Sivalingam, 2002. Data gathering algorithms in sensor networks using energy metrics. *IEEE Trans. Parallel Distrib. Syst.*, 13: 924-935.
- Liu, M., J.N. Cao and G.H. Chen, 2007. EADEEG: An energy-aware data gathering protocol for wireless sensor networks. *J. Software*, 18: 1092-1109.
- Liu, X. and J. Shi, 2012. Clustering routing algorithms in wireless sensor networks: An overview. *KSII Trans. Internet Inform. Syst.*, 6: 1735-1755.
- Liu, X.X., 2012. A survey on clustering routing protocols in wireless sensor networks. *Sensors*, 12: 11113-11153.
- Murthy, C.S.R. and B.S. Manoj, 2004. *Ad hoc Wireless Networks: Architectures and Protocols*. Prentice-Hall, Englewood Cliffs, New Jersey.
- Muruganathan, S.D., D.C.F. Ma, R.I. Bhasin and A. Fapojuwo, 2005. A centralized energy-efficient routing protocol for wireless sensor networks. *IEEE Commun. Mag.*, 43: S8-S13.
- Naeimi, S., H. Ghafghazi, C.O. Chow and H. Ishii, 2012. A survey on the taxonomy of Cluster-based routing protocols for homogeneous wireless sensor networks. *Sensors*, 12: 7350-7409.
- Ran, G., H. Zhang and S. Gong, 2010. Improving on LEACH protocol of wireless sensor networks using fuzzy logic. *J. Inform. Comput. Sci.*, 7: 767-775.
- Roseline, R.A. and P. Sumath, 2011. Energy efficient routing protocols and algorithms for wireless sensor networks-a survey. *Global J. Comput. Sci. Technol.*, 11: 60-67.
- Shen, H., 1999. Finding the k most vital edges with respect to minimum spanning tree. *Acta Inform.*, 36: 405-424.
- Soro, S. and W.B. Heinzelman, 2009. Cluster head election techniques for coverage preservation in wireless sensor networks. *Ad hoc Networks*, 7: 955-972.
- Taheri, H., P. Neamatollahi, O.M. Younis, S. Naghibzadeh and M.H. Yaghmaee, 2012. An energy-aware distributed clustering protocol in wireless sensor networks using fuzzy logic. *Ad Hoc Networks*, 10: 1469-1481.
- Tang, F., I.Y. You, S. Guo, M. Guo and Y. Ma, 2012. A Chain-cluster based routing algorithm for wireless sensor networks. *J. Intell. Manuf.*, 23: 1305-1313.
- Ye, M., C. Li, G. Chen and J. Wu, 2005. EECS: An energy efficient clustering scheme in wireless sensor networks. *Proceedings of the 24th International Performance Computing and Communication Conferences*, April 7-9, 2005, IEEE Press, New York, pp: 535-540.

- Yick, J., B. Mukherjee and D. Ghosal, 2008. Wireless sensor network survey. *J. Comput. Network*, 52: 2292-2330.
- Yoo, S., 2012. Improve Far-zone LEACH protocol for energy conserving. Proceedings of the 8th International Conference on Wireless Communications, Networking and Mobile Computing, September 21-23, 2012, Shanghai, pp: 1-4.
- Younis, M. and K. Akkaya, 2008. Strategies and techniques for node placement in wireless sensor networks: A survey. *Ad hoc Networks*, 6: 621-655.
- Younis, O. and S. Fahmy, 2004. HEED: A hybrid, energy-efficient, distributed clustering approach for *ad hoc* sensor networks. *IEEE Trans. Mobile Comput.*, 3: 366-379.
- Yu, J., Y. Qi, G. Wang, Q. Guo and X. Gu, 2011. An energy-aware distributed unequal clustering protocol for wireless sensor networks. *Int. J. Distrib. Sensor Networks*, 10.1155/2011/202145
- Yu, J., Y. Qi, G. Wang and X. Gu, 2012. A cluster-based routing protocol for wireless sensor networks with nonuniform node distribution. *AEU-Int. J. Electron. Commun.*, 66: 54-61.
- Yuan, H., W. Liu and J. Xie, 2011. Prolonging the lifetime of heterogeneous wireless sensor networks via non-uniform node deployment. Proceedings of 2011 International Conference on Internet Technology and Applications, August 16-18, 2011, Wuhan, pp: 1-4.
- Zungeru, A.M., L.M. Ang and K.P. Seng, 2012. Classical and swarm intelligence based routing protocols for wireless sensor networks: A survey and comparison. *J. Network Comput. Appl.*, 35: 1508-1536.