AN EXTENDED K-MEANS CLUSTER HEAD SELECTION ALGORITHM FOR EFFICIENT ENERGY CONSUMPTION IN WIRELESS SENSOR NETWORKS

Peter Maina Mwangi, John Gichuki Ndia and Geoffrey Muchiri Muketha

School of Computing and Information Technology,
Murang'a University of Technology, Kenya

ABSTRACT

Effective use of sensor nodes' batteries in wireless sensor networks is critical since the batteries are difficult to recharge or replace. This is closely connected to the networks' lifespan since once the battery is used up, the node is no longer useful. The entire network will not function if 60 to 80% of the nodes in it have completely depleted their energy. In order to minimize energy usage and sustain the network for a long time, many cluster head selection algorithms have been developed. However, the existing cluster head selection algorithms such as K-Means, particle swarm selection optimization (PSO), Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and Fuzzy C-Means (FCM) cluster head election algorithm have not fully reduced the issue of energy usage in WSN. The objective of this paper was to develop an extended K Mean Cluster Head selection(CHS) algorithm that uses remaining energy, distance between node and base station, distance between nodes and neighbour nodes, node density, node degree Maximum Cluster size, received signal strength indicator (RSSI) and Signal to Noise Ratio. The algorithm developed was used to enhance the lifespan of WSNs. The performance of the simulated variants of LEACH routing protocols is measured and evaluated using the quantitative research methodology. Utilizing residual node energy, packet delivery ratio, throughput, network longevity, average energy usage, and the number of live and dead node, the suggested approach is contrasted to previous approaches. From the study we observed that the proposed approach outperforms existing actual LEACH, Mod-LEACH and TSILEACH approaches.

KEYWORDS

Cluster head, Cluster Head Selection Algorithms, energy efficiency, LEACH, Sensor node, Parameters, wireless sensor networks

1. INTRODUCTION

Wireless sensor networks (WSNs) are networks of interconnected sensor nodes that communicate wirelessly in order to gather data about their surroundings. Nodes are frequently low-power, ad hoc, and decentralized. Sensor nodes are commonly put in large numbers in WSNs, often in inaccessible locations [1]. This is because changing or charging the batteries in sensor nodes is difficult, and it is also critical to utilise the limited energy of these devices as efficiently as possible. As a result, limiting energy usage at each node to maximize network energy efficiency is one of the most important factors in WSN architecture. When more than a predetermined percentage of the network's nodes die, the network stops working.

One of the most effective methods for designing routing protocols in WSNs is to cluster the network. Tay and Senturk [2],claim that using a clustering strategy can greatly cut energy use. According to [1] and [2], they stated that ”energy usage is decreased by selecting the most appropriate sensor node as group leader based on standards defined within the clustered sensors”. Cluster head election algorithms are in charge of choosing group leaders. Several cluster head election techniques have been proposed[2], [3] and [4]in the deterministic, adaptive, and hybrid categories.

To improve on the classic cluster head election procedures, other cluster head election algorithms have been presented. These algorithms, which use artificial intelligence and adaptive data mining frameworks, includes hybrid optimization method for cluster head selection [5], Hybrid Firefly Algorithm with Particle Swarm Optimization [6] and Energy efficient Cluster Head Selection algorithm based on PSO (PSO-ECHS)[7]. The algorithms can distinguish and aggregate only the genuine values of the gathered information to the base station, which contributes to the elimination of redundant data and the reduction of power usage, thereby increasing the network lifespan of wireless sensor networks[8]. According to [7-11], [16] they stated that “although these algorithms are better than traditional algorithms they still have limitations including low network lifespan, high energy wastage, lack of adaption with heterogeneous networks, poor stability, node death, transmission delay, complexity in handling large-scale WSNs, inadequate consideration of remaining energy nodes, additional overhead and energy coverage, and unbalanced lifetime of nodes”.

One of the promising clustering techniques is the K-Means clustering approach. This is due to its ability to save more energy as compared to other cluster head selection algorithms. In addition, most of the members in K-Means are homogeneously clustered, meaning that most of the nodes in a group are close to each other, hence decreasing communication distance between nodes and cluster head [9]. Although the performance of K-Means is highly rated, there is a need to extend it to further improve energy efficiency. In this study, we proposed an extended K-Means cluster head selection(CHS) algorithm which considers parameters such as remaining energy, distance between node and base station, distance between nodes and neighbour nodes, node density, node degree, maximum cluster size, received signal strength indicator (RSSI), and signal to noise ratio.

The rest of the paper is organized as follows. Section two presents related works, section three presents the methodology, section four presents the extended K-Means cluster head selection algorithm, section five presents the results, section six presents the discussion, and section seven presents the conclusions and future works.

2. RELATED WORKS

In this section, we provide a detailed analysis of the existing cluster head election algorithms, showing how they select cluster heads and their limitations.

In [10], the particle swarm selection optimization (PSO) method is proposed for producing energy-aware clusters by selecting the cluster leaders in the best possible way. The primary benefit of PSO is that it eventually lowers the cost of determining the head nodes' ideal location within a cluster. The study took into account delays, travel distance, and energy usage, but still suffers from low network longevity.

Hybrid Firefly technique with Particle Swarm Optimization (HFA-PSO) [6]is another technique created to enhance the lifespan of wireless sensor network. HFA-PSO increases the number of alive nodes, reduces energy consumption, and improves the global search for fireflies in LEACH-C by utilizing PSO, resulting in optimal cluster head location. Although the HFA-PSO approach increases throughput and residual energy, it is limited in that it does not use the improved searching efficiency of the Hybrid Firefly Algorithm[11].

Pitchaimanickam et al., [6] proposes an Energy Efficient Cluster Head election algorithm based on PSO (PSO-ECHS). This approach has two phases that is group construction and cluster head election. The cluster leaders are chosen using PSO based on distance and leftover power. According to [6] he indicated that "to choose if a sensor node exceeds the threshold energy (i.e., the average energy of the sensor nodes) to qualify for a CH, all of the sensor nodes in the CH selection phase send their positions and remaining energy to the base station at the start of the phase. The PSO-based CH selection technique is then carried out by the base station, followed by the cluster building phase. It derives the weight function for cluster formation based on a number of variables, including distance, energy, and CH node degree". PSO-ECHS minimizes the use of energy and increases network life; nonetheless, its fundamental flaw is that it ignores WSN fault tolerance and energy balancing.

A CH selection algorithm (TabuPSO) based on hybrid PSO and Tabu Search (TS) extends network lifetime while balancing the use of energy [16]. Tabu-PSO CH choosing was deemed critical in order to extend network longevity. The benefits of TS method is that it is used to solve the problem of local optima in PSO-based CH selection. The suggested method increases network lifetime, optimizes routing, and chooses effective cluster heads. Comparing the proposed Tabu-PSO method to multi-hop LEACH with a large number of nodes, the regular packet loss rate was decreased by 27.32% on average. In future there is need to enhance the algorithm further to improve on energy efficiency and enhance fault tolerance.

The Density-Based Spatial Clustering of Applications with Noise (DBSCAN) technique clusters together closely packed data points. To build a dense zone, the algorithm requires a distance measure (eps) and a minimum number of points (minPts). It begins by randomly selecting a data point, and if it is in a dense zone, it groups together any other data points that are within eps of it and have at least minPts of data points within eps[12]. The DBSCAN approach identifies clusters within huge spatial datasets using a single input parameter that takes into account the local density of its constituent elements. Furthermore, the user is given a suggested parameter value, requiring little to no domain expertise. DBSCAN's aim is to group the nodes into distinct clusters and ultimately define the various classes [12]. Its disadvantage is that its performance may degrade if the data is not uniformly distributed owing to its sensitivity to data density and distribution.

Another clustering approach, Fuzzy C-Means (FCM), is proposed [13] Based on location information, the base station computes and distributes sensor nodes into groups, with the cluster leader awarded to the node with the highest remaining energy. FCM grouping methods are strategies for centralized clustering[13]. FCM is a K-Means variant that use fuzzy logic to assign each data point to many clusters with varied degrees of membership. Each data point is assigned a membership degree, indicating how much it belongs to a given cluster. Because data points can belong to more than one cluster, clustering becomes more flexible and robust. The authors found that the FCM algorithm outperformed the classical K-Means algorithm in terms of cluster validity and robustness to noise.

Regional Energy Aware Clustering with Isolated Nodes (REAC-IN) suggested by Lue et al.[14] addresses the issue of node segregation while also increases the longevity and stability of the network. This system has the advantage of being a weighted-based cluster head selection technique that decreases the energy consumption of isolated nodes. Its limitation is that isolated nodes use more energy while communicating with preceding CH nodes.

In 2021, Shyjith et al. [5] suggested hybrid optimization method for CH election. The proposed CH selection consists of three stages: preparation, transmission, and measurement. The energy is being initialized, and the network's nodes are being moved. According to. [5] “the threshold and CH are determined under multi-objective constraints that take into consideration delay, energy, and distance. Before data transfer from CHs to BS commences, the CHs are detected. The leftover power produced by the nodes is eventually updated during the measurement phase”. This algorithm's shortcoming is that it does not take into account cost metrics.

K-Means Cluster head selection algorithm [15] was developed. K-Means grouping approach is the most basic algorithm for unsupervised clustering. This algorithm will split the data set into k groups using the Euclidian distance, maximizing intra-group similarity and minimizing inter-cluster similarity. The nature of k-means is iterative as suggested by [15]. The benefit of K-Means is that it reduces re-clustering and boosts the packet delivery ratio of sensor nodes. Its shortcoming is that various initial partitions may result in multiple end clusters[17].

Table 1. Comparison of various existing cluster head selection algorithms

Algorithms	**Parameters**	**Merits**	**Demerits**
PSO	Residual energy, intra-cluster distance, node degree, head count of the probable cluster heads	It lowers the cost of determining the cluster head nodes.	Suffers from low network longevity.
PSO-ECHS	intra-cluster distance, sink distance and remaining energy	It minimizes the use of energy and increases network life.	It ignores WSN fault tolerance and energy balancing.
DBSCAN	Distance measure (eps) and a minimum number of points (minPts).	Does not require one to specify the number of clusters beforehand. Performs well with arbitrary shaped clusters.	Its performance may degrade if the data is not uniformly distributed owing to its sensitivity to data density and distribution.
REAC-IN	Residual Energy, Number of alive nodes, Number of data received, Average lifetime	Weight-based CH selection. Reduces the energy consumption of isolated node	Its limitation is that isolated nodes use more energy while communicating with preceding CH nodes.

K-Means Cluster head selection algorithm	Residual energy	If we have large number of variables then, K-means would be faster than Hierarchical clustering. On re-computation of centroids, an instance can change the cluster. Tighter clusters are formed with K-means as compared to Hierarchical clustering.	Clustering data of varying sizes and density. It Chooses K manually.

3. METHODOLOGY

3.1. Simulation Environment

To design the proposed cluster head selection algorithm, we used step by step procedures and flowcharts to show cluster heads are selected. We set up simulation experiments using MATLAB R2017a to evaluate the proposed extended K-Means cluster head selection algorithm and compare it with other existing cluster head selection algorithms.

A wireless sensor network of 100 sensor nodes was spread in a 100m x 100m field, with each node having an initial energy of 0.5J. Meters are the units of measurement for X and Y. Table 2 summarizes the simulation parameters.

Table 2. Simulation Parameters

Parameters	**Values**
Sensor deployment area	100 M*100 M
Base Station Location	50M * 50M
Number of nodes	100
Data packet size	100 bytes
Control packet size	25 bytes
Initial energy	0.5J
Maximum number of rounds	2000
Aggregated packet size from cluster head	500 bytes
Electronics energy	50nJ/bit
Free space factor	10, 255 pJ / bit / m2
Multipath factor	0.0013, 0.0050, 0.0063 pJ / bit / m4

A visualization of the simulation parameters of 100 nodes and a base station that are randomly placed in geographical location of X and Y coordinates measured in meters are shown in Figure 1.

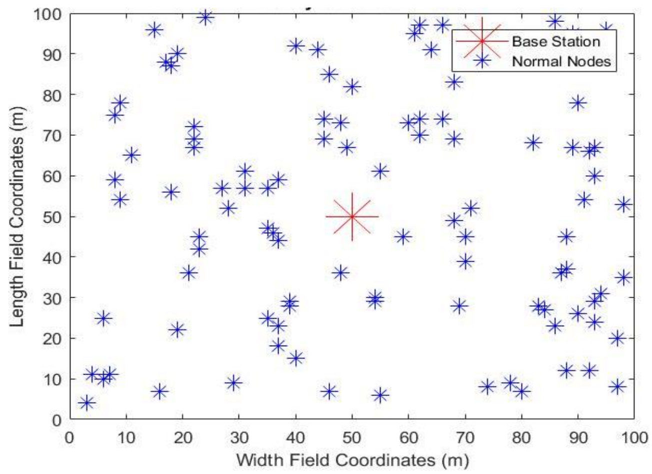

Figure 1. Simulation setup

3.2. Performance Metrics

We used several performance metrics to evaluate the performance of the various cluster head algorithms implemented in LEACH routing protocols, including number of live nodes, number of dead nodes, average energy consumed, network lifetime, throughput, number of packets sent to CH, and remaining energy. These performance metrics have been used by other researchers to compare cluster head selection algorithms [16]. These metrics are described in Table 3.

Table 3. Performance metrics

Metric	Description
Number of live nodes in the network	The amount of nodes that are still alive determines this metrics. A large number of active nodes improves the network's efficiency.
Number of dead nodes in the network	This is given by the number of nodes that have exhausted their energy. The efficiency of the network is determined by fewer number of dead nodes.
Average energy consumption:	This represents the total energy used by nodes during data transmission and reception.
Network lifetime	The period of time during which the network is not ended is known Network lifetime. However, it has also been described as the period of time between the death of the first and last node. More data is transferred as the network remains operational for longer
Throughput:	The typical amount of packets the base station and receives per round.
Packet delivery ratio	The ratio of packets received by a destination node (R1) to packets generated by a source node (R2).

4. PROPOSED EXTENDED K-MEANS CLUSTER HEAD SELECTION ALGORITHM (EKCHS)

The extended k-means cluster head selection algorithm is unsupervised machine leaning algorithm that partitions WSN nodes into clusters using information datasets collected from the nodes. It uses k, which user-specified parameter that will define the number of clusters in a WSN. We used three steps to calculate cluster heads. These steps include:

1) Formation of k centroids using K-Means clustering in WSN;
2)Assignment of live nodes to their clusters; and
3) Cluster head selection within k formed clusters from K-means centroids.
We provide a more detailed description of these three steps in the subsequent sections.

4.1. Forming k Centroids using K-Means Clustering

The first step is forming of k centroids using K-Means. This step entails formulating k centroids using K means clustering this is done by partitioning the WSN into k Clusters. Initial value of k is set and then the algorithm starts by initializing k cluster centroids at random data points. Then, based on a distance measure such as Euclidean distance, each data point is given to the group whose centroid is closest to it. After that, the group centroids are recalculated as the mean of the data points in each cluster. As demonstrated in Algorithm 1, the process of reallocating data points to groups and updating the group centroids is repeated until convergence, which happens when the cluster assignments can no longer change, occurs.

Algorithm 1. Formulating k Centroids using K-Means

Step 1. Filter all live nodes in each round
Step 2. Set the initial value of k
Step 3. Calculate an (x, y) for all filtered live node in each round
Step 4: Use K-Means to generate centroids with reference live node coordinates
Step 5. Repeat step 3 and 4 until convergence
Step 6. Print centroids and clusters

4.2. Assigning Live Nodes to their Clusters

The next step is to initialize live nodes to their clusters using centroids generated in the previous step. To assign live nodes to their clusters, we used the steps shown in Algorithm 2.

Algorithm 2. Assigning live nodes to their clusters

Step 1: Loop through all nodes using a for Loop
Step 2: resetting the minimum distance of all nodes node to centroids to zero
Step 3: Loop through the centroids (centres) and assign the node to cluster whose centroid is closest to the node.
Step 4; Display the nodes and centroid in each cluster
Step 3: Cluster Head election within K formed clusters using K-Means

4.3. Cluster Head Election using K-Means Centroid

The final step is to elect cluster leaders, evaluate the chosen cluster leaders and create a transmission schedule. To elect cluster head using K-Means centroids, we used the steps shown in Algorithm 3.

Algorithm 3. Cluster head election using k-means centroid

Step 1: check node energy whether it is greater than zero (node is live)
Step 2: calculate Euclidian distance between the node and k means centroid
Step 3: compare distances of all the nodes in the cluster.
Step 4: select the node with minimum distance and other parameters
Step5 assign a node that meets step 4 as cluster head

These three algorithms defined above are combined to form the proposed extended K-Means cluster head selection algorithm (Algorithm 4).

Algorithm 4.Proposed extended K-means cluster head selection algorithm.

Step 1: Selection of the values of k clusters
Step2: Set the Random Initial Centroids C
Step3: Assigning Live nodes to their Clusters
Step4: Re-compute the centroid for each cluster
Step5: Repeat step 4 Until convergence (until the centroid don't change)
Step6: Select Cluster head
Step7: Evaluate cluster head

The section below explains each step for the extended K-means cluster head selection algorithm (algorithm 4).

Step 1: K = Selection of the value of k clusters

First the value of k is elected and gets pre-determined. Moreover, the value can be pre-determined randomly set if the anticipated number of clusters is well known.

Step 2: Set the Random Initial centroids

C = {c 1, c 2, c 3 …………………….c n} set of centroids. This is set to mark the initial locations of the centroids. At the end, iterative calculations are performed to optimize the location of the centroids. The operation is halted once the set number of iterations has been completed, or there is no change in the locations of the centroids; the centroids have stabilized.

Step 3: Assigning Live nodes to their Clusters

D = {n 1, n 2 , n 3 ……………………..nx } set the nodes to be used in learning the algorithm. Conventionally, the processed dataset is used with specific set number of data points. A node is allocated to a specific cluster based on node information (Remaining energy, distance between node and base station, distance between nodes and neighbour nodes, node density, node degree Maximum Cluster size, received signal strength indicator (RSSI) and Signal to Noise Ratio) and cluster centre distance. The Euclidean distance is used to measure the distance between nodes and centroid in each round. We used below formula:

$$d = \sqrt{[(x_2 \; - x_1)^2 + (y_2 \; - y_1)^2]}$$

Where d is the Euclidian distance between nodes and centroid and (x1,y1) and (x2, y2) are coordinates of point one and point two respectively (distance between node and centroid).
Step 4: Re-compute the centroid for each group- The cluster centroids are then recalculated as the mean of the data points in each cluster.

$$J(C) = \sum_{J=1}^{K} \sum_{i=1}^{n} \left\| d_i^{(j)} - c_j \right\|^2$$

Where the di -Cj is the Euclidean distance, k is the number of clusters and n is the number cases or observations. The smallest distance is preferred in nodes assignment. The new cluster centre is recalculated using the equation below, and the process is stopped after the reassignment.

$$c_i = \frac{1}{n_1} \sum_{j=1}^{n} d_i$$

Step 5: Repeat - The above steps are repeated until convergence, which occurs when the cluster assignments no longer change.

Step 6: Cluster head selection - The algorithm choses the node near the centroid that meets parameters such as remaining power, distance between node and base station, distance between nodes and neighbour nodes, node density, node degree Maximum Cluster size, received signal strength indicator (RSSI) and Signal to Noise Ratio in each cluster as the cluster head

Step 7: Evaluation - The final step is to evaluate the group leaders (point near the centroid). One popular method is to select the data point with the smallest distance to the centroid as your point near the centroid.

The main advantage of this algorithm is that it clusters and selects nodes based on remaining power, distance between node and base station, distance between nodes and neighbour nodes, node density, node degree, Maximum Cluster size, received signal strength indicator (RSSI), and Signal to Noise Ratio. This guarantees that group leaders are evenly spread throughout the network and have higher energy levels, resulting in more balanced network energy usage. This algorithm also takes other factors into account, such as coordinates. The flow chart shown in Figure 2representsthe proposed algorithm.

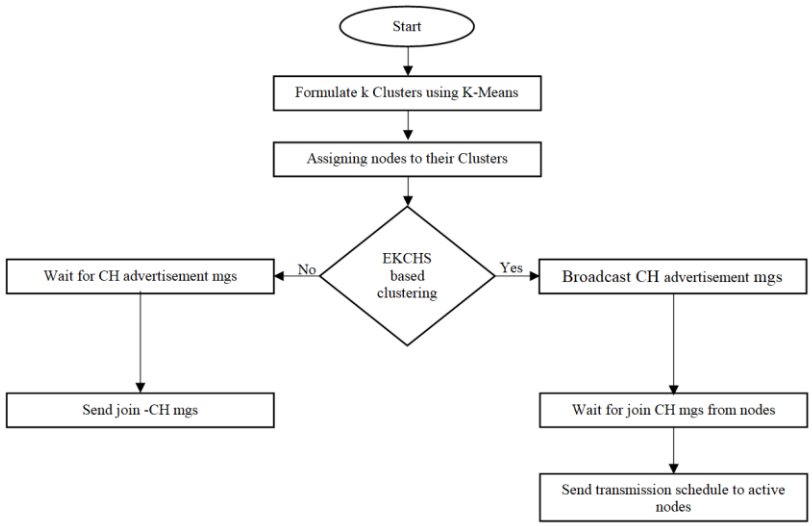

Figure 2. A flowchart for extended K-Means cluster head selection algorithms (EKCHS)

5. RESULTS

In this section, we discuss our results based on the different metrics for evaluating the performance of the proposed Extended K-Means cluster head selection algorithm with other CHS algorithms used in other variants of LEAH.

To depict how nodes are randomly placed with k-means centroid marks. The results display how 100 nodes are randomly arranged in geographical location of X and Y coordinates, where the blue asterisk represents the normal nodes, red asterisk shows the base station and green circle are the centroid markers as shown in Figure 3.

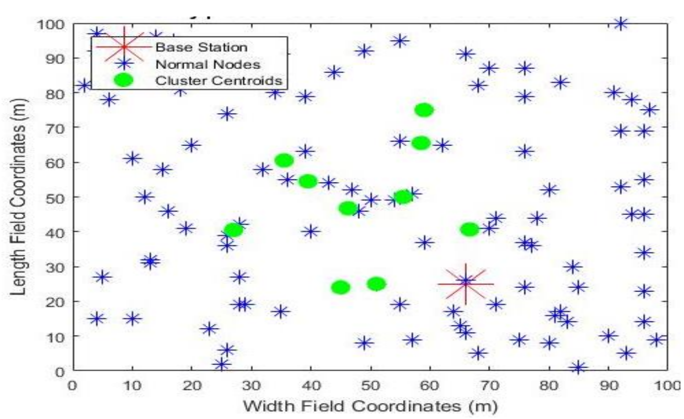

Figure 3. Randomly placed nodes with K centroid markers

To show how clusters are formulated using K-means centroid markers. The results display how the proposed extended K-means CHS algorithm subdivides the network into clusters using K-means centroid markers, where green circles represent cluster centroid and pink borders represents cluster borders as shown in Figure 4.

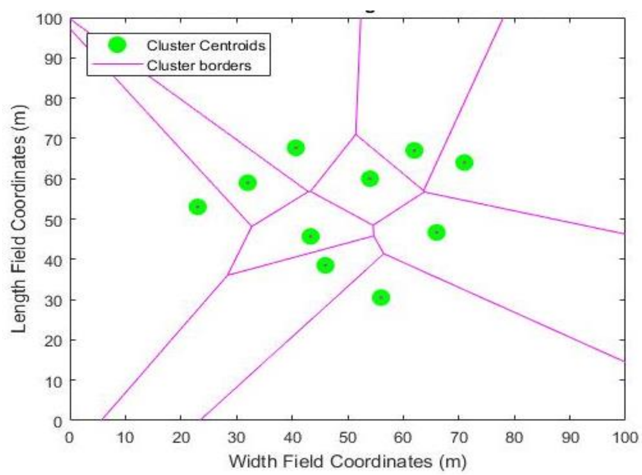

Figure 4. Cluster formulation using extended K-Means centroid

To show how cluster head are elected in the proposed algorithm. The Figure 5 shows normal nodes represented by blue asterisk, the elected cluster heads in each cluster represented by blue circle and their centroids markers represented by green circle. The proposed algorithm elects' node that are near to the centroid markers to be cluster heads in each cluster at a particular round. As the cluster head, the selected cluster leader must meet parameters such as remaining power, distance between node and base station, distance between nodes and neighbour nodes, node density, node degree, Maximum Cluster size, received signal strength indicator (RSSI), and Signal to Noise Ratio.

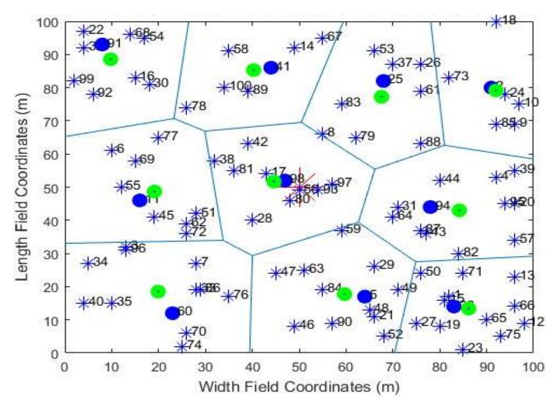

Figure 5. Cluster head election using Extended K-Means cluster head selection algorithm (EKCHS)

To evaluate the proposed model in terms of cluster head election, we used actual Leach, MOD-Leach and TSI-Leach which are some of the existing variants of Leach.

The graph observations reveal that the number of nodes that have died in the suggested work is smaller than in the actual Leach, MOD-Leach, and TSl-Leach. The network lifetime decreases as the number of dead nodes increases. The number of dead nodes in the network at the end of the simulation is shown in Figure 6.

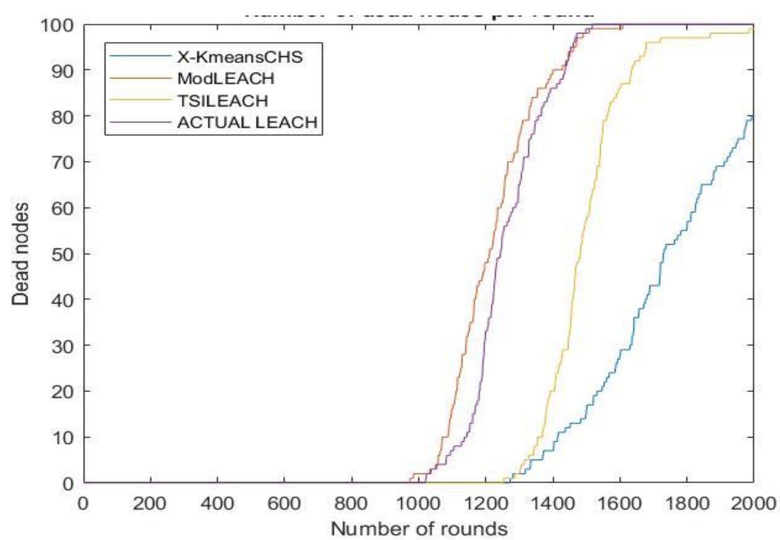

Figure 6. Number of dead nodes per round in Extended K-Means CHS, MOD-Leach, TSI-LEACH and Actual Leach

It is evident that the first node in Extended K-Means CHS algorithm died at approximately 1300 round, MOD-Leach approximately 950 round, TSI-Leach approximately 1250 round and actual Leach 1050 round respectively. At the end of simulation, it is evident that the number of dead nodes in Extended K-Means CHS was 75 and TSI-Leach was 95 respectively while MOD-Leach and actual Leach all the nodes had died. This information is shown in Table 4.

Table 4. Number of dead nodes in the network

Round	**Extended K-MeansCHS**	**Actual Leach**	**MOD-Leach**	**TSI-Leach**
200	None	None	None	None
400	None	None	None	None
600	None	None	None	None
800	None	None	None	None
1000	None	None	5	None
1200	None	35	45	None
1400	5	85	87	26
1600	30	All dead	95	85
1800	60	All dead	All dead	95
2000	75	All dead	All dead	95

At the end of simulation, it was observed that there were at least 25 alive nodes in the proposed extended K-Means CHS. In actual Leach and MOD-Leach all the nodes were dead while in TSI-Leach approximately 7 nodes were alive at the end of simulation. The graph observations reveal that the number of live nodes in the proposed work is more than that in the classic Leach, Mod-Leach, and TSI-Leach. The greater the number of active nodes, the greater the network lifetime. The number of live nodes in the network is represented in Figure 7.

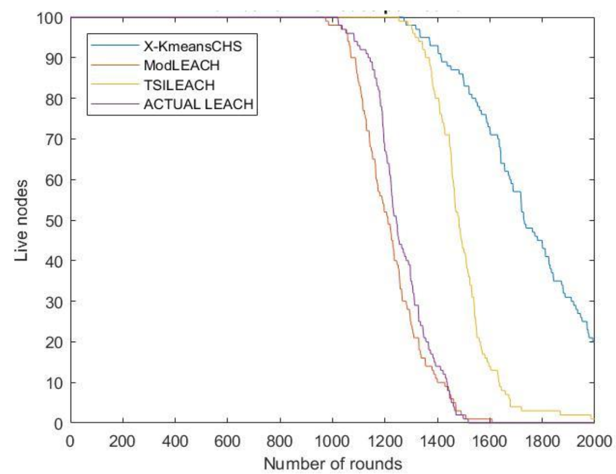

Figure 7. Number of live node per round in proposed Extended K-Means CHS, MOD-Leach, and TSI-Leach and Actual Leach

The four key components of node energy wastage are data transmission, data reception, data fusion, and group negotiation communication. It is observed that at 1000 round the total remaining energy in WSN for the proposed extended K-Means CHS was approximately 25 joules, TSI-Leach followed with approximately 20 joules, MOD –Leach had approximately 15 joules and actual Leach had approximately 17 joules. A comparison of the remaining energy and number of round in Wireless Sensor Network is demonstrated in Figure 8.

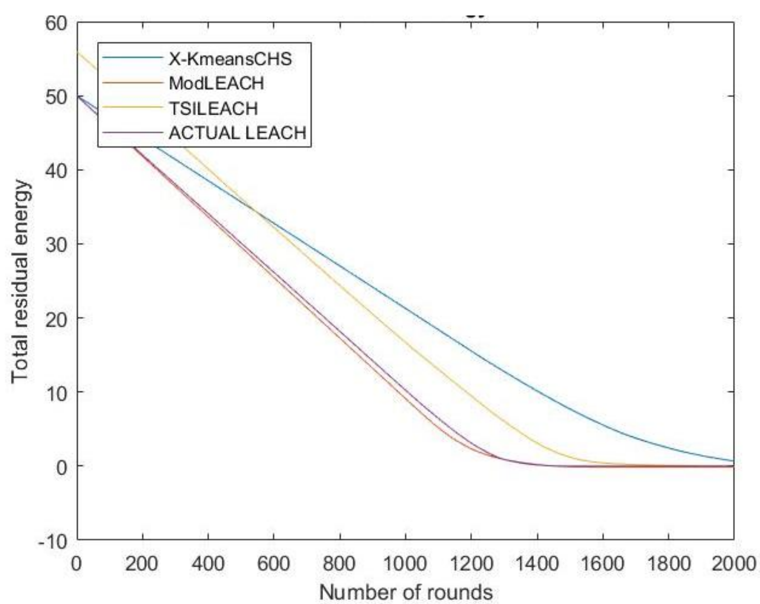

Figure 8. Average remaining energy after a number of rounds extended K-Means CHS, Mod-Leach, TSI-Leach and Actual LEACH

As shown in Figure 8,the energy usage in proposed Extended K-Means CHS is more than MOD-Leach, TSIL-Leach and actual Leach. In terms of the total remaining energy it is observed that at 1000 round the total remaining energy in WSN for the proposed extended K-Means CHS was approximately 25 joules, TSI-Leach followed with approximately 20 joules, MOD–Leach had approximately 15 joules and actual Leach had approximately 17 joules. At the end of simulation, the was approximately 2 joules remaining in proposed K-Means CHS algorithm and all the energy was exhausted in Mod-Leach, TSI-Leach and actual Leach respectively.

After simulation, it is discovered that the energy of the network lifespan of the proposed algorithm runs approximately 2000 communication rounds, and some of the nodes in the network have not exhausted their energy. There are some nodes alive at the end of simulation which is about 14%of the entire network. Figure 9shows dead nodes and remaining live nodes after the final round of communication. Where dead nodes are the nodes with black dots and blue stars are the live nodes.

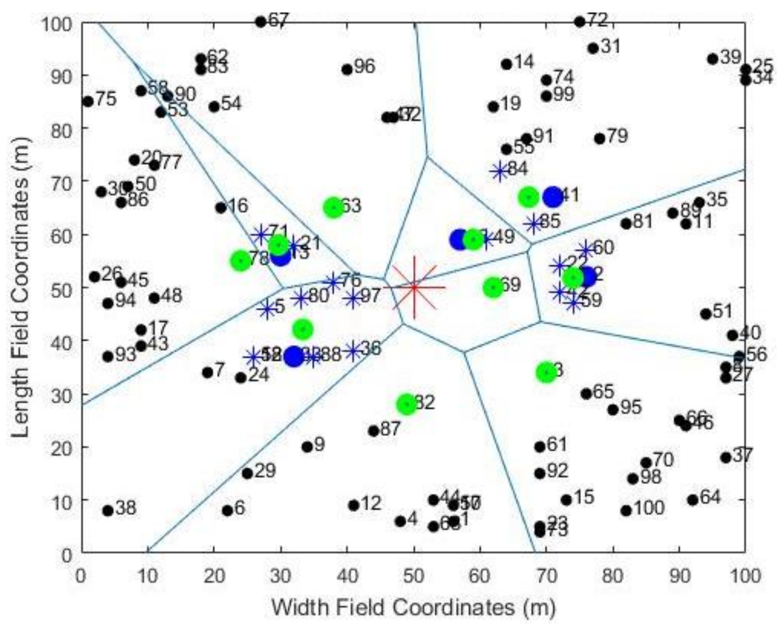

Figure 9. Dead nodes vs live nodes in final round of simulation

At around 1000 rounds it is observed that there were approximately $10*10^4$ total data packets delivered to the CH in Extended K-Means CHS while in Mod-Leach approximately $7.5*10^4$, TSI-Leach $9*10^4$ and actual Leach $7.8*10^4$ respectively. It can be observed that there were more packets that were forwarded to CH in our proposed algorithm compared to actual Leach, MOD-Leach and TSI-Leach. The results for the total data packet sent from nodes to CH of the four presented cluster head election algorithms are shown in Figure 10.

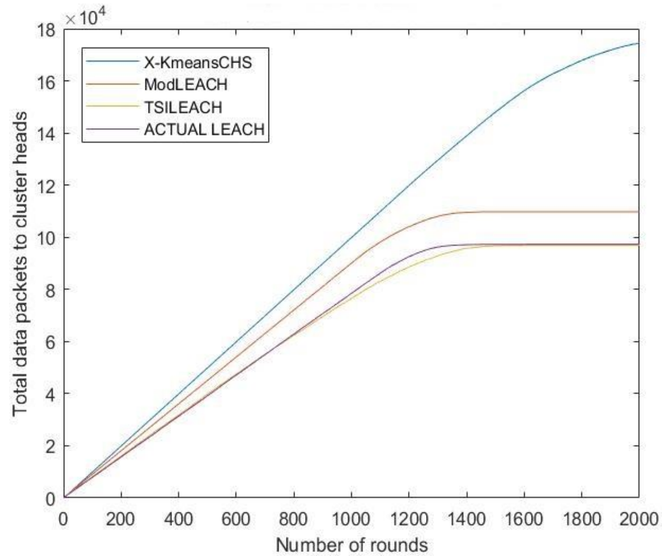

Figure 10. Total Data Packets to Cluster Heads for extended K-Means CHS, MOD-Leach, TSI-Leach and Actual Leach

In terms of throughput, the extended K-Means CHS did better comparing it with, MOD-Leach, TSI-Leach and Actual Leach as shown in Figure 9, where at the end of simulation proposed extended K-Means cluster head selection algorithm had a total throughput $2*10^5$, actual Leach $1.2*10^5$, Mod-Leach $1.1*10^5$and TSI-Leach $1.5*10^5$.The throughput performance of the suggested method and other conventional systems are shown in Figure 11.

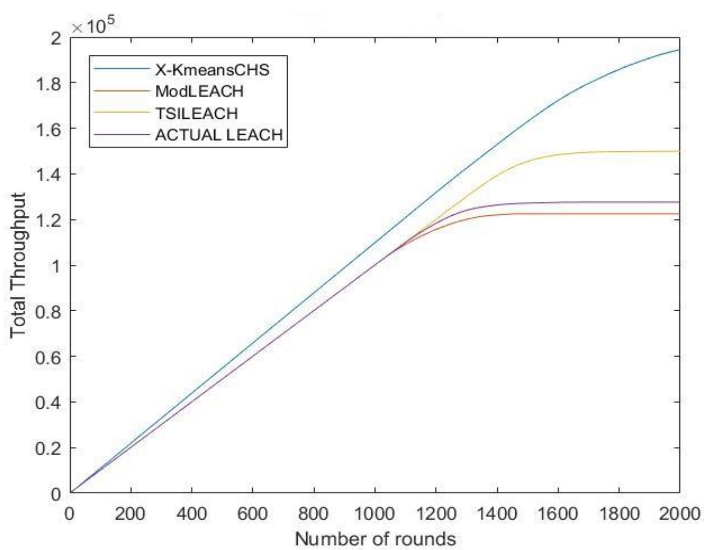

Figure 11. Total Throughput of the network in extended K-Means CHS, MOD-Leach, TSI-Leach and Actual Leach

In terms of cluster head election, the proposed extended K-Means CHS generate uniform number of cluster in every round comparing it with, MOD-Leach, TSI-Leach and Actual Leach the generated non uniform number of clusters were at some point there were so many clusters at particular round and other time very few clusters. The number of cluster heads forms per every round for the simulated algorithms is shown in Figure 12.

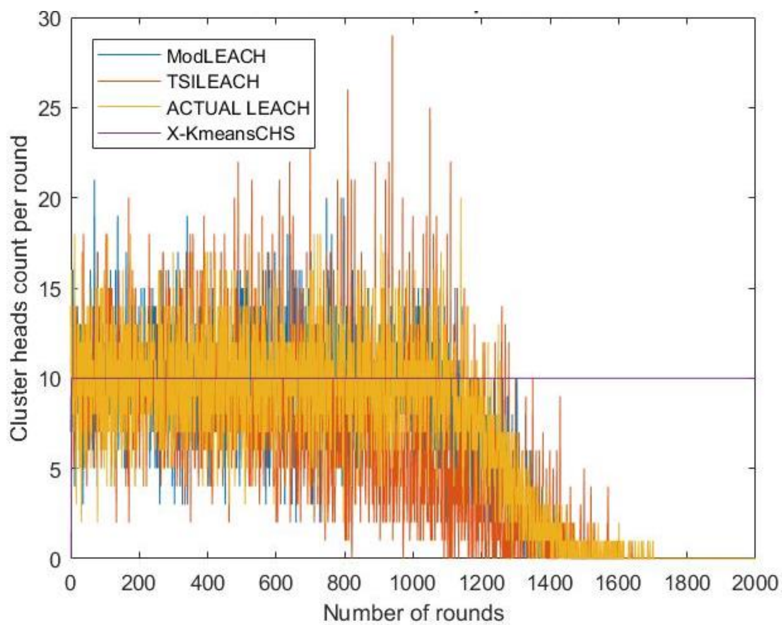

Figure 12. Number of cluster head per round for extended K-Means CHS, MOD-Leach, TSI-Leach and Actual Leach

6. DISCUSSION

We set out to develop an improved cluster head election algorithm for Leach routing protocol that improves the lifespan of WSN. We used simulation experiments to compared our new algorithm with other algorithms used in actual Leach and other variants of Leach such as MOD-Leach and TSL-Leach. To evaluate the proposed algorithm we used various performance metrics such as number of live nodes, number of dead nodes, average energy consumed, network lifetime, throughput, number of packets sent to CH and remaining energy. This section presents a discussion of the results obtained from these comparisons.

In the test for number of dead nodes at the end of simulation, the proposed algorithm outperforms others in that approximately 75 nodes were dead, compared to actual Leach and MOD-Leach where all the nodes had died, and TSI-Leach 95 nodes were dead. It is evident that the first node in Extended K-Means CHS algorithm died at approximately 1300 round, MOD-Leach approximately 950 round, TSI-Leach approximately 1250 round and actual Leach 1050 round respectively. This shows that at the end of simulation the number of dead node is lower compared other simulated algorithms. This also means that the number of live node is much higher compared to traditional Leach, Mod-Leach, TSI-Leach. The extended K-Means cluster head selection algorithm performs better compared to other simulated algorithms in terms of number of dead nodes and number of live node, hence increasing the lifetime of WSNs.

In the test for total remaining energy the results indicated that the energy dissipation rate is lower in the proposed extended K-Means cluster head selection compared to actual Leach, Mod-Leach and TSI-Leach. At the end of simulation, there was approximately 2 joules remaining in proposed K-Means CHS algorithm and all the energy was exhausted in Mod-Leach, TSI-Leach and actual Leach Respectively This show that there is very low energy wastage in the proposed extended K-Means cluster head selection algorithm hence improving the lifetime of WSNs.

In the test of total data packet sent from nodes to cluster head the results indicate that at the end of simulation the proposed algorithm outperforms others, where there were approximately $17*10^4$data packets compared to actual Leach that forwarded approximately $9.5* 10^4$data packets, Mod-Leach $11*10^4$data packets and TSI-Leach $9* 10^4$ data packets sent to cluster head respectively. This means our proposed algorithm is forwarding more data packets to cluster head compared to others.

In the test of throughput of the entire network the results show that at the end of simulation the proposed algorithm outperforms others, where at the end of simulation proposed extended K-Means cluster head selection algorithm had a total throughput $2*10^5$, actual Leach $1.2*10^5$, Mod-Leach $1.1*10^5$and TSI-Leach $1.5*10^5$. This shows that there was more work done in the proposed algorithm hence improving the throughput of the entire network.

In terms of cluster head election, the proposed extended K-Means CHS generate uniform number of cluster in every round that is 10 clusters in each round comparing it with, MOD-Leach, TSI-Leach and Actual Leach that generated non uniform number of clusters were at some point there were so many clusters at particular round and other time very few clusters. This means that the proposed algorithm ensures there is uniform number of clusters in every round.

7. Conclusions and Future Works

In this study, a more effective extended K-Means Cluster head selection algorithm was proposed. This proposed algorithm first forms the k centroids using K-means algorithm, assigns nodes to their clusters and then elect cluster heads within the k formed clusters from k centroids. The proposed K-Means cluster head election algorithm (EKCHS) is compared with existing cluster head election algorithms used in other variants of Leach routing protocol for WSN.

To evaluate the performance of the proposed extended K-Means Clustering we used number of live nodes, number of dead nodes, average energy consumed, network lifetime, throughput, number of packets sent to CH and remaining energy performance metrics. The results show that the proposed extended K-Means clustering algorithm performed better than other proposed approaches. The study has made a contribution in improving the lifetime of WSN network.

In future, we plan to develop a cluster head algorithm that can further improve on energy stabilization and balancing among nodes in WSNs. We also plan to develop technique that will enable nodes to do data processing before forwarding information.

REFERENCES

[1] Tay, M. and Senturk, A.“A New Energy-Aware Cluster Head Selection Algorithm for Wireless Sensor Networks,” Wireless Pers Commun, no. https://doi.org/10.1007/s11277-021-08990-3, p. 2235–2251, 2021.

[2] Raj, B., Ahmedy, I., Idris M. Y. I. and Noor, R. M. “A Survey on Cluster Head Selection and Cluster Formation Methods in Wireless Sensor Networks,” Wireless Communications and Mobile Computing, 2022.

[3] Suganthan, T. and Rizwana, S. “A Review on Cluster Head Selection Algorithms for Wireless Sensor Networks,” International Journal of Computer Sciences And Engineering, pp. 618-622, 2019.

[4] Goyal, A., Mudga , S. and Kumar, S. “A Review on Energy-Efficient Mechanisms for Cluster-Head Selection in WSNs for IoT Application,” International Conference on Applied Scientific Computational Intelligence using Data Science, 2020.

[5] S. M.B, M. C.P and R. V.K., “Optimized and Dynamic Selection of Cluster Head Using Energy Efficient Routing Protocol in WSN.,” Wirel. Pers Commun., p. 577–599, 2021.

[6] Pitchaimanickam, B. and Murugaboopathi, G. “A hybrid firefly algorithm with particle swarm optimization for energy efficient optimal cluster head selection in wireless sensor networks,” Neural Comput. Appl, vol. 32, no. 12, pp. 7709-7723, 2020.

[7] Rao, P., Jana, P. and Banka, H. “A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks,” Wireless Network, vol. 23, no. 7, pp. 2005-2020, 2017.

[8] Javaid, N., Waseem, M. and Khan, Z. “ACH: away cluster heads scheme for energy efficient clustering protocols inWSNs.,” In Saudi International electronics, communications and photonics conference, Piscataway , p. 364–367, 2013.

[9] Khandelwal , A. and Jain, Y. K. “An efficient k-means algorithm for the cluster head selection based on SAW and WPM,” International Journal of Advanced Computer Research, vol. 8, no. 37, pp. 191-202, 2018.

[10] .Singh, B. and Lobiyal, D. K. “A novel energy-aware cluster head selection based on particle swarm optimization for wireless sensor networks.,” Human-Centric Computing and Information Sciences,, vol. 2, p. 1–18, 2012.

[11] Arulmurugan, A., Waris, S. F. and Bhagyalakshmi, N. “Analysis of Cluster Head Selection Methods In WSN,” in 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, , 2021.

[12] Mazin, K. H. and Ali, K. I. “Distributed DBSCAN Protocol for Energy Saving in IoT,” in International Conference on Communication, Computing and Electronics Systems, Singapore, 2021.

[13] Hoang, D. C., Kumar, R. and Panda, S. K. “Fuzzy C-Means clustering protocol for Wireless Sensor Networks,” IEEE International Symposium on Industrial Electronics, pp. 3477-3482, 2010.

[14] Leu, T.-H. Chiang, M. C. Yu and Su, K.W. “Energy efficient clustering scheme for prolonging the lifetime of wireless sensor network with isolated nodes,” IEEE communications letters. , 2015, vol. 19, pp. 259-262, 2015.

[15] P. Sasikumar and S. Khara, “"K-Means Clustering in Wireless Sensor Networks,” in 20112 Fourth International Conference on Computational Intelligence and Communication Networks, Mathura, India, 2012.

[16] Adumbabu,1. and K. Selvakuma, “An Improved Lifetime and Energy Consumption with Enhanced Clustering in WSNs,” Intelligent Automation & Soft Computing, vol. 35, no. 2, pp. 1939-1956., 2023.

[17] Jong-Yong, L. and Daesung, L. “K-means clustering-based WSN protocol for energy efficiency,” International Journal of Electrical and Computer Engineering (IJECE), p. 2371~2377, 2021.

AUTHORS

Peter Maina Mwangi is a Tutorial Fellow at the Department of Computer Science, Murang'a University of Technology, Kenya. He received his BSc. in Computer Science from Busoga University, Uganda in 2010, his MSc in Data Communication and Networks from KCA University, Kenya in 2018. He is currently pursuing PhD in Computer Science from Murang'a University of Technology, Kenya. His Research interest is in Network Security.

John Gichuki Ndia is a lecturer at Murang'a University of Technology, Kenya. He obtained his Bachelor of Information Technology from Busoga University, Uganda in2009, his MSc. in Data Communication from KCA University, Kenya in 2013, and his PhD in Information Technology from Masinde Muliro University of Science and Technology, Kenya in 2020. His Research interests include Software Engineering, Computer Networks Security and AI Applications. He is a Professional Member of Institute of Electrical and Electronics Engineers (IEEE) and the International Association of Engineers (IAENG).

Geoffrey Muchiri Muketha is Professor of Computer Science and Director of Postgraduate Studies at Murang'a University of Technology, Kenya. He received his BSc. in Information Science from Moi University, Kenya in 1995, his MSc. in Computer Science from Periyar University, India in 2004, and his PhD in Software Engineering from Universiti Putra Malaysia in 2011. He has wide experience in teaching and supervision of postgraduate students. His research interests include software and business process metrics, software quality, verification and validation, empirical methods in software engineering, component-based software engineering and network security. He is a Professional Member of Association for Computing Machinery (ACM) and a member of the International Association of Engineers (IAENG).