
Atefeh Heydariyan, Amir Abbas Baradaran, Elham Rezai

Department of Computer Science, Payame Noor University (Tehran branch), Tehran, Iran

ABSTRACT

In this paper an algorithm is proposed based on clustering routing in wireless sensor networks; Moreover, reduction of energy consumption and network lifetime increasing are the most important issues in designing practical programs and wireless sensor network algorithms; meanwhile, clustering of sensor networks plays an effective role of either more suitable management or measurability of these networks, so (IMKREC: Improved k-means algorithm Method for Reduce Energy Consumption in Wireless Sensor Networks) algorithm is a new technique for election of the best Cluster Head in this algorithm, we can eliminate some problems of K-means algorithm and improve them as well, with the network is converted to grids and election center of gravity of each cluster, and considering energy parameters ∞ distance as well, so simulation results show that presented algorithm of lifetime improves effectively measurability and average of network energy consumption in compare of previous mentioned clustering algorithms.

KEYWORDS—Wireless sensor network, Grid, Center of gravity, Energy center, Network lifetime, k-means algorithm

INTRODUCTION

Initially, wireless sensor network consist of a large number of communication nodes within special, capabilities, sensor, processing ∞ calculations, also, wireless sensor network has one or several information saving stations which have been deployed in center or outside of sensor field [T] which is usually called as sink or BS (Base Station), meanwhile, these nodes have been made to find out about the type of an accident, and obtained data are sent to sink for more analysis. [2, 3]; moreover, sensor nodes are widely used for monitoring peripheral parameters, or discovering of some special incidents and example of this is forest firing- monitoring battle field, dynamic purpose, and transferring of observable date to BS or sinks that have been sent through wireless single-hop network is a certain technology for all of people, and sensors of environment are implemented by sensor nodes in width of geographical locations. [5] Although in high measurements these sensors have high efficiency, they have limitations that can be classified in to their low energy power, being impossible of their battery exchange and in majority of cases, little width of their band and radio short wave that all of these limitations are faced with problem in bigger environments within move sensor nodes, [6], so during long period of time, battery energy of wireless sensor nodes are finished, and causes disconnection of network which this factor has caused that energy consumption has an important effect on designing wireless networks; moreover, the purpose of routing in data accumulation is obtaining favorable condition for networks topology caused purpose of routing algorithm in sensor networks is to reduce cost either of transferring or conveying of gathered data by nodes which have been distributed. [7]

Also, regarding mentioned points, many routing methods have been created for these networks that can be classified in to 3 groups.

Data base algorithms and hierarchy based on place [8] among them, hierarchical either methods or clustering take benefit from several specifications, Because of this, the first specification is that they divide network in to several different sections, and each section is monitored by cluster head that causes clustering methods have high measurability another specification is that because 1 n clustering methods, one cluster has received data of its member sensors, and after data accumulation sends them to base station, so it causes that data is decreased. Also in this paper algorithm of distributed clustering routing in wireless sensor networks is proposed, also in presented algorithm firstly network is Girded and with determining center of gravity and K-means algorithm performance on it, then election of cluster Heads and formation of clusters based on parameters of residual energy extent of each sensor, average sensor energy of each cluster, and number of each sensor corners ∞ distance of each sensor to base station, are performed, therefore, simulation results show that presented algorithm of lifetime and measurability improves average of network energy consumption in compare of previous clustering algorithms effectively.

RELATED WORKS

In recent years, many of clustering algorithm and routing protocol based cluster have been devised for wireless sensor networks, so their major purpose is to protect energy consumption of sensor nodes effectively, and decrease date transfer to sink, lengthening of network lifetime [9] Initially, protocol of adaptive clustering hierarchy with low energy LEACH [10] is the first and most famous protocol based on clustering in wireless sensor networks. In LEACH algorithm, creation of clusters is done in distributed form, also relation of nodes to related cluster Head is done based on distance, but nodes that have no cluster head called as normal node which transfer their date to cluster head; moreover, cluster Head nodes in compare of normal nodes need more energy, so LEACH protocol operation is divided in to several rounds, and each round is started with installation phase (formation of clusters) in which clusters are
organized; in addition to installation phase, there is date transfer phase in which normal nodes have transferred their date to cluster heads, and cluster heads send accumulated packet to base station after doing either aggregation or combination of data to decrease amount of information that must be transferred to basis station. Also in LEACH algorithm, timing of sensor data transfer is done by multi-access protocol with either code or time division; Moreover, election of cluster Head is through probability function based on this, each node selects random number between [0,1]; produced number is lower than D(n), then that node introduces itself as CH

\[ D(n) = \begin{cases} \text{PRO} & \text{if nEE} \\ 1 - \text{Pro}(r \mod \frac{1}{\text{pro}}) & \text{else} \end{cases} \quad (1) \]

In which pro is the probability of head cluster in v, number current round and E combination of nodes that in \( \frac{1}{\text{pro}} \) current round haven’t been cluster head [11]. Another algorithm which’s called as HEED in which there are four purposes for network lifetime increasing, finishing of clustering phase, after certain and definite numbers of frequency, and minimizing control suitable of distribution of clusters across network. Then, each node with probability based on its residual energy extent decides if it is cluster head or not, but this decision is short time, and is finalized after several frequencies, so nodes that have elected themselves as cluster head in form their corners, so if each of corners haven, t been member of cluster before, they will be member of this cluster, and if a corner is a member of another cluster before whose residual energy of its head cluster is lower in compare of residual energy of new cluster, corner is joint to new cluster.

In addition, if a corner is head itself decides if it remains as a either CH continuous or transfer to new cluster after in compare of it residual energy extent with its residual energy extent of introduced cluster head, so if each cluster head doesn’t decide for joining to another cluster, amount of its p probability had doubled, and. Again introduces itself as CH to its corner, but if amount of p in node is bigger than [1], then that node produces itself as final CH. In this case, corner of this node will be member of final cluster in which there is no change. Therefore, in this phase of a node hasn’t received any introducing message of cluster, it itself will decide if to be a new cluster or not, (12, 13) There is another algorithm called as EEUC algorithm which is a complete distributed algorithm in which cluster heads are elected through local competition, and election of cluster head at first based on residual energy of each node. Unlike LEACH, EEUC algorithm has neither neighgd frequency nor difference from HEED. [14], also it equilibrates energy consumption on network successfully; moreover, it has an effective improvement on network lifetime; meanwhile, competition limitation of node is decreased to base station the same as its distance; therefore, simulation results show that clusters which are nearer to basis station, either have smaller size or low energy. Consumption for data processing in cluster during this period, also it can keep more energy for delaying inside cluster traffic; more over, trend of this algorithm is that in phase of network using, basis station distributes hello message to all nodes which have special power level with his method each node can calculate it’s approximate distance to BS based on received signal power. Furthermore, this issue not only helps that nodes have special power level for establishment of relationship with basis station, but also it helps us to produce asymmetrical clusters, front figure shows that EEUC Mechanism in clouds circles with asymmetric size that show cluster with in equal sizes.

Table1: parameters

<table>
<thead>
<tr>
<th>CH</th>
<th>Experimental cluster head</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_{\text{com}} )</td>
<td>Competition limitation</td>
</tr>
<tr>
<td>( k )</td>
<td>Constant coefficient between [0,1]</td>
</tr>
<tr>
<td>( d_{\text{max}} )</td>
<td>The minimum and maximum of distance</td>
</tr>
<tr>
<td>( d_{\text{min}} )</td>
<td>Between sensor node and BS</td>
</tr>
<tr>
<td>BS</td>
<td>Base station</td>
</tr>
</tbody>
</table>

According to table, at first several cluster heads are selected for competition in selection of final cluster head experimentally; moreover, each node is converted into experimental cluster head.

With probability which is predefined threshold extent, in this time, other nodes become in active to finish phase for cluster head selection, so this algorithm can increase network life time partly decrease energy consumption [15]. There is another algorithm called as CRCWSSN [16] based on genetic includes Rounds that each round includes two phases called setup and steady in set up phase, clustering operations are performed, and in steady phase is done data transfer, also binary coding system is used to show
normal nodes and CH nodes. Also, selection of CH is based on the distance of each node from center of Gravity in its Grid, and initial energy of each node, meanwhile, trend of this algorithm based on re-clustering that can decrease considerably performance cost of algorithm; moreover, genetic algorithm uses a population consisting of bit strands streams called chromosome (initial population), and applies fitness function in each round on that population and this would be continued until optimal solution will be reached, so the fitness function used in CRCWSN is based on Heinzelman model. (17) Amount of energy consumed of total network is equal to

\[ E = E_a + E_b + E_c + E_d \quad (3) \]

\( E_a \) = energy for sending information from each normal node to CH.
\( E_b \) = energy consumed by CH for receiving information from normal nodes energy for gathering information in CH
\( E_c \) = Energy for gathering information in CH
\( E_d \) = energy required for sending, information from CH to BS

So CRCWSN algorithm can increase network lifetime through re-clustering and decrease energy consumed to network and improves leach and other previous methods [18]

IEACH-SWDN algorithm is one of another method for optimal energy consumed in wireless sensor networks which improves LTACH in which cluster heads are formal based on residual energy of dynamic notes. Which can decrease energy consumption and increase lifetime network Also there is movie, window in order to adapt selection possibility of CH and making the numbers sustainable with 1- information for initial energy from each nodes 2- information average energy of nodes which are never been CH that haven’ been in head cluster network at present.

\[ [0, E_{a,nch}/E_{i,max}, E_{a,nch}] \]

Also moving windows That its high extent produces random numbers C distance between [0,1] is changed dynamically. Also this algorithm includes two phases of: 1- set up 2- steady phase. In first phase, nodes that have the most residual energy, have the most probability for cluster heading, also we threshold by function predict)

In second phase, based on models of first phase and each model is divided in to a large number of gaps according to a number of each cluster nodes, the each node in gap sends packet to cluster head. Moreover Each node that has never been CH in current round send packet that includes residual energy in the last allocated gap to cluster head , the cluster head send it to BS and before starting of next round by BS, then it calculates average of energy nodes that have never CH [Eo-nch] and many send it to all networks. [19]

**Clustering with K-means algorithm**

K-means algorithm divides all information into subset (K) (cluster), so that all components of subset have the nearest distance with center of that subset. Moreover K-means appoint (impute sample) elects as centers of cluster randomly that we obtain points Gravity center of is consider as point K, then other objects (impute samples) relate to suitable cluster based on minimum Oghlidos distance with determined clusters centers, next average of each cluster is calculated again and considered as Cluster’s new center. So these operations are’ repeated a lot until clusters centers not to change at all. Criterion that must be minimized in K-means is that.

\[ E_{K-means} = \frac{1}{C} \sum_{i=1}^{C} \sum_{x_{ij}} ||x_i - C_i||^2 \quad (6) \]

in which C refers to number of clusters and OK (is) K cluster and \( C_i \) is cluster center of OK, while basis station Knows optimal number of clusters and member nodes of each cluster., so next phase is selection of suitable CH for each cluster and allocating of suitable roles to each node.

**Network model**

1- Network includes basis station (BS) and sensory node (SN) [20].
2- Sink node is in outside of monitoring area with unlimited energy.
3- Sensory nodes are constant after deploy mat in a special period of time. They also have calculation power and similar processing [21, 22].
4- Nodes have limited energy and after deploy mat remain unprotected, so there isn’t probability of battery recharging [23]
5- Network’s nodes don’t require GPS for a wiring of its condition. [24].

**Energy model**

Proposal algorithm’s energy model that has been proposed is the same as energy model in Heinzelman.

\[
\begin{align*}
E_{a}(k,d) &= k(E_{elec} + e_d d^2) \quad d < d_{cross} \\
E_{a}(k,d) &= k(E_{elec} + e_p d^4) \quad d \geq d_{cross} 
\end{align*}
\quad (7)
\]

Consumed energy for each transmission of K bit of data is calculated in distance in the following relationship:

\[ E_{R_t}(k) = k(E_{elec}) \quad (8) \]
and energy consumption for each receiving of data K bit is calculated in the following form: also Elect is transferring energy or electronic receiving and k is size of message based on number of bits and d is distance between receiver and transmitter, also $\rho_{m}$ and $\eta_{mp}$ is a required amount of energy for antenna components of transmitter for send of one bit, is kind of hidden modulation that are used from one of these two parameters, distance of transmitter to receiver. Also d is a threshold distance that determines amount of energy decrease crosser in Trans for and calculated its amount from following relation. [25].

$$d_{cross} = \sqrt{\frac{E_{mp}}{\rho_{m}}}$$

(9)

Also data accumulation energy of cluster heads is calculated in this form:

$$E_{ch} = 5nJ \text{ / bit } \text{ / msg}$$

(10)

**Explanation of proposed algorithm**

Proposed algorithm includes of two phases:

**Set up Phase**

In this phase, environment in which nodes have been distributed are Gridded and then in each cell we obtain its Grid’s center of Gravity also with K-means algorithm performance on it clusters are formed and with using of relation, (11) cluster heads are created. Although finalizing of K-means algorithm has been guaranteed, its final answer hasn’t been single and is optimal answerer.

As a whole, it has following problems that can eliminate them in this article. Final answer based on selection of initial cluster there isn’t a specific trend for initial calculation of clusters centers. If in third frequency of algorithm, number of data belonging to a cluster equals to zero, there isn’t any way for changing, and improving of method’s continuation.

Usually in many applications, a number of clusters aren’t definite. Since at first we have Gridded environment in which nodes have been distributed as a matter of fact, we can have the best selection for initial clusters and we can determine either initial center of each cluster or number of clusters from beginning with selection center of Gravity of each Grid.

**Determination of Gravity Center**

In order to determine Gravity Center inside each cell of Grids, we can use from Oghlidos Distance that is based on of distance of each node from (x,y) axis in environment.

$$\text{Center of gravity} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

(10)

**Center of Gravity**

Feature related to points inside each Grid’s cell. Then as explained above, form clusters with using of K-means algorithm. Phase for selection of head cluster in form clusters, there are considerable different parameters in election of head cluster [26], Sow has considered 3 criteria for selection of head cluster:

- Sensory that has the most energy level.
- 2- Sensory that has the least distance to basis station. (Less energy for transmission is consume)
- 3- Third sensory that has the least distance with normal nodes inside each cluster which cause decreasing of energy consumption for relations inside cluster.

Therefore Regarding mentioned criteria, having high energy level in increasing of network life time has an effective role, So, we use of relation (1) that has mentioned criteria for selection of the best head cluster.

$$\text{CH} = \text{REE}_{r} = \frac{\text{REE}_{r}}{\text{REE}_{c}}$$

(11)

We see that in above relation, CH =head cluster, Dist= distance between a node from another node. Nb = Number of bits, TPR = Transmission power, REE$_{vi}$ = residual energy extent of vi node.

In this relation, regarding residual energy extent of each node and (total) of our distance between a node from another node and transmission power also number of bits, we can select the best nodes head cluster.

**Steady Phase**

After formation of clusters and selection of head clusters of each cluster, now we must transmit sensory data by normal nodes to related head clusters. Clusters Head transmit packet data toward basis station after performance of accumulde functions or data combination energy consumed of all nodes is calculated.

When nodes are normal we use of (7) (8) formulas including transmission and receiving energy, and when node cluster. In addition to transmission and receiving energy, another energy is used for obtained data accumulation from normal nodes by head cluster’s node which is that is calculated in the following method as called as accumulation energy which NCH is the number of head clusters.

$$\text{Nch,}E_{ag} = 5nJ/\text{bit/mgs}$$

$$E_{add} = K \times E_{ag} \times Nch$$

(12)

Flow chart of proposed algorithm has mentioned in the following:
In this part, we present simulation results of algorithm of proposed new subject clustering that has been done using Matlab software. Also we will do analysis and comparison of its results with clustering algorithm including LEACH EEUC, CRCWSN, HEED, LEACH, and SWDN. Simulation of proposed algorithm in Matlab software has been done and its energy model is the same as Haizelman model. Initial energy of each node is consumption amount between \([2.0, 5.0]\) if in done simulations, amount of initial energy in algorithms has been, leach, leach-swdn=5j heed, eeuc=4j considered. But simulation results show that proposed algorithms have presented better results.

Simulation parameters are according to following table:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network size</td>
<td>100*100 m</td>
</tr>
<tr>
<td>Number of Nodes</td>
<td>100</td>
</tr>
<tr>
<td>Base station location</td>
<td>0,0 m</td>
</tr>
<tr>
<td>Initial energy for node</td>
<td>rand [0.2,0.5] J</td>
</tr>
<tr>
<td>(E_{elec})</td>
<td>50(\mu)J/bit</td>
</tr>
<tr>
<td>(E_{fs})</td>
<td>10(\mu)pJ/bit/ m(^2)</td>
</tr>
<tr>
<td>(E_{mp})</td>
<td>0.0013(p)J/bit/m(^4)</td>
</tr>
<tr>
<td>Data aggregation energy</td>
<td>5nJ/bit/signal</td>
</tr>
<tr>
<td>(D_{max})</td>
<td>87 m</td>
</tr>
<tr>
<td>Packet Size</td>
<td>1000 bit</td>
</tr>
<tr>
<td>Grid Number</td>
<td>9</td>
</tr>
</tbody>
</table>

In addition, in a network with dimensions 100&100 with 100 nodes that has been distributed in all network, so randomly following diagrams show amount of energy consumption in.

Presented and comparable algorithm as we can see from diagram, in condition of network with dimensions 100& 100 and number of 100 nodes, our presented algorithm had more life time. Although amount of initial energy of its nodes is much less than comparable algorithms.
In following picture’s simulation environment has been shown in which redlines, network Gridding and environment of simulated network is according to figure (1).

The simulated network environment is shown in figure 1. Black points are center of gravity, Blue points are Normal nodes and blue stars are CHs which can be in any point of cluster with this condition that they should be the most optimal node for election as CH based on the above mentioned formulas.

Therefore, we evaluate simulation results in different fields. Following diagram is output of IMKREC algorithm for each 100 node in 100& 100 environments and with B’s condition in 100&120.

So in this figure, also number of dead nodes in each round and diagram related total amount of dead nodes in each round and diagram for total amount of energy consumption in networks life time has been shown.

The following diagram (Figure2) shows if number of dead nodes in each round is less, life time of network will be more:

The following diagram (Figure3) shows if energy consumption is less, network’s life time will be more.

---

**Figure1:** Network environment with 100 nodes

**Figure2:** Life time, number of dead nodes in each round for each 100 node

**Figure3:** Total amount of energy consumption for each 100 Nodes
We evaluate the results of simulation in several aspects: Lifetime: lower energy consumption leads to increase network lifetime and monitoring time. As shown in figure 2, network lifetime is increased compared to similar algorithms. Figure 5 shows: CH: Electing the most suitable node as CH for sending data to BS is very important. Because in IMKREC, CH selection is based on center of gravity and residual energy in each node and node which has more energy and its distance with Gravity center is less will be better for send data to sink.

FND, LND

Time period that takes to remove (eliminate) the first node in network called as FND and time period that takes to remove (eliminate) the last node in network called as LND. These two factors effective effect in network efficiency. And if their amount is increased it will show improvement of presented algorithm, also they have been increased considerably in proposed algorithm in performed comparisons with the other algorithms.

Moreover FND and LND factors are two effective factors in network’s efficiency and increasing of the amount shows improvement for presented algorithm in compare of studied algorithms in this article’ these two factors have been increased well. Diagrams related to FND and LND is shown in figure 6. As shown in diagrams and numbers below, lifetime of the network in IMKREC has been increased.

The following diagram (Figure 7) shows amount of energy consumed in IMKREC algorithm and compared algorithms. As shown in the figure, in a network with size 100*100m IMKREC algorithm has provided a longer lifetime, while the amount of consumed energy is much less than compared algorithms.
Heydariyan et al., 2013

**Figure 7**: Total number of alive nodes in IMKREC and LEACH, LEACH-SWDN, CRCWSN, HEED, EEUC

Figure 8 shows the lifetime of the IMKREC algorithm in terms of different numbers of nodes from 100 to 800 nodes.

**Figure 8**: Lifetime of network in terms of 100 to 800 nodes in 100*100m environment

We have compared IMKREC with initial energy of 3j and HEED, HEEUC algorithms and the output of comparison is as follows:

**Figure 9**: Total number of alive nodes in IMKREC and HEED, HEEUC with initial energy of 3j

**CONCLUSION**

So far, we have been able to design a method for K-means algorithm improvement with selection of criteria such as network Gridding, determination as well center of gravity solve some of its problems partly. Also, we can select the most appropriate node as head cluster to decrease energy consumption very much and increase network lifetime

Therefore, with comparison of this algorithm with LEACH-LEACH-SWDN-CRCWSN, EEUC, HEED algorithms, we have shown that not only efficiency and network lifetime are increased very much in the proposed algorithm in comparison with comparable algorithms, but also the amount of energy consumption of the total network will be also decreased.

**REFERENCES**