# Data Aggregation in Wireless Sensor Networks Using Modified Voronoi Fuzzy Clustering Algorithm

**Nadia Adnan Shiltagh**
Assistant professor Dr.
Computer Engineering/ Baghdad University
E-mail:dr-nadiaat@hotmail.com.

**Maab Alaa Hussein**
Computer Engineering/Baghdad University
E-mail:maab_alaa@yahoo.com

## ABSTRACT

Data centric techniques, like data aggregation via modified algorithm based on fuzzy clustering algorithm with voronoi diagram which is called **modified Voronoi Fuzzy Clustering Algorithm (VFCA)** is presented in this paper. In the modified algorithm, the sensed area divided into number of voronoi cells by applying voronoi diagram, these cells are clustered by a fuzzy C-means method (FCM) to reduce the transmission distance. Then an appropriate cluster head (CH) for each cluster is elected. Three parameters are used for this election process, the energy, distance between CH and its neighbor sensors and packet loss values. Furthermore, data aggregation is employed in each CH to reduce the amount of data transmission which lead to extend the network lifetime and reduce the traffic that may be accrue in the buffer of sink node. Each cluster head collected data from its members and forwards it to the sink node. A comparative study between modified VFCA and LEACH protocol is implemented in this paper and shows that the modified VFCA is more efficient than LEACH protocol in terms of network lifetime and average energy consumption. Another comparative study between modified VFCA and K-Means clustering algorithm is presented and shows that the modified VFCA is more efficient than K-Means clustering algorithm in terms of packets transmitted to sink node, buffer utilization, packet loss values and running time. A simulation process is developed and tested using Matlab R2010a program in a computer having the following properties: windows 7 (32-bit operating system), core i7, RAM 4GB, hard 1TB.

**Keywords:-**voronoi fuzzy, fuzzy c-means ,clustering algorithm, data aggregation.

## تجميع البيانات في شبكات المتحسسات اللاسلكية باستخدام خوارزميه التجميع المعدلة VFCA

**أ.م.د. نادية عدنان شلتاغ**
هندسة الحاسبات/جامعة بغداد

**مآب علاء حسين**
هندسة الحاسبات/جامعة بغداد

## الخلاصة

تقنيات مركزية البيانات مثل تجميع البيانات عن طريق خوارزمية معدلة على أساس **خوارزمية التجميع الـ FCM , مع مخطط الرسم البياني voronoi** ستوضح في هذا العمل. من خلال **الخوارزمية المعدلة (VFCA)**. في الخوارزمية المعدلة، المنطقة المراقبة تقسم إلى عدد من الخلايا تسمى **voronoi cells** من خلال تطبيق مخطط **voronoi**، على المنطقة المراد مراقبتها. كل خلية ممثلة بواحدة من أجهزة الاستشعار (**العقد**) التي وزعت بشكل عشوائي في المنطقة المستشعرة. يتم تقسيم هذه الخلايا الى مجاميع باستخدام طريقة التجميع (**FCM**) لتقليل المسافة التي تقطعها المعلومة للوصول إلى الهدف وتقليل إرسال البيانات المكررة من العقد المتجاورة. ثم يتم اختيار العقدة المناسبة لتمثل كل مجموعة (**رئيس المجموعة**). واستخدمت ثلاثة

معايير لعملية اختيار رؤساء المجاميع وهي: **الطاقة، بعد المسافة بين رئيس المجموعة والعقد المجاورة له , و قيم الخسارة من مخزن العقدة.** علاوة على ذلك، يعمل تجميع البيانات عند كل رئيس مجموعة على تقليل كمية البيانات المنقولة خلال الشبكة وهذا يؤدي إلى تمديد عمر الشبكة وتقليل حركة المرور التي قد تتراكم في المخزن الخاص بالـ **Sink Node**. رئيس كل مجموعة يجمع البيانات من أجهزة الاستشعار المنتمية اليه ويحولها إلى الـ **(Sink Node)**.

أظهرت نتائج المحاكاة لدراسة مقارنة بين الخوارزمية المعدلة **VFCA** والبروتوكول **LEACH** أن **VFCA** المعدلة هي أكثر كفاءة من البروتوكول **LEACH** من حيث عمر الشبكة و من حيث متوسط تبديد الطاقة ومعدل استهلاك الطاقة. وتبين دراسة أُخرى للمقارنة بين الخوارزمية المعدلة **VFCA** و **K–Means** أن **VFCA** المعدلة هي أكثر كفاءة من **K–Means** من حيث الحزم المرسلة للعقدة المسؤولة عن الشبكة **(Sink)**. و من حيث قيم الخسارة من مخزن كل عقدة ومن حيث والوقت المستغرق لإتمام العملية. ، تم تطوير عملية المحاكاة والاختبار باستخدام برنامج **MATLAB R2010a** في حاسوب يحمل المواصفات التالية: ويندوز **7** (**32** بت نظام التشغيل)، كور **i7**، ذاكرة الوصول العشوائي **4GB**، القرص الثابت **1TB**.

# 1. INTRODUCTION

Nodes in a WSN sense the environment and send the collected data to the base station. Many nodes report similar readings as data in a WSN are correlated. Thus, a large amount of energy is spent during the transmission of thousands of redundant data. Furthermore, as nodes transmit sensed values to the base station by transiting through intermediate nodes, significant energy is spent in communication. One technique used to decrease the number of redundant messages transmitted and thus pro-long the network lifetime is data aggregation. During data aggregation, intermediate nodes merge the data received from other nodes into a single representative value. The energy of the network is conserved through a reduction in the number of messages being exchanged among nodes, **Khedo, et al., 2010**. The general data aggregation algorithm works as shown in **Fig.1,** Patil**,2010**.

# 2. DATA AGGREGATIO IN CLUSTER BASED NETWORKS

Various protocols are being used for the clustering of the sensor networks like Direct Communication(DC), Minimal Transmission Energy(MTE), LEACH, K-Means, and Fuzzy C-Means. In DC, the sensor nodes far away from the base station get depleted with its energy at a faster rate, whereas in MTE, the nodes near to the base station get depleted. In LEACH, the selection of cluster heads is based on predetermined probability; other nodes choose the cluster by computing the nearest distance to the cluster heads elected. However, there is a randomized selection of cluster heads, where some of the clusters are devoid of cluster heads, resulting in poor clustering. In K-Means clustering, also known as hard clustering, the nodes near the boundary region are affected since the degree of belongingness is described in terms of either zero or one, the edge nodes may have the same degree of belongingness to more than one clusters, resulting in poor cluster formation. In the Fuzzy C-Means algorithm, which is an unsupervised, nondeterministic iterative method, there is an optimal cluster formation. The sensor nodes are assigned to a cluster, based on the degree of belongingness to different cluster in the deployed area, but the degree of belongingness needs to be computed in each and every round for every sensor node, **Keerthi and Babu,2012**.

## 3- THE MODIFIED VORONOI FUZZY CLUSTERING ALGORITHM(VFCA)

The main purpose of the Voronoi Fuzzy Clustering Algorithm (VFCA) is to increase the network lifetime. Sensors are placed to detect an event occurring in the sensing area the sensed data need to be transmitted to the sink node for acquiring knowledge. The process of transmitting of sensed data from sensor nodes to the sink node may lead to consume the energy of nodes to transmit and receive data. Data aggregation helps the nodes in the network to save their energy by fusion of data from different sensors to eliminate redundant transmissions to the base station. This significantly reduces the amount of data traffic as well as the distances over which the data needs to be transmitted and increase the network life time, data aggregation and modified Voronoi Fuzzy Clustering Algorithm are proposed in this algorithm. Initially voronoi method is applied to construct voronoi diagram for the network by partitioning the sensed area into subareas each of which controlled by one sensor node these subarea called voronoi cell, then clustering process is starting by applying FCM then data transmission and aggregation is implemented. The modified VFCA is shown in **Fig.2**.

### 3.1 The Voronoi Method

In the modified algorithm, all nodes are assumed to be stationary and any sensor node position is projected by **x** and **y** coordinate. That is for any sensor node $\mathbf{p_i=s=(x_i,y_i)}$ the calculation of distance between nodes to construct voronoi diagram is done by using Euclidean distance function to given by Eq. (1), **Nithyakalyani and Kumar, 2013**:

$$dis(s,r) = \sqrt{(x_s - x_r)^2 + (y_s - y_r)^2} \quad (1)$$

,Where **dis (s,r):** is the Euclidean distance between node (s) and set of all neighbor nodes (r).
**P** is a set of distinct points (sites) in the plane which is represents a set of points as $\{p_1, p_2, \dots, p_n\}$. We subdivide the plane into **n** cells so that each cell contains exactly one site. An arbitrary point **(x,y)** is in a cell corresponding to a site $\mathbf{p_i}$with coordinates $\mathbf{(x_{pi}, y_{pi})}$ if and only if for all $\mathbf{p_j}$with $\mathbf{j \neq i, 1 \leq i,j \leq n}$**.** That is, the Euclidean distance from **(x,y)** to any other site is greater than the distance from **(x,y)**to $\mathbf{p_i}$ , **Mumm ,2009.**

$$\sqrt{(x - X_{pi})^2 + (y - Y_{pi})^2} < \sqrt{(x - X_{pj})^2 + (y - Y_{pj})^2} \quad (2)$$

### 3.2 Setup phase of VFCA

Formation of clusters and cluster head election are the main goal of setup phase. After the division of the sensed area into voronoi cells the network of sensors should be divided into clusters and this process is called “Clustering”. The clustering process for nodes is formed by using Fuzzy C-Means clustering algorithm, the first requirement of clustering is to the selection of cluster head node and then assign nodes to the best cluster head to become as members for cluster. Each cluster contains a set of nodes and the number of nodes is not necessary equal in the clusters, a cluster head is initially elected in each cluster. The node nearest to the center selected as cluster head . This is done only at the first round but for the next rounds the selection of cluster head based on the distance and energy for the nodes. **Fig.3** shows the five clusters. Which shows the location of final centers when choosing number of clusters for example (5-clusters).

### 3.3. The Steady-State phase of VFCA

Once the clusters are created and cluster heads are elected, data transmission can begin. The data transfer is begin starting from each node in the cell or cluster to communicate with the base station. This communication doesn’t pass directly, the node communicates with CH, and all heads transfer data directly to the main base station. Every member in each cluster transmits the sensed data $(V(p_i)(d))$ to its corresponding cluster head, $(V(p_i)(d))$ comprises value of sensed data for cell $V(p_i)$ . The cluster head nodes received the data until it's buffer full then the data aggregation computing using one of the following equations, **Nithyakalyani and Kumar, 2013**:

$$\boldsymbol{Avg} = \frac{\sum_{i=1}^{n} di}{n} \quad (3)$$

$$\boldsymbol{Max} = \max(D) \quad (4)$$

$$\boldsymbol{Min}=\min(D) \tag{5}$$

Where : di : is the data sensed by node i.
D: is the total sensed data received in each cluster head.
Avg: average value of ($d_i$).
Max: maximum value of ($d_i$).
Min: minimum value of ($d_i$).
By taking the aggregation function for the received data in each cluster head the cluster heads send the aggregated data which is represent the maximum degree for each cluster to the sink node which is also take the same aggregation function for the received data. For each sensor node, the energy is consumed due to the receiving and transmitting for the data. Depending on the distance between transmitter and receiver. After sending data, the dissipated energy for each node and each cluster head to transmit and receive data should be calculated in order to use it as parameter behind the distance to select new cluster head. If the energy of cluster head reach to zero this means that the cluster head will be die and new node from that cluster should be chosen with highest remaining energy and nearest to the center. The energy dissipated for transmitting and receiving data for normal nodes is calculated as in the following Eqs. (6) and (7).

$$ETX\_node = E * distance(i,j) \tag{6}$$

$$ERX\_node = E * distance(i,j) \tag{7}$$

Where:
ETX_node: Transmitter Electronics Energy dissipated for each node.
ERX_node : Receiver Electronics Energy dissipated for each node.
distance(i,j) : Distance between node i and its cluster head j.
E : Initial Energy for each node.

Then the remaining energy for each node can be calculated by subtracting dissipated energy for transmitted and received from the energy of that node.

The energy dissipated for data transmitting and receiving data for cluster head nodes is calculated as in the following Eqs. (8) and (9).

$$ETX\_CH = ETX\_CH – (Sum(ETX\_node(cluster))) \tag{8}$$

$$ERX\_CH = ERX\_CH - (sum(ERX\_node(cluster)) \tag{9}$$

Where:
ETX_CH: Transmitted energy dissipated for each cluster head.
ERX_CH : Received energy dissipated for each cluster head.

From the above equations one can conclude that the cluster head node has dissipated energy in transmitting data to sink and dissipated energy to receiving data from its members, while the normal nodes has only dissipated energy for sending data to its CH, and sink node has dissipated energy to received data from cluster heads.

- The selection of cluster head node usually depends on several parameters, this paper was used energy as parameters as explained above and then was used packet loss value from the buffer of cluster head node which can be calculated as explained in Eq.(10).Each cluster head after receiving data from its members should check if there is packets was lost from its buffer as in Eq.(10)

$$Pl = BUF - P \tag{10}$$

Where:
Pl: packet loss.
BUF: size of the buffer.
P: Number of the received packets.
whereby each time cluster head receive sensed data from its members the cluster head subtract received data packets from its buffer size to see if there is space in its buffer for the receiving packets or not and if it is find that some of the received packets was loss it calculate the value of packet loss and then re-cluster the network again in the second round and chose the new cluster head nodes with low packet loss and nearest to the final centers of FCM clustering.

## 4- SIMULATION SETUP AND RESULTS

In this section the simulation of the modified VFCA in the network environment and comparison with LEACH protocol in WSN are presented. Simulation is implemented with GUI (Graphical User Interface) in Matlab and it is achieved with n-sensor nodes that are randomly distributed within a (100x100) m region. These nodes are assumed to be stationary, the proposed algorithm is simulated in a wireless sensor network and is compared with LEACH protocol. All simulation parameters in the modified algorithm are presented in Table(1).
To evaluate the performance of the modified VFCA, the number of sensor nodes should be inserted for example 100-stationary sensor nodes are randomly placed in a sensed area also number of clusters should be insert for example 5 - clusters. **Fig.4** illustrates the model of the 100-nodes simulation network after applying voronoi diagram and then FCM clustering for the first round.

The running time taken to construct voronoi diagram represented in **Fig.5**. From the figure one can see that voronoi diagram take less time for constructing and increasing in low rate with increased the number of nodes.

The running time taken by FCM for 100 sensor nodes divided to different number of clusters represented in **Fig.6**. From the figure one can see that FCM take minimum time to divide the voronoi cells to clusters and this time increasing in low rate with increased the number of clusters.

Then after divide the voronoi cells to clusters should press the next button illustrated in **Fig.4** and second button illustrated in **Fig.7** to see the information about each cluster these information include the distance between each node in the cluster and the cluster head ,the sensed data for each node in each cluster represented by temperature degrees and the energy for each node in the cluster all these information are shown in **Fig.8**.

The modified VFCA choosing the optimal cluster head node at each new round depending on the division of the voronoi cells that cover all the sensed area to clusters using FCM and the process of election of cluster head depending on the highest node energy, this may be lead to enhance the Performance of the applying modified VFCA to the WSN from hand of running time of the network and the energy dissipated and network lifetime to better than LEACH protocol as seen following.

### 4.1 Network Lifetime

The lifetime of the network for the modified VFCA and, LEACH protocol with respect to alive nodes in the number of rounds is shown in **Fig.9**. As compared with LEACH, the modified VFCA has an improvement over the stability period as compared with LEACH protocol. As it is clear from **Fig.9**, the first node dies in LEACH protocol at round 932 and last node dies at round 1312, whereas in modified VFCA, the first dead node is at round 977 and last node die at round 1520 as illustrated in Table(2). This shows an improvement due to the use of the voronoi diagram. This help the network to cover all area and reduce transmission distance. Also by using Fuzzy C-Means clustering (FCM) with voronoi diagram will be reduced the transmission distance. The modified VFCA is chose the best node as the cluster head node depending on the cluster node with highest energy and this will be lead to increase the network lifetime.

### 4.2 Average Energy Consumption

For any routing protocol, lesser energy consumption is an important issue so that the network stays alive for longer period or for more number of rounds. **Fig.10**shows the average energy dissipation for all nodes with respect to number of rounds. The rate of energy consumption of the modified VFCA is lower than the rate of LEACH. Nodes in the modified VFCA deplete their energy in very slow rate as compared to LEACH. Eq. (11) shows how to calculate the average energy conception for the network, **Anitha andKamalakkannan, 2013.**

$$E_a = \frac{\sum_{K=1}^{N} (E_{iK} - E_{ck})}{N} \qquad (11)$$

Where, $E_{ik}$ = The initial energy level of a node.
$E_{ck}$ = The current energy level of a node.
N = Number of nodes in the simulation.

## 5- PERFORMANCE OF THE MODIFIED VFCA WITH K-MEANS ALGORITHM

The simulations run with the parameters mentioned in Table (1). To evaluate the performance of the modified VFCA, 20-stationary sensor nodes are randomly placed in a sensed area with one sink node. **Fig.11** illustrates the model of the 20-nodes simulation network. The numbers of elected cluster heads from all nodes in the network is chosen randomly. In the modified VFCA the role of CH node rotates between all sensor nodes in the network if the selected cluster head has packet lost from its buffer because the congestion traffic in its buffer the cluster head sends the packets that are in its buffer and all other cluster heads also send the aggregated data in their buffers and then the process of clustering repeated and chose new cluster head nodes depending on the value of packet loss and the distance between sensor nodes and the centers that gets it from FCM clustering algorithm.

### 5.1 Total Number of Packets at Buffer of Sink Node

The total number of packets can be computed according to the packets reached to the buffer of the sink node. The network model that is used for implementing varying the number of cluster heads consists of 400 nodes. **Fig.12**. illustrates the compression between the modified VFCA and K-Means algorithm in terms of the total number of packets received at the sink node in the network, when varying the number of clusters in the network (5-10-15-20-25) clusters, this may lead to increased number of packets arrived at sink node as shown in Table (3).

From **Fig.12** one can conclude that the modified VFCA divides the area into cells and this will lead to optimization in distribution and aggregation of data from the network then the sink node may reach the data that needed from the sensed area without losses and congested. In K-Means algorithm ,although the sink node received more data than modified VFCA when increasing the number of clusters but this may be lead to congestion at sink node and losses some of the important data, where as in the modified VFCA the congestion and losses not accrue. Therefore, the VFCA is better than K-Means algorithm.

### 5.2 Buffer Utilization (BU)

The buffer utilization of sink node can be computed according to the buffer of sink node when increasing number of nodes and fixing number of clusters once and increasing number of clusters and fixing number of nodes in the other once. as described in Eq. (12).

$$BU = \frac{P_{total}}{BU_{\sin k}} \tag{12}$$

Where, BU: Buffer Utilization.
$P_{total}$: total packets received at sink buffer.
$BU_{sink}$: size of the sink buffer.

**a. (BU)When Increasing the Number of Nodes and Fixing Number of Clusters:**
**Fig.13** denotes the Buffer Utilization of sink node after applying Eq.(12) , by increasing number of nodes in the network and fixing numbers of clusters to 5. One can see from **Fig.13** every time number of nodes increase the buffer utilization is increased because when increasing number of nodes and fixing the number of clusters this lead to increase the number of nodes in each cluster and then when cluster head aggregate the packets that is received from its members, the sink received values from only five nodes represent data of its members. The best value of buffer utilization should be less than or equal to one but not greater than one from the **Fig.13** one can see that the buffer utilization for K-Means algorithm begin smaller than one when number of nodes equal to 20-nodes and when the number of nodes increased the value of BU increased to value greater than one. But the value of BU when apply modified VFCA does not arise greater than one. This mean that the modified VFCA is better than K-Means algorithm because the modified VFCA managing the buffer accurately so that no losing in packets are happened. Table (4) shows the values of buffer utilization in the sink node.

**b. (BU)When Increasing Number of Clusters And Fixing Number of Sensor Nodes:**

**Fig.14** denotes the Buffer Utilization of the sink node by increasing number clusters in the network and fixing number of sensor nodes to 400. One can see from **Fig.14** when the number of clusters increase the buffer utilization approaches to one because when increasing number of clusters the numbers of cluster heads sending their data packets to sink node will be increased and this leads to increase the number of data packets arrived at sink node and then approximately all space in the CH buffer will be filled by data packets.

The behavior of the modified VFCA is better than k-Means algorithm. This is because the modified VFCA managing the buffer of sink more efficient than the K-Means algorithm. Table (5) shows the values of buffer utilization of the sink with respect to the number of clusters in the network.

### 5.3 Packet loss value

**Fig.15** illustrates packet loss value for each node in the network that consists of 20 sensor nodes divided to 20 voronoi cells and these cells are divided to five clusters for the first round. One can see from the figure that only cluster head nodes has packet lost from its buffer because the CH received data from its members and these data may be huge lead to packet loss from CH buffer. From **Fig.16** one can see that the modified VFCA is better than K-mean because the cluster head approximately received all data from its member. Table (6) shows the cluster head nodes, number of received data at each CH and packet loss values for both VFCA and K-Means algorithm for the first round.

### 5.4 Running Time

**Fig.16** shows the Comparison between modified VFCA and K-means algorithm in terms of running time for the network in **Fig.12**. The number of nodes is changed at each time of execution and evaluate the time in both algorithms, the number of clusters constant for both algorithms to find the running time. The number of cluster heads equal to 5. **Fig.16** shows that the running time of VFCA less than running time of K-Means algorithm. This is because the method of computing the membership function in the modified VFCA which is repeated in a single voronoi cell, so that the execution of this function is just in the cell and not in all nodes in the cell, and hence the running time will be reduced.

## 6- CONCLUSION

In this paper, a modified VFCA for data aggregation is presented. In order to reduce the average energy dissipated in the sensor network and increased the network lifetime. At first the voronoi diagram is applied for the sensed area based on energy of each node once and packet loss value for other once. Consequently, the cluster head node is selected by applying FCM clustering algorithm. Then every node send its sensed data to its cluster head. Cluster head aggregate the received data by taking maximum value and send the aggregated data to sink node. Finally, the results shows that the behavior of modified VFCA is better than both LEACH protocol and K-Means clustering algorithm.

## REFRENCES


- Anitha R. U. and Kamalakkannan D. P.,2013, *Enhanced Cluster Based Routing Protocol for Mobile Nodes in Wireless Sensor Network*, Proceedings of the International

Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME), IEEE, India, 20-22 February, PP.187-193.

- Keerthi, M, and Babu, B.S., 2012, *An Improved FCM's Clustering Protocol for Wireless Sensor Networks*, International Journal of Electronics and Communications (IJEC), Vol.1, Issue 1, PP. 69-73.
- Khedo, K., Doomun, R., and Aucharuz, S.,2010, *READA: Redundancy Elimination for Accurate Data Aggregation in Wireless Sensor Networks*, Wireless Sensor Network, Vol. 2 No. 4, 2010, pp. 300-308.
- Mumm M.,2009, *Voronoi Diagrams*, The Montana Mathmatic Enthusiast (TMME), Vol 1, no.2, pp.44.
- Nithyakalyani, S., and Kumar S. S. ,2013, *An approach to Data Aggregation in Wireless Sensor Network Using Voronoi Fuzzy Clustering Algorithm*, Journal of Scientific & Industrial Research,Vol. 72, PP. 287-293.
- Patil, N. S., and Patil, P. R. ,2010, *Data Aggregation in Wireless Sensor Network*, ICCIC International Conference on Computational Intelligence and Computing Research, B.V. Tamilnadu college of Engineering coimbatore, India, 28-29 Dec.2010, PP. 162-168.

**List of symbols**

$(x_i,y_i)$=X-coordinate and Y-coordinate of sensor node.

BU: Buffer Utilization.

BUF: size of the buffer.
$BU_{sink}$: Size of the sink buffer.
D: is the total sensed data received in each cluster head.
di = The data sensed by node i, $m^2$.
distance(i,j) : Distance between node i and its cluster head j, $m^2$.
E : Initial Energy for each node, Joule.
Ec = The current energy level of a node, Joule.
Ei = The initial energy level of a node, Joule.
ERX_CH : Received energy dissipated for each cluster head, Joule.
ERX_node : Receiver Electronics Energy dissipated for each node, Joule.
ETX_CH: Transmitted energy dissipated for each cluster head, Joule.
ETX_node: Transmitter Electronics Energy dissipated for each node, Joule.
N = Number of nodes in the simulation.
P: received packets.
$p_i$= The location of sensor node.
Pl: packet loss.
$P_{total}$: Total packets received at sink buffer.
s= senson node.
V(pi) = The voronoi cell of node ($p_i$).

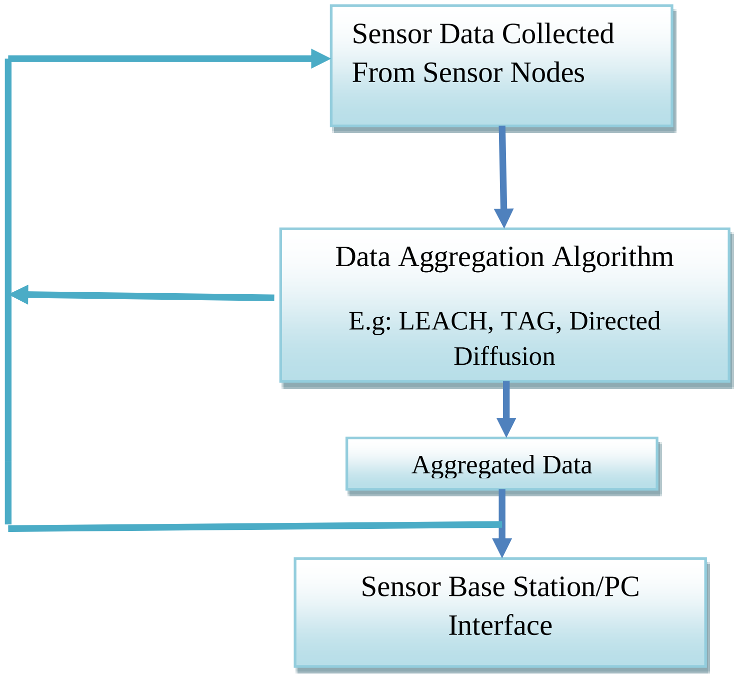


**Figure 1.** General architecture of the data aggregation algorithm

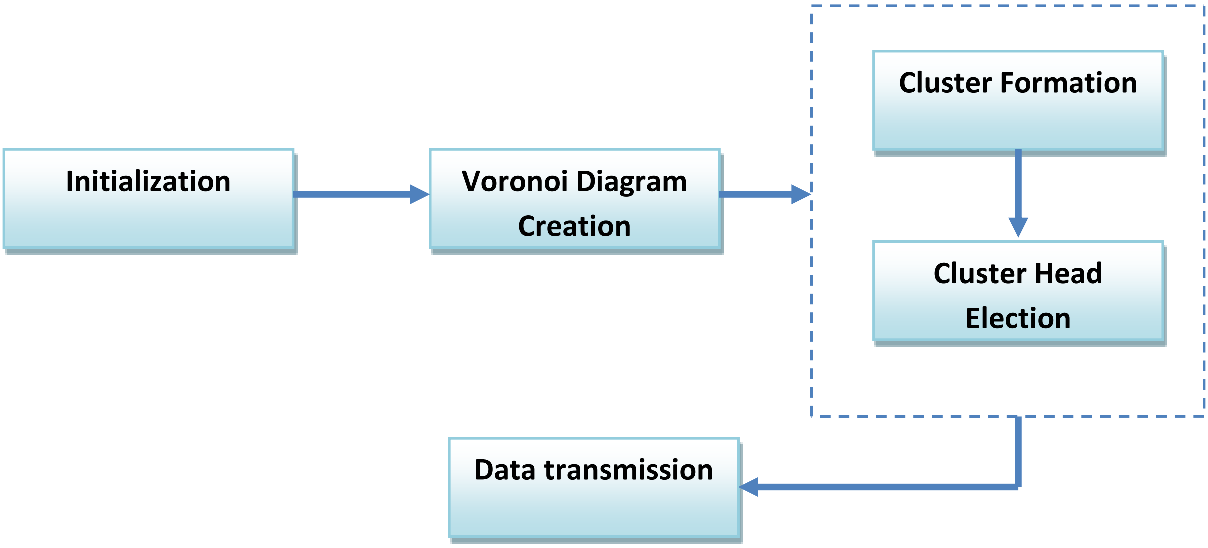


**Figure 2**. Modified voronoi fuzzy clustering algorithm

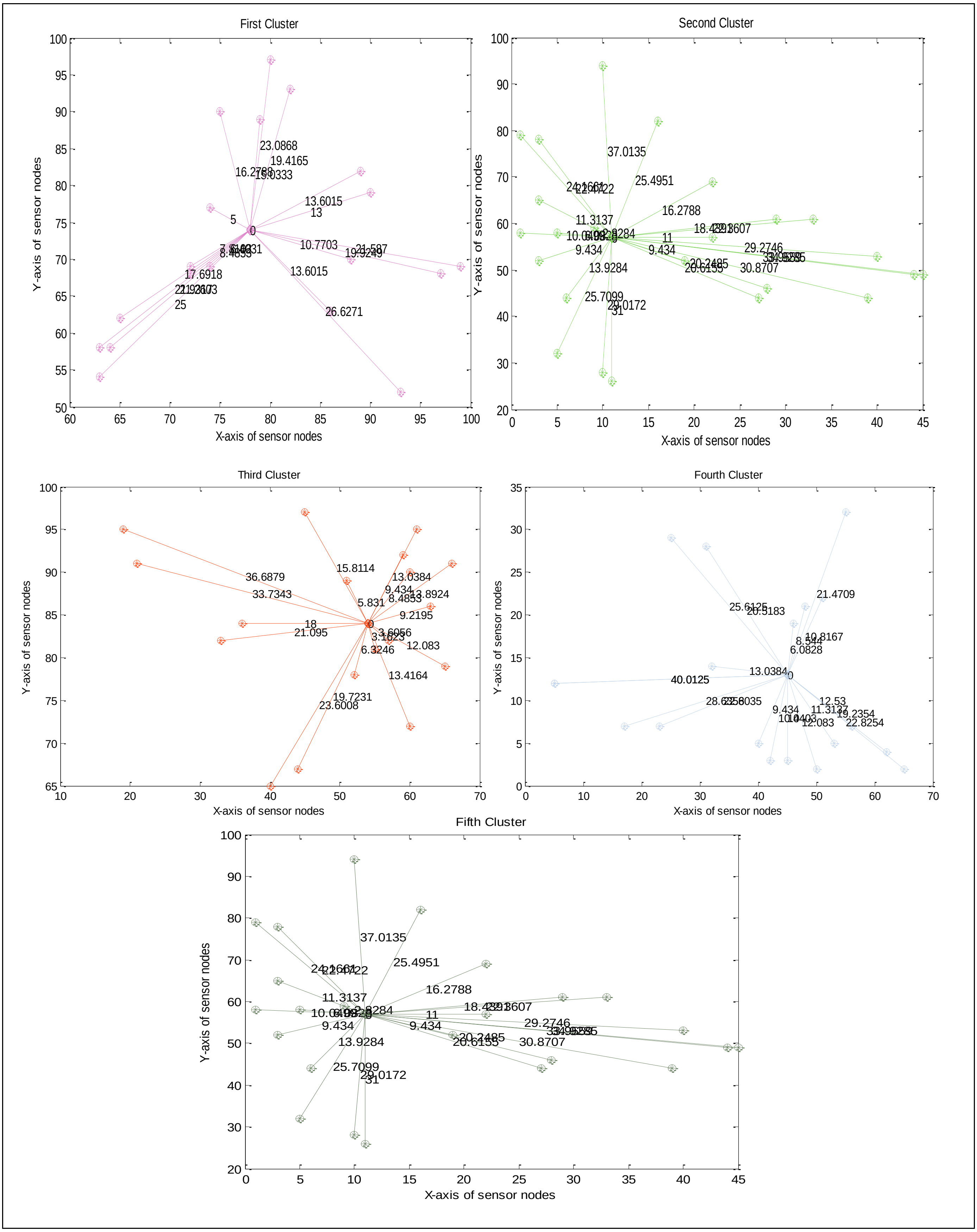


**Figure 3.** Sensors after applying fuzzy c-means clustering

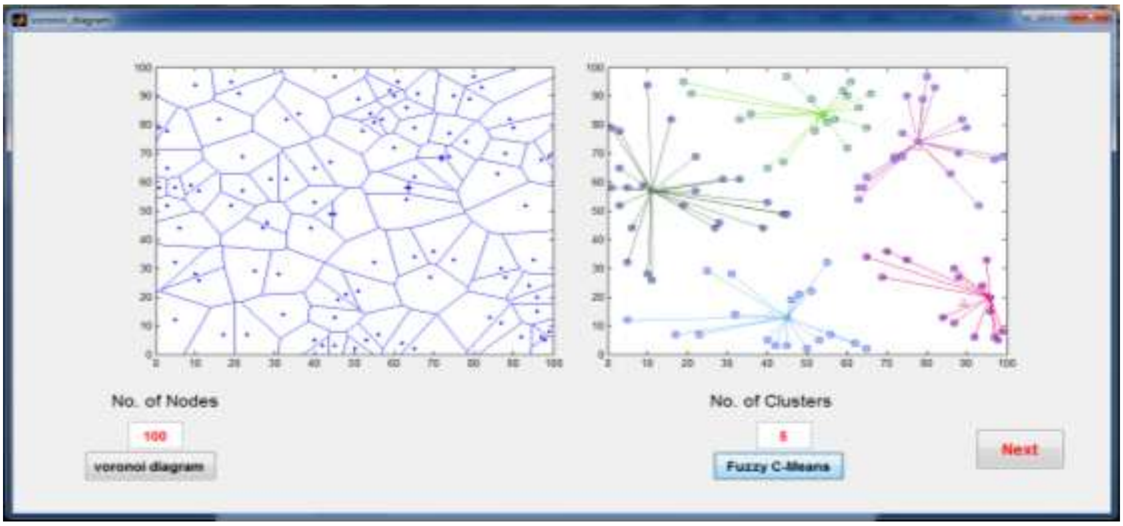


**Figure 4.** Sensor nodes after applying voronoi diagram and FCM

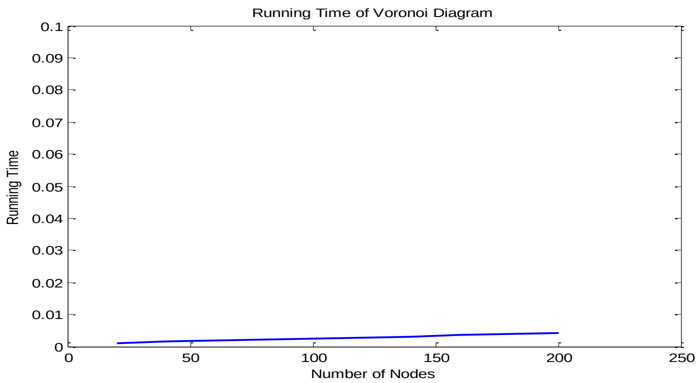


**Figure 5.** Running time of voronoi diagram vs number of nodes

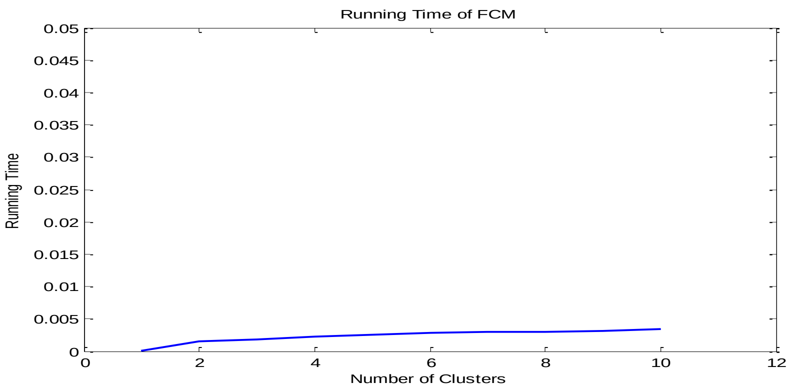


**Figure 6.** Running time of FCM vs number of clusters


**Figure 7.** Represent the button to run the second and third round

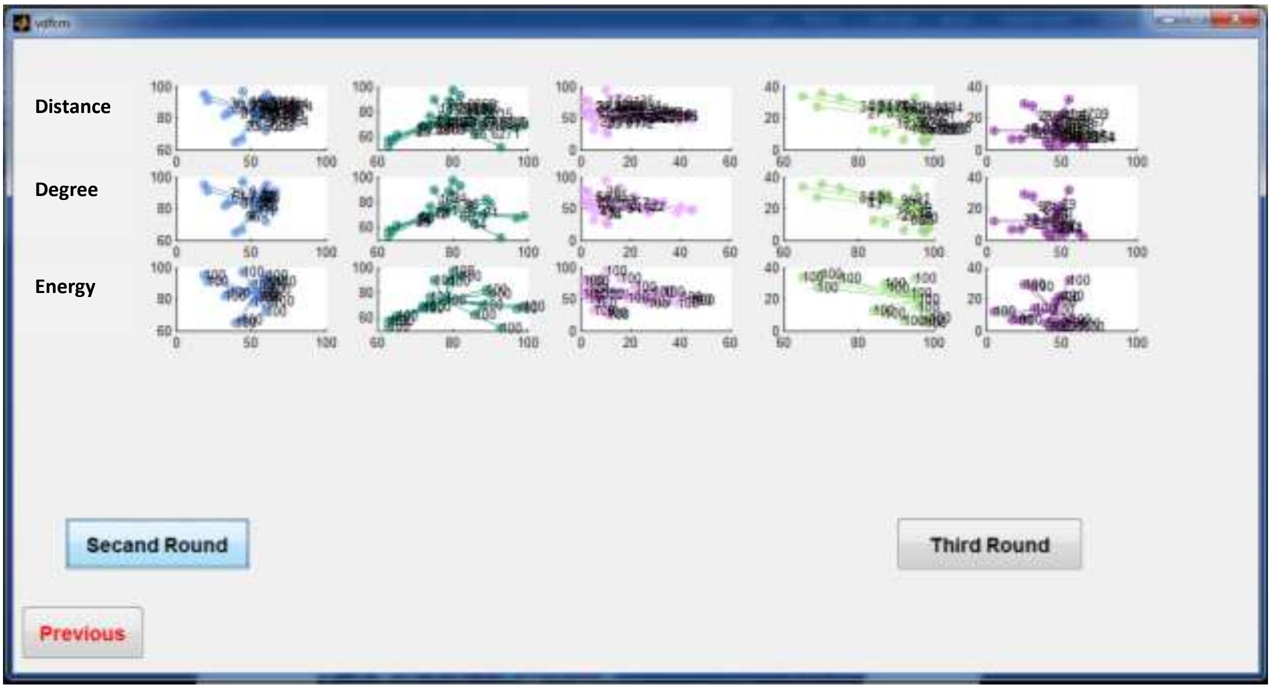


**Figure 8.** Information for each node in each cluster for first round

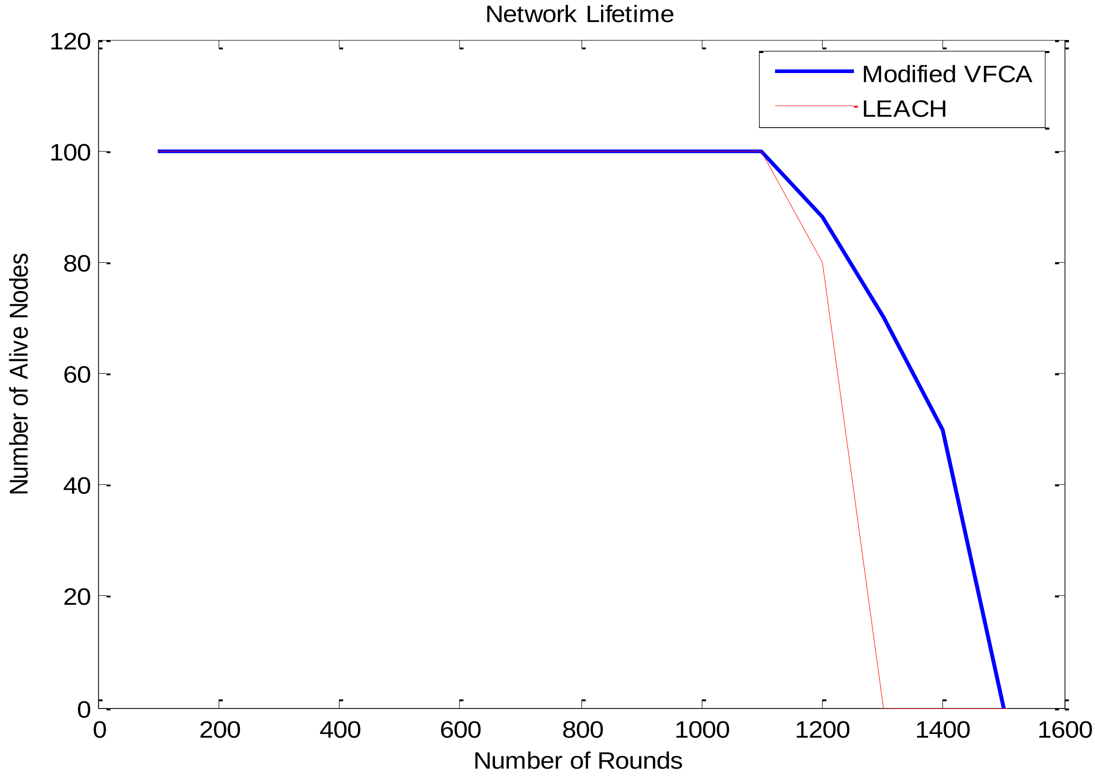


**Figure 9.** Network lifetime

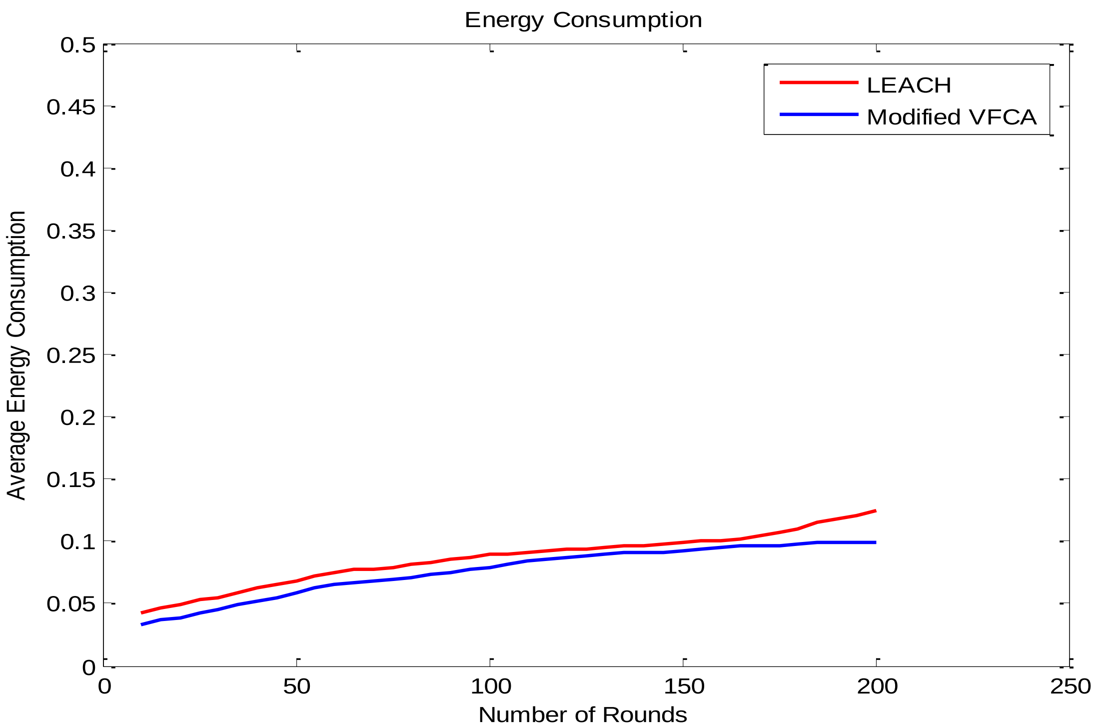


**Figure 10.** Average energy consumption for the network

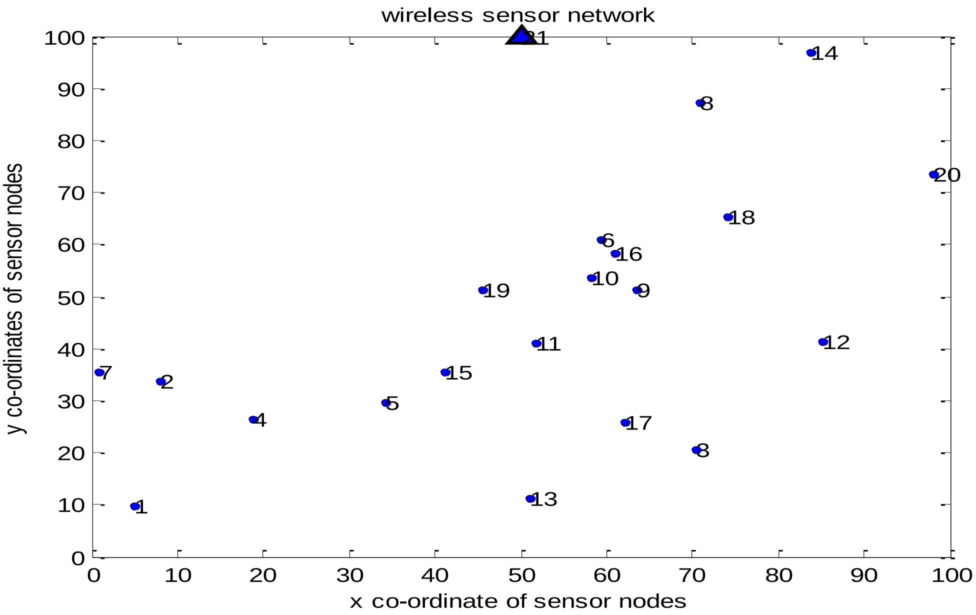


**Figure 11.** The nodes distribution over the sensed area

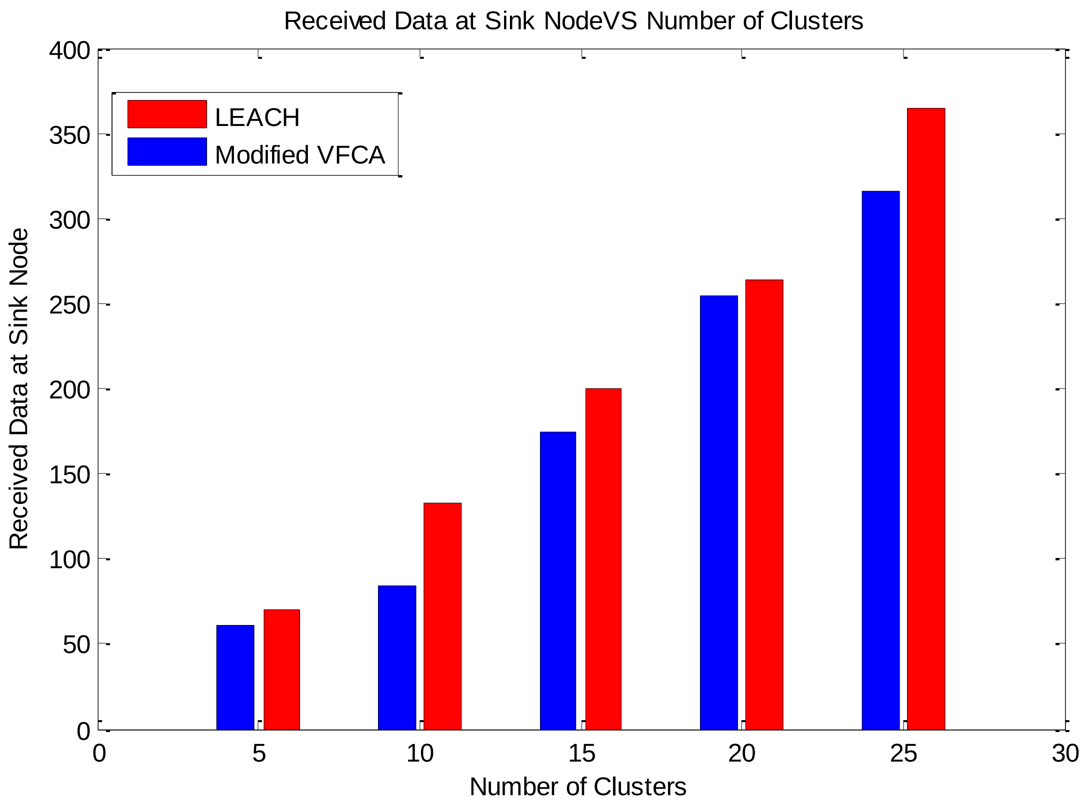


**Figure 12.** Received data at sink node vs number of clusters

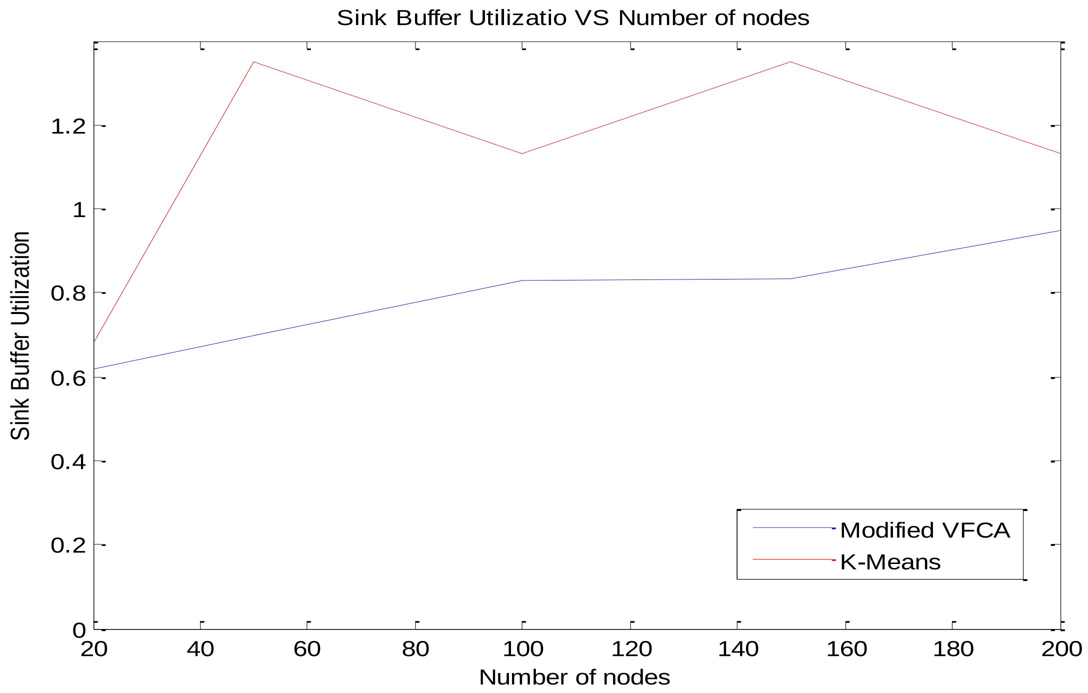


**Figure 13.** Buffer utilization of sink node vs number of nodes

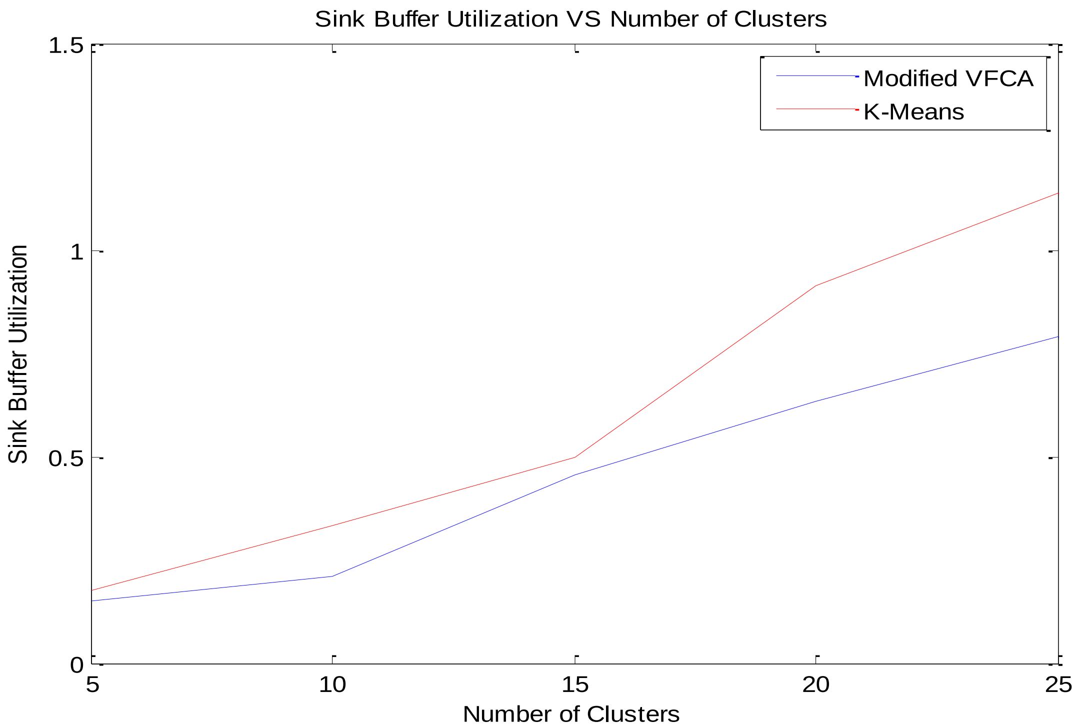


**Figure 14.** Buffer utilization of sink vs Number of clusters

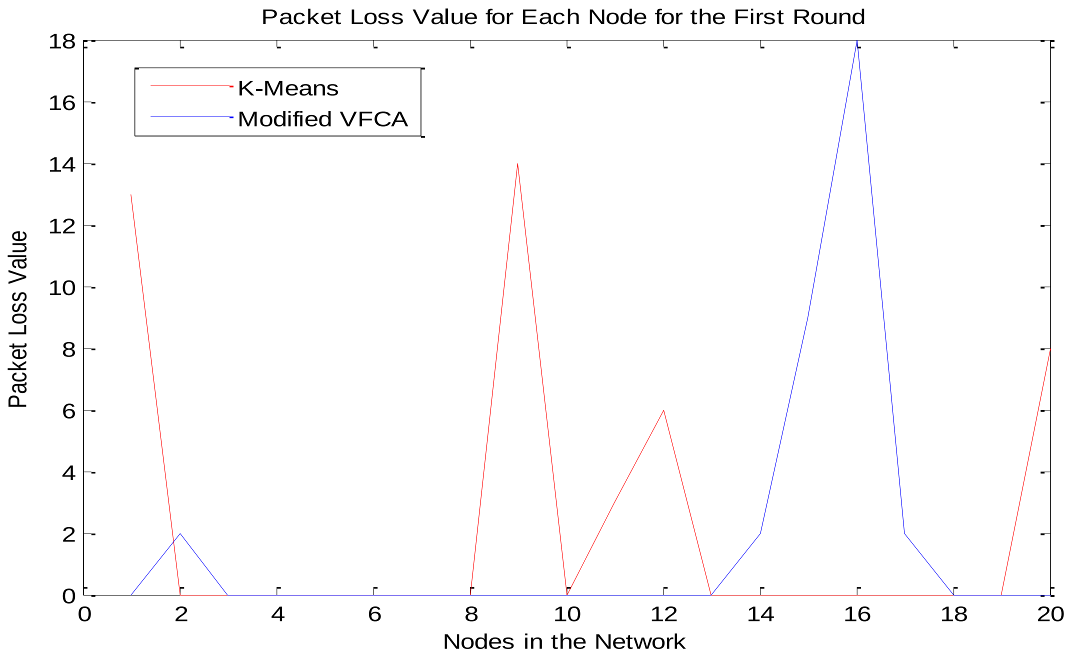


**Figure 15.** Packet loss value for each node in the network

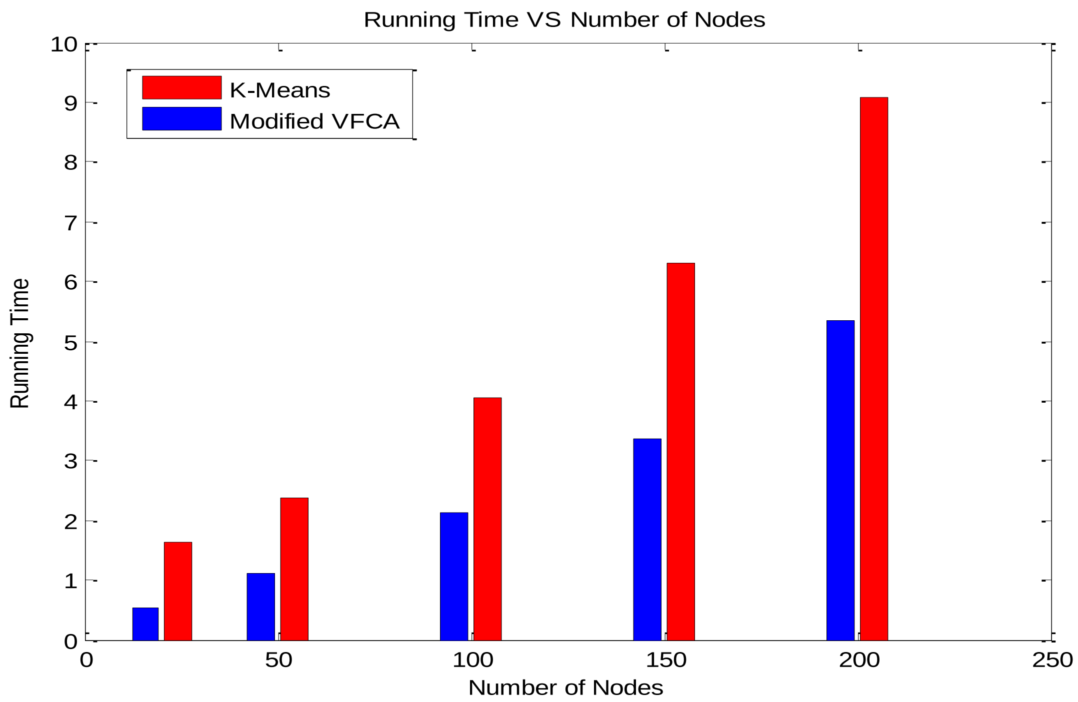


**Figure 16.** Running time vs number of nodes

**Table 1.** The simulation parameters

| Parameter | Value |
|---|---|
| **Network size** | 100mx100m |
| **Number of nodes** | n=100 |
| **Number of sinks** | 1 |
| **Behavior of nodes** | Static |
| **Base Station position** | [50,100] |
| **Number of clusters** | k=5 |
| **Iteration steps** | 1000 |
| **Termination condition** | 0.00001 |
| **Buffer size of nodes** | 35 packet |
| **Buffer size of sink** | 100 packet |
| **Initial energy** | 100 nJ/bit |
| **ETX=ERX** | 50nJ/bit |

**Table 2.** Alive nodes in both LEACH and VFCA when every time increasing rounds (Network lifetime)

| No. on Rounds | Alive nodes in LEACH | Alive Nodes in VFCA |
|---|---|---|
| 100 | 100 | 100 |
| 200 | 100 | 100 |
| 300 | 100 | 100 |
| 400 | 100 | 100 |
| 500 | 100 | 100 |
| 600 | 100 | 100 |
| 700 | 100 | 100 |
| 800 | 100 | 100 |
| 900 | 100 | 100 |
| 1000 | 100 | 100 |
| 1100 | 100 | 100 |
| 1200 | 80 | 88 |
| 1300 | 23 | 70 |
| 1400 | 0 | 50 |
| 1500 | 0 | 0 |

**Table 3.** Received packets at sink node vs number of clusters

| Number of Clusters | Number of Received Packets at Sink Node (VFCA) | Number of Received Packets at Sink Node (K-Means) |
|---|---|---|
| 5 | 61 | 70 |
| 10 | 84 | 133 |
| 15 | 175 | 200 |
| 20 | 254 | 264 |
| 25 | 316 | 365 |

**Table 4.** Buffer utilization in sink node vs number of nodes

| Number of nodes | Sink Buffer Utilization (VFCA) | Sink Buffer Utilization (K-Means) |
|---|---|---|
| 20 | 0.6167 | 0.6833 |
| 50 | 0.7 | 1.35 |
| 100 | 0.83 | 1.133 |
| 150 | 0.8334 | 1.35 |
| 200 | 0.95 | 1.133 |

**Table 5.** Buffer utilization of the sink vs number of clusters

| Number of clusters | Sink Buffer Utilization (VFCA) | Sink Buffer Utilization (K-Means) |
|---|---|---|
| **5** | 0.1525 | *0.176* |
| **10** | 0.21 | *0.3325* |
| **15** | 0.4575 | *0.5* |
| **20** | 0.635 | *0.9125* |
| **25** | 0.79 | *1.1375* |

**Table 6.** Packet loss value for each cluster head node

<table>
<tr><th colspan="2">Cluster head nodes</th><th>Packet loss Value</th></tr>
<tr><td rowspan="5">K-Means</td><td>Node 2 - cluster head of first cluster</td><td>13</td></tr>
<tr><td>Node 9 - cluster head of second cluster</td><td>14</td></tr>
<tr><td>Node 11 - cluster head of third cluster</td><td>3</td></tr>
<tr><td>Node 12 - cluster head of fourth cluster</td><td>6</td></tr>
<tr><td>Node 9 - cluster head of fifth cluster</td><td>8</td></tr>
<tr><td rowspan="5">Modified VFCA</td><td>Node 2 - cluster head of first cluster</td><td>2</td></tr>
<tr><td>Node 2 - cluster head of first cluster</td><td>2</td></tr>
<tr><td>Node 2 - cluster head of first cluster</td><td>9</td></tr>
<tr><td>Node 2 - cluster head of first cluster</td><td>18</td></tr>
<tr><td>Node 2 - cluster head of first cluster</td><td>2</td></tr>
</table>