K-means Based Energy Aware Clustering Algorithm in Wireless Sensor Network

Anand Gachhadar, Om Nath Acharya

Abstract — In this article, an energy efficient novel clustering scheme is designed in order to provide low energy consumption, reducing overload on sensor nodes and increase network lifetime of wireless sensor network. The cluster based technique is one of the major approaches in reducing energy consumption in wireless sensor networks. The main idea of this article is to reduce data transmission distance of sensor nodes in wireless sensor networks by using clustering concepts. However, clustering concepts faces several challenges in selection of cluster head, rotating the role of cluster head in cluster and optimal data routing in the network. So, we propose a protocol which provides energy efficient clustering and optimal data routing to increase the network longevity.

Index Terms — Wireless sensor network, k- means algorithm, LEACH, LEACH-C, Cluster head, clustering, base station.

— — — — — — — — — ◆ — — — — — — — — —

1. INTRODUCTION

THE Wireless sensor networks are mostly deployed in remote and promiscuous locations where physical monitoring is almost impossible. With the recent technological advances in wireless communications, processor, memory, radio, low power, highly integrated digital electronics and Micro Electromechanical Systems (MEMS) it becomes possible to significantly develop tiny and small size, low power, and low cost multi-functional sensor nodes. A wireless sensor network consists of thousands of these micro sensors which are largely deployed in a hostile environment which has the capability of sensing, computations and wireless communication. Sensor nodes collect data and are then transmitted to the Base Station (BS). Wireless sensor nodes are most popularly deployed in applications like battle field, farming, agriculture, environment monitoring system, vehicle monitoring systems etc. Figure 1. depicts the infrastructure of wireless sensor network. Sensor nodes can be fixed or mobile whereas the Base Station remains fixed. Energy is one of the major concerns in wireless sensor networks. Since, the environment in which the sensor nodes are located is hostile it is not possible to manually replace the battery. So we have to find ways to conserve energy. A wireless sensor node consumes most of it energy in transmitting and receiving packets from neighboring nodes. The only source of energy for sensor nodes is battery which is finite and draining of its battery may make its sensing area uncovered. So, in order to increase the network life time and efficiency of the network we need to design an energy efficient wireless sensor network which is possible through clustering schemes [1-3]. Energy Efficient Clustering Virtual Area Partition [12] is energy efficient clustering which is based on vir

tual area partition in heterogeneous power of each node may be different. VAP-E can balance the load between clusters, enhance the energy efficiency of sensor nodes, prolong the life time of networks, and improve the efficiency of communications. K-Hop Overlapping Clustering Algorithm (KOCA) [13] is based on K-hop overlapping which is used to overcome the problem of overlapping multi-hop clustering for Wireless Sensor Network (WSN). Goal of KOCA is generating connected overlapping clusters that cover the entire sensor network with a specific average overlapping degree. Energy Efficient Unequal Clustering (EEUC) [14] is distance based scheme and it requires global identification such as location and distances to the base station. Low Energy Adaptive Clustering Hierarchical (LEACH) Protocol [2, 3, 16] is randomized, self configuring and adaptive cluster formation, local control for data transfers and low energy media access control and application specific data processing. In LEACH, the large number of sensor nodes will be divided into several clusters. For each cluster, a sensor node is selected as a cluster head. The selection of cluster head nodes is based on a pre-determined probability. Other noncluster heads choose the nearest cluster to join by receiving1 the strength of the advertisement message from the cluster head nodes. A noncluster head node can only monitor the environment and send data to its cluster head node. The cluster head node is responsible for collecting the information of noncluster head nodes in the cluster. Then, it processes data and sends data to the BS. As a non-cluster head node cannot send data directly to the BS, the data transmission distance of the sensor node is shrunk. Therefore, the energy consumption is reduced in the wireless sensor networks. However, the random selection of the cluster head node may obtain a poor clustering setup, and cluster head nodes may be redundant for some rounds of operation. The distribution of cluster head nodes is not uniform, thus some sensor nodes have to transfer data through a longer distance and the reasonable energy saving is not obtained in wireless sensor networks. Low Energy Adaptive Clustering Hierarchical-Centralized (LEACH-C) is proposed as an improvement of LEACH which uses a centralized clustering algorithm to create the clusters [3]. In LEACH-

- *Anand Gachhadar, Asst. Professor in Electrical and Electronics Engineering, Kathmandu University, Dhulikhel, Nepal, Ph- +977-9841746512. E-mail:anand.gachhadar@ku.edu.np*
- *Om Nath Acharya, Lecturer in Electrical and Electronics Engineering, Kathmandu University, Dhulikhel, Nepal, Ph- +977-9851170208. E-mail:acharya.om@ku.edu.np*

C, the BS collects the information of the position and energy level from all sensor nodes in the networks. Based on this information, the BS calculates the number of cluster head nodes and configures the network into clusters. In [15], the authors propose the Hybrid Energy-Efficient Distributed (HEED) clustering protocol to prolong the network lifetime and support scalable data aggregation. In this protocol, the cluster heads are probabilistically selected based on their residual energy and the sensor nodes join the clusters according to their energy level. However, the critical problem is that the cluster size is also not uniform in these schemes.

In this paper we propose a K-means Based Energy Aware Clustering (KEAC) Algorithm for single-hop wireless sensor network to provide less energy consumption throughout the overall network. Cluster Heads are elected based on remaining energy and distance from their k-means and cluster heads are uniformly distributed within the network based on their weight in a single-hop environment. The life time of wireless sensor network is extended by using the uniform cluster location and balancing the network loading among the clusters. The major bene_t of this proposed scheme is the energy consumption and better network life-time. In this approach we assume that the radio range of each sensor node can be adjusted in function of distance to the receiver [6]. This way, the transmitter sensor node can reach the sink node directly, without necessarily forwarding its packets by multi-hop transmission. The use of clusters for transmitting data to the base station leverages the advantages of small transmit distances for most nodes, requiring only a few nodes to transmit far distances to the base station.

2. SYSTEM MODEL

The system infrastructure is composed of a BS and number of sensor nodes within a sensing area. We divide nodes into non cluster heads, and cluster head nodes. The non cluster head operate in a sensing mode to monitor the environment information and transmit data to the cluster head. Cluster heads are responsible for transmitting it to the BS. This is a centralized system election of cluster head takes place in base station and is elected cluster heads are announced to all the other nodes. This process reduces the control overhead in sensor nodes and hence conserves energy. This approach is not only inefficient in terms of energy consumption but also unrealistic. The sink or base station is usually located far away from the sensing area and not directly reachable to the nodes due to various propagation problems. So, most realistic approach is multi-hop path among the clusters head between the sink and the sensing nodes which provides more energy efficient.

Basic elements of Wireless Sensor Network:-

Sensor Node: A sensor node is the core component of a WSN. Sensor nodes can take on multiple roles in a network, such as simple sensing, data storage, routing, and data processing.

Clusters: Clusters are the organizational unit for WSNs. The dense nature of these networks requires the need for them to be broken down into 3 clusters to simplify tasks such a communication.

Cluster heads: Cluster heads are the organization leader of a cluster. They often are required to organize activities in the cluster. These tasks include but are not limited to data-aggregation and organizing the communication schedule of a cluster.
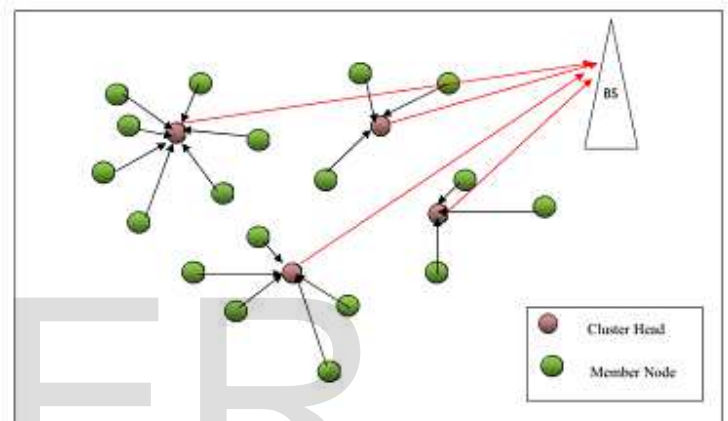

Figure 1: Single hop sink.

Base Station: The base station is at the upper level of the hierarchical WSN. It provides the communication link between the sensor network and the end-user.

End User: The data in a sensor network can be used for a wide-range of applications. Therefore, a particular application may make use of the network data over the internet, using a Personal Digital Assistant (PDA), or even a desktop computer. In a queried sensor network (where the required data is gathered from a query sent through the network). This query is generated by the end user.

In wireless sensor networks, data communications consume a large amount of energy. The total energy consumption consists of the average energy dissipated by data transmission of the non-cluster head nodes and cluster head nodes. In addition, the energy consumption for data collection and aggregation of cluster head nodes is considered. Figure 2 Illustrates the radio energy dissipation model in wireless networks [2]. In this model, to exchange a B-bit message between the two sensor nodes, the energy consumption can be calculated by Figure 2: Radio Energy Dissipation Model.
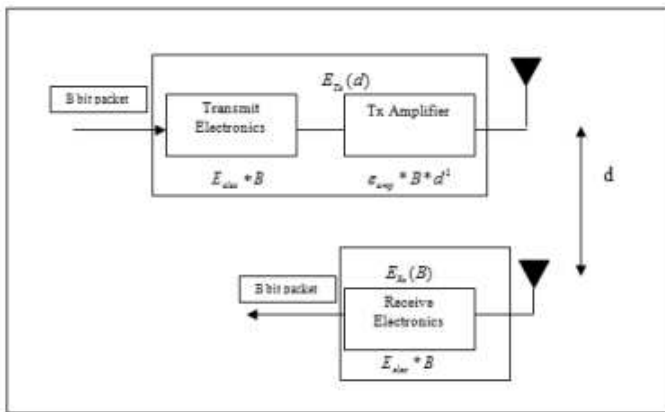
Figure 2: Radio Energy Dissipation Model.

$$E_{Tx}(d,x) = E_{elec} * B + \varepsilon_{amp} * B \qquad (1)$$

$$E_{Rx}(B) = E_{elec} * B \qquad (2)$$

Where d is the distance between the sensor nodes, $E_{Tx}(d,x)$ is the transmitter energy consumption and $E_{Rx}(B)$ is the receiver energy consumption. E_{elec} is the electronics energy consumption per bit in the transmitter and receiver sensor nodes. ε_{amp} is the amplifier energy consumption in transmitter sensor nodes, which can be calculated by

$$\varepsilon_{amp} = \varepsilon_{fs} * d^2 \text{, for } d \leq d_0 \qquad (3)$$

$$= \varepsilon_{mp} * d^4 \text{, for } d \geq d_0 \qquad (4)$$

Where, d_0 is the threshold value. If the distance d is less than x0, the free space propagation model is used. Otherwise, the multi-path fading channel model is used. ε_{fs} and ε_{mp} are communication energy parameters.

TABLE 1
Radio model parameters

Operation	Value
Transmit Electronics ($E_{Tx\ elec}$) Receiver Electronics ($E_{Rx\ elec}$) $E_{Tx,elec} = E_{Rx,elec} = E_{elec}$	50nJ/ bit
Transmit amplifier ε_{amp}	100pJ/bit/m²

Consider a system with N number of total wireless sensor nodes. All the system parameters are clearly defined in Table 1 and Table 2. Let, d_{CH} be the distance from node to cluster head and d_{BS} be the distance from Cluster Head (CH) to BS.

Energy consumed by a node for transmission of B bits to its Cluster Head is given by

$$E_{node} = E_{init} - E_{Tx}(B, d_{CH}) \qquad (5)$$

Energy consumed by a cluster head for transmitting it to the BS.

$$E_{CH} = E_{init} - E_{Rx}(B) - E_{da} - E_{Tx}(B, d_{BS}) \qquad (6)$$

Where, E_{node} is energy consumed by a node and E_{CH} is the energy consumed by a cluster head, E_{init} is the initial energy of each node.

3. PROPOSED ALGORITHM

In order to reduce energy consumption and extend lifetime of the sensor nodes in wireless sensor networks, efficient energy saving algorithm must be developed and designed. We propose a centralized clustering architecture, KEAC algorithm to provide energy efficient and prolong network lifetime in wireless sensor networks. In this proposed scheme, we assume that the BS receives the information of location and residual energy from each sensor nodes with which average residual energy can be calculated. Node with higher residual energy becomes a candidate for election as a cluster head. We modify k-means algorithm [9] to make an ideal distribution for cluster by taking into account the information of location and residual energy for all sensor nodes. The proposed algorithm consists of three phases.

1. **Initialization Phase**: In this phase BS broadcasts Initialization Request (IREQ) message to all sensor nodes within a sensing area. Upon receiving IREQ each nodes replies with Initialization Reply (IREP) message to the base station. The IREP message consists of remaining energy and its current location.

2. **Setup Phase:** The main purpose of this phase is to find cluster heads. Information regarding location and residual energy is received in Initialization phase. Based upon this information a suitable initial means of points for clusters can be obtained. Base station now calculates the center location for all sensor nodes. If there are n sensor nodes in wireless sensor networks, center location M (m_x, m_y) can be calculated by

$$M = \frac{\sum_{i=1}^{n} X_i}{n} \qquad (7)$$

Where, X_i is the co-ordinate of sensor node i.

Let, D be the average distance between M and all other sensor nodes, which can be calculated by

$$D = \frac{\sum_{i=1}^{n} |X_i - M|}{n} \qquad (8)$$

According to M and D, the locations of initial means of point $M_i(M_{ix}, M_{iy})$ for cluster i can be calculated by

$$M_{ix} = D * \cos(2\pi(i-1)) + M_x \qquad (9)$$

$$M_{iy} = D * \sin(2\pi(i-1)) + M_y \tag{10}$$

Where k is the number of clusters and i=1, 2, 3...., k. and its value is decided at the beginning of set up phase which is given by

$$k = \left[\frac{\sqrt{n}}{\sqrt{2\pi}} \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}} \frac{M}{d^2_{BS}} \right] \tag{11}$$

where, M is the side of the given square field. The d_{BS} is the average distance from the cluster head nodes to the BS which is defined in LEACH-C.

- **Cluster Head Selection Phase**: Based on k-means algorithm n number of nodes is partitioned in k clusters and each node belongs to the cluster with nearest mean of point. If there are k clusters in the system, the k-means function can be expressed as:

$$avgs \min \sum_{i=1}^{n} \sum_{X_j \varepsilon S_i} \left| X_j - M_i \right|^2 \tag{12}$$

Where, Si is the cluster i, Xj is coordinate of sensor node j and Mi is the coordinate of mean of point. In order to create uniform distributed clusters, the minimal distance between the means of points and all sensor nodes is calculated. Then, these sensor nodes are classi_ed into the cluster according to the minimal distance. If the Xj is the closest to the Mi the sensor node j will join the cluster i. After all nodes join their clusters based upon k -mean point then cluster head is calculated using

$$W_i = c1 * E_i + c2 * D_i \tag{13}$$

where, c1 and c2 are constant values. E_i is the residual energy, D_i is the distance between sensor nodes from the mean point. The node with minimum weight W_i within a cluster S_i is now elected as a cluster head and the remaining nodes of cluster S_i acts as the members of cluster head. The cluster head is a sensor node which is closer to the mean point and has higher residual higher than other nodes in the same cluster. The BS now broadcast this information to all sensor nodes and similarly updates their routing table.

4. RESULTS AND DISCUSSION

In this section, we evaluate the performance of our proposed clustering scheme with LEACH and LEACH-C clustering schemes. The simulation is done in MATLAB. The assumptions for our simulation study are as follows.

- The system model consists of sensor nodes and a base station.

- The BS is fixed and is located far away from the sensing nodes.
- The location of each sensing nodes are randomly distributed within a sensing area.
- The cluster head gather all the information and then forward it towards the Base station.
- The sensor nodes can be static as well as mobile
- Sensor nodes are assumed to be homogeneous.

TABLE 2
Parameters used in simulation model

PARAMETERS	VALUE
Electronics energy (E_{elec})	50 nJ/bit
Energy for data aggregation(E_{da})	5nJ/bit/signal
Initial energy of node (E)	2J
Packet size (B)	2000bits
Number of nodes (N)	100
Position of BS (X,Y)	(100,175)
Sensing area (MxM)	200 x200
ε_{fs}	10pJ/bit/m²
ε_{mp}	0.0013pJ/bit/m⁴

All the parameters on which the simulation is based are listed in Table 2 [2]. Performance Analysis is done on the basis of Energy Consumption and the node which die out first. Analysis is performed in each round by calculating the total network energy consumed during each round and is plotted in Figure 3. Figure 3 shows the decrease in total network energy is each round. Energy Consumption is measured for LEACH, LEACH-C and KEAC and it was found that energy consumption in KEAC is highly decreased.

Figure 3 shows the total network energy when the number of nodes is equal to 100 and the sensing area is taken to 200m x 200m. It is evident that energy conserved by our proposed algorithm is higher than the rest of the algorithms.
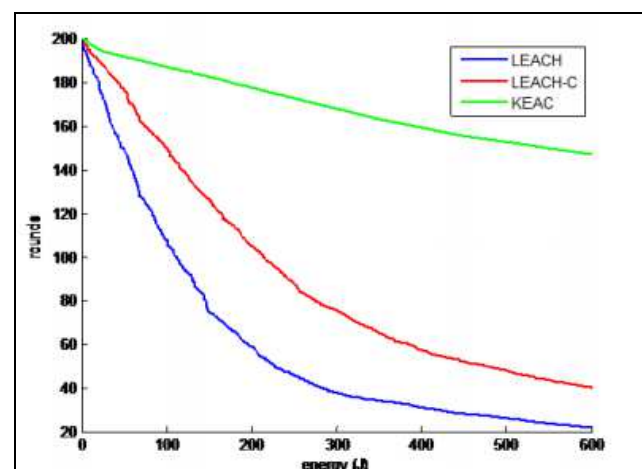

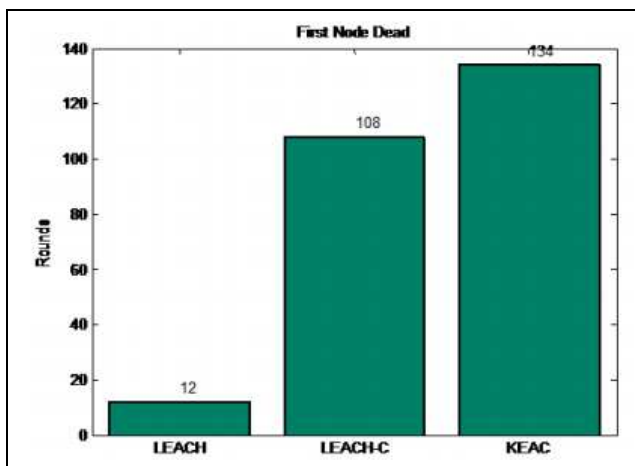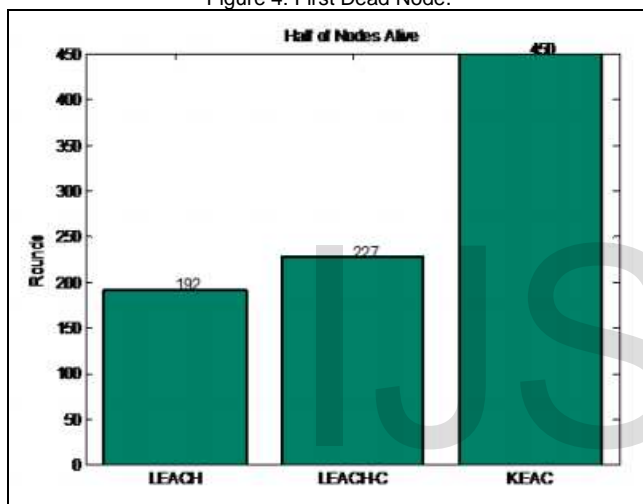Figure 3: Network Energy Consumption.

Figure 4: First Dead Node.
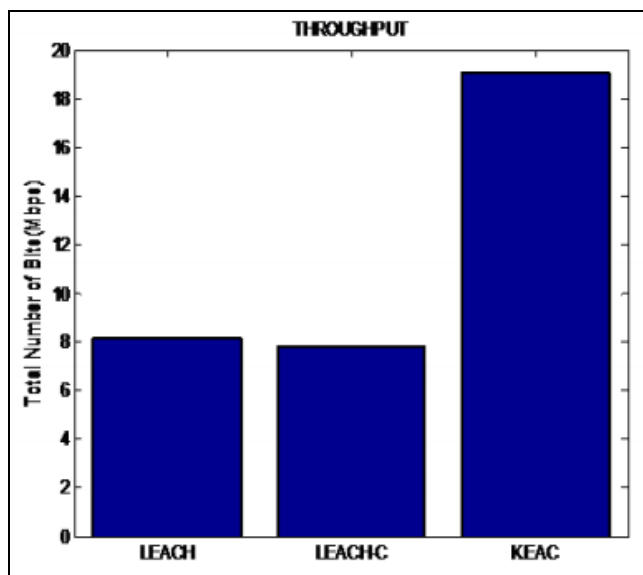

Figure 5: Half of Nodes Alive.

Figure 6: Throughput of Network.

Figure 4, 5 depicts when the first nodes dies and half of the sensor nodes alive for three method. The energy consumption in a single hop network increases due to larger distances from BS to sensor nodes and large number of cluster heads. So, in LEACH and LEACH-C clustering algorithm the election of cluster head is not uniform and thereby can be large number of cluster heads which may drain the total network energy. Simulation parameters are assumed 10 to be same in both the cases and simulation result shows KEAC is better than rest of the two schemes.

The reason for this behavior is that the data communications consume a large amount of energy in wireless sensor networks. However, the transmission distance between non-cluster head node and cluster head node is suitable by using our scheme. The transmission power of non-cluster head nodes is reduced. Figure 6 shows the throughput of the entire network. Throughput of network has been measured for simulation time and total throughput after each round has been calculated. Finally, the total number of packets received by the BS is then simulated for each clustering schemes. Figure 6 shows that the throughput of KEAC is higher than LEACH, LEACH-C. The reason for high throughput is more number of sensor nodes die out in LEACH and LEACH-C due to higher energy consumption in data transmission from node to the BS. So, the total throughput reduces as compared to that of KEAC. .

5. CONCLUSION

Energy saving and prolonging network life time is a major issue in wireless sensor network. We have taken a small step in providing energy efficient network by calculating the average distance between network nodes and taking into account the residual energy. Finally, weight for each node is calculated based on average distance and residual energy and node with minimum weight acts as a cluster head. This approach has highly increased network lifetime by uniform cluster distribution and balancing the network loading among the clusters. The results show that our proposed scheme achieves low energy consumption and prolongs network life time.

REFERENCES

[1] A. Chamam, S. Pierre, "On the planning of wireless sensor networks: energy efficient clustering under the joint routing and coverage constraint".IEEE Trans Mob Comput.8(8), 10781086 (2009).

[2] W.B. Heinzelman, P. Chandrakasan, H. Balakrishnan, "Energy-e_cient communication protocol for wireless microsensor networks".inProc 33rd Hawaii International Conference on System Sciences, pp. 110 (January2000).

[3] W.B. Heinzelman, P. Chandrakasan, H. Balakrishnan,"An application-specific protocol architecture for wireless microsensor networks". IEEE Trans Wirel Commun. 1(4), 660670 (2002). doi:10.1109/TWC.2002.804190.

[4] C. JauYang, J. Pei-Hao," An e_cient cluster based power saving schemes for wireless Sensor Networks".EURASIP journal on wireless communica-tion and Networking 2012, 2012:172.

[5]   M. Mou_da ,Z. Houda and L. Francis ," Cluster based Routing proto-cols for Energy E_ciency in Wireless Sensor Networks". CRAN Laboratory, Nancy University, CNRS France.

[6]   A.F. Santos, M.O. Marques, I.S. Bonatti, Power Aware model for wire-less Sensor Nodes. July 11, 2008.

[7]   A. Ezzati, B. SAID, B.H. Abderrahim, L. Moulay, Advanced Low En-ergy Adaptive Heirarchy.International Journal on Computer Science and Engineering, vol. 02, No. 07, 2010, 2491-2497.

[8]   A.K.F. Mortaza, A.J.J Mohammed, Modify LEACH Algorithm for Wireless Sensor Network.International Journal of Computer Science Is-sues (IJCSI), vol8. Issue 5, No 1, September 2011, ISSN: 1694-0814

[9]   J. Zhu, H. Wang,"An improved K-means clustering algorithm". inProc 2nd IEEE International Conference on Information Management and Engi-neering (ICIME), pp. 190192 (2010) doi:10.1186/1687-1499-2012-172.

[10]   M. Tubaishat, S. Madria,"Sensor networks: An Overview".IEEE Po-tentials.22(2), 2023 (2003).

[11]   J.N. Al-Karaki, A.E. Kamal, "Routing techniques in wireless sensor networks: A Survey". IEEE Wirel Commun.11(6), 628 (2004). doi:10.1109/ MWC.2004.1368893.

[12]   R.Wang, L. Guozhi, and C. Zheng, A clustering algorithm based on virtual area partition for heterogeneous wireless sensor networks. In In-ternational Conference on Mechatronics and Automation, 372376.

[13]   M. Youssef, A. Youssef, and M. Younis, Overlapping multi-hop clus-tering for wireless sensor networks. IEEE transactions on parallel and distributed systems, 20, 12, 18441856.

[14]   C. Li, M. Ye, G. Chen and J. Wu, An energy-efficient unequal cluster-ing mechanism for wireless sensor networks. In IEEE International Con-ference on Mobile Ad-hoc and Sensor Systems, 604611.