ENERGY-EFFICIENT IMPROVED OPTIMAL K-MEANS: DYNAMIC CLUSTER HEAD SELECTION BASED ON DELAYING THE FIRST NODE DEATH IN MWSN-IOT

Awatef Chniguir and Zouhair Ben Jemaa

Laboratoire LRSIC, National Engineering School of Tunis, University of Tunis El Manar, Belvedere, State or Province, BP 37, 1002, Tunisia

ABSTRACT

The Internet of Things (IoT), which attaches dynamic devices that can access the Internet to create a smart environment, is a tempting research area. IoT-based mobile wireless sensor networks (WSN-IoT) are one of the major databases from which the IoT collects data for analysis and interpretation. However, one of the critical constraints is network lifetime. Routing-based clustered protocols and cluster head (CH) selection are crucial in load balancing and sensor longevity. Yet, with clustering, sensor node mobility requires more overhead because the nodes close to the center may get far and thus become unsuitable to be a CH, whereas those far from the center may get close and become good CH candidates, influencing energy consumption. This paper suggests an energy-efficient clustering protocol with a dynamic cluster head selection considering the distance to the cluster center, remaining energy, and each node's mobility degree implementing a rotation mechanism that allows cluster members to be equally elected while prioritizing those with the minimum weight. The advised algorithm aims to delay the first node death (FND) and thus prolong the stability period and minimize energy consumption by avoiding re-clustering. The performance of the proposed protocol exceeds that achieved in our previous work, Improved OK-means, by 11%, in first dead node lifetime maximization, 43% in throughput, and 44% in energy efficiency.

KEYWORDS

Mobile WSN-IOT, K-means clustering, energy efficiency, CH selection, weight function, FND, lifetime.

1. INTRODUCTION

Wireless networking has profoundly transformed our lifestyle. The consistent demand for communication everywhere, especially in domains with higher activity, such as healthcare and agriculture, is growing significantly. The Internet of Things (IoT), one of the emerging technologies, allows current objects to connect over disparate networks [1]. It permits monitoring in miscellaneous applications and gathering tremendous amounts of data by integrating loads of wireless devices and sensors. Wireless sensor networks (WSN) are one of the major databases from which the IoT collects data for analysis and interpretation. Indeed, WSNs consist of numerous sensors equipped with irreplaceable batteries randomly dispersed in a monitoring area to perform different tasks. Therefore, the sensors are usually hard to reach and cannot be substituted. Low power and a short transmission range are two characteristics of sensor nodes. Sensors can move around after being initially placed by the environment (water or air) or mechanical techniques such as self-propelled wheels or springs, or they can be attached to automobiles, animals, or robots [2]. These devices collect data by setting up wireless networks and using embedded wireless radios. They are connected to external networks or the Internet to

render sending the data they gather to remote users possible. The hotspot is one of the biggest problems with stationary WSN-IoT systems. The closer the sensor nodes are to the sink, the briefer they last as they have to use energy to send data from the sensors farther away to the sink. Mobile sensors may utilize their batteries more efficiently and smoothly, as it minimizes the number of transmission hops necessary to reach the BS and eventually increases the lifetime [3]. Thus, a primarily interconnected static WSN-IoT network can become a collection of disconnected sub-networks due to link failure or energy depletion; however, mobile nodes can help reorganize the network. Two methods introduce mobility. The easiest option is to use static sensor nodes and move the sink nodes. The second strategy is to keep static sinks while sensor nodes are movable, e.g., when monitoring animals. The proposed approach follows the second method.

For sensor networks, hierarchical or clustered routing is more advantageous than flat routing. Clustering, a valuable solution for energy efficiency, is one of the strategies to preserve the energy supply of these nodes and thus increase the network lifetime [4]. Clustering aims to divide all nodes into groups and assign super functions to some elected nodes as a master named Cluster Head (CH) responsible for collecting data from the other nodes in the same cluster, Cluster Members (CM). Data aggregation can reduce latency because it reduces traffic and delays accordingly. Nodes along the path to the sink - CH in clustered routing schema- perform data fusion to reduce the data they receive. Transmission by several CH allows load balancing in the network and disburses energy between them. Therefore, CH selection is elemental and could significantly decrease energy consumption in the network and delay the death or disconnection of some sensors. Additionally, the nodes can also be used in a Round-Robin process and set a data transfer time. Consequently, retransmission is avoided, data redundancy in the covered region is decreased, and medium access collision is avoided [5]. According to preliminary research, the closest nodes to the BS are elected as CH to decrease energy consumption, notably when that node differs each round, as in mobile WSN-IoT compared to a static network. Indeed, it has been demonstrated that a WSN lifetime may be successfully prolonged if the energy consumption of each sensor decreases and the network energy load is carefully balanced. The distribution of nodes in the field achieves the best results when sensor localization is well studied, even if sensors are initially randomly deployed [6]. In traditional clustering, nodes opt to be CH based on an a priori probability. CH then overflows their neighbors with a message including a CH role assignment, and the recipients choose the closest one to serve as their own. Random clustering techniques are readily easy to use and keep control packets from going through numerous rounds before settling on a single cluster head per cluster. The uneven size and shape of the clusters, which results in a significant disparity in energy usage, is one of the main drawbacks of random clustering. An improvement over random clustering, some methods "grow" clusters around these nodes by first randomly assigning cluster head positions to a few of the network nodes. In contrast to random clustering, other approaches like k-means clustering invert the phases of the clustering process, i.e., they identify the clusters and their members before selecting one to serve as the cluster leader. As the most important unsupervised learning task, clustering entails discovering a structure in an unlabeled dataset as with all other problems of this type. In unsupervised 'teacherless' learning, only input is used. The purpose is to model the structure or distribution of the data to learn about certain aspects of the data. The method predicts from the dataset without advanced knowledge of its true labels [7]. The K-means method begins by arbitrarily centering K points in space. Then, the iterative repetition of the subsequent phases occurs: (1) Each instance is assigned to the cluster whose center is geographically nearest, and (2) each cluster center is moved to the instances' respective means. The process is continued until no instances within the cluster experience membership changes.

We suggest in this paper an Energy-Efficient version of Improved Optimal K-means (**EEIOK-means**) that involves residual energy and mobility degree in the CH selection process.
Our contributions can be summarized as follows:

- A strategy for making an MWSN-IoT routing protocol efficient on energy using an unsupervised learning method called K-means and a CH selection based on a cost function.
- A rotation mechanism to avoid re-clustering and reduce the overhead caused by the control packets.
- A cost function for the CH selection mechanism that includes node battery, node mobility, and the distance separating each node from the center of the germane cluster.
- A study of the influence of weighting coefficients and node velocity on the throughput and energy efficiency of the protocol.

The paper is organized as follows: The related work is described in Section 2. Section 3 describes the system model, radio model, and mobility model. Section 4 presents the CH selection process, flowchart, and pseudo-code of each part of the proposed protocol. The performance evaluation is given in Section 5. Section 6 presents a comparative analysis. Eventually, Section 7 concludes the paper and highlights some future works.

2. RELATED WORK

This section introduces some of the existing energy-aware clustering algorithms for MWSN-IoT and gives an overview of CHs selection metrics and techniques.

As previously stated, our proposed protocol is an extension of OK-means, a recently suggested improved node localization approach based on LEACH [8] and the K-means clustering algorithm presented by [9]. The CH selection process is probabilistic, and the CH number is taken from 10\% of the whole network. The closest nodes to k-means clustering centers are picked in the initialization stage for the CH role in their clusters, and the clusters and membership are maintained until all nodes die. A rotation mechanism based on CH placement will assign it to a different node in the same cluster. The second-closest sensor to k-means centers, not a CH in the previous round, would be elected in the next round. Compared to LEACH, Optimal K-means (OK-means) have reduced energy consumption by about 28% and avoided re-clustering. The lack of node mobility and the disregard of residual energy in CH selection are the main drawbacks of this algorithm. As a first step toward addressing these deficiencies, an improved version using the residual energy in the CH selection process is presented by [10]. The WSN lifetime protracts in a static context. Our proposed protocol expounds the previous work and implements the Improved OK-means [10] by providing further mobility for the nodes to promote their usage in agricultural applications.

Zafar et al. [11] have introduced two routing protocols for MWSN based on the HHCA algorithm [12]: Mobility-aware Centralized Clustering Algorithm (MCCA) and Mobility-aware Hybrid Clustering Algorithm (MHCA). The MCCA [11] uses two levels of mobility-aware centralized Fuzzy C-Means gridding. The algorithm consists of setup and steady-state steps. The development of the layer-2 clusters and the selection of Grid Head 2 (GH2) in the setup step comes first. The BS chooses GH2(s) from the nearest nodes with the optimum residual energy and the lowest velocity, and Layer-1 cluster development and GH1 selection come next. The BS relays the position, residual energy, and non-GH2 sensor node velocity to the GH2(s), which uses fuzzy gridding to determine cluster centers and construct lower-tier clusters. The GH2s choose layer-1 GH1s with the lowest pace and highest residual energy. The fuzzy C-Means technique

based on centralized gridding and LEACH-Mobile [13] is used in MHCA [11]. Layer-2 grids are produced, and GHs are chosen in the same way as in MCCA. Layer-1 cluster generation and CH selection use distributed clustering. The nodes, not GHs, are selected as CHs depending on LEACH, node energy, and velocity. Based on the simulation results, both methods aim at reducing data loss caused by mobility while improving performance, i.e., energy consumption, cluster count, and packets received at the BS. Unfortunately, considerable energy is lost due to CH double selection, causing premature sensor death and eventually impacting latency. Even when CHs number is stable for a few rounds, it remains high and unstable for extended periods. Therefore, our proposed algorithm calculates a fixed and optimal number of clusters by the Elbow method [14], which minimizes the exchange of information in the network.

Zhang et al. [15] have presented the Centralized Energy-Efficient Clustering Routing Protocol for Mobile Nodes (CEECR) to reduce energy consumption and increase the packet delivery ratio. This enhancement was fulfilled by a weight function that decides the CH selection based on the node velocity, initial energy, transmission range, and distance. Two types of nodes, mobile nodes (MN) and fixed nodes (FN), are used as candidates with low velocity and high energy that are suitable to be elected as CH, resulting in fewer detached nodes. However, the method overlooks lifetime and dead nodes. Inspired mobility degree consideration in the CH selection process, our algorithm uses mobile nodes.

Umbreen et al. [16] have introduced the Energy-Efficient Mobility-Based Cluster Head Selection (EEMCS), where the suggested cluster head selection procedure considers sensor node residual energy, node mobility level, the number of nodes connected, and their distance from the BS. The EEMCS outperforms CRPD [17], LEACH, and MODLEACH [18] in energy consumption by employing multi-hop intercluster communication in data transmission and maintaining cluster formation stability until re-clustering. Adaptive clustering and the weighting function in CH selection are the algorithm strengths, although the death of the first node occurs early, which is the thrust of our proposed algorithm.

Aydin et al. [19] have presented two energy-efficient mobile routing algorithms based on the LEACH protocol for agriculture. The algorithms are scheduled for mesh networks of Bluetooth Low Energy (BLE) sensors to acquire higher-quality and more efficient products in the field by examining soil moisture, temperature, and mineral-like data. They have used two different techniques for clustering: greedy and artificial neural network methods. The CH and SCH nodes are determined by their proximity to the cluster center and the remaining energy they retain. Although the nodes are stationary and only the sink is mobile, the two clustered protocols achieve adequate energy dissipation and lifetime results, and the sink assumes mobility via a genetic algorithm. Nevertheless, this method does not consider node mobility and packet delivery ratio.

Table 1 summarizes the evaluated clustering algorithms along with their drawbacks. The primary issue with the approaches examined in this part is their short stability period, lack of node mobility, and high energy dissipation. The suggested technique compensates for these limitations by using mobile nodes and introducing a cost function in CH selection.

3. SYSTEM MODEL

The proposed algorithm is an Energy-Efficient version of the Improved Optimal K-means (EEIOK-means) that includes residual energy and mobility level in the CH selection procedure.

3.1. Network Model

In this study, 150 homogeneous sensors with the same amount of energy were randomly deployed in an area of 100 m by 100 m to collect data and send it to the BS. Sensors make a random walk every round with a moderate velocity (0.2 m/s–5.0 m/s). Moreover, sensors are aware of their coordinates, the remaining energy in their batteries, and the distance to the BS. The BS collects information and calculates the weight of each node according to Formula 9 (Section 4.2). In its cluster, the node with the lowest weight serves as CH. The BS performs K-means clustering to designate K groups of sensors. K is first calculated in the algorithm with the elbow method and depends only on the network density. In addition, the BS is in the middle of the field, and the TDMA schedule ensures that nodes never go without data to send. The proposed network model is a centralized, mobile-aware, clustering, homogeneous routing protocol based on k-means.

Table 1. Summary of related works.

Mechanism	Approach	Network type	Mobility	Metrics of CH selection	Weakness
[8]	OK-means	centralized	No	Distance to the Cluster Center	Lack of mobility throughput not supported
[9]	Improved Ok-means	centralized	No	Distance to the Cluster Center & Residual Energy	Lack of mobility and didn't support throughput
[10]	MCCA and MHCA	centralized	Yes with mobile nodes	Velocity & Residual Energy	A short period of stability
[14]	CEECR	Centralized & hybrid	Hybrid (fixe and mobile nodes)	Initial Energy, transmission range, Velocity & Distance	Lifetime not supported
[15]	EEMCS	centralized	Yes with mobile Nodes	Residual Energy, Mobility Level, Density of neighbors & distance to BS	Short period of stability
[19]	Greedy&GA and ANN&GA	centralized	Yes with mobile Sink	Distance to the Cluster Center and Residual Energy	throughput and mobility of nodes not supported

3.2. Radio Model

The radio energy model [20] provides a suitable Signal-to-Noise Ratio (SNR) in L-bit message transmission over a distance d. Equation 1 shows the energy used by the radio transmission module,

$$E_{Tx} = \begin{cases} LE_{elec} + L\epsilon_{fs}d^2, & d \leq d_0 \\ LE_{elec} + L\epsilon_{amp}d^4, & d > d_0 \end{cases}$$

(1)

Where E_{elec}, ϵ_{fs}, and ϵ_{amp} are the energy dissipated per bit by the transmitter or receiver equipment, the energy dissipated for free-space transmission, and the energy dissipated for multi-path transmission, respectively. d_0 is the threshold distance given by 2.

$$d_0 = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{amp}}} \tag{2}$$

A sensor can be a CH or not in a hierarchically clustered network. If $d \le d_{0}$, the energy dissipated by a CH (E_{dCH}) and the energy dissipated by a CM (E_{dCM}) node are given by 3:

$$E_{dCH} = (\frac{n}{K} - 1).L.E_{elec} + \frac{n}{K}L.E_{DA} + LE_{elec} + L\epsilon_{fs}d^2_{toBS} \tag{3}$$

$$E_{dCM} = L.E_{elec} + L\epsilon fsd^2 toCH \tag{4}$$

1. n is the number of nodes,
2. K is the number of clusters
3. E_{DA} is the energy consumed in data aggregation.

The sum of both dissipated energies is:

$$\tag{5}$$

$$E_{cluster/round} = E_{dCH} + (\frac{n}{K})E_{dCM}$$

$$E_{total/round} = L(2nE_{elec} + nE_{DA} + \epsilon_{fs}(Kd^2_{toBS} + nd^2_{toCH})) \tag{6}$$

- $E_{cluster/round}$ is the whole energy expended in a single cluster every round.
- $E_{total/round}$ is the total energy consumed in the network per round.

The optimal number of clusters is computed using the following formula according to [8], [9], and [20]:

$$\tag{7}$$

$$k_{opt} = \begin{cases} \sqrt{\frac{n}{2\pi}}\frac{2}{0.765} & d < d_0 \\ \sqrt{\frac{n}{2\pi}}\sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}\frac{M}{d^2_{toBS}} & d > d_0 \end{cases}$$

3.3. Mobility Model

After their first deployment, sensor nodes are mobile. Therefore, extreme mobility can break up the structure of a network, leading to clusters falling apart and nodes becoming disconnected from their cluster heads, and the result is a significant data loss that hurts data rates and energy consumption. Mobility computation depends on modifying the random node positions in the network. Indeed, we use a random number generator to choose the node's new location [21]. By calculating the difference between the node's prior and new positions, the mobility level is:

$$DM = \sqrt{(x_{new} - x_{old})^2 + (y_{new} - y_{old})^2} \tag{8}$$

We applied the Random Waypoint Mobility (RWM) to our network following [11]. Both a pause time and a motion period exist in this model. A node remains in its present location throughout the pause period. The node picks a random direction throughout the mobility phase and travels in

that direction at a random speed then enters the pause phase again when it reaches the new location. The pause time is 1 s while the motion speed ranges from 0.5 to 2 m/s in the proposed algorithm.

4. PROPOSED WSN MANAGEMENT APPROACH

4.1. Clustering

In the algorithm, K points are initially randomly assigned to serve as centers of the 100-m-by-100-m field composed of 150 sensors. The following stages are then repeated, i.e., each center is moved to the appropriate means of instances once each instance is allotted to the cluster with the nearest center. The process is reiterated until no instance modifies its membership.

In Algorithm 1, X and Y are the vectors corresponding to the coordinates of the nodes forming a dataset required for clustering:

$$X = \begin{vmatrix} X1 \\ X2 \\ . \\ . \\ Xn \end{vmatrix}, Y = \begin{vmatrix} Y1 \\ Y2 \\ . \\ . \\ Yn \end{vmatrix}$$

In this work, in addition to the number of clusters K generated by the Elbow method, we considered N=150 yielding Kopt=4. The output is thus the K cluster centers and K vectors of labels corresponding to CMs. K-means forms clusters and their CH.

Algorithm 1 K-means Cluster Center Finding Algorithm

Require: K, X,Y
Ensure: centers,labels

1: Random selection of K Centers ;
2: **while** true **do**
3: Find new centers by computing the cluster's gravity center;
4: centers= gravity centers;
5: **if** centers didn't change **then**
6: Break
7: **end if**
8: Keep the last K centers.
9: **end while**

4.2. Cluster Head Selection

In this work, the selection of the cluster head is based on the node distances to the cluster center, their residual energies, and mobility degrees through the computation of the weight of each node as follows:

$$\omega_i = d^2_{itocc} + \frac{\alpha}{RE_i} + \frac{\beta}{DM_i} \qquad (9)$$

- *ωi* is the weight of a node "i"
- *d2ito CC* is the distance from node "i" to its Cluster Center.
- *REi* is the Remaining Energy in node "i"
- *DMi* is the Mobility Degree of the node "i" computed by equation (8)
- *α* and *β* are energy and mobility coefficients, respectively. As mentioned in [22] and [23], $\alpha + \beta = 1$.
- We value the localization of each CM in the cluster, so we did not assign a coefficient for the d^2_{itocc} parameter.

Clusters are constructed, and members are allocated for the first round when the WSN starts. Algorithm 2 ensures K cluster heads, and each K cluster head has the minimum weight compared to the other members of its cluster.

Algorithm 2 CH Selection Algorithm

Require: labels, RE,α,β,N,K
Ensure: X, Y

1: Random generation of node velocity.
2: Compute the degree of mobility (DMi)as (8)
3: **for** label ← 1 to K **do**
4: Compute squared distance of nodes to cluster center(d^2_{itocc})
5: Calculate the weight(ωi) as (9)
6: Eliminate dead nodes.
7: Take the node having the minimal ωi as CH.
8: **end for**

4.3. Data Transmission

After establishing the clusters and electing the CH, a multi-hop communication mechanism will be used for inter-cluster communication, depending on the distance to the BS. Cluster members give data to their CH, which aggregates the information and sends it to the sink. The CH sends the acquired data in one hop if no nearby node is closer to the CH than the sink.

4.4. Proposed Algorithm

Figure 1 presents the Flowchart. Algorithm 1 accomplishes the k-means clustering, Algorithm 2 fulfills CH selection, and Algorithm 3 resumes all the steps of the proposed protocol.

Functions send( ) and receive( ) are used to calculate the dissipated energy, as described in the radio model section, while the function sendtoBS( ) computes the dissipated energy when transmitting data to the BS. We suppose that there is data for transmission in each round.

Algorithm 3 EEIOK-means Algorithm

Require: X,N, (xSink, ySink), $E_0,E_{DA},\varepsilon_{fs}$, ,$E_{ELEC}$,L, α,β,$d_{thres}$,maxrounds
Ensure: Alivenode,RE

1: **function** send(d)
2: $d_{thres} = (\xi fs/\xi amp)^{1/2}$
3: **if** d > dthres **then**

4: RE = RE − ((L ∗ E_{ELEC}) + (L ∗ E_{ELEC} ∗ d4));
5: RE = RE − ((L ∗ E_{ELEC}) + (L ∗ E_{ELEC} ∗ d2)) ;
6: **end if**
7: **end function**
8: **function** receive()
9: RE = RE − ((L ∗ E_{ELEC}) + (L ∗ E_{DA}));
10: **end function**
11: **function** sendtoBS()
12: dthres = (ε_{fs} / ε_{amp})^{1/2}
13: **if** d > dthres
14: RE = RE − ((L ∗ E_{ELEC} + E_{DA}) + (L ∗ E_{ELEC} ∗ d_{toBS}^{4}));
15: RE = RE − ((L ∗ E_{ELEC} + E_{DA}) − (L ∗ E_{ELEC} ∗ d_{toBS}^{2})) ;
16: **end if**
17: **end function**
18: Get random coordinates for nodes
19: Populate all nodes into list Nodecollection(N=150)
20: Node parameter set-up
21: Calculate Koptimal according to Elbow method
22: K-means Cluster Center(Call of Algorithm 1 with K = Koptimal)
23: **for** r ← 1 to maxrounds **do**
24: computation of total energy
25: Check remaining energy of each node
26: Cluster Head Selection (Call of Algorithm 2)
27: **for** n ← 1 to Nodecollection **do**
28: n send(d) to the corresponding CH;
29: CH receive L data;
30: **end for**
31: **for** K ← 1 to Koptimal **do**
32: the CH sendtoBS(;)
33: the BS receive L data(;)
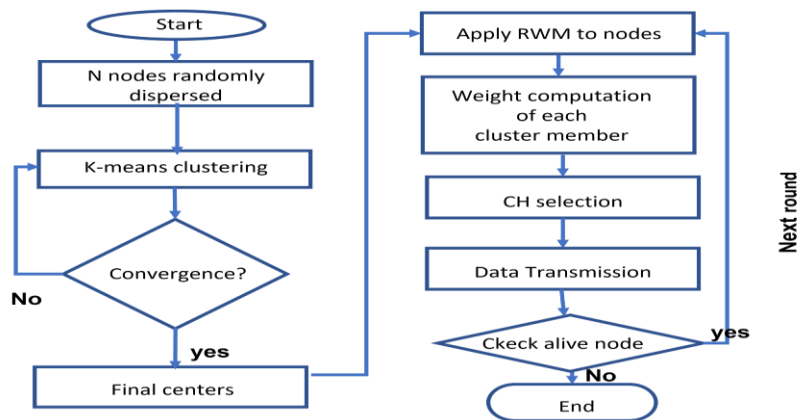34: **end for**
35: **end for**

Figure 1. Flowchart of the proposed algorithm.

5. PERFORMANCE EVALUATION

5.1. Performance Metrics

The following measures are considered to assess the performance of the suggested algorithm:

Network Lifetime: Several definitions are used to optimize the network lifetime in routing problems [6]. In this work, we considered the network lifespan as the duration of the operational state of a node, i.e., the time needed for a network to be functional from its creation to its shutdown.

Stability Period: The time until the first node death (FND) is discovered. A node is dead when its energy supply is depleted.

Throughput: The number of packets received by the BS in each round.

Energy Efficiency: The number of packets received by the BS divided by the total dissipated energy. Equation 10 gives the Energy Efficiency (E) [24], where N packets represent the number of packets and CE is the consumed energy in total.

The number of packets received by the BS divided by the total dissipated energy.

$$ E = \frac{N_{packets}}{CE} $$

(10)

5.2. Simulations Results

This section summarizes the findings from the python simulation of the proposed scheme. The results are based on the network settings listed in Table 2. The monitoring field contains homogeneous sensors dispersed randomly. Experiments proceed through three phases. First, we investigate the influence of the weighting coefficients used in computing the weight of each node to determine the CH's role. Second, we vary the node velocity to see its effect on the lifetime. Third, we raise the number of nodes to test the scalability of our suggested approach.

Table 2. Simulation Parameters.

Parameters	Symbol	Value
Number of nodes	N	150
Deployment space	MxM	100m x100m
Sink position	(x_{sink}, y_{sink})	(50,50)
Initial energy of nodes	E_0	1.5 J
Packet Length	L	4000bit
Power Energy	E_{elec}	50 nJ/bit
Data aggregation energy	E_{DA}	5 nJ/bit/message
Energy for free space	ε_{fs}	50 pJ/bit/m4
Energy for multi-path	ε_{amp}	0.13 bit/m4

5.2.1. The Effect of the Weighting Coefficients (A&B)

As cited in formula 9, the sum of α & β is 1. Thus, several simulations with possible combinations reveal the best simulation that benefits the FND and Energy Efficiency parameters. CH selection depends on the nodes' residual energy and mobility degree, which could govern their closeness to the Cluster Center. The node localization in the network depends heavily on the Mobility Degree because the position of nodes becomes dynamic. The results endorse this notion; the best energy weighting parameter value (α) equals 0.2 and the degree of mobility is 0.8, as seen in Figure 2. This combination yields an efficient result because of the low energy consumption and high E value (851.611) due to the significant number of packets received by the BS. Moreover, FND occurs much later, around 3549.
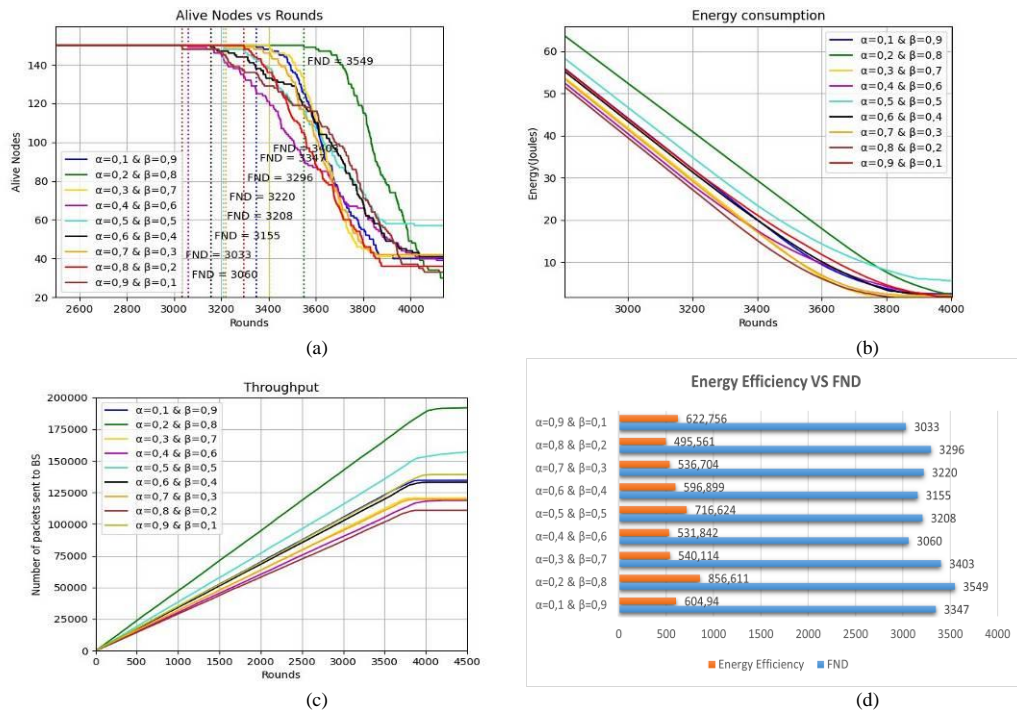

Figure 2. The effect of the weighting coefficients: (a) The number of alive nodes, (b) Energy consumption, (c) Throughput, and (d) Energy efficiency.

5.2.2. The Influence of Nodes Velocity

Here, we have V1, V2, V3, and V4 which are (0.2, 5) m/s, (5, 10) m/s, (10, 15) m/s, and (15, 25) m/s, respectively. The results show that V2 has the highest FND, about 3566 per 150 nodes, as shown in Figure 3, but with the lowest E (Energy Efficiency) and the smallest number of packets received by the base station. The total energy consumption is shown in Figure 3b, and the nodes with V2 speed consume much less, but those with high speed (V3) perform better in terms of throughput, as shown in Figure 3c. Hence, the suggested algorithm performs efficiently and much better with high velocities that do not exceed 15 m/s if based on Figure 3d. This result explains the high value of the weight coefficient β (0.8) f found previously.
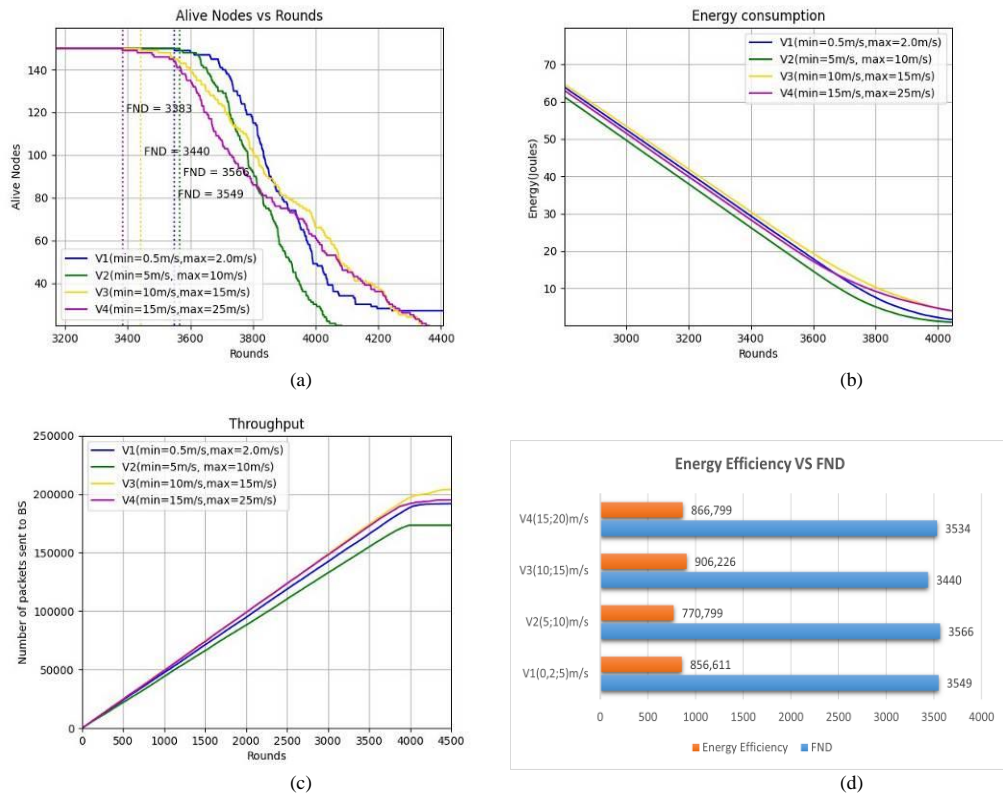
Figure 3. The influence of node velocity on: (a) the number of alive nodes, (b) energy consumption, (c) throughput, and (d) energy efficiency.

5.2.3. The Scalability

We raise the number of nodes from 50 to 1000 and observe the stability period oscillating between 2792 and 3549 rounds. The Half Node Death (HND) occurs at around 4000 rounds, as shown in Figure 4. Hence, our approach can be used efficiently for small and medium-scale networks. However, on a large scale, multiple nodes would cause many computations that could disturb BS function as our approach is centralized. This issue is worth being addressed in future work.
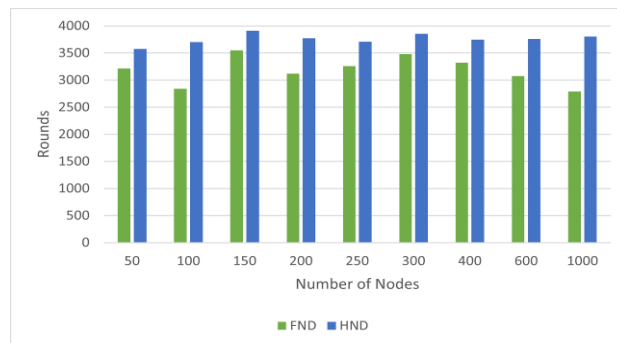

Figure 4. The FND for different node densities.

6. COMPARATIVE ANALYSIS

In this section, we first compared the proposed protocol with the benchmark OK-means and the improved version (Improved OK-means) presented in [10]. We focus on the stability period, energy consumption, throughput, and energy efficiency in a mobile context.

While CH selection using the OK-means algorithm relies on the proximity to the cluster center, the improved version of the OK-means algorithm considers residual energy in addition to the node distance to the center. The results shown in Figure 5a reveal a much better stability period, proving that our algorithm outperforms its predecessors. For instance, improved OK-means FND occurs after 3181 rounds, whereas OK-means algorithm FND happens after 2389 rounds. Here, the delayed death of the first node is due to CHs rotation. Indeed, when we evaluated the CH selection with the membership weight based on residual energy, mobility degree, and distance to the center as our proposed approach, the FND got much better (3549) in the same density network (150 nodes). Therefore, our algorithm tops the OK-means and Improved OK-means due to the weight of the potentially elected node considering residual energy and mobility. After the stability period, the sensors die much more slowly than in the case of the Improved OK-means, which is explainable by the fact that after the stability period, almost all nodes are CH and tired and thus begin dying together. The FND has indeed been delayed, but the network tends to collapse more quickly as soon as the instability phase starts. Figure 5b shows higher energy consumption in EEIOK-means compared to OK-means and Improved OK-means due to mobility, and Figure 5c shows a high throughput. Mobility increases the lifetime, throughput, and energy efficiency, as shown in Figure 5d.

Next, we make a second comparison under the same simulation conditions, i.e., node mobility and homogeneity. Understanding that the weight function evaluates the residual energy in each cluster member (CM) while considering their mobility degree, the nodes close to the center are good candidates if they still have almost fully charged batteries. Thus, there is a load balance between sensors over a very long period, delaying the death of the node. EEMCS and MCCA are energy-efficient, centralized, mobile-aware, clustering, and homogeneous routing protocols as our proposed algorithms. EEMCS uses a weight function to select the CH that considers mobility level, energy level, neighbour density, and distance to the BS. MCCA uses the same concept but ignores the number of connected nodes as we did in our proposed algorithm. The weight function aims at an optimal CH choice, a CH participating in the load balance of the network. Reinforced by a rotation mechanism in each cluster for CH election, our method surpasses EEMCS and MCCA in terms of the number of alive nodes and FND, and the simulations are done with 100 nodes. Following the comparison works, the simulations are performed with initial energies of 0.1 J in EEMCS and 2 J in MCCA and data lengths of 4000 and 500 bits in EEMCS and MCCA, respectively. The results are exhibited in Figures 6 and 7 where we compare the lifetime in our approach to EEMCS and MCCA, respectively. Our algorithm is much better than the previous protocols in terms of FND, even with a lower density due to the rotation mechanism. The first CH elected in the cluster has the minimum weight, and then all CMs will be elected in turn, delaying the death of the first node and extending the lifetime.
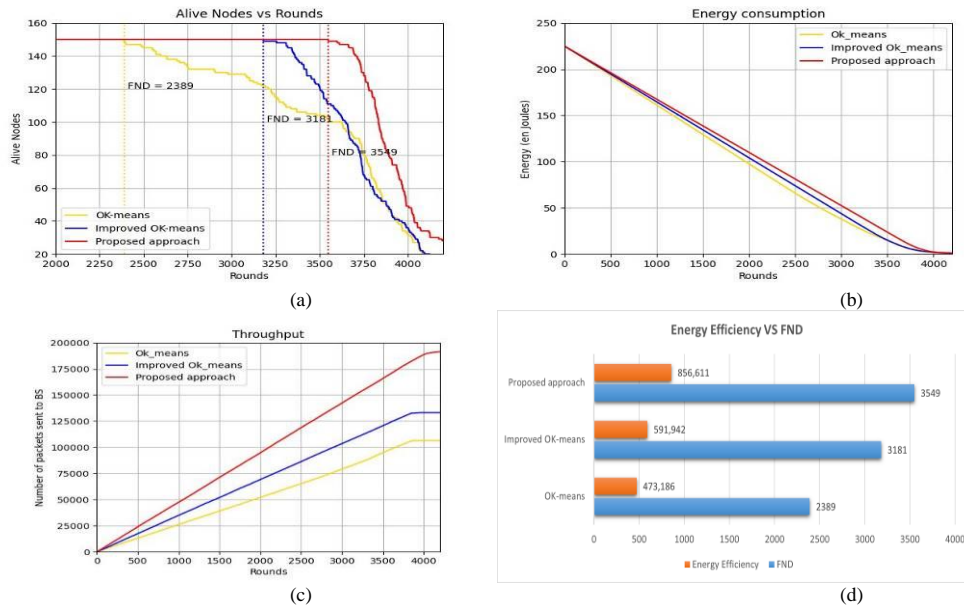
Figure 5. Comparison of the proposed algorithm to OK-means and Improved OK-means: (a)The number of alive nodes, (b)Energy consumption, (c)Throughput, and (d)Energy efficiency.
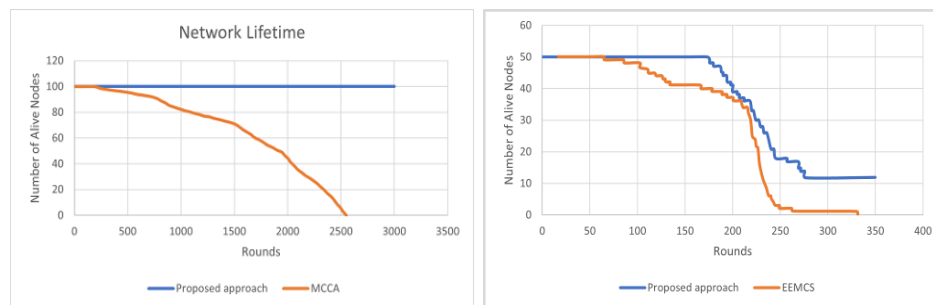

Figure 6. Comparison between the proposed approach and recent works in terms of the lifetime: (a) Lifetime in MCCA and EEIOK-means, (b) Lifetime in EEMCS, and EEIOK-means.

Figures **7a** and **7b** show energy consumption in our approach compared to other protocols, revealing that it occurs more slowly and proves CH choice effectiveness. This finding is attributable to CH selection which consumes much energy in clustering; hence, a good choice will minimize it. Our approach responds well to this theory for two reasons: First, we used clustering once, which allowed us to reduce the number of exchanges between the nodes and thus minimize the delay. Secondly, we considered the weight of each node based on its residual energy, location in the cluster, and mobility step.
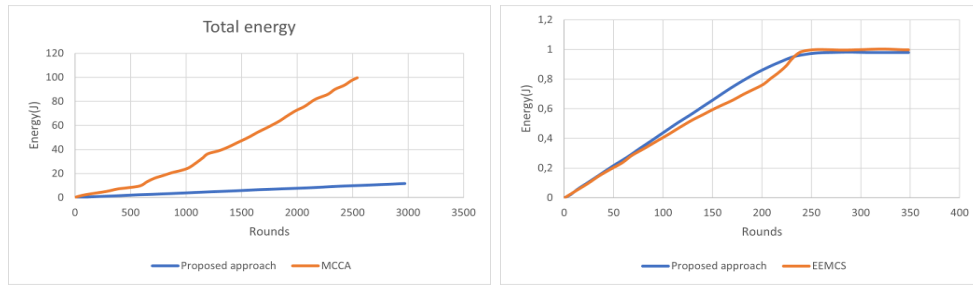
Figure 7. Comparison between the proposed approach and recent works in terms of energy consumption: (a)Energy consumption in MCCA and EEIOK-means, (b) Energy consumption in EEMCS and EEIOK-means.

7. DISCUSSIONS AND LIMITATIONS

Given the centralized nature of our proposed approach, the BS performs all cost function computations for CH selection, which requires full knowledge of the network topology or information about the amount of the remaining energy, or both. Accordingly, extra work is due regularly or when a cluster head fails. The key idea in this work is to minimize this cost by avoiding re-clustering and implementing an automatic rotation mechanism in each cluster to guarantee optimal CH selection. This approach suits large-scale WSNs, as the base station has more resources than the sensor nodes. Based on [10] results, the centralized hierarchical clustering outperforms hybrid clustering in mobile settings, we apply it to mobile nodes wherever distributed architecture remains the best choice considering energy efficiency. We plan to scale to intelligent sensors, use Q-learning for a more efficient routing protocol, and let nodes be autonomous as they must keep records of every other node in their K-hop neighborhood, such as the remaining energy, node velocity, node mobility, or other metrics. Another issue is the brief period after the death of the first node, which means that the network lifetime is practically equal to the FND. Hence, we intend to use the machine learning technique to make the network operational until the death of the last node.

8. CONCLUSIONS

In the mobile context, the energy efficiency of the network is significantly affected by CH selection. The main contribution of this work lies in providing a cluster selection technique that maximizes network lifetime. Indeed, the CH selection consumes a lot of energy, and an ideal option will minimize this consumption. The proposed protocol is based on residual energy, distance to the cluster, and mobility step. The strategy consists in maximizing the first dead node lifetime. It allows a maximum number of nodes to be involved in a CH. We found that the stability of the network has significantly improved and that the nodes die successively and quickly just after the stability period. This result is caused by the fact that almost all the nodes have fulfilled the role of CH, and thus, they will be exhausted shortly. Our advised method (EEIOK-means) improves the lifetime by 48% better than OK-means and 11\% compared to Improved OK-means, 80% than OK-means, and 43% than Improved OK-means in terms of throughput, 81% more than OK-means and 44% compared to Improved OK-means in terms of energy efficiency. The cost function used to choose the CH and node mobility ensures that the selected nodes are as close as possible to the base station, lowering energy consumption. Besides, as throughput and energy consumption are inversely proportional, the ratio between them increases. Thus, the results of our methodology in a rise of 44% in energy efficiency compared to Improved OK-means. Besides, the proposed approach outperforms the current energy-efficient clustering routing protocols (EEMCS & MCCA) in FND. Our findings highlight the need for

further work in re-clustering to improve the instability period, which will increase latency. In cluster-structured WSNs, message exchange limits are required to preserve sensor energy. Then, we need to self-establish sensors in the network and let them learn the optimal routings methods while saving energy. Nowadays, reinforcement learning can be a promising field to meet our needs : minimize consumption and latency while increasing throughput.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1]   Gulati, K., Boddu, R. S. K., Kapila, D., Bangare, S. L., Chandnani, N., & Saravanan, G. (2022). A review paper on wireless sensor network techniques in Internet of Things (IoT). Materials Today: Proceedings, 51, 161-165.

[2]   Sara, G. S., & Sridharan, D. (2014). Routing in mobile wireless sensor network: A survey. Telecommunication Systems, 57, 51-79.

[3]   Nguyen, L., & Nguyen, H. T. (2020). Mobility based network lifetime in wireless sensor networks: A review. Computer networks, 174, 107236.

[4]   Soua, R., & Minet, P. (2011, October). A survey on energy efficient techniques in wireless sensor networks. In 2011 4th joint IFIP wireless and mobile networking conference (WMNC 2011) (pp. 1-9). IEEE.

[5]   Farahani, G. (2019). Energy Consumption Reduction in Wireless Sensor Network Based on Clustering. International Journal of Computer Networks & Communications (IJCNC) Vol, 11.

[6]   Yetgin, H., Cheung, K. T. K., El-Hajjar, M., & Hanzo, L. H. (2017). A survey of network lifetime maximization techniques in wireless sensor networks. IEEE Communications Surveys & Tutorials, 19(2), 828-854.

[7]   Kulkarni, R. V., Förster, A., & Venayagamoorthy, G. K. (2010). Computational intelligence in wireless sensor networks: A survey. IEEE communications surveys & tutorials, 13(1), 68-96.

[8]   Heinzelman, W. R., Chandrakasan, A., & Balakrishnan, H. (2000, January). Energy-efficient communication protocol for wireless microsensor networks. In Proceedings of the 33rd annual Hawaii international conference on system sciences (pp. 10-pp). IEEE.

[9]   El Khediri, S., Fakhet, W., Moulahi, T., Khan, R., Thaljaoui, A., & Kachouri, A. (2020). Improved node localization using K-means clustering for Wireless Sensor Networks. Computer Science Review, 37, 100284.

[10]   Chniguir, A., & Jemaa, Z. B. (2022, May). Lifetime maximization of the first dead node in a WSN. In 2022 International Wireless Communications and Mobile Computing (IWCMC) (pp. 378-383). IEEE.

[11]   Zafar, S., Bashir, A., & Chaudhry, S. A. (2019). Mobility-aware hierarchical clustering in mobile wireless sensor networks. Ieee Access, 7, 20394-20403.

[12]   Zhu, B., Bedeer, E., Nguyen, H. H., Barton, R., & Henry, J. (2020). Improved soft-k-means clustering algorithm for balancing energy consumption in wireless sensor networks. IEEE Internet of Things Journal, 8(6), 4868-4881.

[13]   Kim, D. S., & Chung, Y. J. (2006, June). Self-organization routing protocol supporting mobile nodes for wireless sensor network. In First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS'06) (Vol. 2, pp. 622-626). IEEE.

[14]   Bholowalia, P., & Kumar, A. (2014). EBK-means: A clustering technique based on elbow method and k-means in WSN. International Journal of Computer Applications, 105(9).

[15]   Zhang, J., & Yan, R. (2019). Centralized energy-efficient clustering routing protocol for mobile nodes in wireless sensor networks. IEEE Communications Letters, 23(7), 1215-1218.

[16]   Umbreen, S., Shehzad, D., Shafi, N., Khan, B., & Habib, U. (2020). An energy-efficient mobility-based cluster head selection for lifetime enhancement of wireless sensor networks. Ieee Access, 8, 207779-207793.

[17]   Wang, S., Yu, J., Atiquzzaman, M., Chen, H., & Ni, L. (2018). CRPD: a novel clustering routing protocol for dynamic wireless sensor networks. Personal and Ubiquitous Computing, 22, 545-559.

[18]   Mahmood, D., Javaid, N., Mahmood, S., Qureshi, S., Memon, A. M., & Zaman, T. (2013, October). MODLEACH: a variant of LEACH for WSNs. In 2013 Eighth international conference on broadband and wireless computing, communication and applications (pp. 158-163). IEEE.

[19]   Aydin, M. A., Karabekir, B., & Zaim, A. H. (2021). Energy efficient clustering-based mobile routing algorithm on WSNs. IEEE Access, 9, 89593-89601.

[20]   Smaragdakis, G., Matta, I., & Bestavros, A. (2004, August). SEP: A stable election protocol for clustered heterogeneous wireless sensor networks. In Second international workshop on sensor and actor network protocols and applications (SANPA 2004) (Vol. 3).

[21]   Belabed, F., & Boouallegue, R. (2017, June). Clustering approach using node mobility in wireless sensor networks. In 2017 13th International wireless communications and mobile computing conference (IWCMC) (pp. 987-992). IEEE.

[22]   Wang, J., Gu, X., Liu, W., Sangaiah, A. K., & Kim, H. J. (2019). An empower hamilton loop based data collection algorithm with mobile agent for WSNs. Human-centric Computing and Information Sciences, 9(1), 1-14.

[23]   Sadrishojaei, M., Navimipour, N. J., Reshadi, M., & Hosseinzadeh, M. (2021). A new preventive routing method based on clustering and location prediction in the mobile internet of things. IEEE Internet of Things Journal, 8(13), 10652-10664.

[24]   Rault, T., Bouabdallah, A., & Challal, Y. (2014). Energy efficiency in wireless sensor networks: A top-down survey. Computer networks, 67, 104-122.