Indonesian Journal of Electrical Engineering and Computer Science
Vol. 3, No. 1, July 2016, pp. 143 ~ 150
DOI: 10.11591/ijeecs.v3.i1.pp143-150

An Energy Efficient Clustring Algrithm Based on DEEC Protocol and K-mean Method

Elahmadi Cheikh*, Chakkor Saad, Baghori Mostafa, Hajraoui Abderrahmane,
Department of Physics, Communication and Detection Systems Laboratory
University of Abdelmalek Essaâdi, Faculty of Sciences, Tetouan,
MoroccoCentre of Renewable Energy, Laayoune, Morocco
*Corresponding author, e-mail: Elahmadi13@gmail.com

Abstract

In WSN the sensor nodes are usually powered by batteries and thus have very limited lifetime if no power management is performed. Because of this major limitation, energy-efficient techniques are main research challenge in this context to solve it. Dividing the network in clusters is an effective technique to achieve this goal. This algorithm is based on creating virtual sub-groups of sensor nodes in order to minimize routing calculations and to reduce the size of cluster head data aggregation. Nowadays, a lot of heterogeneous clustering protocols for WSN are created. Nevertheless, these protocols need to find the optimal clusters formation in the network that conserve CHs and theirs member nodes energy consumption. A new approach is proposed combining between an efficient clustering algorithm K-means and our proposed DEEC. This approach has been employed to enhance DEEC protocol performances. Numerical simulation proves that the improved protocol entitled KM-DEEC achieves a satisfactory results compared to others DEEC protocol versions.

Keywords*: K-means, heterogeneous wireless sensor networks, Energy-efficiency, TDEEC protocol, network lifetime.*

1. Introduction

Great advances in communication technology have led to the development of intelligent, lightweight, low cost sensor nodes that are small in size and cooperatively collect data from the place of deployment [1].

They help us by extending our capability to perfectly monitor, study, control objects and environments of various scales and conditions such as human bodies, habitats, and security surveillance [3, 4]. Thus WSN suffers from several challenges such as network protocols, data gathering and distribution, energy management, fault detection and security [7]. Generally, designing the WSNs was very difficult because of the limited computation capability, limited power and small memory size [2]. In fact, the energy consumption is the most important factor among these three factors, because the battery is not changeable if once the sensor nodes are deployed [5].The energy is also the major consideration in designing the routing of the WSNs. In the literature there is a lot of methods and protocols aims energy consumption [6]. Clustering is one of these methods used to manage network energy consumption efficiently [8]. This approach is based on the grouping of similar objects and a clustering of a set is a partition of its elements that is chosen to minimize some measure of dissimilarity [9]. Eventually nodes are divided into some clusters and some nodes based on some parameter are the selected as cluster heads (CH) [10]. These cluster heads swap data with the Base station (BS) which costs the most energy of the nodes [11]. The optimal selection of the cluster heads with high energy that scattered evenly in the area can be seen as a great challenge in hierarchical routing protocols mainly for heterogeneous sensor networks [12].

Actually, clustering gives the best results this is the reason why we adopted this approach due to its efficient to improve the scalability of the network and to achieve the energy efficient routing [15]. So With these advantages, partitioned clustering can afford promising distributed implementation of deterministic approach [14]. Then it can be used as the optimal routing protocol for heterogeneous wireless sensors network. So there are a lot of centralized methods as well-known and well-studied distributed deterministic partitioned clustering such as

Received February 2, 2016; Revised May 25, 2016; Accepted June 10, 2016

K-mean which features simple, highly reliable, and fast-convergent iterations and re-clustering during failure states [13]. In this paper we introduce the k-mean algorithm to the routing protocol PROPOSED DEEC [13] in order to release efficient energy consumption and extend the network's lifetime.

Our paper is organized as follows: first, in section II we present an analytical related works and their features offered to clustering techniques. In section III we resume the problem formulation. Then, section IV deals with the centralized way of clustering nodes using k-means algorithm. The section V is dedicated to the PROPOSED DEEC algorithm. In the section VI we show through the numerical simulation the efficient of our proposed routing protocol. Finally, we present the conclusion and future work perspectives.

2. Related Work

Clustering is made to connect similar nodes and saves necessary energy dissipated in direct data transmission to the base station. Nodes organize themselves into hierarchical tier-frameworks. Within each cluster, data aggregation and forwarding are executed at cluster-head to decrease the amount of data transmitting to the base station. Generally, the remaining energy of sensor nodes and sensor's proximity to base station are the main factors in cluster formation and cluster heads designation. Automatically other nodes choose their appropriate cluster head just after deployment and transmit sensed information to the base station through the cluster head. Furthermore, many different traditional clustering algorithms for wireless sensor networks have been proposed. Often, in WSNs, the primary concern is the energy efficiency in order to extend the utility of the network [17]. Several cluster-based protocols have been proposed in the literature, with the objective of increasing the sensor network efficiency. The LEACH protocol presented in [18] is a distributed cluster-based protocol in which the cluster heads are selected with some probability mostly for the homogeneous networks. Though this protocol does not agreement that the desired number of cluster heads is selected and cluster heads are not evenly positioned across the network. In PEGASIS [20], further improvement on energy-conservation is suggested by connecting the sensors into a chain which each cluster member searches for a neighbor closer than the cluster head within the cluster to set up an energy-saving and delay-adaptive data relay link. On the other hand the protocol HEED [21] extends LEACH by incorporating communication range limits and intra-cluster communication cost information.

In fact, the heterogeneous protocols are more energy efficient than the homogeneous ones. Q. Li et al. have proposed Distributed Energy Efficient Clustering Protocol (DEEC) [19]. This protocol is based on multi-level and two level energy heterogeneous schemes. The cluster heads are selected using the probability utilizing the ratio between residual energy of each node and the average energy of the network. The epochs of being cluster-heads for nodes are different according to their initial and residual energy. Thus the aim problem of DEEC that it doesn't takes into account the cluster size and density of nodes in the clusters to elect their heads. When multiple cluster heads are randomly selected within a small area, a large additional loss of energy occurs because the election of cluster heads is performed periodically where the density of nodes is high and therefore the distance between these nodes will be very small. The amount of lost energy is approximately proportional to the number of cluster heads in this area. We had resolve this problem in PROPOSED DEEC [13] based on modified energy efficient algorithm for choosing cluster heads that exclude a number of low energy levels nodes due to their distribution density and their dimensions area. There are others approach enhance the lifetime of network and improve the performances of routing protocols in heterogeneous networks such as DDEEC [22], SDEEC [23], TDEEC [24], ETDEEC [25].

3. Problem Formulation

As mentioned in [13], it was noticed that in the cluster head election procedure the ratio between the nodes residual energy and the total energy of the network is a decisive criterion. Now a serious question occurs in a mind which is that: how to choice a optimal CH position which maintains energy consumption of the nodes? To overcome this problem, we have leaning to employ k-means clustering algorithm in this paper. This method relies on the centroid random partitioning principle for the network clustering formation.

4. K-means Algorithm for WSN Clustering

The K-means clustering algorithm is the most widely used clustering approach that attempts to find the cluster's centre point by minimizing the distance between the points assigned to that cluster and this virtual center. Then K-means algorithm is based essentially on the Euclidian distances and cluster head selection [12]. As illustrate in Figure 2

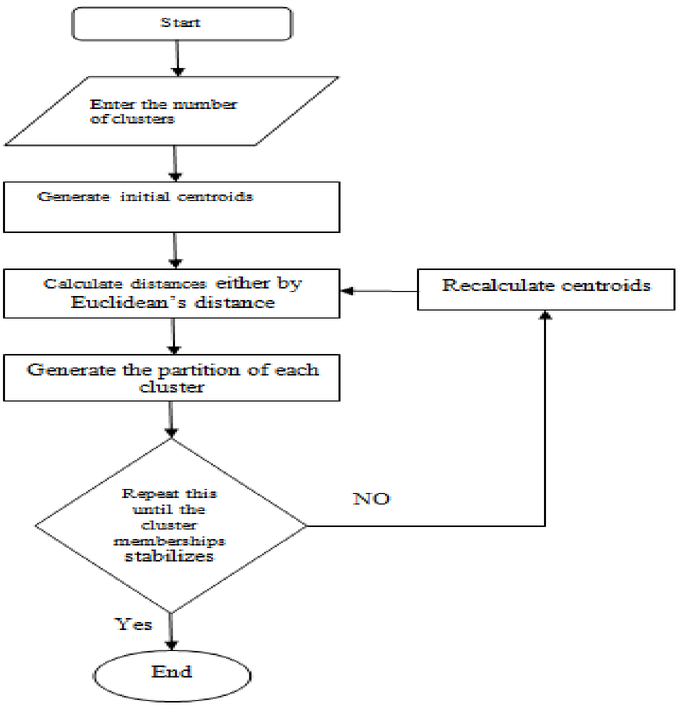


Figure 1. Flowchart of Algorithm K-means

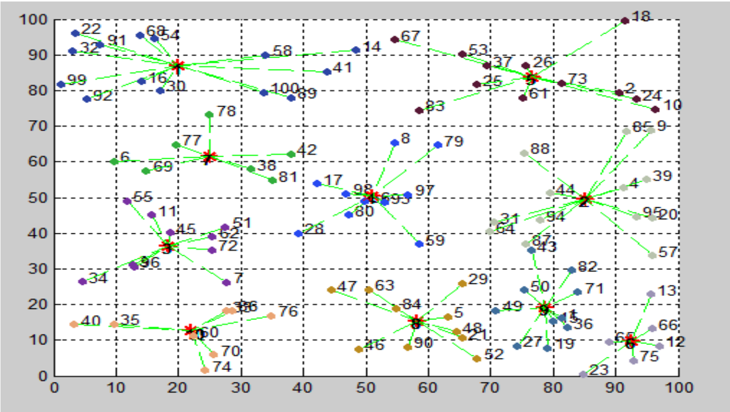


Figure 2. Formation of Clusters by K-mean

Step 1: Start Phase

Division clusters into k non empty subsets and compute seed points as the centroids $C_j(i)$ j=1 to k are the index of centroids and i=1 to mc where i denotes the nodes mc the number of the nearest nodes. Each of the remaining nodes decides its CH nearest to it according to the Euclidean distance.

Step 2: Re-clustering

After that the centroid of each cluster is calculated. With the new CH in each cluster, this operation is recursively executed until the CH is not changed any more.

Step 3: election of cluster head

After the clusters are created, an ID number is assigned to each node of a cluster according to the distance from the centroid, assigning smaller number to the closer one. Figure 2 shows the repartition of the sensor nodes with the ID number. The ID number of a node indicates the order to be chosen as the CH. Therefore, the ID number plays an important role in the selection of a node as CH.

The connectivity of the network is retained by checking the residual energy of the CH every round. If this energy is smaller than the threshold, the node in the next order is selected as a new CH. The newly elected CH informs other nodes of the change of the CH.

5. Heterogeneous Network Model

In our model, we assume that there are N sensor nodes, which are evenly scattered within a M×M square region and organized into clusters hierarchy for aggregate data by cluster heads to base station. That is located at the center of this region. Nodes have low mobility or stationary as being assumed at [15, 16]. In the two-level heterogeneous networks advanced nodes fraction m with a time more energy than the others which have an initial energy E0. The total energy is assumed as follow:

$$\begin{aligned} E_{total} &= N(1-m)E_0 + NmE_0(1+a) \\ &= NE_0(1+am) \end{aligned} \quad (1)$$

In multi-level heterogeneous networks, the clustering algorithm should consider the discrepancy of initial energy, Etotal is expressed by:

$$E_{total} = \sum_{i=1}^{N} E_0(1+a_i) = E_0\left(N + \sum_{i=1}^{N} a_i\right) \quad (2)$$

ni denotes the number of rounds to be a cluster-head for the node si, and we refer to it as the rotating epoch. In DEEC protocol, we choose different $n_i = \frac{1}{p_{opt}}$ which is based on the residual energy of Ei(r) node si at round r. If nodes have different amounts of energy, pi of the nodes with more energy should be larger than popt. Let Ē(r) denotes the average energy at round r of the network, which can be obtained by:

$$\bar{E}(r) = \frac{1}{N}\sum_{i=1}^{N} E_i(r) \quad (3)$$

To calculate Ē(r) we have:

$$p_i = p_{opt}\frac{E_i(r)}{\bar{E}(r)} \quad (4)$$

Where G is the set of nodes that are eligible to be cluster-heads at round r, ni is chosen based on the residual energy Ei(r) at round r of node si as follow:

$$n_i = \frac{1}{p_i} = n_{opt}\frac{\bar{E}(r)}{E_i(r)} \quad (5)$$

When the networks are heterogeneous, the reference value of each node should be different according to the initial energy. In the model of multi-level heterogeneous networks, the weighted probability shown as:

$$p(s_i) = \frac{p_{opt} N (1 + a_i) E_i(r)}{\left(N + \sum_{i=1}^{N} a_i\right) \overline{E}(r)} \quad if \ s_i \in G \tag{6}$$

Thus we can estimate the average energy Ē(r) of rth round as follow:

$$\overline{E}(r) = \frac{1}{N} E_{total} \left(1 - \frac{r}{R}\right) \tag{7}$$

Where R denotes the total rounds of the network lifetime. Let Eround denotes the energy consumed by the network in each round. R can be approximated as follow:

$$\overline{E}(r) = \frac{E_{total}}{E_{round}} \tag{8}$$

The total energy dissipated in the network during a round is equal to:

$$E_{round} = L(2NE_{elec} + NE_{DA} + k\varepsilon_{mp} d_{toBS}^4 + N\varepsilon_{fs} d_{toCH}^2) \tag{9}$$

$$d_{toBS} = \frac{M}{\sqrt{2\pi k}}, \quad d_{toCH} = 0.765 \frac{M}{2} \tag{10}$$

Where k is the number of clusters, EDA is the data aggregation cost expended in the cluster-heads evolving. The lost energy is proportional to the number of cluster heads noted by s in this area. Thus the probability threshold (5), which each node si uses to determine a cluster-head in each round, becomes in our proposed approach as follow:

$$T(s_i) = \begin{cases} \dfrac{p_i}{1 - p_i \left(r \bmod \left(\dfrac{1}{p_i} - \dfrac{\alpha}{N \times p_i} \right) \right)} & if \ s_i \in G \\ 0 & Otherwise \end{cases} \tag{11}$$

Where α is the number of nodes that are excluded from the cluster head threshold selection due to their location and distribution density reason, with an initial value of 0. When s increases, T(si) increases as well, therefore the chances of nodes, that are eligible to be cluster heads, decreases. Indeed, with this algorithm we can save the lost energy caused by the election of these cluster heads excluded and extend the lifetime of the network. The analytical method to calculate the number s is a perspective of this work.

6. Simulation Results

The proposed approach has been implemented in MATLAB and the performance has been evaluated by simulation, the lifetime of the network is measured in terms of rounds when the first sensor node dies. The base station is assumed in the center of the sensing region. All the parameters values including the first order radio model characteristic are mentioned in the Table 1 below. To compare the performance of the proposed approach with our first contribution, the effect caused by signal collision and interference in the wireless channel is ignored, a multi-level heterogeneous network is considered. In this simulation, the value of multi-level heterogeneity is fixed in amax=3. Once a node runs out of energy, it is considered as dead and cannot transmit or receive data. The radio parameters used in our simulations are shown in Table 1.

Table 1. Parameters used in Simulations

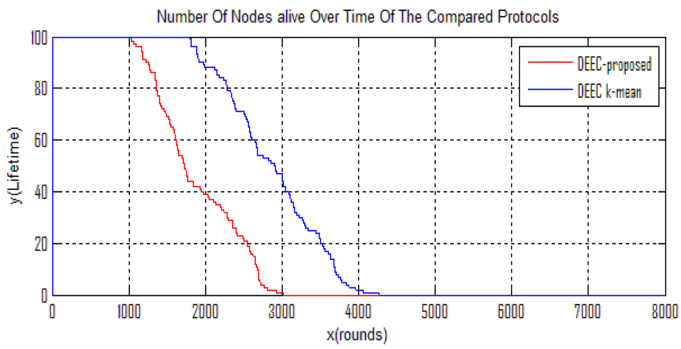
Parameter	Value
Network area	100 m×100 m
Number of nodes	100
E_0	0.5 J
E_{elec}	50 nJ/bit
ε_{fs}	10 pJ/bits/m^2
ε_{mp}	0.0013 pJ/bit/m^4
$E_{Tx}=E_{Rx}$	50 nJ/bit
E_{DA}	50 nJ/bit/message
$d_0=\sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}}$	70 m
Packet Size	4000bits
P_{opt}	0.05

When the new proposed protocol is used, we note that the network is subdivided into ten clusters that are correctly distributed over the sensing area. There is no intersection between the different clusters.

In the simulation results Figure 3 and 4 gives the curves of the number of nodes alive and nodes dead over time for the two compared protocols, whereas Figure 5 shows the number of packet messages received in the base station per round for each scenario. The Tables 2 provides statistics on the number of dead nodes per rounds as well as the percentage increases in the lifetime of the network for the proposed approach which is compared to the first contribution.

This figure shows an improvement of energy efficiency in terms of enlarging the round number of the first node dies. Furthermore, our proposed protocol presents a most significant improvement of the two compared protocols. In fact, it is clear that the stable time of the new method is extended for the whole network compared to the first proposition.

Here, it is shown that the first node death occurs at the 1743 round by using our approach whereas this value is about 937 for other algorithm used. Thus in first protocol all nodes die early at the 3388 round on the contrary of the proposed approach in which all nodes die tardily at the 4755 for all studied scenarios. Moreover we can see that the unstable region of our algorithm is larger than the ones of the other protocol.



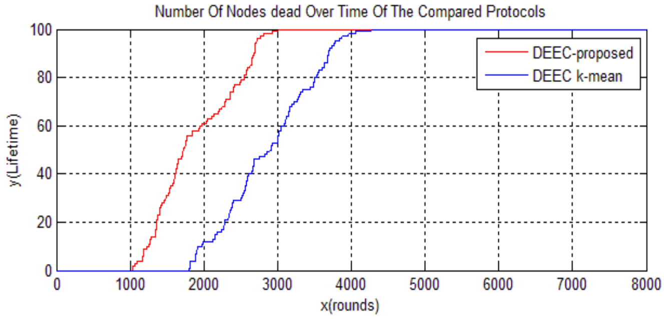


Figure 4. Number of Dead Nodes Over Time

In other hands, referred to Table 2 it is very clear that the algorithm DEEC-Kmean gives a significant performance improvement in terms of energy and lifetime gains, compared to our first contribution. This energy efficiency reach for the first dead 86,4% and 28,74% for the all dead.

Table 2. Number of Dead Nodes per Rounds

	DEEC-k-mean	First proposition	Increase
First dead	1743	937	86,34 %
All-dead	4755	3388	28,74 %

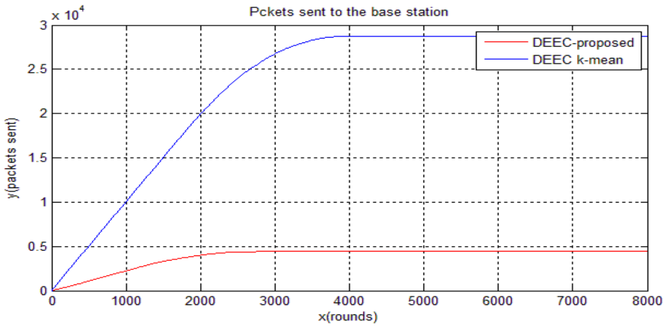


Figure 5. Number of Packet Transmit to BS

If we consider number of packets sent to BS then the DEEC-Kmean is clear winner. Later has sent highest number of packets to BS. This may be due to uneven and large number of cluster formation. High numbers of clusters are not preferred but a fair amount of data is desired in which balanced k-means seems to perform well.

7. Conclusion

In this paper we have showed the energy limitation of the first protocol when clusters heads are selected closest to others an important part of energy is dissipated which reducing the performance of the network. In this paper, we put forward an improved routing algorithm Balanced based K-Means for heterogeneous networks. This centralized algorithm can avoid the complexity of charge of treatments generated by all network nodes. The simulation results show that Balanced k-means by making equi-distribution of space in the nodes and take into account the residual energy of each nodes at the election step. Finally, this technique to improve the network performances by saving more energy and extending more efficiently the network lifetime.

As a perspective, we will study other metrics with k-mean technique which may be more efficient in terms of quality of service.

References

[1] Marcos Diogenes. Survey on Wireless Sensor Network Devices. *IEEE.* 2003.
[2] V Raghunathan, C Schurgers, Park S, MB Srivastava. Energy aware wireless micro-sensor networks. *IEEE Signal Processing Magazine.* 2002; 19(2): 40-50.
[3] Rui Tan, Al. *Quality-driven volcanic earthquake detection using wireless sensor networks*. RTSS. San Diego, CA, USA. 2010: 271-280.
[4] Chakkor Saad, Baghouri Mostafa, Hajraoui Abderrahmane. Fuzzy logic approach to improving Stable Election Protocol for clustered heterogeneous wireless sensor networks. *Journal of Theoretical and Applied Information Technology*. 2013; 53(3).
[5] S Hussain, E Shakshuki, AW Matin. Agent-based system architecture for wireless sensor networks. *Advanced Information Networking and Applications*. 2006; 2: 18-20.
[6] Elahmadi cheikh, Chakkor Saad, Baghouri Mostafa, Hajraoui Abderrahmane. Energy Efficient Enhancement of TDEEC Wireless Sensors Network Protocol Based on Passive RFID Implementation. *International Journal of Advanced Research in Computer and Communication Engineering*. 2014; 3(5).
[7] Elahmadi cheikh, Chakkor Saad, Baghouri Mostafa, Hajraoui Abderrahmane. Desert Nomads Healthcare Monitoring Application Based on BSN and Wi-Max. *International Journal of Advanced Research in Computer and Communication Engineering*. 2014; 3(12).
[8] Lin SHEN, Liqin WANG. *A Clustering Algorithm Based Geographic Location Information for Wireless Sensor Networks*. *Sensors & Transducers*. 2013; 157(10): 147-152.
[9] Maneesha V Ramesh. *Real-time wireless sensor network for landslide detection*. In Sensorcomm. Athens, Greece. 2009: 405-409.
[10] V Raghunathan, C Schurgers, Park S, MB Srivastava. Energy aware wireless micro-sensor networks. *IEEE Signal Processing Magazine.* 2002; 19(2): 40-50.
[11] I Akyildiz Al. A Survey on Sensor Networks. *IEEE Communication Magazine.* 2002; 40(8): 102-114.
[12] D Bhattacharyya, T Kim, S Pal. A Comparative Study of Wireless sensor networks and their routing protocols. *Sensors*. 2010; 10: 10506-10523.
[13] JM Shea, JP Macker. Automatic Selection of Number of Clusters in Networks using Relative Eigenvalue Quality. *IMCC*. *IEEE.* 2013; 1: 131-136.
[14] BEN ALL S, Abdellah E, Abderrahim BH, Moulay LH. Hierarchical Adaptive Balanced energy efficient Routing Protocol (HABRP) for heterogeneous wireless sensor networks. *IEEE*. 2010.
[15] D Pelleg, A Moore. *X-means: Extending Kmeans with efficient estimation of the number of clusters*. In Proceedings of the 17th International Conf. on Machine Learning. Morgan Kaufmann, San Francisco, CA. 2000; 727-734.
[16] E Cheikh, C Saad, B Mostafa, H Abderrahmane. New approach to improving lifetime in heterogeneous wireless sensor networks based on clustering energy efficiency algorithm. *Journal of Theoretical and Applied Information Technology*. 2014; 62(2).
[17] Intanagonwiwat C, Govindan R, Estrin D. *Directed diffusion: Ascalable and robust communication paradigm for sensor networks*. Proceedings of the 6th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom00). Boston, MA, USA. 2000: 56-67.
[18] Heinzelman WR, Chandrakasan A, Balakrishnan H. *Energy-efficient communication protocol for wireless microsensor networks*. Proceedings of the 33rd Hawaii International Conference on System Scienc.
[19] Li Qing, Qingxin Zhu, Mingwen Wang. DEEC: Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks. *Computer Communications.* 2006; 29: 2230-2237.
[20] S Lindsey, CS Raghavenda. *PEGASIS: power efficient gathering in sensor information systems.* in: Proceeding of the IEEE Aerospace Conference. Big Sky, Montana. 2002.
[21] O Younis, S Fahmy. HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Transactions on Mobile Computing*. 2004; 3(4): 660-669.
[22] Elbhiri Brahim, Saadane Rachid, Driss Aboutajdine. Stochastic Distributed Energy-Efficient Clustering (SDEEC) for heterogeneous wireless sensor networks. *ICGST-CNIR Journal*. 2009; 9(2): 11-17.
[23] Elbhiri Brahim, Saadane Rachid, Sanaa El fkihi, Driss Aboutajdine. *Developed Distributed Energy-Efficient Clustering (DDEEC) for heterogeneous wireless sensor networks*. IEEE Communications and Mobile Network (ISVC), 5th International Symposium on. 2010.
[24] Parul Saini, Ajay K Sharma. Energy Efficient Scheme for Clustering Protocol Prolonging the Lifetime of Heterogeneous Wireless Sensor Networks. *International Journal of Computer Applications*. 2010; 6(2).
[25] Suniti Dutt, OS Khanna. An Enhanced Energy Efficient Clustering Scheme for Prolonging the Lifetime of Heterogeneous Wireless Sensor Networks. *International Journal of Computer Applications*. 2013; 76(8).