AN INTEGRATED ALGORITHM TO ACHIEVE THE BEST CLUSTERHEADS IN WIRELESS SENSOR NETWORKS

M. Kazemzadeh[1] , A. Khatibi Bardsiri[2*]

[1] Computer Engineering Department, Islamic Azad University, Kerman Branch, Kerman, Iran
[2] Computer Engineering Department, Islamic Azad University, Bardsir Branch, Bardsir, Iran

ABSTRACT

Due to the increasing applications of wireless sensor networks (WSN) and the introduction of a new challenge in this regard, offering an algorithm is necessary to maintain and extend the life of network. To this end, offering a mechanism to determine the appropriate cluster head and timely switching is highly important. Since energy is an important challenge in WSNs, using clustering models can be considered as a solution. Selecting appropriate head cluster leads to considerably reduced energy consumption in these networks, leading to increased network lifetime. Different from previous studies, this article aims to employ a creative method and fuzzy logic to maintain and select the best head clusters. The simulation results showed that the proposed method was associated with condition improvement in most cases. This algorithm is also capable of mixing with all three algorithms including Low Energy Adaptive Clustering Hierarchy (LEACH), weighted energy efficient clustering (WEEC), and General Low Energy Adaptive Clustering Hierarchy (G-LEACH) which is considered as a noticeable advantage. Therefore, using this algorithm leads to delayed death of the first node. As a result, network energy is appropriately consumed. The number of network lost packets is far lower than other algorithms due to cluster head death prevention.

KEYWORDS: *Terms-wireless sensor networks; clustering; head cluster; fuzzy logic*

1.0 INTRODUCTION

Recent advances in the field of electronics and wireless communications have made the production of multi-function sensors possible with low energy consumption and cost. The sensors are able to communicate with each other over short distances. A node is a very small sensor with a sensing equipment, data processing, and wireless communication. A WSN, in fact, is a set of many sensor nodes which are spread in the environment. Each node autonomously pursues a specific objective in collaboration with other nodes. Nodes are close to each other and every node can communicate with another node in another group and give their information in order to report the environment condition to a central point. WSNs consist of many small sensors utilized to collect data and recognize events.

These networks also have some limitations. One of the main constraints is energy limitation with a noticeable effect on network lifetime. In order to reduce energy

[*] Corresponding Email: a.khatibi@srbiau.ac.ir

consumption and increase network lifetime, multiple methods are employed. Clustering is one of these methods (Akyildiz et al., 2002). There are two classifications of studies in this field: first, the studies which consider energy parameter as the head cluster (Al-Karaki & Kamal, 2014). This consideration is not sufficient for complex networks because a high energy node is located in a corner in which case all other nodes must spend a lot of energy to transmit information. The second group refers to those studies which consider more parameters such as distance and motion (Gautam et al., 2009, Gautam & Pyan, 2010, Han et al., 2014). This article works on three areas as follow:

- Current solution in selecting head cluster in WSN.
- Weak points of clustering methods in WSN.
- Proposed method to reduce energy consumption in clustering-based WSN.

The proposed method is a mixture of a creative method and fuzzy logic to maintain and select the best head cluster in network. In this method, remaining energy and distance from the base station are taken into account to select the next head cluster. The simulation results showed that the proposed method yields improvement over its counterparts. Another feature of the proposed algorithm is the capability of integrating with other algorithms including Low Energy Adaptive Clustering Hierarchy (LEACH), weighted energy efficient clustering (WEEC), and General Low Energy Adaptive Clustering Hierarchy (G-LEACH). This is considered as a noticeable advantage. Using this algorithm leads to delayed death of the first node ; consequently, the network energy is appropriately used. The number of dropped packets is far less than other algorithms due to the prevention of head cluster death. Multiple studies were conducted on the application of WSNs. As a result, many algorithms were introduced in the field of clustering and head cluster selection. Some of these algorithms are as follow: LEACH algorithm was introduced in 2000. LEACH is a centralized and self-organizing protocol with dynamic clustering. It uses a random method to distribute balanced energy consumption among nodes. In this method, nodes organize themselves into local groups. Then, one node takes the responsibility of local base station (cluster head). LEACH algorithm benefits from random rotation of high energy cluster heads in order not to immediately discharge a certain node. In this algorithm, each node selects a random number ranging between 0 to 1. If the random number is less than the threshold value, the node is chosen as cluster head. Threshold value is calculated through Equation (1):

$$T(n) = \begin{cases} \dfrac{P}{1 - P \times (r \bmod \frac{1}{p})}, & if\ n \in G \\ 0, & otherwise \end{cases} \tag{1}$$

P is the desired % of cluster heads, *r* is the current round, and *G* is the set of nodes that have not been clustering hierarchy (CH) in the last 1/*P* rounds. This algorithm, however, has some drawbacks:

1. Since the number of cluster heads is not constant in each round, this number might become more or less than the optimal number.
2. Cluster head might be located in the corner, so high energy needs to be consumed to communicate with this node.
3. Each node in each round needs to calculate a random number and a threshold value.

Centralized Low Energy Adaptive Clustering Hierarchy (C-LEACH) is one of the works in this field (Pinto et al., 2012). The work is centrally performed. In centralized method, all nodes send the information to one node. Based on the information taken from other nodes, this node will elect the cluster head. This will eventually lead to higher energy consumption. AHP is another related work in this field (Xu et al., 2014). In this algorithm, three parameters are taken into account including energy, motion, and distance of a node to cluster center. The base station makes the final decision regarding the CH election. Tang et al (2013) found that each node calculates its distance from neighbors. The central node is selected as CH. Such lection causes other nodes to use higher amount of energy to send the data to this node.

2.0 CLUSTERING MECHANISM IN THE PROPOSED NETWORK

In the proposed method, base station informs all network nodes by a message. After base station announcement, nodes are required to send a checking message to base station. This is mainly done to evaluate the number of present nodes in the network and distance. Nodes will estimate their distance from the sink through Equation (2).

$$\Delta_x = V.\Delta_t \tag{2}$$

where Δ_x is the distance between the n node to sink which is $D_{(n)}$ in fuzzy relations. V is a constant which is considered equal to the speed of light. The speed of signal equals the speed of light. Δ_t is, however, the difference between the time of sending and receiving a packet. Higher Δ_t shows increased distance between the sender and receiver. In the next step, since the base station is considered in the corner of the network and proposed network scale is 100 by 100 meters, this sink elects the closest distance from a candidate sensory node. Figure 1 shows the location of the elected node.

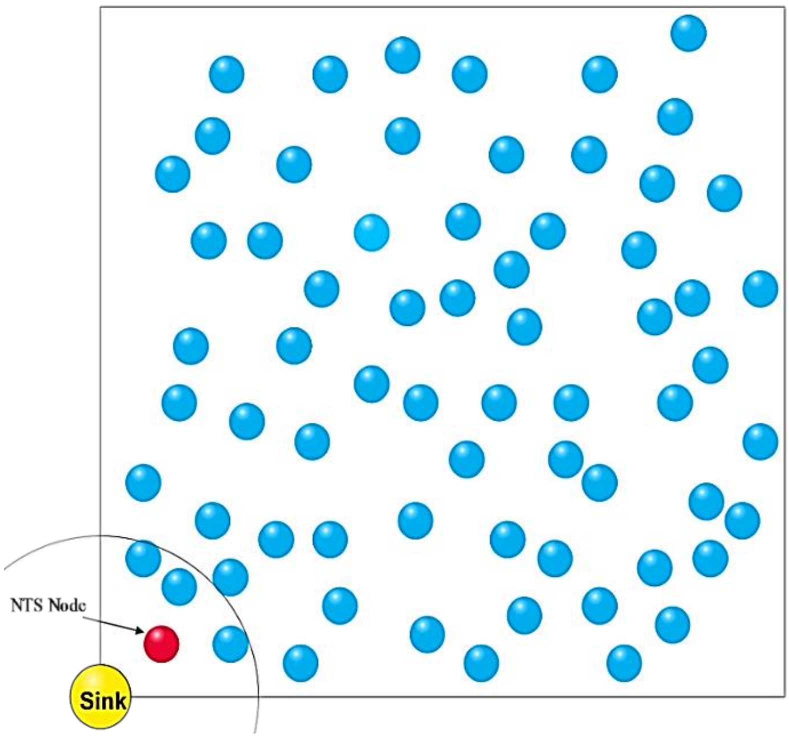

Figure 1. Location of the first elected sensory node as CH

Then, CH elects two other nodes within 25 % of the diameter of the network. Second elected CH and other nodes are proposed. Longer distance nodes are the objective of this algorithm which prevent the two nodes to be located next to each other. Figure 2 shows the location of the elected nodes.

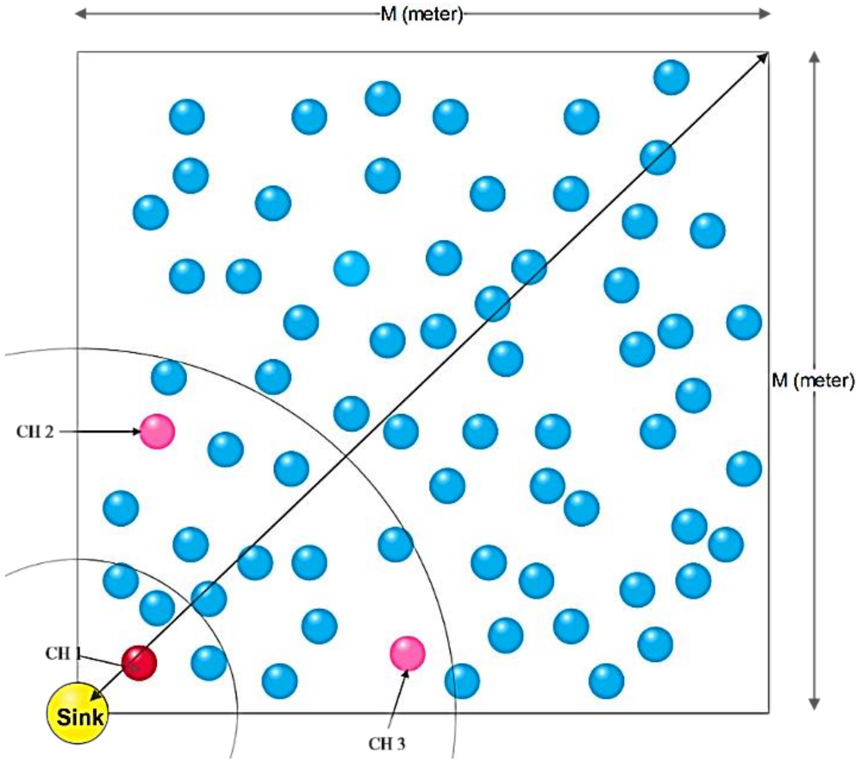

Figure 2. Location of second and third CHs

According to the following equation, a node is elected in the farthest distance from the base station. The mentioned node covers the area up to X from the sink. If the network side is M = 100, then:

$$D > (90\%)(\sqrt{M^2 + M^2}) \quad (3)$$

$$D > (90\%)(\sqrt{20000})$$

D>127.27 m

The fourth CH covers all nodes up to 56.568 meters from the sink. Figure 3 shows the numerical value in Equation (4), the location of the fourth CH, and edge detection.

$$D_{member} < (40\%)\left(\sqrt{M^2 + M^2}\right) \quad (4)$$

$$D_{member} < 56.568\ m$$

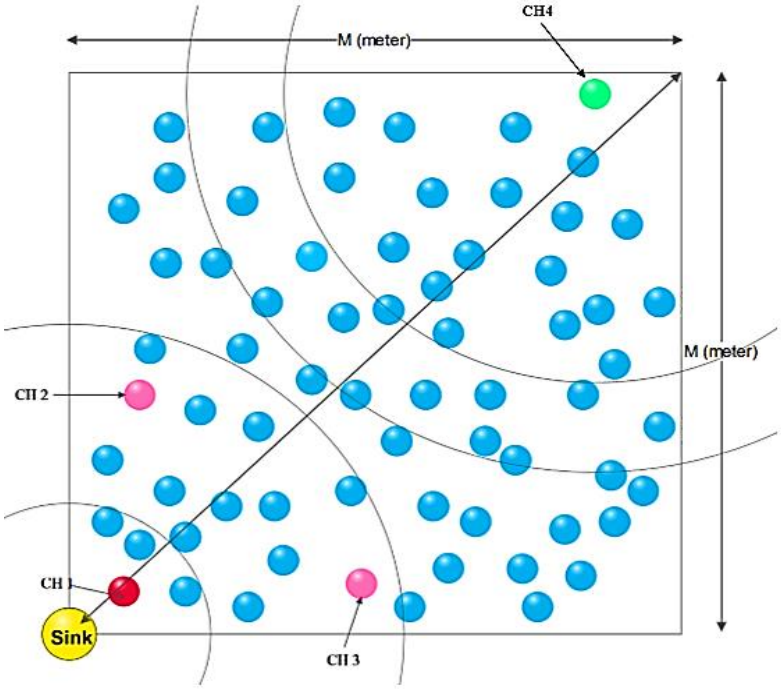

Figure 3. Location of fourth CH and edge determination

An area is selected called edge. In this area, two other nodes will be elected. This election is performed through Equation (5).

$$(50\%)\left(\sqrt{M^2 + M^2}\right) > D_{Border} > (40\%)\left(\sqrt{M^2 + M^2}\right) \quad (5)$$

Here, two other nodes will be elected as the fifth and sixth CHs. Figure 4 and 5 show the final network structure.

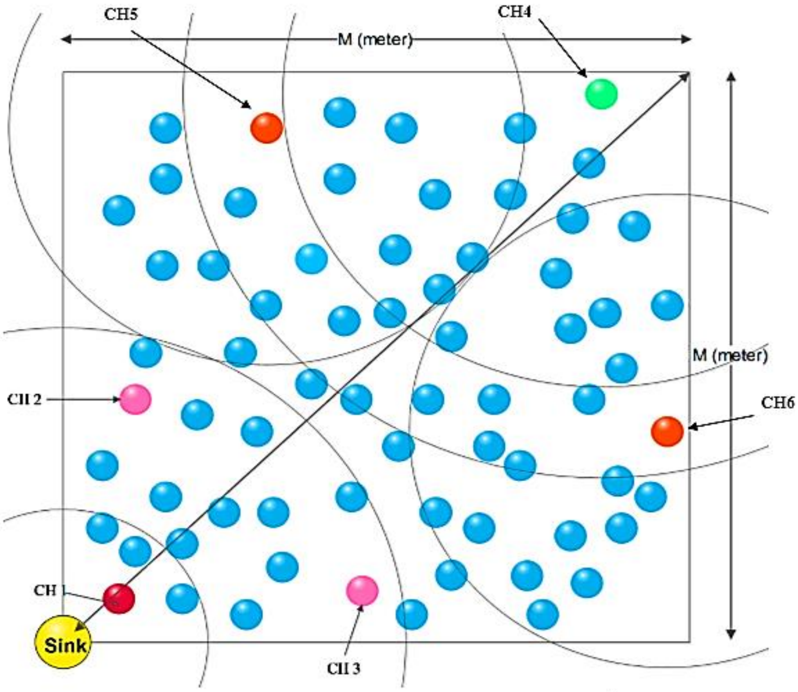

Figure 4. Fifth and sixth CH

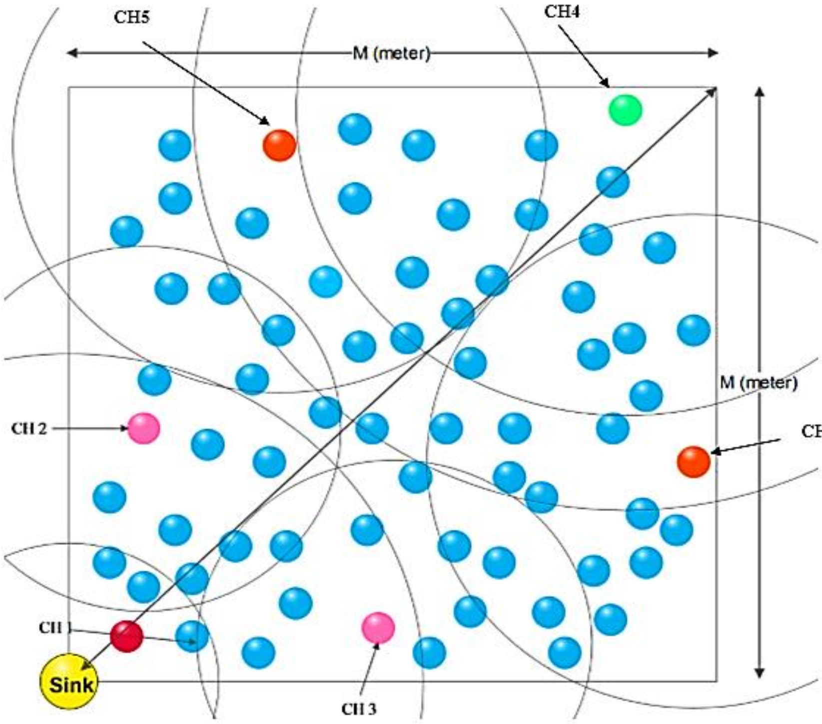

Figure 5. Location of all HCs and the zone of each

3.0 CH ELECTION MECHANISM (STABLE STATE)

Second phase: fuzzy clustering protocol implementation
In the second phase, internal round, the proposed idea is performed (Figure 6). The objective outlined here is to elect the appropriate CH with the highest remaining energy and minimum distance to base station node considered fuzzy system input:

1. Remaining energy of sensory node (first fuzzy system input)
2. Minimum distance to the sink (second fuzzy system input)

As a result, each node will yield a certain and clear condition in the unit of time consisting of two parameters: remaining energy and distance to the base station. These two inputs will change into fuzzy output considered node cost, $NC_{(n)}$. It is calculated by Equation (6):

$$NC(n) = \frac{\sum_{i=1}^{n} U_i \times C_i}{\sum_{i=1}^{n} U_i} \tag{6}$$

Finally, the highest $NC_{(n)}$ sensory node is elected as the best candidate node for HC in internal round. As a result, each CH will plan for the next round so that internal nodes in each node are required to report their remaining energy and distance to the base station to HC at the end of each internal round. CH node is shown in Table (1). In this cluster, CH node consists of 17 indicators with a set of $\{6,22,73,99,18,4\}$. The numbers in the table are samples. Therefore, the node with 73 and $NC_{(n)} = 0.48$ has the highest level of CH chance in the next internal round.

Table 1. A sample of calculation table in CH with sample numbers

Node ID	Remaining Energy(J)	Distance to BS(m)	$NC_{(n)}$
17	4.17	67	0.32
6	3.23	76	0.28
22	4.1	53.3	0.39
73	4.3	48.6	0.48
99	3.5	66.09	0.31
18	2.3	80.01	0.44
4	3.8	49	0.37

The proposed fuzzy system was designed as follows:

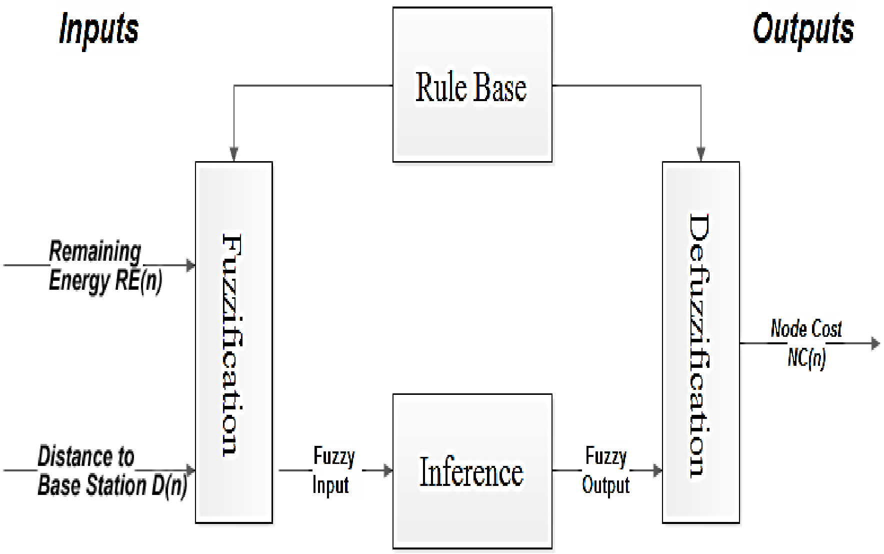

Figure 6. Proposed fuzzy system (two inputs including remaining energy and distance to base station and one output (node cost))

As can be seen in the fuzzy logic and triangular model, each input parameter consists of a triangulation chart. Based on the equal and clear triangulation, parameter behavior can be coordinated to x and y axes. Each point in x-axis has two corresponding points in y-axis. One of the charts is the network remaining energy. This energy can be located in five levels containing Very High, High, Medium, Low, and Very Low. The remaining energy of each sensory node can be in one of these levels or maximum of two levels in a row. Higher remaining energy will increase the chance of CH in the next round.

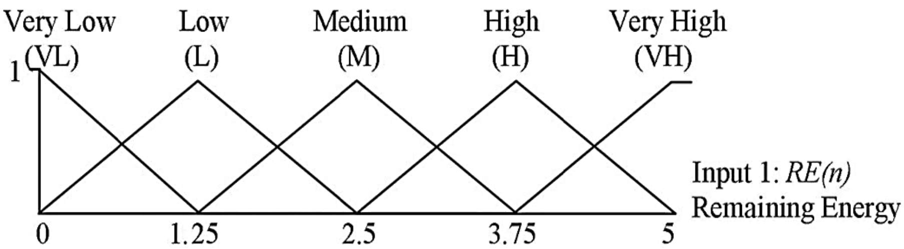

Figure 7. Remaining energy fuzzy diagram of node in different levels

The second fuzzy system input is the distance from the sensory node to base station. Based on the static feature of network nodes, this is always considered as a constant value. For example, the total sum of network sides is considered as the maximum distance. If the network is simulated in 100×100 m2 network, the maximum distance to base station can be the ambient diameter. $Max_{(D)} = \sqrt{a^2 + b^2}$ where a and b are network sides. In this network, $Max_{(D)}$ is reported 141.43. In this phase, a proportionality for distances was created in order to include the distances within 10 units in x-axis (Figure 8). Higher distance to the base station will create less priority for CH election in the next internal round.

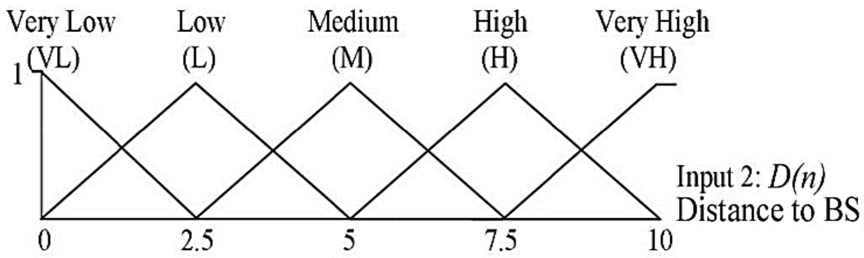

Figure 8. Fuzzy diagram of distance to base station in different levels

According to Equation (6), each fuzzy input is integrated in the form of fuzzy diagram and then each sensory node cost in [0, 1] interval was calculated. The higher cost will increase the HC chance. Finally, the output will be the $NC(n)$ function input. Fuzzy system acts based on Equation (7) to calculate $NC(n)$.

$$NC = \frac{\sum Rule_i \times C_i(NC)}{\sum Rule_i} \tag{7}$$

Based on fuzzy rules, each obtained $NC(n)$ value needs to be multiplied by cost function coefficient. This is done through Figure 9. For example, a node with High fuzzy rule has C_i= 0.75.

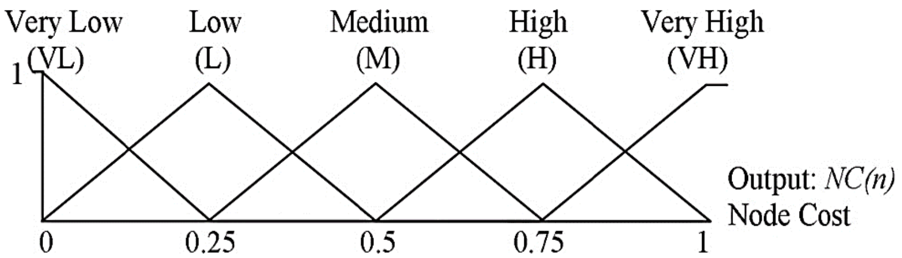

Figure 9. Fuzzy diagram to calculate each sensory node cost coefficient (the nose consists of remaining energy and distance)

For better understanding of above mentioned fuzzy calculation, an example is provided. Assume that A node has remaining energy of 3 J and proportional distance of 6. According to Figure 7, remaining energy is on two Medium and High triangles. It will also be in Low and Medium triangles for distance. For remaining energy of 3, there are two values on vertical axis. Multiplying Slope and x-axis value, axis values was obtained. Concerning distance, similar action was taken. Therefore,

$RE = 3,$ Medium & High $\Longrightarrow M = 0.4$, $H = 0.6$
$D = 6,$ Medium & High $\Longrightarrow M = 0.7$, $H = 0.3$

Now, for each fuzzy of remaining energy, two distance values were obtained. As a result, 22 states were achieved.

$M_{RE} \times M_D$	$M_{RE} \times H_D$	$H_{RE} \times M_D$	$H_{RE} \times H_D$

Each of above states shows a rule. *NC* is calculated based on Table 2 and 3. The rules are calculated based on Table 2 as follows:

Table 2. Rules for calculating *NC*

$\boldsymbol{Rule_{13}}$	$\boldsymbol{Rule_{14}}$	$\boldsymbol{Rule_{18}}$	$\boldsymbol{Rule_{19}}$
$M_{RE} \times M_D$	$M_{RE} \times H_D$	$H_{RE} \times M_D$	$H_{RE} \times H_D$
0.4×0.7	0.4×0.3	0.6×0.7	0.6×0.3
Medium	Low	High	Medium
$Ci_{(n)} = 0.5$	$Ci_{(n)} = 0.25$	$Ci_{(n)} = 0.75$	$Ci_{(n)} = 0.5$

These values will be on x axis from Figure 9 in order to obtain the corresponding C_i. Now, the following numerical condition in Equation (7) is:

$$NC = \frac{Rule_{13} \times C_{i_{(n)}} + Rule_{14} \times C_{i_{(n)}} + Rule_{18} \times C_{i_{(n)}} + Rule_{19} \times C_{i_{(n)}}}{Rule_{13} + Rule_{14} + Rule_{18} + Rule_{19}}$$

$$NC = \frac{[(0.4 \times 0.7)(0.5)] + [(0.4 \times 0.3)(0.25)] + [(0.6 \times 0.7)(0.75)] + [(0.6 \times 0.3)(0.5)]}{0.28 + 0.12 + 0.42 + 0.18}$$

$$NC = \frac{(0.14) + (0.03) + (0.315) + (0.09)}{1} = 0.575$$

Each node value in CH is calculated by two parameters including energy and distance. CH node elects the next node according to the maximum value of the next node. In fuzzy system, fuzzy rules are taken advantage as Table 3.

Above rules are IF-THEN, of which the relationship between fuzzy input variables and output variable are described using verbal variables by fuzzy sets and fuzzy operator. The table shows the rules with 52= 25 states and fuzzy rules. This is performed by AND rule. For example, if $RE_{(n)}$ equals Very High and $D_{(n)}$ equals Very Low, then $NC_{(n)}$ equals Very High.

Table 3. Rules for the proposed fuzzy system

No	Antecedent		Consequent
	Remaining Energy $RE_{(n)}$	**Distance to BS** $D_{(n)}$	**Node Cost** $NC_{(n)}$
1	Very Low	Very Low	Low
2	Very Low	Low	Very Low
3	Very Low	Medium	Very Low
4	Very Low	High	Very Low
5	Very Low	Very High	Very Low
6	Low	Very Low	Medium
7	Low	Low	Medium
8	Low	Medium	Low
9	Low	High	Low
10	Low	Very High	Very Low
11	Medium	Very Low	High
12	Medium	Low	Medium
13	Medium	Medium	Medium
14	Medium	High	Low
15	Medium	Very High	Low
16	High	Very Low	Very High
17	High	Low	High
18	High	Medium	High
19	High	High	Medium
20	High	Very High	Medium
21	Very High	Very Low	Very High
22	Very High	Low	Very High
23	Very High	Medium	Very High
24	Very High	High	High
25	Very High	Very High	High

4.0 THE DIAGRAM OF THE PROPOSED METHOD

Figure 10 shows the diagram of the proposed method.

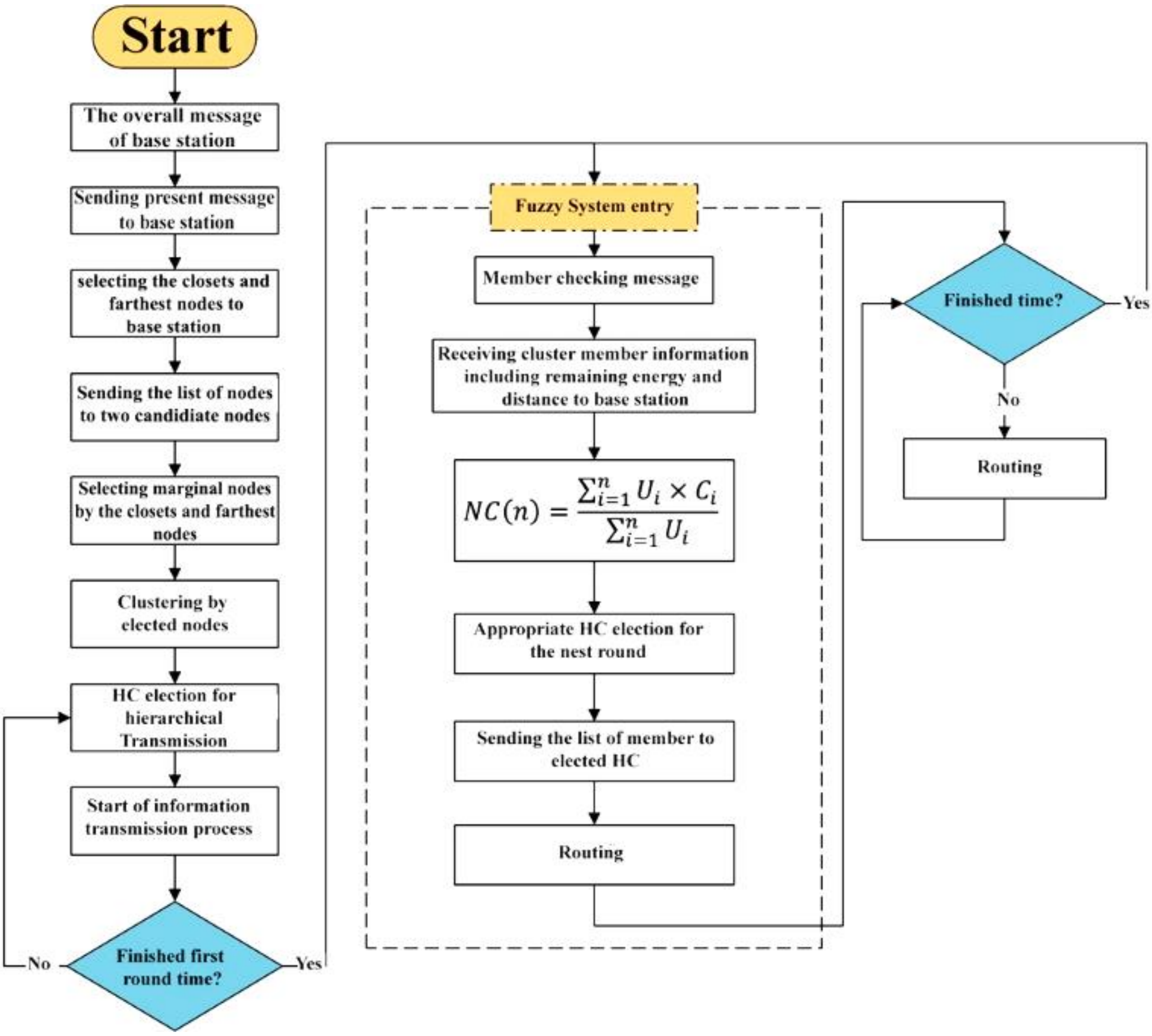

Figure 10. Proposed diagram

5.0 RESULTS AND DISCUSSION

5.1 Remaining Energy Test

One of the fundamental and debatable parameters in network efficiency is remaining energy in simulation interval and energy consumption slope in network. Figure 11 and 12 show network remaining energy after 100 seconds of simulation and energy consumption rate in 10-second intervals, respectively. As can be seen, in equal conditions of 100 nodes, energy consumption of the proposed protocol gives positive and acceptable results compared to LEACH, WEEC, and LEACH-G protocols. Network load distribution and clustering mechanism, as well as hierarchical routing lead to improved information packet transmission, information packet volumes, and, in turn, improved network general efficiency. The clustering protocol tries for balanced load distribution by electing appropriate number of cluster members known as CHEDP. Increased number of nodes in the proposed protocol leads to better display of its capabilities.

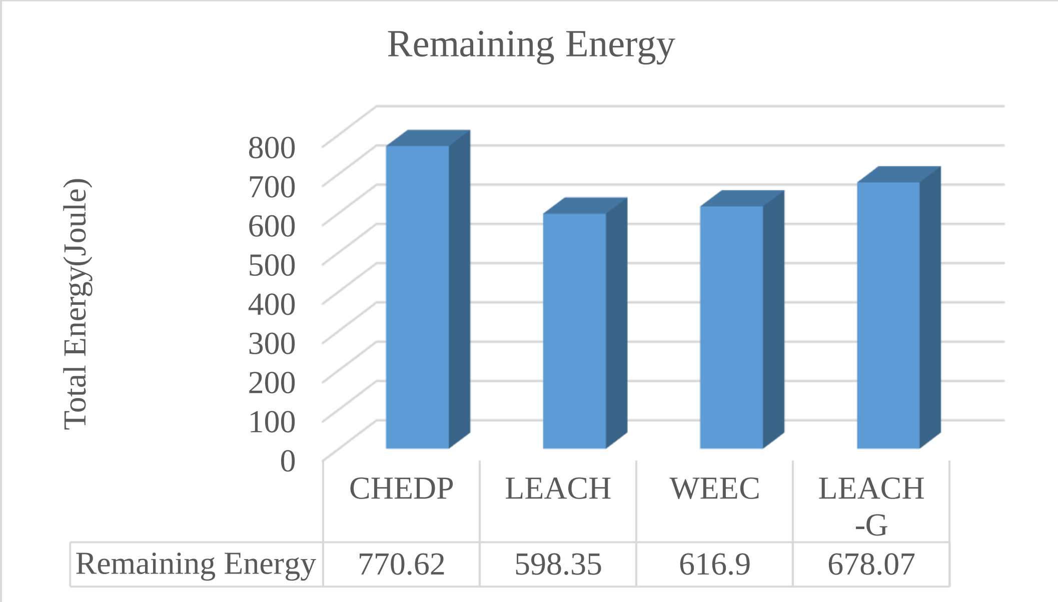

Figure 11. Network remaining energy after 100 seconds of simulation

5.2 Network Energy Consumption Rate Test

Energy consumption rate test shows the energy consumption management method. Low slope shows better management. Reduced external costs such as clustering repetition and repeated messages can lead to reduced energy consumption slope. As can be seen in Figure 11, total network energy is declined by only 1200 J at 100 seconds after simulation. Similar protocols, however, report higher values.

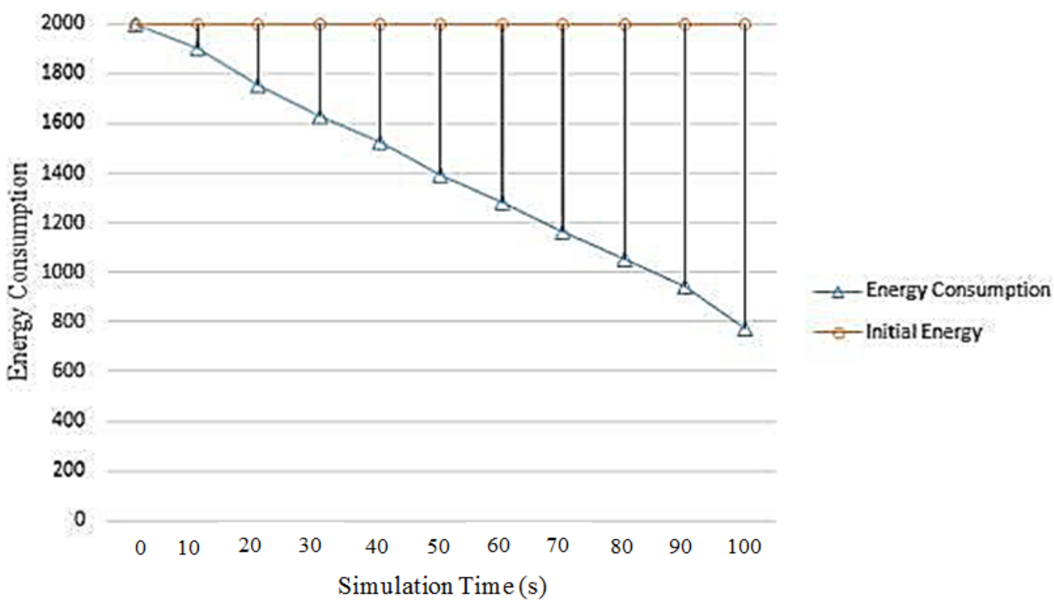

Figure 12. Network energy consumption after 100 seconds of simulation

5.3 Routing Packet Test

This parameter is utilized to evaluate the proposed protocol over the counterparts in routing packet rate and shows the number of sent and received packets. This is mainly done to conclude the extent of which this evaluation was able to healthily give produced packets to destination. Closer distance between sent and received packets means more optimal protocol. Figure 13 shows the number of routing packets in network after 100 seconds.

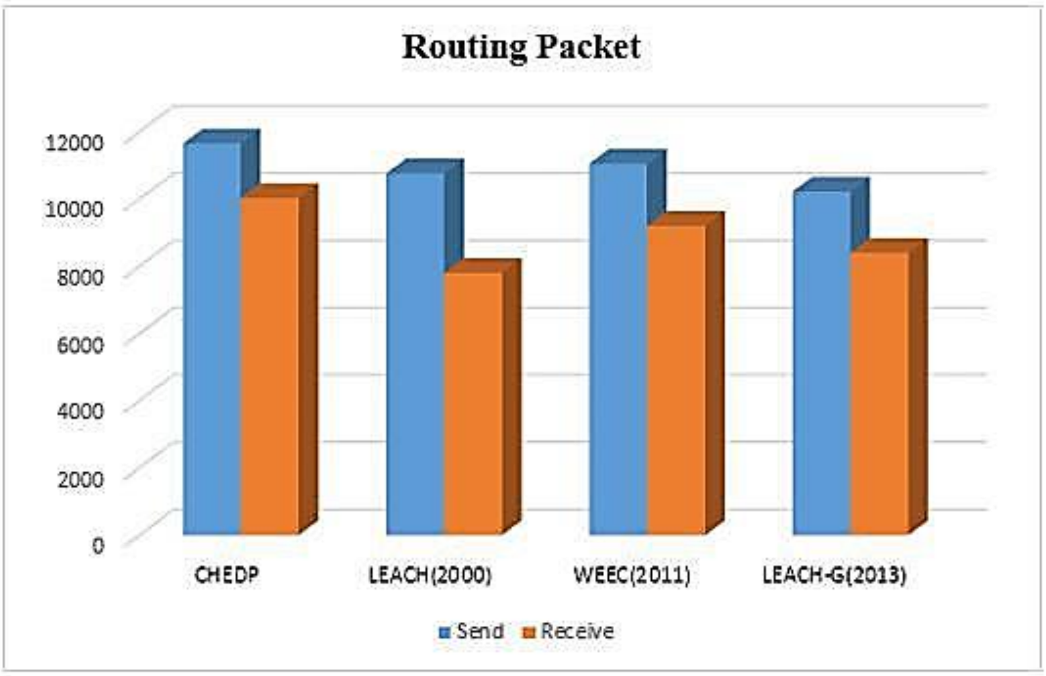

Figure 13. Routing packet in network after 100 seconds of simulation

5.4 Drop Packet Rate

Another parameter to evaluate network efficiency is the number of drop packets. Lower number of dropped packets shows better performance. Figure 14 shows the number of dropped packet after 100 seconds.

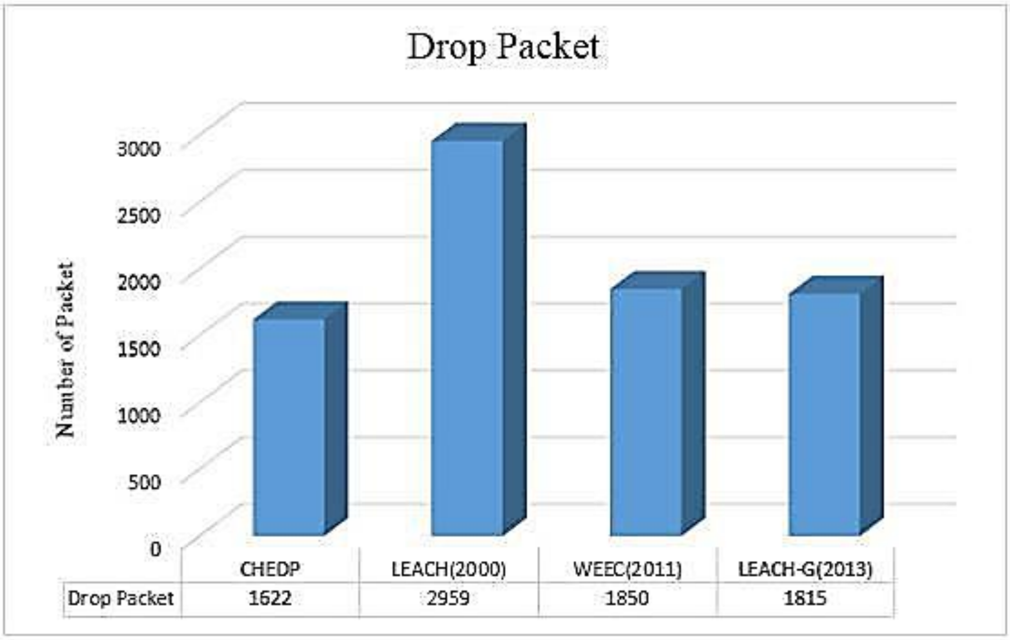

Figure 14. Drop packet routing after 100 seconds of simulation

5.5 Packet Delivery Ratio

It is calculated by Equation (8) :

$$PDR = \frac{Data\ Recieved}{Data\ Sent} \times 100 \quad (8)$$

In other words, received-to-sent data ratio is calculated by percentage. Higher percentage shows better network performance. Figure 15 shows packet delivery ratio after 100 seconds.

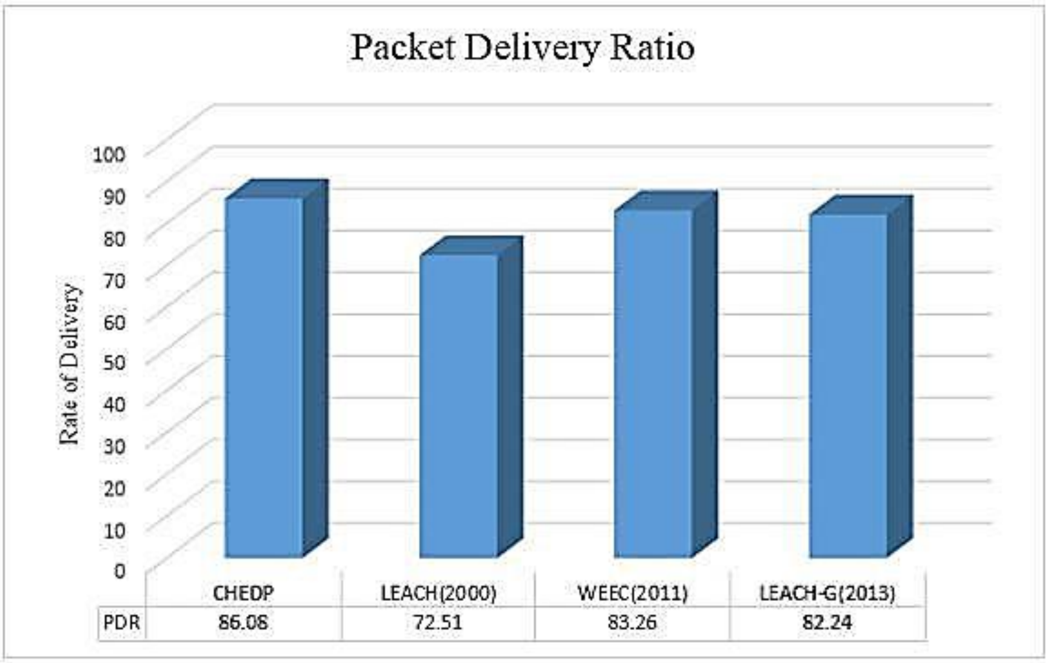

Figure 15. Packet delivery ratio after 100 seconds

5.6 Normalized Routing Load Test

Another parameter to evaluate simulation results is Normalized Routing Load. Lower Normalized Routing Load shows faster routing protocol loading. The simulation results show that the proposed protocol loads by double over the base protocol, leading to faster routing. This parameter is important and effective in network routing calculated by Equation (9):

$$\mathrm{NRL} = \frac{\text{Sent Data+Forward Packet}}{\text{Recieved Data}} \times 100 \tag{9}$$

In other words, received-to-sent data ratio is calculated (%). Lower ratio shows better performance. Figure 16 shows Normalized Routing Load rate after 100 seconds of simulation.

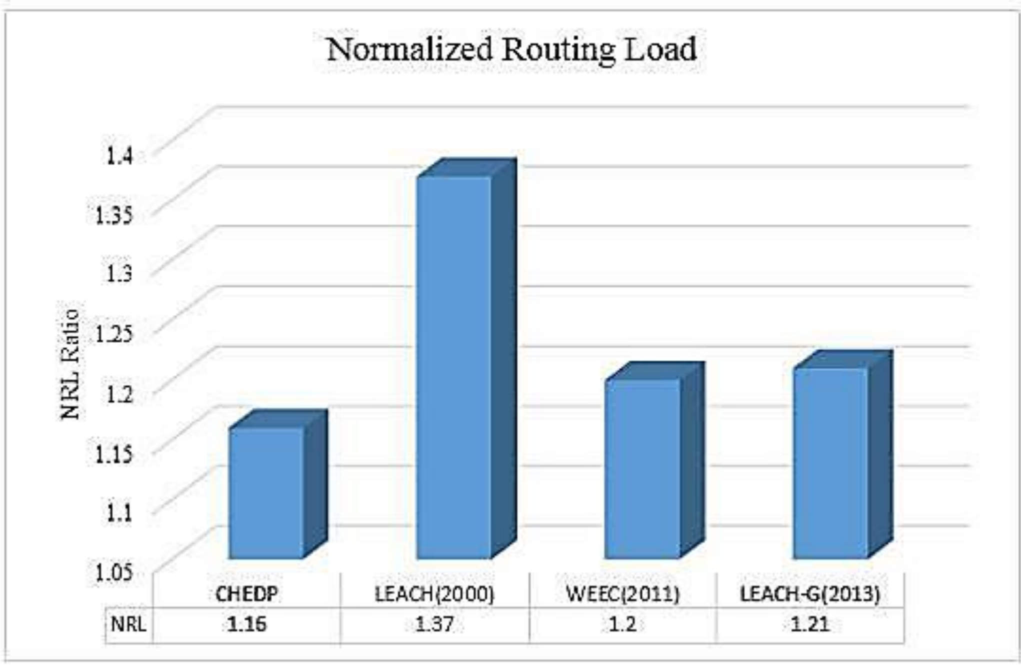

Figure 16. Normalized Routing Load after 100 seconds of simulation

5.7 First Node Death Test

Figure 17 shows the results of another test to evaluate the time of death for the first and last sensory node in four protocols including CHEDP, LEACH, WEEC, and LEACH-G in equal conditions with 100 normal nodes at time of 100 seconds. It shows longer lifetime of sensory nodes in the proposed protocol. The death of the first and last node is one of the most important parameters to evaluate network lifetime. The proposed protocol gives highly successful results. At time of 100 seconds, the proposed protocol did not have any nodes facing death. Therefore, the time was changed to 200 seconds. At t=140 s, the first node (finished energy) was lost.

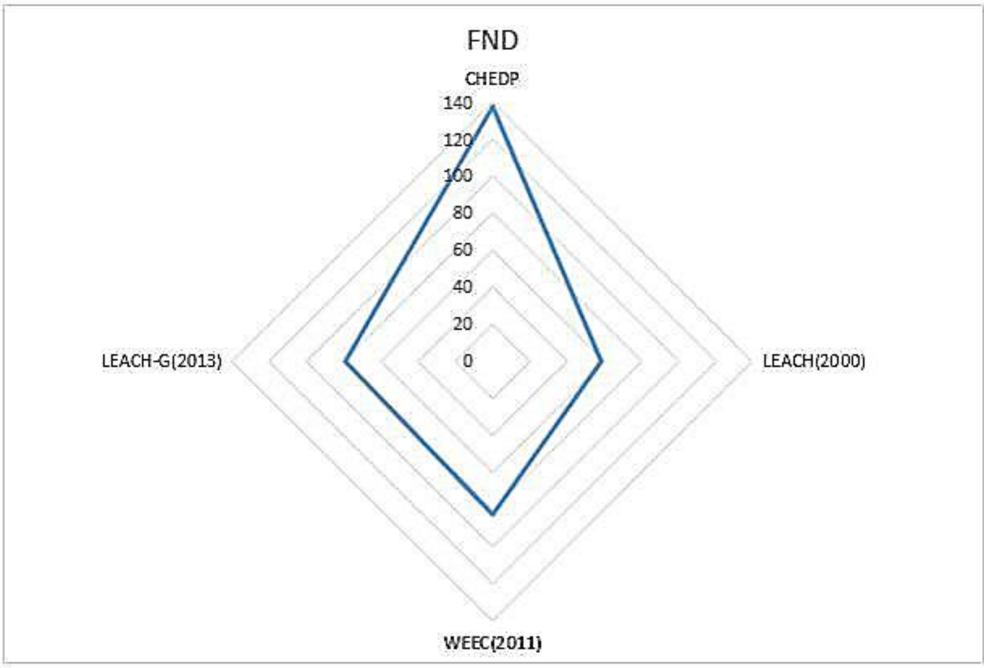

Figure 17. Time of the first node death after 100 seconds of simulation

6.0 CONCLUSION

The wireless sensor networks offer many benefits despite having paramount challenges. One of these challenges is the energy consumption of network nodes. In the process of clustering it tries to use the appropriate slope to reduce energy consumption. This paper presents an innovative protocol that combines an innovative way and fuzzy logic for maintaining and selecting the best head in the action network introduced. The proposed method sets the criteria for the residual energy parameters and the distance of a cluster head node selection to be used for future research. The simulation results show that the proposed method constitutes the main improvement in the network protocols. The proposed method has a major advantage of which it is able to combine all the three algorithms: LEACH, WEEC and LEACH-G. Therefore, using this algorithm postpones death of the first network node and the required energy is used appropriately. It also prevents death due to missing batch number of nodes in the networks which are far lower than other protocols.

ACKNOWLEDGEMENT

The authors thank the Islamic Azad University for supporting the reseaarch.

REFERENCES

Akyildiz, I. F., Su, W., Sankarasubramaniam, Y. & Cayirci, E. (2002). Wireless Sensor Networks :A Survey. *Computer Networks*, 38(4): 393-422.

Al-Karaki, J. N. & Kamal, A.E. (2004). Routing Techniques in Wireless Sensor Networks: A Survey. *IEEE Wireless Communications*, 11(6): 6-28.

Gautam, N., Lee, W. I. & Pyun, J.Y. (2009). *Track-Sector Clustering for Energy Efficient Routing in Wireless Sensor Networks*. 9th IEEE International Conference on Computer and Information Technology, 2:116-121.

Gautam, N. & Pyun J.Y. (2010). Distance Aware Intelligent Clustering Protocol for Wireless Sensor Networks. *Journal of Communication and Networks*, 12(2): 122-129.

Han, Z., Wu J., Zhang, J., Liu, L & Tian, K. (2014). A General Self-Organized Tree-Based Energy-Balance Routing Protocol for Wireless Sensor Network. *IEEE Transactions on Nuclear Sciences*, 61(2): 732-740.

Pinto, A. R., Poehls, L. B., Montez, C., & Vargas, F. (2012). *Power Optimization for Wireless Sensor Networks*. Intech Open Science.

Xu, Z., Yin, Y., Wang, J., & Kim, J. U. (2014). *A Density-based Energy-efficient Clustering Heterogeneous Algorithm for Wireless Sensor Networks*. 4th International Workshop IEEE, 368-372.

Tang, B., Wang, D. & Zhang, H. (2013). *A Centralized Clustering Geographic Energy Aware Routing for Wireless Sensor Networks*. International Conference on Systems, Man and Cybernetics, 1-6.