Routing Protocol LEACH-K Using K-Means Algorithm in Wireless Sensor Network

Chapter · March 2020

DOI: 10.1007/978-3-030-44038-1_27

3 authors, including:

Rahma Gantassi
Chonnam National University (CNU), South Korea.

Bechir Bengouissem
École Nationale d'Ingénieurs de Tunis

Routing Protocol LEACH-K Using K-Means Algorithm in Wireless Sensor Network

Rahma Gantassi[1(✉)], Bechir Ben Gouissem[1], and Jalel Ben Othmen[2]

[1] Communication System Laboratory Sys'Com (ENIT), University of Tunis El Manar (UTM), Tunis, Tunisia
rahma1231@gmail.com, bechir.gouissem@enit.rnu.tn
[2] Centrale Supeec University of Paris 13, Paris, France
jalel.benothman@l2s.centralesupelec.fr

Abstract. Nowadays, the tendency in WSN research is the use of machine learning to improve efficiency in energy conservation. The purpose of the research is to reduce energy consumption to boosts network lifetime. To minimize nodes energy consumption, researchers have taken interest in the use of K-Means especially in large scale networks for applications where the controlled area extensive. This research did not pay much attention to the impact of K-Means on network performance and quality of service metrics such as throughput, Energy, latency, etc. In our case, we applied K-Means algorithm on LEACH routing protocol before CH election in order to minimize energy consumption. In the present work we applied K-Means before the selection of CH and study the impacts of K-Means on the several the quality of service criteria. Hence the use of K-Means before the election of CH, which divides the network into K cluster where all the nodes of each cluster are very close to the centroid location, which makes nodes closer to the CH. Therefore, our work has reduced energy consumption and latency time, increase network stability time, network life time and throughput.

1 Introduction

Implementation of WSNs has become a hot area of research in recent years due to the vast potential of sensor networks to enable applications that connect the physical world to the virtual world. According to the topology of network, wireless sensor networks routing protocols can be divided into flat and hierarchical routing protocols. In wireless sensor networks, the tiny, low cost and low power sensors are able to communicate within a short range and work together to form a sensor network for gathering data from a field. These sensors have data processing and communication capabilities. They also have enabled us to monitor and collect data in any environment. According to the topology of network, wireless sensor networks routing protocols can be divided into flat and hierarchical routing protocol where LEACH (Low Energy Adaptive Clustering

L. Barolli et al. (Eds.): WAINA 2020, AISC 1150, pp. 299–309, 2020.
https://doi.org/10.1007/978-3-030-44038-1_27

Hierarchy) is the simplest hierarchical protocol that possesses clustering approach [1]. LEACH is a protocol that tends to reduce energy consumption in a wireless sensor network. It combines the efficiency of energy consumption and the quality of access to the media, and is based on the division into groups, with a view to enable the use of the concept of data aggregation for better performance in terms of lifetime [2].

To provide energy efficiency and prolong network life time in wireless sensor networks, many of the researchers are heading to Machine Learning to improve LEACH efficiency. Some researchers use the genetic algorithm and others use neural networks, SVM, K-Means, etc. Many researchers integrated K-Means on LEACH in order to minimize energy consumption and prioritize the election of CH over the cluster function of K-Means. We have noticed that if we apply K-Means on the initial network of K-Means, our network will be divided into K groups where the nodes of each group are placed very close to the centroid. Hence, we find that all network nodes are close to the centroid of the initial network. Since the sink is in the center of the area, we find, therefore, that the distance between nodes and the cluster head is tiny and the distance between each cluster head and the sink is equally very small. Consequently, we minimize energy consumption. Besides, latency time of each packet will be reduced, which leads to the increase of the throughput, and the reduction of energy consumption is the outcome of the increase of the stability and life time of network.

This paper is divided as follows: Sect. 2 describes related works. Section 3 explains our innovative idea (LEACH-K). Section 4 talks about the simulation results and analysis and finally in Sect. 5 we conclude our work.

2 Related Work

Much research has been done to measure the effect of K-Means on wireless sensor networks performance. Some works have only investigated the impact of K-Means algorithm on network lifetime and energy consumption.

As in [3], Gayatri Devi et al., used K-Means clustering algorithm in WSN. Among its advantages is the reduction of energy consumption of WSN as well as increasing the number of clusters. However, the weakness of this work lies in the use of small parameters. For instance, in the simulation stage, he worked on 2 clusters and 10 nodes. In another simulation test, he worked on 100 nodes.

In [4], the researchers worked on the performance analysis of LEACH with machine learning algorithms in WSN. The researcher applied and studied K-Means algorithm and Gauss algorithm in LEACH. This DBI-based K-Means work offered simpler parallel-distributed computing of K and allowed faster auto-classification of sensor nodes. In this manner, he created a balanced energy of clusters and guaranteed distributed energy consumption. Therefore, the outcome of this idea reduced the computation time thanks to K-Means and minimized energy consumption thanks to Gauss. He compared this work with LEACH and LEACH-C where the objective was to check the performance of energy consumption. Among the weaknesses of this work is that it was tested on a 100-node network in a time duration of 200 rounds.

Sukhchandan Randhawa et al. was implemented in [5] both centralized and distributed K-Means clustering algorithms in network simulator, where they alternated between two major steps in K-Means clustering algorithm, assigning observations to clusters and computing cluster centers so that a criterion is reached.

Whereas, in our work we studied network performance in several quality of service metrics in addition to energy consumption as many wireless sensor networks applications are with QoS constraints.

3 Protocol Low Energy Adaptive Clustering Hierarchy-KMeans Approach (LEACH-K)

LEACH-K (Low Energy Adaptive Clustering Hierarchy-KMeans approach) is executed in “rounds” which are predetermined time intervals. Each round consists of two phases: an initialization phase and a transmission phase. Our intervention targets the initialization phase in order to minimize the dissipated energy during the initialization phase. This protocol is an extension of LEACH which aims at increasing the network lifetime and improving the QOS (Quality of Service) metrics. We utilize the proposed K-Means-based clustering method to provide an improved version of the LEACH protocol before the election of Cluster heads. We assume that sensor nodes are distributed randomly.

So, after sink gets the position of all nodes, it can create the required number of clusters by using K-Means clustering algorithm based on the number of network nodes. Then, it broadcasts nodes list to all sensors of the network. After this, the sensor nodes declared as Cluster Heads, begin to operate as Cluster Heads and broadcast their presence on the network.

In LEACH-K before applying the first phase (initialization phase) of the LEACH protocol, we use the K-Means function [6].

1. Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
2. Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the K centroids.
4. Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

The rest of the proposed protocol is the same as LEACH, where nodes of each K-Means group join the closest CH, and CH broadcasts the TDMA scheduling for nodes. By using clustering algorithm, symmetric clusters can be approximately created.

Our proposed solution is implemented in MATLAB to be evaluated and compared with references routing protocol LEACH and TEEN. We choose MATLAB because of its ease of interface and availability of pre-programmed necessary functions. This section provides the extensive simulation results of the LEACH-K, LEACH protocol and TEEN in the MATLAB software. In these simulation scenarios, we evaluate the following parameters in the LEACH, TEEN protocol and our proposed solution.

4 Simulation Results

Our proposed solution is implemented in MATLAB to be evaluated and compared with references routing protocol LEACH and TEEN. We choosed MATLAB because of its ease of interface and availability of pre-programmed necessary functions. This section provides the extensive simulation results of the LEACH-K, LEACH protocol and TEEN in the MATLAB software. In these simulation scenarios, we evaluate the following parameters in the LEACH, TEEN protocol and our proposed solution (Table 1):

In these simulation scenarios, nodes are randomly distributed in the WSN and have 0.5 J battery power. We evaluate the following parameters in the LEACH, TEEN protocol and our proposed protocol:

Table 1. Simulation parameters

Settings	Values
Initial energy	Eelec = 0.5 J/noeud
Electronic dissipation energy (sending, receiving)	Eelec = 50 nJ/bit
Data aggregation energy	EDA = 5 nJ/bit/m2
Transmit amplifier if dtoBS $\leq$ d0	Efs = 10 pJ/bit/m2
Transmit amplifier if dtoBS $\geq$ d0	Emp = 0.0013 pJ/bit/m4
Simulation area	(100 * 100) m
Sink location	(0.5, 0.5)
Number of nodes	100 nodes
EDA	0.00130 * 00000000001 J

K: Number of initial clusters of K-Means in each round
N: Initial nodes number in each round

The simulation scenario is divided into three scenarios:

The first scenario: we simulate LEACH-K protocol in different k values and we compare it in every K case with LEACH, TEEN protocol until obtaining the optimal value of K.

The second scenario: we study our proposed protocol, apply it in the optimal case of k and vary, every time, the number of nodes until we reach 1000 nodes.

Scenario 3: Study of the QOS criteria for our idea in the case of the optimal value of K when increasing the number of sensor nodes in the network to evaluate the node density impacts in different QoS parameters.

4.1 First Scenarios

Figure 1 shows the energy in each round of LEACH, TEEN and LEACH-K protocols in the first scenario in which we have utilized the proposed K-Means that is based on the clustering method before the Cluster heads election for 3000 rounds and 100 nodes.

As indicated in this figure, our solution reduces sensor nodes energy consumption, which increases network lifetime. Also, our solution increases the throughput.

Figures 1(a) and (b) show the variation of the residual energy and throughput for K = 2. The graphs illustrated in Figs. 1(a) and (b) depict that by the integration of an unsupervised algorithm as K-Means before Cluster Heads election permit the increase of total residual energy and throughput. The sensor network can be coated with an acceptable amount of energy (more than 41,507 J in comparison to 0 of LEACH and 0.013022 of TEEN during 3000 rounds) and provides a considerable throughput (more than 30008 packets in comparison to 2009 for LEACH and 2001 for TEEN during 3000 rounds).

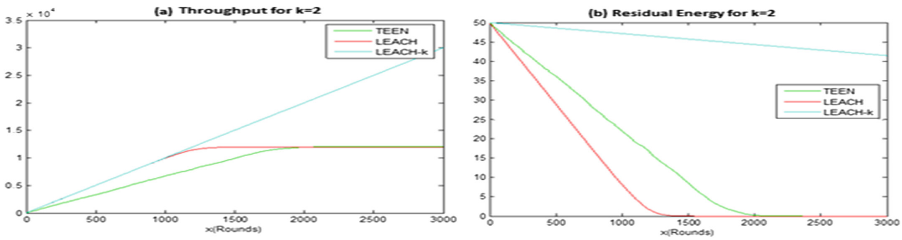

Fig. 1. Comparison of LEACH, TEEN and LEACH-K with a small variation of K.

The use of K-Means before the election of CHs for LEACH increases the number of clusters which implies the increase of CHs in the network and consequently the number of packets transmitted to the base station increases, signifying the increase of throughput.

Figures 2(a), (b), below show respectively the variation of the total consumed energy, throughput, according to variation cluster numbers K within a duration of 3000 rounds in order to determinate the optimal K value.

The histograms illustrated in Figs. 2(a) and (b) summarize the value of residual energy and throughput according to the variation of K value for 100 sensor nodes. For a network of 100 nodes when increasing the number of K automatically, the number of CHs and the number of packets transmitted to the SINK increases and thereafter the throughput increases, too. Otherwise, the purpose of K-Means is to group the location

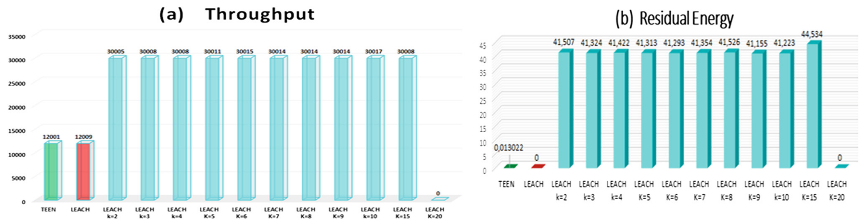

Fig. 2. Percentage of network throughput gain rate and residuals Energies with variation of K

of nodes around the centroid and subsequently the distance between the cluster heads and the base station decreases since the SINK is in the middle of the network, which reduces the loss of energy. This result explains the increase in residual energy.

We notice that for K = 20, the throughput and the residual energy rate drop to zero. This is logical since when we use K-Means before the election of CH, the network will be divided into clusters. Each cluster of K-Means will be divided into other clusters according to LEACH. In the above scenario, when K = 20, the number of clusters becomes very high, which makes the number of clusters exceed 5%. According to formula and the parameters of LEACH [1] the optimal probability for nodes to become cluster heads does not exceed 5% because CH nodes consume a lot of energy.

This increases energy consumption, which ultimately reduces network lifetime to zero. Table 2 summarizes the Throughput, Residual Energy, Stability, Latency time and the number of CHs according to the variation of the K (number of K-Means cluster).

Table 2. Comparison of the network lifetime, the stability, throughput, latency time and number of CHs between the LEACH, TENN and LEACH-K with variation of K

	Throughput	Residuals energy	Latency time (ms)	Stability (rounds)	Number of CHs
LEACH	12009	0	194.865	799	18202
TEEN	12001	0.013022	153.74	1078	11949
K = 2	30005	41.507	189.66	1311	19685
K = 3	30008	41.324	118.8765	1468	19230
K = 4	30008	41.424	139.159	1311	19052
K = 5	30015	41.317	143.724	1345	19647
K = 6	30015	41.213	100.325	1384	18990
K = 7	30012	41.354	99.585	1354	18780
K = 8	30014	41.526	101.86	1257	18657
K = 9	30005	41.155	98.618	1317	18756
K = 10	30017	41.497	96.1602	1399	19543
K = 20	0	0	0	0	0

The results described in Table 2 depict that when we apply our idea, the latency time decreases while the throughput, residual energy, stability as well as the number of cluster head increase in comparison to the value yielded by LEACH and TEEN. The obtained values conserve their increase together with the increase of K yet with different values.

4.2 Second Scenario

In this scenario: we study our proposed protocol in the case of prédetermined optimal K and vary, every time, the number of nodes until we reach 1000 nodes.

we have chosen two simulation case to show the effect of the number of nodes on our protocol which are illustrated in Figs. 3(a) and (b) below and which show respectively the variation of the total consumed energy, throughput for optimal value of K, K = 10 and, N = 50. The chart illustrated in Figs. 3(a) and (b) depict that by increasing the number of nods, total residual energy increases.

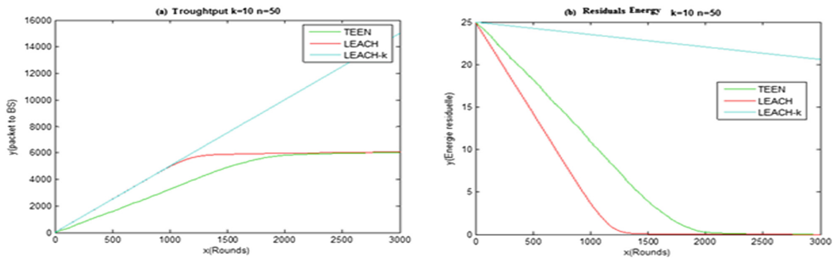

Fig. 3. Comparison of LEACH, TENN and LEACH-K with a small variation of K and N

Figures 4(a), (b), below show respectively the effect of the variation of the number of nodes on total residuals energy and throughput with optimal value of K, K = 10. The histograms illustrated in Figs. 4(a) and (b) depict that by increasing the number of nodes, residual energy and throughput increase for every value of N because the

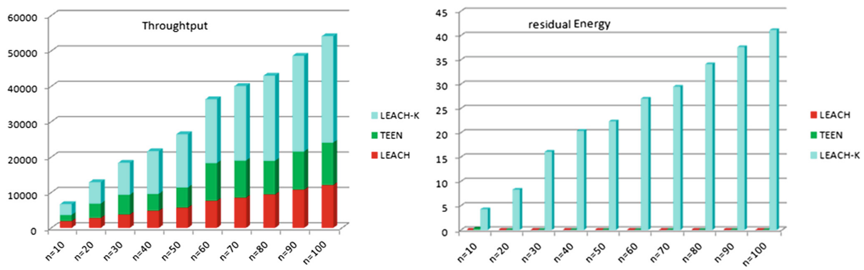

Fig. 4. Percentage of network throughput gain rate and residual Energies with variation of N when k = 10

number of packets received by the sink (throughput) and the energy consumed by the network depend essentially on the number of nodes deployed in the network.

They depend especially on the number of cluster initials for the network configuration described above, a better tradeoff is found for all K values, between residual energy and quality of service metrics with 3000 rounds.

This dependence of the initial number of clusters for the network configuration described above is illustrated as a better tradeoff that is found for all K values between residual energy and quality of service metrics with 3000 rounds. However, energy falls to zero at 2500 round for all the values of N for LEACH and TEEN, whereas for LEACH-K the energy increases with the increase of N.

Table 3 summarizes the throughput, residual energies, Stability and latency time depending on the variation of the N and optimal value of K, K = 10.

The results described in Table 3 depict that when we apply our idea, the latency time decreases while the throughput, residual energy, stability as well as the number of cluster head increase in comparison to the value yielded by LEACH and TEEN. The obtained values conserve their increase together with the increase of K yet with different values.

4.3 Scenario 3

Scenario 3: Study of the QOS criteria for our idea in the case of the optimal value of K when increasing the number of sensor nodes in the network

As indicated in these histograms, the solution we provide improves and increases the throughput. Also, it improves the residual energy of the sensor nodes and can decrease the energy consumption of the sensor nodes, throughput and the simulation, which will result in more network lifetime.

Figure 5 illustrates the Throughput and residual Energies depending on the variation of the N (number of initial nodes with n = 100 to n = 1000) when K = 10 (The optimal value of K) the number of rounds = 10000.

The above figures show respectively the variation of the residual energy and throughput of K = 10 according to the variation of the number of nodes up to 1000 nodes. We notice that if we increase the number of nodes, the residual energy increases well as the throughput.

The results of this histogram confirm the previous results when we choose the optimum value for K = 10 and we test the effect of the increase of the number of nodes and the number of rounds. We note that LEACH-K gives good results for QOS during 1000 rounds (Fig. 5).

Table 3. Comparison of the network lifetime (Energy residual), the stability, throughput, llatency time and number of CHs between the LEACH, TENN, and LEACH-k with variation of N (number of nodes).

	Throughtput			Residuals energy			Latency time (ms)			Stability (rounds)			Number of CHs		
	LEACH	TEEN	LEACH-K	LEACH	TEEN	LEACH-K	LEACH	TEEN	LEACH-K	LEACH	TEEN	LEACH-K	LEACH	TEEN	LEACH-K
N = 10	1775	1887	3002	0	0.324	48.25	120.5	103.8	50.04	512	978	1203	1887	2698	3002
N = 20	2761	4084	6000	0	0	40,27	120.45	107.2	42.5	595	986	1220	2761	4087	5670
N = 30	3735	5607	9004	0	0	40,08	120.77	109.56	51.55	744	994	1211	5607	3735	9004
N = 40	4794	4818	12002	0	0	40,354	117.52	105.01	57.84	758	1009	1222	7213	4818	12002
N = 50	5700	5700	15020	0	0	42,325	116.35	110.86	50.38	874	1041	1235	7184	7012	18006
N = 60	7630	10679	18006	0	0	44,005	115.45	117.02	46.44	886	1062	1254	7136	10679	18006
N = 70	8509	10851	21005	0	0	46,45	117.84	157	62.7	984	1107	1257	8383	12431	21005
N = 80	9433	10532	24005	0	0	44,08	112.88	158	77.54	1067	1172	1282	9532	14205	24005
N = 90	10738	10844	27013	0	0	47,58	110.1	164.02	85.002	1090	1197	1285	14305	9532	27013
N = 100	12074	12025	30010	0	0	41,08	96.865	153.74	88.86	1097	1224	1299	16038	10738	30010

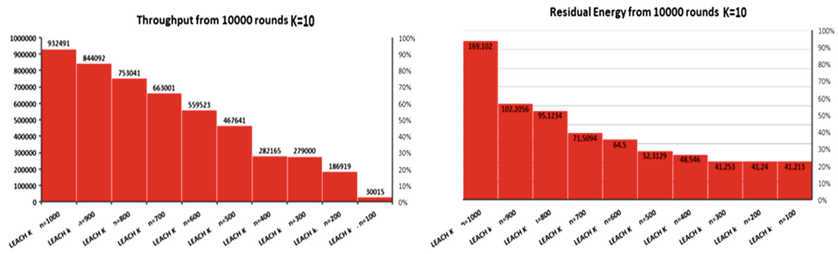

Fig. 5. Network throughput gain rate and residual energies with variation of N when k = 10

5 Conclusion

In this paper, we have compared LEACH-K with LEACH and TEEN routing protocol using Matlab simulator for energy consumption, throughput and latency time with reference to the number of nodes and the number of clusters with K-Means function to determine the best configurations that offers the most suitable tradeoff between energy conservation and QoS metrics for each pattern

The use of the K-Means function with the LEACH protocol exactly before the election of the CHs gives us good results, an increase of the network lifetime, latency, throughput, number of CHs and a decrease in latency and energy consumption

However, the mechanisms of aggregation and Machine Learning used by LEACH-K degrade the quantity of information that generates low stability and latency time

We can take care of the problem of low stability, latency time and mobile nodes that allow a gain in terms of energy and all metrics of quality of service essentially in terms of throughput since mobility of SINK will be the topic of our next work

References

1. Heinzelman, W.R.: Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd Annual Hawaii International Conference on System Sciences. IEEE (2000)
2. Zhang, M., Wang, S.: An energy efficient dynamic clustering protocol based on weight in wireless sensor networks. J. Netw. **6**(7), 1057–1064 (2011)
3. Devi, G., et al.: The K-Means clustering used in wireless sensor network. Int. J. Comput. Sci. Eng. (IJCSE) **8**(4), 106–111 (2016)
4. Rabiaaa, E., Nourab, B., Adnenec, C.: Improvements in LEACH based on K-means and gauss algorithms, In: The International Conference on Advanced Wireless, Information, and Communication Technologies (AWICT 2015) (2015)

5. Randhawa, S., Jain, S.: Performance analysis of LEACH with machine learning algorithms in wireless sensor networks. Int. J. Comput. Appl. **147**(2), 7–12 (2016)
6. Kanungo, T., Mount, D.M., Netanyahu, N.S.: An efficient k-means clustering algorithm: analysis and implementation. IEEE Trans. Pattern Anal. Mach. Intell. **24**, 881–892 (2002)