



Journal of Applied Sciences

ISSN 1812-5654

science
alert

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

A Distributed Energy-aware Clustering Algorithm for Life Time Enhancement of Wireless Sensor Network

Javad Memariani, Zuriati Ahmad Zukarnain,
Azizol Abdullah and Zurina Mohd. Hanapi

Department of Communication Technology and Network, Universiti Putra Malaysia (UPM),
43400 Serdang, Malaysia

Abstract: In distributed sensor networks, a large number of small sensors are deployed to create a network which cooperates together to set up a sensing network. To design applications and protocols for sensor networks, the maximum network lifetime and minimum energy consumption should be considered as important parameters. This study presented a distributed clustering algorithm where cluster heads are elected following a two-way message exchange between each sensor and its corresponding neighbors. Sensor's eligibility to be elected cluster head is based on some parameters such as energy, centrality, density, in addition distances between nodes and the sink. The main goal of our clustering algorithm is to distribute the energy load among all sensor nodes to reduce the energy consumption and increase the network lifetime of wireless sensor networks. The proposed algorithm, in the worse case, has an algorithmic complexity of $O(N)$ at each sensor node.

Key words: Clustering, border node detection, unequal size, wireless sensor networks

INTRODUCTION

The development of micro devices and wireless communication technology created the tiny sensor device called sensor node. Over the past few years, wireless sensor networks for various applications including target tracking (Rezaii and Tinati, 2009), military surveillance (Liu *et al.*, 2010), car parking system (Idris *et al.*, 2009), etc. have been used. Major benefits are requested for the sensor technology is that it has high capability for wireless communications, sensing and data processing despite its low price. However, these sensor nodes restricted with the small memory, limited energy power and constrained processing capability. Due to their energy limitation, generally wireless sensors have a restricted transmission range hence, for routing data toward the sensor nodes multi-hop transmission is used rather than one-hop since it is more energy-efficient. As depicted in Fig. 1, data collected by sensors is transmitted multi-hop to a special node equipped with higher energy and processing capabilities called sink.

Clustering in wireless sensor networks has been substantiated as energy efficient tool for maximizing network lifetime (Chong and Kumar, 2003; Arboleda and Nasser, 2006). One of the consequential motivations

in wireless sensor networks is to design energy-efficient and scalable clustering protocols since the non exchangeable and irreplaceable power energy characteristics of the sensor nodes (Memariani *et al.*, 2011). The key benefit of the clustering scheme is to reduce the distance of transferring data by communicating with cluster heads, which are obviously having shorter distance against transmission data to the sink directly (Abidoye *et al.*, 2011). Furthermore, it reduces unnecessarily repetitive transmissions by minimizing the network traffic towards the sink that impacts on energy consumption (Guo *et al.*, 2010). In the nature of sensor networks, cluster heads process, filter and aggregate data

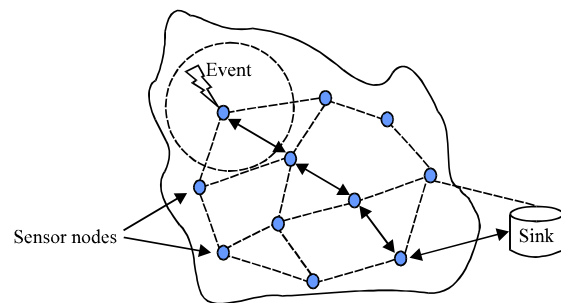


Fig. 1: Multi-hop routing of observed data

sent by the cluster member, which causes to diminish the network load and relieve bandwidth (Xin *et al.*, 2008).

In addition, regular sensors are not involved in routing and relaying data and transmissions are only performing by cluster heads, hence considerable energy can be saved by clustering (Karl and Willig, 2007; Arboleda and Nasser, 2006). Yet, since Cluster Heads (CHs) consume more energy in aggregating and routing messages (Liang *et al.*, 2011), it is significant to have an energy-efficient mechanism for CH selection and rotation. In order to minimize energy consumption, both optimal number of CHs and their optimal election should be pondered.

A wide range of energy-efficient clustering algorithms for wireless sensor networks have been addressed in literature. Low-Energy Adaptive Clustering Hierarchy (LEACH) (Chandrakasan and Balakrishnan, 2000), one of the first clustering algorithms proposed for sensor networks, is a distributed, self-adaptive and clustered routing protocol. The clusters in the network are formed based upon the received signal strength and uses local CHs as routers to the sink. Every sensor nodes makes its own decision to become a CH based upon a pre-determined value of optimal percentage of CHs in the network and also their previous experiences of being a CH. By applying a random rotation through the CHs, LEACH provides a balance of energy dissipation.

Hybrid Energy-efficient Distributed Clustering (HEED) (Younis and Fahmy, 2004) is a distributed clustering protocol that uses a hybrid criterion for cluster head election. HEED introduces two factors to elect CHs. For the first parameter it considers the remaining energy of the node and a secondary parameter is the node's degree. HEED prolongs the network lifetime by ensuring a uniform distribution of CHs across the network and adjusts the probability of CH-election to ensure inter-CH connectivity. Unlike LEACH, it does not elect CHs randomly.

Weighted Clustering Algorithm (WCA) (Chatterjee *et al.*, 2002) is a reactive clustering algorithm where cluster head selection is based on the evaluation, for every node, of a score function called "combined weight". This function is a weighted linear combination of the node's degree, transmission power, mobility and remaining energy of the node. Any node sends its combined weight to its neighbors and the node having the lowest weight is elected as CH. The algorithm also restricts the number of nodes in a cluster so that the performance of the MAC protocol is not decreased.

A distributed Energy-efficient Clustering Protocol (EECF) (Chamam and Pierre, 2010) proposed a novel distributed clustering algorithm where cluster heads are

selected following a three-way message exchange between each sensor and its one-hop neighbors. In each cluster one CH is elected based on its remaining energy and its degree. The performance of EECF is analyzed using network lifetime and ratio of elected CHs.

Fault Tolerant, Energy Efficient, Distributed Clustering (FEED) (Mehrani *et al.*, 2010) proposed a clustering algorithm which prolongs network lifetime by using energy, density and centrality factors and also the distances between nodes for making clusters. They assume a supervisor node for every cluster head which is to be its replacement when the cluster head fails. This property causes an increase in network lifetime and also helps the network to be fault tolerant. FEED has been shown to outperform HEED and LEACH, so it is chosen as a comparative base in our performance evaluation.

In the present study, a distributed clustering protocol for sensor networks, called Enhanced Energy-aware Clustering (EEAC) is proposed. This supposes that all nodes can be either a CH or member of cluster that every member has to be linked to a CH. Also, it is supposed different cluster size in the network based on the location information which is the distance between a cluster head and the sink. In our algorithm, the set of CHs is elected following a two-way message exchange between every sensor and its one-hop neighbors. CH election is based on remaining energy, density, centrality, in addition distance between nodes and the sink. Because the nodes located in the center of regions usually get higher structural significance than those positioned in the border (Oliveira *et al.*, 2010) and to avoid "border effect", hence, a border node detection algorithm to prevent the boundary node elected as cluster head is presented. Run on a flat network, our clustering algorithm completes with a clustered sensor nodes configuration that guarantee the maximum network life time.

PRELIMINARIES

Clustering algorithms can be based upon such criteria as the battery power of nodes, distance, density etc. In our work five parameters have been considered which are as follows:

Energy: One of the most significant research interests is energy efficiency in wireless sensor networks because it settles the lifetime of the sensor network. Residual energy is crucial to cluster heads because cluster heads suffer heavier burden than general cluster members (Chatterjee *et al.*, 2002). The energy will drain quickly as cluster head needs to not only collect data from its cluster members but also process, data aggregation and then

transmit messages to the sink. To ensure that the cluster heads perform their task without interrupt, the nodes are more eligible than the other nodes in terms of residual energy that have the maximum remaining energy.

Density: In a real network most of the nodes may focus in certain areas, and in other areas may exist quite sparse or even no node. The degree was originally proposed by Gerla and Tsai (1995) and Parekh (1994) in which it is calculated based on its distance from other nodes around a node. The algorithm should select cluster head from the areas containing high density with maximum number of neighbors (Chatterjee *et al.*, 2002; Yi *et al.*, 2009). In this study, the number of nodes around a node which are placed in the transmission range are considered as the value for density factor.

Centrality: Sometimes density of a node is high, but most of the neighbors are in one side of that node. The nodes located in the center of regions usually get higher structural significance than those positioned in the border (Oliveira *et al.*, 2010) hence, whenever there exists data following, central nodes play a significant role to pass data to next hop. Therefore, cluster heads are preferred to locate at the center of their corresponding neighbors that this causes more load-balancing (Kim *et al.*, 2008).

Distance between nodes: The distance between nodes is considered to keep the balance between clusters (Lee and Jeong, 2011). The clustering algorithm should try to minimize the transmit distances which causes to consume less energy (Chandrakasan and Balakrishnan, 2000; Wang *et al.*, 2011). In proposed clustering algorithm, a cluster head is selected by other nodes when it is located close to other respective nodes.

Distance to the sink: More network lifetime is achieved when the overall CHs' energy consumption is less. This parameter is directly related to the CH's proximity to BS (Torghabeh *et al.*, 2010) by observing that closer distance save more energy for forwarding messages between CH and sink (Gao *et al.*, 2010). The nodes which are closer to the sink should be elected most often than the nodes far from the sink since a great difference between energy levels of a near node and a far node would not occur (Aminit *et al.*, 2007).

Network model: In this study, a sensor network with N sensors that are randomly deployed to monitor temperature or light etc is considered. Sensors are ceaselessly sensing their local area and periodically send observed data to their own CH. The received data from

each sensor nodes can be aggregated into a single data packet and can be forwarded to the sink through other CHs or nodes in a multi-hop fashion. The network model described by Bandyopadhyay and Coyle (2003) to compute the total energy consumption by all the nodes in the network is used. For the sake of clearness, we now summarize the main assumptions behind our algorithm. It is assumed that nodes are deployed according to a Poisson process with parameter λ . There is only one sink in the network and is situated out of the network. Also, there are no mobile nodes. All nodes know their location information using location services (Bruck *et al.*, 2009; Mao *et al.*, 2007) and no GPS receiver is required. All nodes have same transmission range r_t . There is free of collision and error since well-scheduled communication is applied.

Border node detection: The objective of finding border nodes is to provide an algorithm that allows sensor nodes to recognize themselves as being placed on the border field. In some WSN applications, identifying border nodes is vital for tracking and guiding (Kroller *et al.*, 2006), geographic routing (Fang *et al.*, 2006; Kroller *et al.*, 2006) and topology discovery (Funke, 2005; Kroller *et al.*, 2006; Wang *et al.*, 2006).

Travencolo *et al.* (Bruno *et al.*, 2009) used the recently introduced concept of diversity (Bruno and Luciano Da, 2008; Costa, 2008) in order to quantify the potential of accessibility from and to each node while considering a specific dynamics, in the present case self-avoiding random walks. Nodes with more balanced transition probabilities to other nodes, expressed in terms of entropy are understood as being more internal to the network, while the other nodes are associated to the network borders. They compared their results of border detection with other measurements such as node density, centrality and average shortest path length through geographical and non-geographical networks. Zhang *et al.* (2009) developed an algorithm for boundary node detection based on two novel computational geometric techniques called localized Voronoi and neighbor embracing polygons. Their method needs only one-hop neighbors' information, which guarantees the scalability and energy efficiency of the detection algorithms and it requires only a limited number of simple local computations (Tong and Tavanapong, 2006) and presented the concept of "orifice" as an arc in the circle with the radius of node range which is not covered by other nodes. With union of orifices, it is possible to define a contour for both border of the network and the holes inside it. Orifices border identification scheme based on where the node's communication range intersections and overlaps in addition connectivity through nodes.

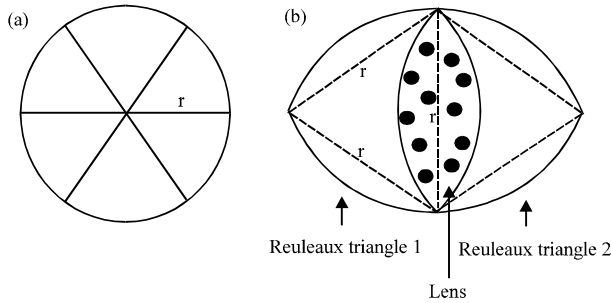


Fig. 2(a-b): (a) Slicing of sensing disk and (b) Adjacent Reuleaux triangles

```

Procedure Slice_Processing ()
Begin
  Randomly slice a sensing disk into six
  Reuleaux triangle, RT (i), 1 ≤ i ≤ 6
  i ≤ 0
  While i ≤ 6 Do
  Begin
    If RT (i) contains k active node Then
    Begin
      Update_Slice (i, "Connected")
    End
    i = i+1
  End
  If Total_Connected_Slice ≤ Threshold Then
  Begin
    Return ("Border Node Detected.")
  End
End
    
```

Fig. 3: Border node detection algorithm

Recent work (Ammari and Das, 2011) has proposed a solution for problem of minimum connected k -coverage (MMC_k) in heterogeneous and homogeneous wireless sensor networks, where each point in the field is covered or sensed by at least k active nodes while reducing the necessary total number of active nodes for energy-saving purposes and guaranteeing connectivity between them. To solve the MMC_k problem, they characterize k -coverage based on the intersection of k sensing disks and using the Reuleaux triangle model. For more detail of the their approach see the paper (Ammari and Das, 2011). Such a basic principle of the k -coverage idea is utilized in the present work in order to identify the border nodes in wireless sensor networks, in the sense that the coverage and connectivity would be directly related to the internality of the sensor nodes. The idea underlying this approach is conceptually simple and powerful. Note that, in our node detection algorithm it is needed to measure the level of connectivity of nodes rather than coverage. In more clearly, the border nodes have not been fully covered by their nodes around themselves, resulting in

low level of connectivity. On the other hand, non-border nodes are covered by the internal and peripheral nodes around themselves in the network, resulting in high level of connectivity. To facilitate the detection of border nodes using the characterization of k -coverage, firstly, it is worth mentioning the theorem 1 from (Ammari and Das, 2011).

Theorem 1: A field is guaranteed to be k -covered with a minimum number of nodes if all active nodes are selected from the lenses of adjacent Reuleaux triangles of width r .

The sensing disk of a nodes is k -covered by a minimum number of nodes if all active nodes are placed in the inner and/or border lenses of the six Reuleaux triangles of width r as shown in Fig. 2a and Fig. 2b. Therefore, a region is guaranteed to be k -covered with a minimum number of nodes if all lenses of the sensing disks of the nodes are k -covered only by nodes situated in them.

Border node detection algorithm: As mentioned before, the aim of this algorithm is to evaluate the level of connectivity which helps us to identifying the border nodes. Based on Theorem 1, each node i randomly separate its sensing disk into six overlapping Reuleaux triangle of width r and verify whether each one of them includes at least k active nodes. For our algorithm it just required each Reuleaux triangle of the sensing disk to be consisted at least one active ($k = 1$). The nodes whose the number of overlapping Reuleaux triangle which contains at least one active nodes (level of connectivity) is below a threshold are classified as border node. The advantage of the algorithm is each node evaluates itself, whether it is a border node or not, based merely on the local information it has about the neighbors, i.e., nodes located in its range. The algorithm proposed to classify the border nodes is presented in Fig. 3.

The algorithm is fully distributed with the meaning that only the position information of one-hop neighbors and a minimum number of simple local computations are required. This algorithm is performed by making holes in the original networks, so, that the border nodes and the original borders created by the holes are identified precisely. In addition, it is important to note that the thresholds used to obtain these results were chosen from one to six meaning that one and six are the smallest and highest level of connectivity.

Figure 4a-f present different threshold for the same network topology and the boundary nodes were classified in red color. According to Fig. 4a-c, for thresholds between one and three, almost no border

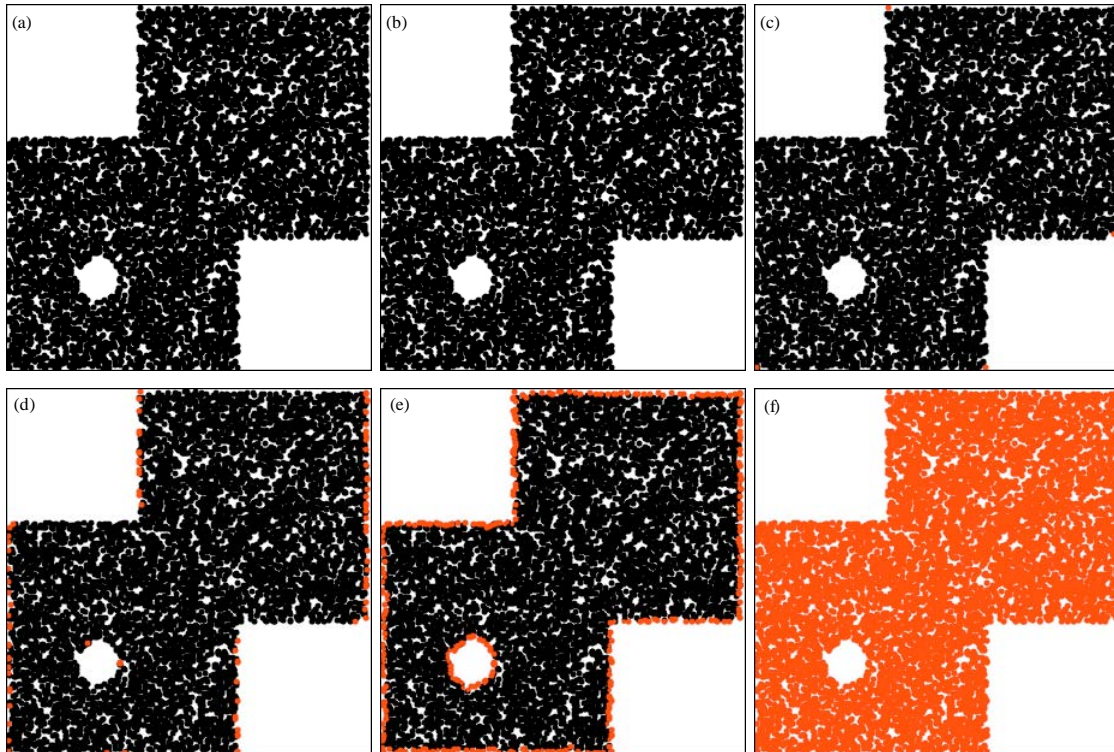


Fig. 4(a-f): Border detection using the concept of minimum connected coverage (MMC_t), (a-f) Different threshold used ranging from one to six

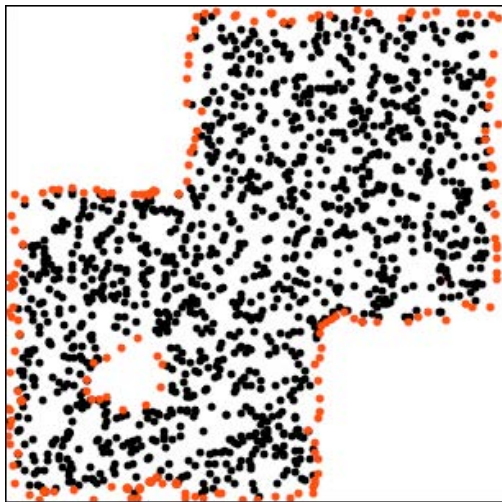


Fig. 5: Border detection for a sparse network. This network has the same shape of the network shown in Fig. 4e, but five times sparser. The boundary nodes were classified (red nodes) and when compared with the results acquired for the denser network, Fig. 4e, it can be seen that the obtained results are similar

nodes are detected. Figure 4d shows, four as threshold, the border nodes that are identified are not complete. Figure 4f depicted that the best case is five because the marginal nodes in the border of the network and the hole are detected precisely. Six is worst case, Fig. 4e, where almost all of the nodes identified as border nodes.

Figure 5 shows the network with low density, five times sparser, for the same network topology. For this network the threshold is considered five and the same border detection methodology is applied. As it is depicted, the border nodes are classified successfully. In sparse network there is no formal definition of the borders of networks, but as it can be seen, the nodes that are close to margin of the area are identified as border nodes which is the right definition.

In addition to verify the proposed approach of detecting boundary nodes, two measurements are used, density and centrality (Bruno *et al.*, 2009). The result of border node detection obtained by using density and centrality are shown in Fig. 6a and 6b, respectively. Among these two measurements, the best result was found for density that many of border nodes detected but still some nodes were not classified as border node. The same issue occurs, in a worse manner, when the centrality

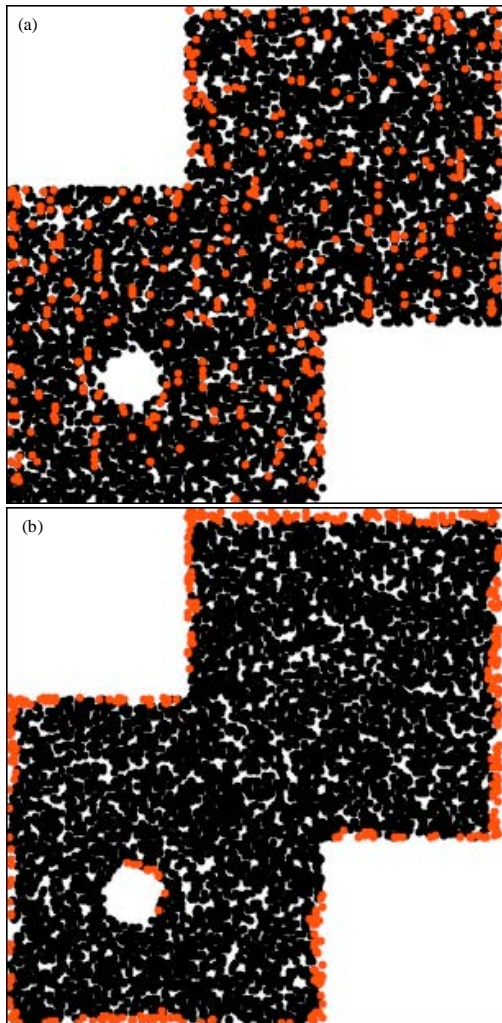


Fig. 6(a-b): (a) Border detection using the node degree. Although, almost of the border was correctly identified but some nodes were not also considered as border and (b) Border detection using the centrality and many nodes were misclassified

is used. In order to investigate the relationship between these two measurements and the level of connectivity for networks depicted in Fig. 6, their Pearson correlation coefficients have been calculated that are 0.468 and 0.475 for centrality and density as depicted in Fig. 7a and b, respectively. It can be noted from these coefficients that a strong correlation was found with regards to considered network. This can be explained, when the density of a network increases there are more nodes to deserve the connectivity between nodes. When a node that is placed in the center of region there is better connectivity because the neighbors of the node are placed more uniformly than

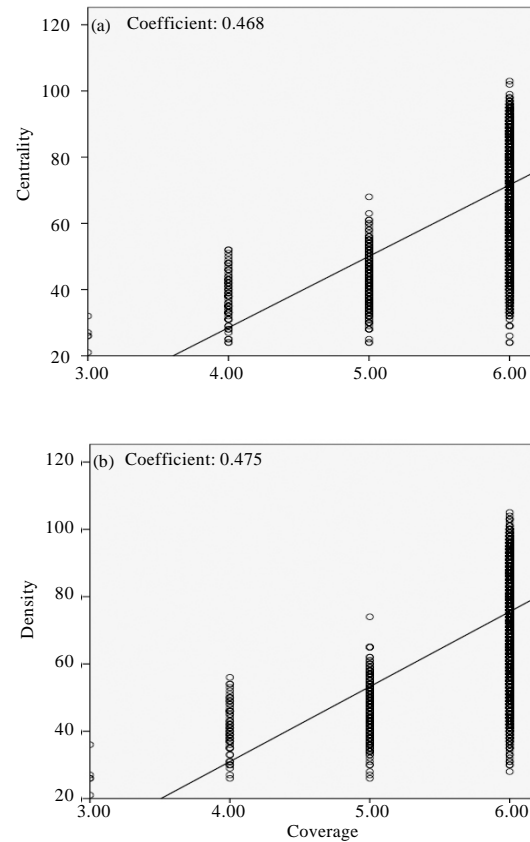


Fig. 7(a-b): Pearson correlation coefficients between the level of connectivity and the (a) Centrality and (b) Density, A strong correlation between these measurements was identified

the border nodes because most of the neighbors are in one side of it.

Cluster size: One of the most efficient utilization of the restricted energy resources of the sensor nodes is to organize wireless sensor networks into clusters. Nevertheless, the unbalanced energy consumption problem exists and it is closely linked to the location and the role of the individual nodes in the network. In wireless sensor networks, where many applications require a many-to-one traffic pattern (Atiq-ur-Rahman *et al.*, 2011), unequal distribution of energy becomes an extremely significant issue, as a “hot spot” is created around the sink (Perillo *et al.*, 2005). The nodes placed in this “hot spot” are needed to forward an unbalanced high amount of traffic and generally lose all their energy at a very early stage. Oftentimes the network is organized into clusters of equal size, but such equal clustering results in uncontrolled number of clusters and unequal load on the cluster head nodes. The cluster size is a significant

parameter that decides the number of clusters in the network and total energy consumption in the cluster. By increasing the size of cluster, the number of clusters diminishes. Therefore, the energy consumption by inter-cluster communication reduces, but the energy consumption by intra-cluster communication increases in proportion to the cluster size.

In unequal clustering, the network is separated into clusters with different sizes. The clusters adjacent to the sink are smaller than the clusters that are far from the sink. Unequal clustering size configuration model for network, causes more balance energy depletion among the cluster head nodes, therefore network lifetime increases (Soro and Heinzelman, 2005). By applying this not only the “hot spot” problem considerably would alleviate, as well it increases the energy efficiency of the network, since data is sent over many fewer hops on average (Perillo *et al.*, 2005).

The unequal clustering algorithm aims to create the clusters with larger size as increasing distance from the sink. To solve the “hot spot” problem, (Soro and Heinzelman, 2005) propose an unequal clustering algorithm. The network area is separated into cirques. Clusters in the same cirque have the same size, while clusters in the different cirques have different sizes. Deploying some high-energy nodes take on the cluster head role to control network operation and make sure that the energy depletion of these cluster heads is balanced. The algorithm can effectively extend network lifetime. However, some high-energy nodes are needed to play as cluster heads, and the positions of cluster heads must be calculated previously, which may not be true for the real situation. Min *et al.* (2010) proposes a cirque-based static clustering algorithm for multi-hop wireless sensor networks. Similar to reference (Soro and Heinzelman, 2005), the circular network field is divided into cirques. Clusters closer to the BS have smaller sizes, so that a part of the cluster heads’ energy can be preserved for forwarding. Furthermore, a continuous working mechanism is given for reducing the energy consumption of cluster reform and prolonging the network lifetime. Lee *et al.* (2011) tackled the equal clustering problem and presented the theoretical analysis of optimal cluster size based on the distance between the CH and the sink. In their clustering algorithm, each cluster has a different cluster size based on its location information which is the distance between a cluster head and a sink. Jiang *et al.* (2010) proposed PSO algorithm that selects an optimal group of high-energy nodes as the cluster heads and divides all nodes in the network into

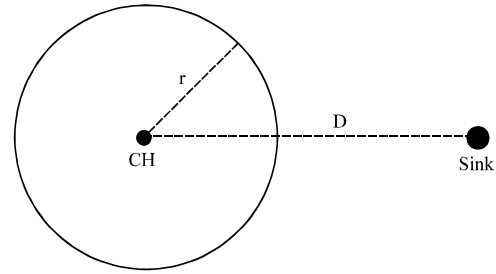


Fig. 8: A simple model for a single cluster head

different size clusters. This algorithm minimizes the intra cluster distance between the cluster head and its members and also optimizes the energy management of the network by making the clusters closer to the sink having smaller size.

The analytical model of cluster size proposed by Lee *et al.* (2011) is used which is based on the distance from the sink. Each sensor computes a cluster size based on its location information. As mentioned previously, due to a trade-off between inter- and intra cluster communication, the effect of clustering is reduced as the distance from the sink increases. However, adopting the unequal clustering approach in our clustering algorithm, the equal clustering problem and also generating better-cost clustering with respect to the optimal number of clusters can be mitigated.

Let r_c , r and D denote the transmission range, the cluster size and the distance between the CH and the sink, respectively as shown in Fig. 8. The model (Lee *et al.*, 2011) is given as follows:

$$r = \left[\frac{1}{r_t} \sqrt[3]{\frac{3D}{\lambda\pi}} \right] \tag{1}$$

The Eq. 1 will use in the first step of our algorithm to adjust the transmission range of the sensor nodes.

DESCRIPTION OF PROPOSED CLUSTERING ALGORITHM

In clustering algorithm, cluster head acts as the local control center and is burdened with transmitting data from other cluster heads through multi-hop, thus, the energy dissipation of the cluster head is much more than that of the general nodes. Obviously, to maintain the connectivity of the entire network, it is very important that the cluster heads keep alive as long as possible for the inter-cluster communication. The proposed clustering algorithm makes local decisions to determine transmission radius and to elect cluster-heads. The boundary node detection algorithm lets not to select CHs from the border

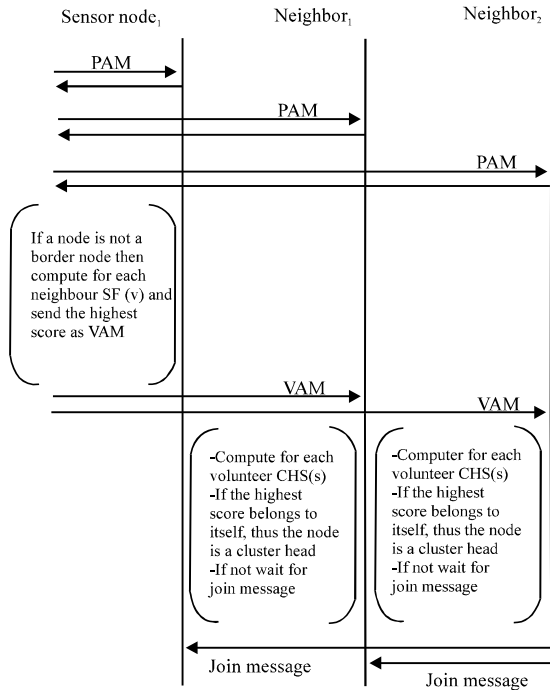


Fig. 9: Description the EEAC protocol, executed by each sensor node

of the area. Our algorithm is fully distributed. As shown in Fig. 9, it consists of a number of actions and message exchanges executed by each sensor, which leads to the election of a set of CHs and ensures that each regular sensor (not CH) is connected to a CH. Here is the description of our proposed clustering algorithm:

- After the sensors are deployed, each sensor i can compute its distance from the sink via location services such as (Bruck *et al.*, 2009; Mao *et al.*, 2007) and no GPS receiver is required. Let r_i be the cluster size, in terms of the number of hops, when sensor i become the CH. Note that r_i can be derived by Eq. 1. A cluster size r_i is used for forwarding bound thus, nodes adjust their transmission range with size r_i . Then, each node disseminates to all perspective nodes around itself a message called Preliminary Advertisement Message (PAM). This message contains the node's ID and its geographical coordinate. Each node computes the distance between itself and neighbors by knowing the information about their coordinates
- After receiving a PAM from all its neighbors, each node i stores its neighbors into set $N(i) = \{v \in \text{Neighbors}(i)\}$. Because the aim of our clustering algorithm is to prevent the border nodes to participate in CH election, thus, the nodes that are

located in the border only wait for receiving cluster head join message from the nearest cluster head and will not proceed from this step. To identifying the border nodes, the algorithm described in "border node detection" section is utilized. All nodes are aware about their residual energy. Subsequently, each node calculates for each of its neighbors a score using the linear scoring function $SF(v)$, representing the relative preference that sensor v has for sensor i as its potential CH, according to each neighbor's distance which is given as follows:

$$SF(v) = (w_1E) + (w_2DE) - (w_3CE) - (w_4Dis_{bs}) - (w_5Dis_{iv}) \quad (2)$$

where, w_1, w_2, w_3, w_4 and w_5 are the weighting coefficients for the energy (E), density (DE), centrality (CE), distance between nodes and the sink (Dis_{bs}) and distance between nodes i and v (Dis_{iv}), respectively. The density and centrality factors can be calculated according to the proposed algorithm (Mehrani *et al.*, 2010). Far nodes in a region are given a negative score, thus, they have a little chance to become a cluster head for that region. Afterwards, the node with the highest score is elected from the set $N(i)$ and introduce to other neighbors as Volunteer Advertisement Message (VAM) which includes node's ID and score. A node is a "volunteer" when is not placed in the border of the network and has the highest score $SF(v)$ at each node.

- After receiving a VAM from all its neighbors, each node stores it in the set $N_{vt}(i) = \{v \in \text{Neighbors}(i)\}$. Subsequently, node i computes, for each $z \in N_{vt}(i)$, a function $CHS(z)$, representing the volunteers final score at each sensor i . The function $CHS(z)$ expected as follows:

$$CHS(z) = ps \quad (3)$$

where, p determines the number of nodes selected a volunteer and s is the sum of received scores for a volunteer. As it can be seen, the number of voters is very significant for volunteers to become a cluster head. For the remaining of this phase, each node selects a node with a high score from the $N_{vt}(i)$. If the highest score belongs to itself, thus, the node is a cluster head then introduces itself to neighbors as cluster head by a join message holding its ID, otherwise it waits until a predefined timeout to receive the join message from the near cluster head. Let us note here that each sensor stores the received join message to its cluster head list. Finally regular

nodes decide to join to the closest cluster head. If a sensor i has no neighbors and has no elected CH within its range, it promotes itself CH. Then, the members go to steady phase and periodically sense information from the environment and send the data to their cluster head.

Proposition 1: The proposed algorithm has an algorithmic complexity of $O(N)$ at each sensor node.

Proof: The execution of our clustering algorithm at each node i consists of a sequence of three steps: (1) Broadcasting the local information and ID of each node through the PAM message (2) calculating the SF (v) function for each neighbor v in $N(i)$ (after all PAM messages are received from neighbors) (3) calculating the CHS (z) for each volunteer z as the final score. Processing time at Steps (2) and (3) is proportional to the number of i 's neighbors which, in the worst case, is equal to N . Hence, the overall complexity is $O(N)$.

PERFORMANCE EVALUATION

Simulation setup: In our clustering algorithm, CHs are elected based on density, centrality, distance between nodes and distance between nodes to the sink in addition remaining energy at the instant the clustering protocol is executed. Hence, different residual-energy maps can lead to different sets of CHs. The simulation starts from the initial phase, which every node sends the PAM message to neighbors including its ID and coordinate. Then, each node tries to identify itself whether a boundary node is or not. The border nodes are prevented to take part in CHs election. Afterwards, for each volunteer a score is computed that present the relative preference that a sensor has for each of its neighbors as a potential CH. The neighbor with highest score is chosen as a candidate to be voted by other nodes in the region. Subsequently, the ID and score of the candidate will reveal to others as a VAM message. Here the number of voters that vote a volunteer is important. Then, each sensor for each volunteer computes a score as the final score. If the highest score belong to current node then nodes introduce itself to others as a CH otherwise it waits to receive a join message. When executed once, the clustering algorithm selects a set of CHs that suppose to increase the lifetime of the network, as if this configuration will be retained for the entire lifetime of network.

However, since sensors remaining energies are progressively drained during network operation, our algorithm will be run periodically to discover the new CHs from rest of the nodes. Every pre-determined period of time T , a new set of CHs will be elected, based on the actual residual energies, density and centrality at the

instant the algorithm is run. Without loss of generality, $T = 1$ unit is chosen (Chamam and Pierre, 2010). A simplified form of the log-distance path loss channel model is utilized as described by Heinzelman *et al.* (2002). For both regular sensors and CHs a free-space fading model is used. Therefore, the energy consumed by a sensor to transmit l bits over a range R is: $E_{Tx}(l,R) = l \cdot E_{elec} + l \cdot \epsilon \cdot R^2$ where E_{elec} is the amount of energy consumption per bit to run the transmitter or receiver circuitry. ϵ is the amplifier's energy needed to transmit 1 bit of data over a proportionately short distance which is as follows:

$$\epsilon = \frac{1}{G_t G_r} \left(\frac{4\pi}{\lambda} \right)^2$$

where, G_t is the gain of the transmitter's antenna, G_r is the gain of the receiver's antenna and λ is the wavelength. The amount of energy consumption in receiving a packet with l bits can be calculated like follow: $E_r(l) = l \cdot E_{elec}$. Hence, the energy depleted by a CH to transmit l bits of data over its range R_{CH} and to receive and fuse an l -bit length message from each of the regular nodes connected to it, is:

$$E_{CH} = l \cdot E_{elec} + l \cdot \epsilon \cdot (R_{CH})^2 + N_{CH} \cdot l \cdot E_{elec} + N_{CH} \cdot l \cdot E_A$$

where, E_A is the energy dissipated in data aggregation and N_{CH} is the number of regular sensors connected to the CH. For the simulations, the same parameter values is considered as Heinzelman *et al.* (2002), that are: $\epsilon = 10$ pJ/bit/m², $E_{elec} = 50$ nJ/bit and $E_A = 50$ nJ/bit/signal. Also, a message of 100 bytes for all messages is taking into account.

In order to verify the proposed clustering algorithm, a series of computer simulations are run. We developed our event-driven simulation program in C++. The proposed clustering algorithm is compared to a recently published clustering protocol FEED (Mehrani *et al.*, 2010). As performance metrics, the network lifetime and the ratio of CHs with respect to the total number of sensors are considered. The efficiency of our clustering algorithm is investigated in terms of the minimum number of CHs needed to fulfill a distributed clustered network. The values used for simulation were $w_1 = 0.5$, $w_2 = 0.2$, $w_3 = 0.1$, $w_4 = 0.1$ and $w_5 = 0.1$.

Evaluation of proposed algorithm: To measure the network lifetime generated by our algorithm and FEED, these two protocols periodically are run, at instants $k.T$ ($k \in \mathbb{N}$) and T as a measure unit of the network lifetime is used. The proposed algorithm under various network environments via simulation experiments is evaluated. Sensors are randomly deployed in a square area. We

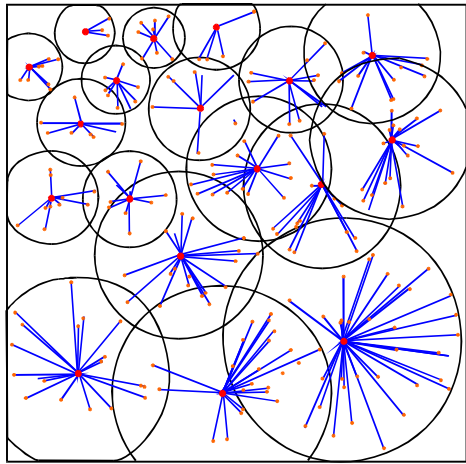


Fig. 10: Illustration of clustering formation by EEAC

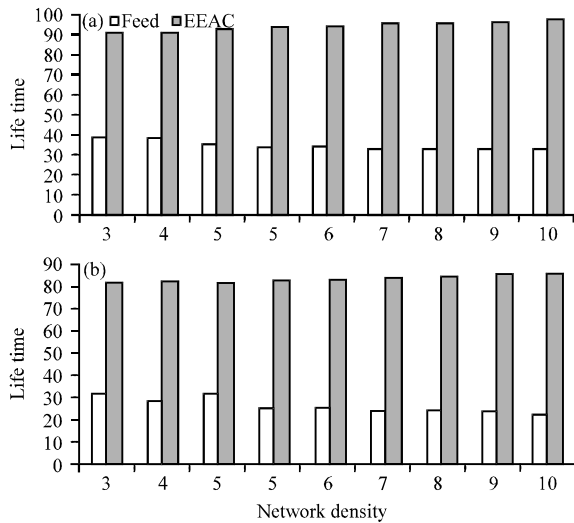


Fig. 11(a-b): Network lifetime generated by EEAC vs. FEED, for various network size; (a) 300 and (b) 500

assume that there is no transmission delay. The data packets are generated twice in every one second by members after the cluster election has finished. The network lifetime under various network densities ranging from 3 to 10 is measured.

Figure 10 shows an example clustering result generated by our proposed algorithm in a square area of size $A = 300$, where the network density is 4. Note that a circle represents forwarding bound of messages with size r_1 . From Fig. 10, it can be seen that the cluster head election distributed through the network and each cluster has a different cluster size according to its location. As the distance from the sink increases, a cluster has relatively large cluster size. It implies that the cluster has

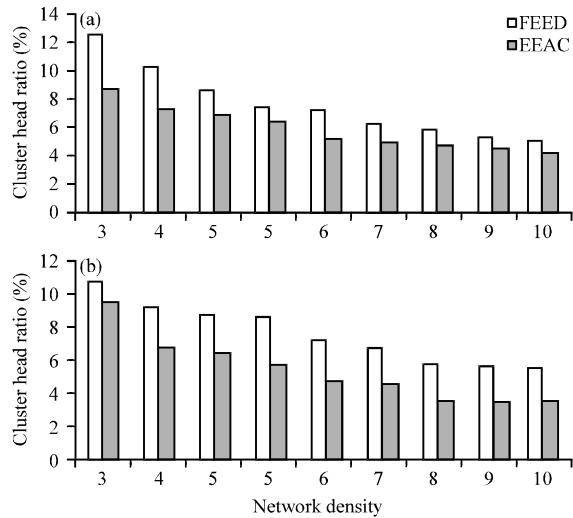


Fig. 12(a-b): Ratio “Number of CHs/Total number of sensors” generated by EEAC vs. FEED, for various network size, (a) 300 and (b) 500

more cluster member and it collects more data and transmits the aggregated data to the sink. Figure 11 depicts the network lifetime generated by EEAC and FEED, for two different network topologies, $A = 300$ and $A = 500$. EEAC outperforms FEED by more than 125% for $A = 300$, Fig. 11a and approximate 160% for $A = 500$ (Fig. 11b). This could be explained by less message exchange of cluster head election phases and unequal clustering formation of the network. It means that the equal clustering of FEED cannot adopt the various network conditions and may waste energy.

As mentioned earlier, the cluster size significantly affects the total energy consumption thus, unequal cluster size of EEAC can relax the degradation of clustering effect according to the distance from the sink and balance the energy consumption between intra- and inter-cluster communications. Besides, since EEAC prevents border nodes become cluster head and they are chosen from the center of the region with high density and centrality therefore, this causes more balance of energy consumption among the whole network. The ratio “Number of CHs/Total number of sensors” expresses the efficiency of the protocol in terms of the number of CHs required to cluster the network. Fig. 12 depicts the ratio of CHs for EEAC vs. FEED, for $A = 300$, Fig. 12a and for $A = 500$, Fig. 12b. It is shown that, for large-scale configurations ($\lambda = 4$ and more), EEAC outperforms FEED by approximately 30% for $A = 300$ and 22% for $A = 500$. This means that our protocol is able to find acceptable configurations with less CHs, which is very coherent with

the results concerning the network lifetime, because using less CHs will, in average, increase the network lifetime.

CONCLUSION

In this study, an energy-efficient and distributed clustering algorithm is presented for wireless sensor networks that aims at prolonging the network lifetime. Proposed algorithm can dynamically adjust itself with the ever network. Our proposed algorithm organizes sensor nodes into clusters based some significant parameters. In addition an algorithm for border node detection proposed which used to prevent the boundary nodes become cluster head. In our clustering algorithm far cluster heads have bigger size than the cluster heads near the sink that causes more balance energy through the network. It showed that our algorithm, in the worse case, has an algorithmic complexity of $O(N)$ at each sensor node. Performance evaluation showed that our algorithm provides a better network lifetime and a better ratio "Number of CHs/Total number of sensors" than FEED, a recently published clustering protocol for wireless sensor networks.

REFERENCES

- Abido, A.P., N.A. Azeez, A.O. Adesina and K.K. Agbele, 2011. UDCA: Energy optimization in wireless sensor networks using uniform distributed clustering algorithms. *Res. J. Inform. Technol.*, 3: 191-200.
- Aminit, N., M. Fazelit, S.G. Miremadi and M.T. Manzuri, 2007. Distance-based segmentation: An energy-efficient clustering hierarchy for wireless microsensor networks. *Proceedings of the 5th Annual Conference on Communication Networks and Services Research*, May 14-17 2007, Frederlcton, NB., pp: 18-25.
- Ammari, H.M. and S.K. Das, 2011. Scheduling protocols for homogeneous and heterogeneous-covered wireless sensor networks. *Pervasive Mob. Comput.*, 7: 79-97.
- Arboleda, L.M.C. and N. Nasser, 2006. Comparison of clustering algorithms and protocols for wireless sensor networks. *Proceedings of the Canadian Conference Electrical and Computer Engineering*, May 2006, Ottawa, pp: 1787-1792.
- Atiq-ur-Rahman, H. Hasbullah and O.U. Khan, 2011. Energy holes mitigation techniques in sink's proximity using sensor deployment in wireless sensor networks. *Res. J. Inform. Technol.*, 3: 167-180.
- Bandyopadhyay, S. and E.J. Coyle, 2003. An energy efficient hierarchical clustering algorithm for wireless sensor networks. *Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications*, March-30-April, 3, IEEE Societies, pp: 1713-1723.
- Bruck, J., J. Gao and A. Jiang, 2009. Localization and routing in sensor networks by local angle information. *ACM Trans. Sen. Netw.*, 5: 1-31.
- Bruno, A.N.T. and F.C. Luciano Da, 2008. Hierarchical spatial organization of geographical networks. *J. Phys. A: Math. Theor.*, 41: 224004-224004.
- Bruno, A.N.T., V.M. Palhares and C.L. Da Fontoura, 2009. Border detection in complex networks. *New J. Phys.*, 11: 063019-063019.
- Chamam, A. and S. Pierre, 2010. A distributed energy-efficient clustering protocol for wireless sensor networks. *Comput. Electr. Eng.*, 36: 303-312.
- Chandrakasan, A. and H. Balakrishnan, 2000. Energy-efficient communication protocol for wireless microsensor networks. *Syst. Sci.*, 2: 10-10.
- Chatterjee, M., S.K. Das and D. Turgut, 2002. WCA: A weighted clustering algorithm for mobile ad hoc networks. *Cluster Comput.*, 5: 193-204.
- Chong, C.Y. and S.P. Kumar, 2003. Sensor networks: Evolution, opportunities and challenges. *Proc. IEEE*, 91: 1247-1256.
- Costa, L.D.F., 2008. Inward and outward node accessibility in complex networks as revealed by non-linear dynamics. *Preprint cond-mat/0801.1982*.
- Fang, Q., J. Gao and L.J. Guibas, 2006. Locating and bypassing holes in sensor networks. *Mob. Netw. Applied*, 11: 187-200.
- Funke, S., 2005. Topological hole detection in wireless sensor networks and its applications. *Proceedings of the Joint Workshop on Foundations of Mobile Computing*, September 2, 2005, Cologne, Germany.
- Gao, H., H. Li and Y. Cheng, 2010. A hybrid relative distance based cluster scheme for energy efficiency in wireless sensor networks. *Proceedings of the IEEE Global Telecommunications Conference*, December 6-10, 2010, Miami, FL USA., pp: 1-5.
- Gerla, M. and J.T.C. Tsai, 1995. Multiclustor, mobile, multimedia radio network. *Wireless Networks*, 1: 255-265.
- Guo, L., B. Wang, Z. Liu and W. Wang, 2010. An energy equilibrium routing algorithm based on cluster-head prediction for wireless sensor networks. *Inform. Technol. J.*, 9: 1403-1408.
- 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.

- Idris, M.Y.I., E.M. Tamil, N.M. Noor, Z. Razak and K.W. Fong, 2009. Parking guidance system utilizing wireless sensor network and ultrasonic sensor. *Inform. Technol. J.*, 8: 138-146.
- Jiang, C.J., W.R. Shi, M. Xiang and X.L. Tang, 2010. Energy-balanced unequal clustering protocol for wireless sensor networks. *J. China Univ. Posts Telecommun.*, 17: 94-99.
- Karl, H. and A. Willig, 2007. *Protocols and Architectures for Wireless Sensor Networks*. John Wiley and Sons, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, England.
- Kim, J.M., S.H. Park, Y.J. Han and T.M. Chung, 2008. CHEF: Cluster head election mechanism using fuzzy logic in wireless sensor networks. *Proceedings of the 10th International Conference on Advanced Communication Technology*, February 17-20, 2008, Gangwon-Do, pp: 654-659.
- Kroller, A., S.P. Fekete, D. Pfisterer and S. Fischer, 2006. Deterministic boundary recognition and topology extraction for large sensor networks. *Proceedings of the 17TH Annual ACM-SIAM Symposium on Discrete Algorithm*, January 22-26, 2006, Miami, Florida, USA .
- Lee, C. and T. Jeong, 2011. FRCA: A fuzzy relevance-based cluster head selection algorithm for wireless mobile ad-hoc sensor networks. *Sensors*, 11: 5383-5401.
- Lee, S., H. Choe, B. Park, Y. Song and C. Kim, 2011. LUCA: An energy-efficient unequal clustering algorithm using location information for wireless sensor networks. *Wireless Pers. Commun.*, 56: 715-731.
- Liang, W., J. Xu and Y. Liu, 2011. Towards energy saving and load balancing data aggregation for wireless sensor networks. *Inform. Technol. J.*, 10: 409-415.
- Liu, Z., B. Wang and L. Guo, 2010. A survey on connected dominating set construction algorithm for wireless sensor networks. *Inform. Technol. J.*, 9: 1081-1092.
- Mao, G., F. Baris and D.O. Anderson, 2007. Wireless sensor network localization techniques. *Int. J. Comput. Telecommun. Networking*, 51: 2529-2553.
- Mehrani, M., J. Shanbehzadeh, A. Sarrafzadeh, S.J. Mirabedini and C. Manford, 2010. FEED: Fault tolerant, energy efficient, distributed Clustering for WSN. *Proceedings of the 12th International Conference on Advanced Communication Technology*, February 7-10, 2010, Young Researchers Club, Islamic Azad Univ., Dezful, Iran, pp: 580-585.
- Memariani, J., Z.A. Zukarnain, A. Abdullah and Z.M. Hanapi, 2011. Distributed, energy-efficient, fault tolerant Weighted Clustering. *Inf. Eng. Inf. Sci.*, 253: 568-577.
- Min, X., S. Wei-Ren, J. Chang-Jiang and Z. Ying, 2010. Energy efficient clustering algorithm for maximizing lifetime of wireless sensor networks. *AEU-Int. J. Electronics Commun.*, 64: 289-298.
- Oliveira, E.M.R., H.S. Ramos and A.A.F. Loureiro, 2010. Centrality-based routing for wireless sensor networks. *Proceedings of the Wireless Days (WD)*, October 20-22, 2010, Dept. of Comput. Sci., Fed. Univ. of Minas Gerais, Belo Horizonte, Brazil, pp: 1-5.
- Parekh, A.K., 1994. Selecting routers in ad-hoc wireless networks. *Proceedings of the SBT/IEEE International Telecommunication Symposium*, August 22-25, 1994, Siio Paulo, Brazil.
- Perillo, M., Z. Cheng and W. Heinzelman, 2005. An analysis of strategies for mitigating the sensor network hot spot problem. *Proceedings of the 2nd Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous'05)*, July 17-21, IEEE Computer Society, San Diego, USA., pp: 474-478.
- Rezaii, T.Y. and M.A. Tinati, 2009. Improved joint probabilistic data association filter for multi-target tracking in wireless sensor networks. *J. Applied Sci.*, 9: 924-930.
- Soro, S. and W.B. Heinzelman, 2005. Prolonging the lifetime of wireless sensor networks via unequal clustering. *Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium*, April 4-8, 2005, IEEE Computer Society Press, San Francisco, USA., pp: 236-240.
- Tong, B. and W. Tavanapong, 2006. On discovering sensing coverage holes in large-scale sensor networks. *Technical Report TR 06-03*, Computer Science, Iowa State University.
- Torghabeh, N.A., M.R.A. Totonchi and M.H.Y. Moghaddam, 2010. Cluster head selection using a two-level fuzzy logic in wireless sensor networks. *Proceedings of the Computer Engineering and Technology, 2010 2nd International Conference on*, April 16-18, 2010, Department of Computing Engineering, Islamic Azad University, Mashad, Iran, pp: 357-361.
- Wang, W., Z. Liu, X. Hu, B. Wang, L. Guo, W. Xiong and C. Gao, 2011. CEDCAP: Cluster-based energy-efficient data collecting and aggregation protocol for WSNs. *Res. J. Inform. Technol.*, 3: 93-103.
- Wang, Y., J. Gao and J.S.B. Mitchell, 2006. Boundary recognition in sensor networks by topological methods. *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking*, September 24-29, 2006, Los Angeles, CA, USA.

- Xin, G., W.H. Yang and B. DeGang, 2008. EEHCA: An energy-efficient hierarchical clustering algorithm for wireless sensor networks. *Inform. Technol. J.*, 7: 245-252.
- Yi, G., S. Guiling, L. Weixiang and P. Yong, 2009. Recluster-LEACH: A recluster control algorithm based on density for wireless sensor network. *Proceedings of the 2nd International Conference on Power Electronics and Intelligent Transportation System*, December 19-20, 200, Shenzhen, pp: 198-202.
- Younis, O. and S. Fahmy, 2004. Distributed clustering in ad-hoc sensor networks: A hybrid, energy-efficient approach. *Proceedings of the Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, Department of Computing Science, March 7-11, 2004, Purdue University, West Lafayette, In, USA.
- Zhang, C., Y. Zhang and Y. Fang, 2009. Localized algorithms for coverage boundary detection in wireless sensor networks. *Wireless Netw.*, 15: 3-20.