

Energy-Efficient Hybrid K-Means Algorithm for Clustered Wireless Sensor Networks

Aziz Mahboub¹, Mounir Arioua², El Mokhtar En-Naimi³

^{1,3}LIST Laboratory, Department of Computer Science, Faculty of Sciences and Techniques, Tangier, Morocco

²National School of Applied Sciences-Tetouan, Abdelmalek Essaâdi University, Morocco

Article Info

Article history:

Received Dec 11, 2016

Revised Mar 10, 2017

Accepted Mar 24, 2017

Keyword:

Clustering

Energy optimization

K-means algorithm

LEACH

Wireless sensor network

ABSTRACT

Energy efficiency is the most critical challenge in wireless sensor network. The transmission energy is the most consuming task in sensor nodes, specifically in large distances. Clustered routing techniques are efficient approaches used to lower the transmission energy and maximize the network's lifetime. In this paper, a hybrid clustered routing approach is proposed for energy optimization in WSN. This approach is based on K-Means clustering algorithm and LEACH protocol. The simulation results using MATLAB tool have shown that the proposed hybrid approach outperforms LEACH protocol and optimizes the nodes energy and the network lifetime.

Copyright © 2017 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Aziz MAHBOUB,

LIST Laboratory, Department of Computer Science,

Faculty of Sciences and Techniques Abdelmalek Essaâdi University,

Street of charf B. P 410, Tangier – Morocco.

Email: amahboub@uae.ac.ma

1. INTRODUCTION

Wireless sensors network (WSN) is composed of group of wireless sensors deployed in field with a high density; enables detection of physical phenomena such as thermal, optical, vibration and temperature or pressure [1], [2]. WSN are being increasingly used for several application areas, such as environmental monitoring, military target tracking, healthcare, seismic sensing, etc. Each sensor node is assembled with a transducer, microcomputer, transceiver and power source. These wireless nodes are usually powered by limited capacity batteries which replacement is delicate in hostile environment where hundreds of them are randomly deployed. The major limitations encountered in WSN are energy resources and high cost of transmission [3], [4]. Therefore, energy limitation is the key challenges in wireless sensors network, since the batteries are not rechargeable usually which limits the network lifetime [5-7]. To extend the wireless network lifetime, the energy of sensor nodes should be efficiently used. Energy required in communications is the important consuming part compared to processing process in sensor nodes [8]. A variety of techniques have been developed in order to reduce the communication energy consumption. Optimized routing techniques such as clustering algorithms have been widely adopted in WSN due to their energy effectiveness.

This paper investigates the energy efficiency of clustering in routing protocol. We propose a substantial K-Means clustering algorithm based on LEACH protocol which consists of using jointly K-Means and LEACH approaches. The proposed hybrid algorithm is proved to be energy efficient than LEACH protocol and significantly extends the sensor network lifetime.

The remainder of this paper is organized as follows. Section 2 discusses the outline of K-Means algorithm. The improved K-Means algorithm is described in section 3. Simulations and results are discussed in section 4. The final section concludes the paper by setting the direction of future work.

2. K-MEANS ALGORITHM

K-Means is one of the promising and effective clustering algorithms [9]. It consists of regrouping nodes in the network into several clusters, the foundation of clusters is based on two parameters; the first parameter is the number of wish clusters, the second parameter is the Euclidean distance is used for locate the closest cluster for each node [10]. The selection of cluster heads for the k-Means clustering algorithm is based on two things; the cluster heads position must be the cluster center, and the residual energy of the node.

In wireless sensor network, the k-Means clustering method is based on an iterative optimization of the distance between the nodes to classify. The algorithm generates K clusters from a set of N nodes. The objective function of K-Means algorithm is [11], [12]:

$$F = \sum_{r=1}^k \sum_{x_i \in C_r} (x_i - ch_r)^2 \quad (1)$$

Where C_r is the set of nodes that belong to cluster r , The K-Means clustering uses the Euclidean distance:

$$d(x_i, CH_r) = \sqrt{(x_i - ch_r)^2} \quad (2)$$

Therefore the K-Means algorithm only seeks to find the global minimum of ch_r . Where x_i is a node of a cluster, ch_r the cluster head [12-16].

2.1. Algorithmic Steps for k-Means Clustering

The progress of the K-Means clustering algorithm is done in four phases, the first step is choosing the desired number of clusters to generate in wireless sensors network, then for each cluster on chosen the cluster head randomly ch , After the attribute the closest cluster to each node using the Euclidean distance [17].

The $\{x_1, x_2, x_3, \dots, x_n\}$ represented the wireless sensors deployed in field and $\{ch_1, ch_2, \dots, ch_k\}$ is the clusters head chosen initially randomly.

1. Initialize the middle of the clusters randomly ch_r , where: $r = 1, \dots, k$ and $k < n$
2. Attribute the closest cluster to each data point using the Euclidean distance

$$C_j = j: d(x_j, ch_r) \leq d(x_j, ch_l), l \neq r, j = 1 \dots n \quad (3)$$

3. For all clusters k , the cluster middles ch_i are updated using:

$$ch_i = \left(\frac{1}{c}\right) \sum_{j=1}^{c_i} x_i \quad (4)$$

3. ENERGY MODEL

A sensor uses its energy to carry out three main tasks: acquisition, data processing and communication. However, the energy used for acquisition is not prominent as much as in the communication operation. Likewise, the energy consumed in the data processing operation is less important than communication energy [18].

The proposed approach uses the first order model adopted by LEACH and SEP protocols, the radio model is shown in Figure 1.

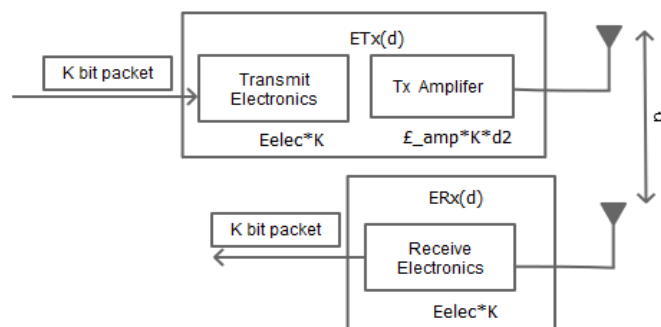


Figure 1. Energy model in wireless sensor network[18]

However, the consumed energy in data processing operation is less important than the energy of communication (energy required to send or receive data to another node).

The distance between transmitter and receiver influences the quantity of the consumed energy to send data into destination. To broadcast a K -bit for a distance d , the energy expended by the system can be calculated by the Equations (5) and (6):

The expended energy can be schematized as follows:

$$E_{Tx}(k, d) = E_{Tx-elec} + E_{Txamp}(k, d) \quad (5)$$

The energy expended in the transmit electronics for free space propagation ETx-fs is described by:

$$E_{Tx-fs}(k, d) = K * (E_{elec} + \epsilon_{fs} * d^2) \quad \text{if } d < d_0 \quad (6)$$

The energy expended in the transmit electronics for free multi-path propagation ETx-mp is given by:

$$E_{Tx-mp}(k, d) = K * (E_{elec} + \epsilon_{amp} * d^4) \quad \text{if } d \geq d_0 \quad (7)$$

To receive a message of k bits, the energy consumed by the receiver is given by:

$$E_{Rx}(k) = E_{Tx-elec}(k) = k * E_{elec} \quad (8)$$

Where: E_{Tx} is the electrical energy required to transmit a K -bit message over a distance d , E_{elec} corresponds to the energy per bit required in transmitting and receiving electronics to process the information. ϵ_{fs} and ϵ_{amp} are constants corresponding to the energy per bit required in the transmission amplifier to transmit an L -bit message over a distance d^2 and d^4 for free space and multi-path propagation modes, respectively. By equating formula (6) and (7), we determine the distance $d=d_0$ when the propagation transitions from direct path to multi-path:

$$d_0 = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{amp}}} \quad (9)$$

If the distance between the transmitter and the receiver is larger than the crossover distance d_0 , the multi-path model is employed. Otherwise, the free space model is adopted to measure the energy dissipation.

4. HYBRID K-MEANS APPROACH

The proposed routing approach aims mainly to optimize the energy consumed in data communications. Using K-means clustering algorithm along with LEACH protocol enable member nodes to be attached to the closest CH in the network. By adopting the proposed scheme, the transmission energy is minimized and the network lifespan is improved.

By considering n nodes randomly generate in area with dimension $M * M$, Our approach, the k-Means clustering algorithm is used initially for classify the nodes in wireless sensors network in groups, The classification takes place as follows: the initialization of k , take k number of groups, the chosen k centroids initially at random places of groups which will be constructed or $k < n$.

The next step consists of grouping the nodes into clusters using the Euclidean distance. Each node is attached to the closest centroid in the network. Then the recalculate the position of centroid in each group, if the position of the centroid is changed then returns to step the join of each node in the new nearest centroid; If the position of the centroids does not change.

Secondly, the Leach protocol is used in each group for the selection of clusters head and their members. The optimal number of CHs is estimated to be 5% of the total number of nodes [18]. The node elects itself to become a cluster head by some probability and broadcasts its status to the group members.

A larger number of nodes may elect themselves as cluster heads than the desired number of cluster heads. This decision is made by the node by choosing a random number between 0 and 1. The node becomes a CH for the current round if the number is less than the threshold $T(s)$, the node declares CH according to the Equation following threshold:

$$T(s) = \begin{cases} \frac{p}{1-p \left(r \bmod \left(\frac{1}{p} \right) \right)} & \text{if } s \in G_k \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

Where p is the desired percentage of CH nodes in the sensor population and r is the current round number, where G_k is the group of nodes that have not been CHs in the last $1/p$ rounds [18].

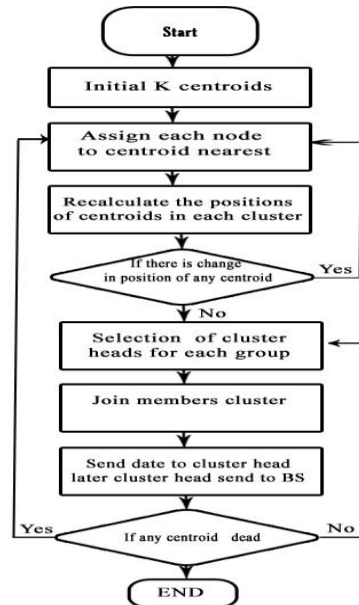


Figure 2. Flowchart of the cluster formation process of proposed approach

5. RESULTS AND DISCUSSION

All simulations were performed using MATLAB. We started by a series of simulations to ascertain the experimental optimal value of the number of classified groups that can provide energy efficiency in the sensor network. We tested the performed network with different clusters number ($k=2$, $k=3$, $k=4$ and $k=5$). We performed simulation for two algorithms (LEACH and proposed approach), we compare and discuss the obtained results. The simulated network consists of 100 nodes distributed randomly in an area of $100\text{ m} \times 100\text{ m}$. The base station is located at position (50m, -50m). Every node transmits a 4000-bit message per round to its cluster head. The probability of being cluster head is set to 0.05, about 5% of nodes per round become cluster head. Figure 3 presents the number of live nodes versus transmission rounds for Hybrid K-mean approach for different classification number (K). The best energy performance of the proposed approach was obtained when $K=3$. The stability period was sufficiently extended and the network lifetime was enormously lengthened. In consequence, we choose to perform the Hybrid K-Mean approach with the optimal value of classification ($K=3$).

In order to evaluate the proposed approach performance, several plots were generated to characterize energy consumption and the distribution of live and dead nodes in the hybrid K-Means network. We calculate the total WSN energy, and distribution of alive nodes for each method (LEACH and Hybrid K-Means) on common plots to compare network performance. The total system energy and number of nodes versus transmission round are shown in Figure 4 and Figure 5 respectively.

s. Simulation total energy of the network using LEACH protocol and the proposed Hybrid K-Means algorithm is carried out in order to evaluate the algorithms energy performance. Table 1 show the simulation parameters used in this work.

The comparison we carry out in this work between the proposed approach and LEACH protocol is based on some key performance parameters of wireless sensors network: firstly, the network lifetime that describe the time interval between the start of the network operation and the death of the last sensor node. The second parameter is the number of alive nodes which is the total number of the sensor nodes that have not yet depleted all of their energies. The third parameter is the number of dead nodes which is the total number of sensor nodes that have consumed all of their energy and are not able to do any kind of functionality.

The obtained results have demonstrated that the proposed approach reduces the energy consumption by round, and extend the network lifetime. According to Figure 4, the energy consumption rate of the proposed approach is less than the LEACH protocol. At approximately round 2300 and 2400 for the LEACH protocol, there is no energy in the WSN; On the other hand the exhaustion of system energy in proposed approach begins at 4000 rounds, therefore the proposed approach provided the longest lifetime of WSN due to the Leach protocol.

Table 1. Simulation parameters

Simulation Area	100 m * 100 m
BS Location	(50 m , -100 m)
Number of Nodes	100
Transmission amplifier free space ϵ_{fs}	10 pJ/bit/m ²
Transmission amplifier, multipath ϵ_{mp}	0.0013 pJ/bit/m ⁴
Data Aggregation Energy	5 nJ/bit/msg
Transmission Energy E_{Tx}	50 nJ
Receiving Energy E_{RX}	50 nJ
Network Dimension	100m*100m
Signal wavelength λ	0.325 m
Antenna gain factor G_t, G_r	1
Message size (bits)	4000 bits
Initial node energy	0.5 J

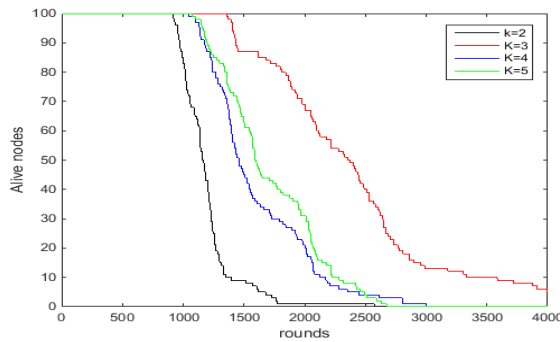


Figure 3. Nnumber of nodes alive versus transmission round for proposed approach In different k

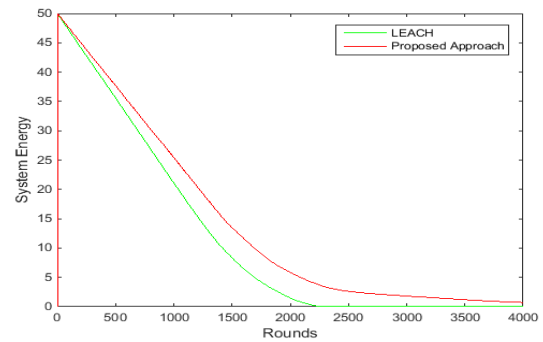


Figure 4. Total system energy versus transmission round for LEACH and proposed approach

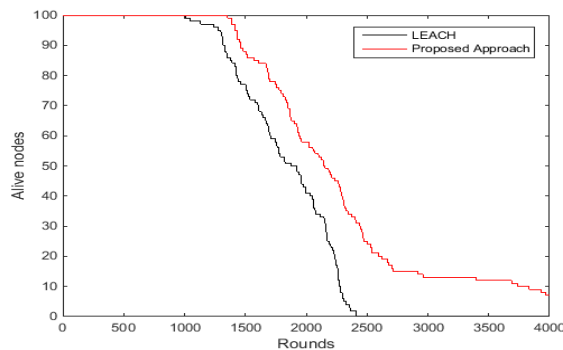


Figure 5. Nnumber of nodes alive versus transmission round for LEACH and proposed approach

6. CONCLUSION

The main objective of this work is to propose a novel clustering routing approach based on k-Means algorithm and LEACH protocol applied for wireless sensor networks. The proposed approach minimizes the energy consumption, extends the network lifetime of the sensor nodes. The evolution and enhancement of the presented routing algorithm should be done in the future to support multipath and heterogeneous modes.

REFERENCES

- [1] VehbiÇağrıGüngör and Gerhard P. Hancke, "Industrial Wireless Sensor Networks: Applications, Protocols, and Standards", CRC Press 83, pp: 1027-40, 2013.
- [2] I. Akyildiz, "A Survey on Sensor Networks", IEEE Commun. Mag., vol. 40, no. 8, pp. 102 -114, 2002.
- [3] Heinzelman W, Chandrakasan A, Balakrisham H, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks", In Proceeding of the 33rd Annual Hawaii International Conference on System Sciences, pp. 3005-14, 2000.
- [4] Aziz Mahboub, Mounir Arioua, EL Mokhtar En-Naimiand ImadEzzazi, "Multi-zonal Approach for Clustered Wireless Sensor Networks", 2nd International Conference on Electrical and Information Technologies, ICEIT 2016, May 4-7, 2016.
- [5] Aziz Mahboub, Mounir Arioua, EL Mokhtar En-Naimi, Imad Ezzazi, "Performance Evolution of Energy-Efficient Clustering Algorithms in Wireless Sensor Network", *Journal of Theoretical and Applied Information Technology*, 2016.
- [6] Peng Zhu, "A New Approach to Sensor Energy Saving Algorithm", *TELKOMNIKA*, Vol. 11, No. 5, pp. 2485 ~ 2489, 2013.
- [7] Nishi Sharma, Vandna Verma, "Energy Efficient LEACH Protocol for Wireless Sensor Network", *International Journal of Information & Network Security (IJINS)*, Vol. 2, No. 4, pp. 333~338, 2013.
- [8] Mechta D, Harous S Alem I Khebbab D, "LEACH-CKM: Low Energy Adaptive Clustering Hierarchy protocol with K-Means and MTE", 10th International Conference on Innovations in Information Technology, 2014.
- [9] Kanungo T, Mount DM, Netanyahu NS, Piatko CD, Silverman R, Wu AY, "An Efficient k-Means Clustering Algorithm Analysis and Implementation", *IEEE T Pattern Anal*, 24:881-92, 2002.
- [10] Celebi ME, Kingravi HA, Vela PA, "A Comparative Study of Efficient Initialization Methods for the k-Means Clustering Algorithm", *Expert Syst Appl*, 40:200-10, 2013.
- [11] Velmurugan T. "Performance based analysis between k-Means and Fuzzy C-Means clustering algorithms for connection oriented telecommunication data". *Appl Soft Comput*,19:134-46, 2014.
- [12] Harb H.a., Makhoul A, Laiymani D.a, Jaber A.b.,Tawil R.b., "K-Means based Clustering Approach for Data Aggregation in Periodic Sensor Networks", 10th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, pp. 434-41, 2014.
- [13] Nithyakalyani S.a, Gopinath B.b., "Analysis of Node Clustering Algorithms on Data Aggregation in Wireless Sensor Network", *Journal of Scientific and Industrial Research*, 74 Issue 1, pp. 38-42, 2015.
- [14] Kumar P.a., Chaturvedi A.b., "Life Time Enhancement of Wireless Sensor Network using Fuzzy c-Means Clustering Algorithm", International Conference on Electronics and Communication Systems, 2014.
- [15] Tan L.ab, Gong Y.a., Chen G.a., "A Balanced Parallel Clustering Protocol for Wireless Sensor Networks using K-Means Techniques", 2nd International Conference on Sensor Technologies and Applications, 2008.
- [16] Thai T. Vu, Viet D. Nguyen2, Hien M. Nguyen, "An Energy-Aware Routing Protocol for Wireless Sensor Networks Based on K-Means Clustering", AETA 2013: Recent Advances in Electrical Engineering and Related Sciences, pp. 297-306, 2014.
- [17] Dianhu Cheng , Xiangqian Ding , Jianxin Zeng , Ning Yang ,"Hybrid K-means Algorithm and Genetic Algorithm for Cluster Analysis", *TELKOMNIKA Indonesian Journal of Electrical Engineering*, Vol.12, No.4, pp. 2924 ~ 2935, 2014.
- [18] Aziz Mahboub, Mounir Arioua, EL mokhtar En- naimi, Imad Ezzazi and Ahmed El Oualkad. "Multi-zonal approach Clustering based on Stable Election Protocol in Heterogeneous Wireless Sensor Networks". IEEE 4th Edition of the International Colliquium on Information Sience and Technology CiSt'16, 2016.

BIOGRAPHIES OF AUTHORS



Aziz Mahboub is a PhD student in LIST (Laboratoire d'Informatique, Systèmes et Télécommunications) Laboratory, Department of Computer Science. Faculty of Sciences and Techniques, Tangier, Morocco, at the University of Abdelmalek Essaâdi. His research interests include wireless sensor networks, wireless networking and communication. He is an autor of several articles.



Mounir Arioua received his Ph.D degree in Telecommunications and Computer Science from Cadi Ayyad University of Marrakech in 2012. In 2012, he joined the National School of Applied Sciences of Marrakech as research assistant and assistant professor. Currently, he is professor at National School of Applied Sciences in the department of Engineering Technologies, University of Abdelmalek Essaadi. His research interests include wireless sensor networks, Internet of things, wireless networking and communication, wireless communications and mobile computing, real-time processing and embedded systems.



El Mokhtar En-Naimi is a Professor in Faculty of Sciences and Technologies of Tangier, Department of Computer Science. He is responsible for a Bachelor of Sciences and Technologies, BST Computer. Actually, he is a Head of Computer Science Department, since October 2016. He is also a member of the LIST Laboratory (Laboratoire d'Informatique, Systèmes et Télécommunications), the University of Abdelmalek Essaâdi, FST of Tangier, Morocco. He is an Autor and/or Co-Autor of several Articles in Computer Science, in particular, in SMA, Cases Based Reasoning (CBR), eLearning, Big DATA, Datamining, Wireless sensor network, VANet, MANet, etc. In addition, he is an Associate Member of the ISCN - Institute of Complex Systems in Normandy, the University of the Havre, France.