The Seybold Report

ISSN 1533-9211

DOI: 10.5281/zenodo.8251884

OPTIMIZATION OF LEACH PROTOCOL WITH K-MEANS ++ CLUSTERING METHOD ON WIRELESS SENSOR NETWORK (WSN)

I GUSTI AGUNG GEDE ARYA KADYANAN*[1] and

IDA BAGUS GEDE DWIDASMARA[2]

[1, 2] Informatics, Faculty of Mathematic and Natural Sciences, Udayana University South Kuta, Badung, Bali, Indonesia.
Email: [1]gungde@unud.ac.id (*Corresponding Author), [2]dwidasmara@unud.ac.id

Abstract

Objectives: This research paper absolutely discussed about Wireless Sensor Network is a group of nodes that use batteries as their energy source, so energy use is one of the important points in building a WSN. LEACH protocol is one solution to solve energy problems in WSN, but LEACH has several drawbacks, namely it does not consider the remaining energy distance of the cluster head at the base station. **Methodology:** Based on these problems, this study uses the K-Means development algorithm, namely K-Means ++, to modify LEACH to KM-LEACH to further optimize the selection of the cluster head to be selected based on the remaining energy. The testing phase in this study was carried out by comparing 3 test parameters, such as the total energy consumption (joules), the number of dead nodes and nodes alive which were compared with previous studies using the K-Means algorithm. **Results:** The results were obtained by KM-LEACH that more energy efficient with a total energy consumption of 40.86 joules versus 48.40 joules. Ultimately, the result of the number for dead nodes is less, 63 nodes compared to 81 nodes with LEACH with the K-Means algorithm. **Types of research:** Empirical

Keywords: Wireless Sensor Network, K-Means++ clustering, LEACH Protocol, Energy Efficiency, Routing Protocol

1. INTRODUCTION

In contrast, the technology development in the field of communication today is more fast than before. One of the reason is technologies are widely applied, which is the Wireless Sensor Network (WSN). Eventually, Wireless Sensor Networks can consist with the hundreds to thousands of nodes, where each node has the ability for sensing, data processing and wireless communication. Each node uses a battery as the sensor supply node, which causes energy use to be an important point in building a WSN. Battery replacement at sensor nodes becomes difficult to do considering the sensor nodes are placed randomly in the network and under certain conditions; it is difficult to determine the position of all sensor nodes on the WSN. Of course, this causes disruption of network stability on the WSN and affects network performance in sensing or data processing [12]. Therefore, another solution is needed to overcome the energy limitations of WSN, namely by choosing an energy efficient routing protocol.

Routing protocol is a method for determining the closest path in the data transfer process from the node to the base station. Many studies have been carried out to analyze the use of appropriate routing protocols in optimizing energy use in WSNs. The first algorithm developed is the Low-Energy Adaptive Clustering Hierarchy (LEACH) routing protocol. However, the

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.8251884

LEACH algorithm has several shortcomings such as the remaining energy of the nodes not considered in the cluster head (CH) selection process. Therefore, CH failure due to low residual energy cannot be avoided. Then in choosing, the CH also did not consider the distance between the CH and the Base Station. So that the power consumption on CH located far from, the Base Station is higher than that near the Base Station [7].

In previous research [8] using the K-Means clustering algorithm to optimize cluster head selection in a cluster to create a uniform cluster and reduce the distance between the cluster head and other nodes. The results of this study indicate that the resulting clusters are symmetrical and significantly reduce the distance between the nodes and the cluster heads, which have a balanced node energy consumption with an increase in the network life span.

Therefore, this study aims to optimize the LEACH protocol by using the K-Means development algorithm, namely by using the K-Means ++ algorithm to become a new LEACH algorithm, namely KM-LEACH. KM-LEACH is an optimization of the LEACH protocol which optimizes the K-Means ++ Clustering algorithm to further optimize the selection of the cluster head that will be selected based on the remaining energy. So it is hoped that by optimizing the cluster head selection, it can reduce the problem of uneven distribution of cluster head, uneven energy consumption and improve energy stability in the Wireless Sensor Network (WSN).

2. RESEACH METHODS

1. Data Collection

The data to be used in this study were obtained from the results of comparison simulations in the MATLAB 2015a application with several test parameters to obtain data:

a. Amount of energy consumption (joule)
b. Number of nodes alive for 400 rounds
c. Number of dead nodes for 400 rounds

Where the data obtained uses a test scenario based on parameters from previous research [8]. Where later the simulation parameters will be used to become a simulation scenario to obtain primary data in this study. Where the scenario that is designed will use 100 nodes spread at coordinate points x and y with an area of 100 x 100 m^2 to create a WSN network topology scheme.

2. Methods Used

In this study, using 2 methods, namely the LEACH algorithm with K-Means and K-Means ++. After that, the performance of the LEACH protocol with K-Means and the LEACH protocol with K-Means ++ will be compared to obtain the results of the study. The following is an overview of the program that will be made in general which is represented in the following figure:

ISSN 1533-9211
DOI: 10.5281/zenodo.8251884

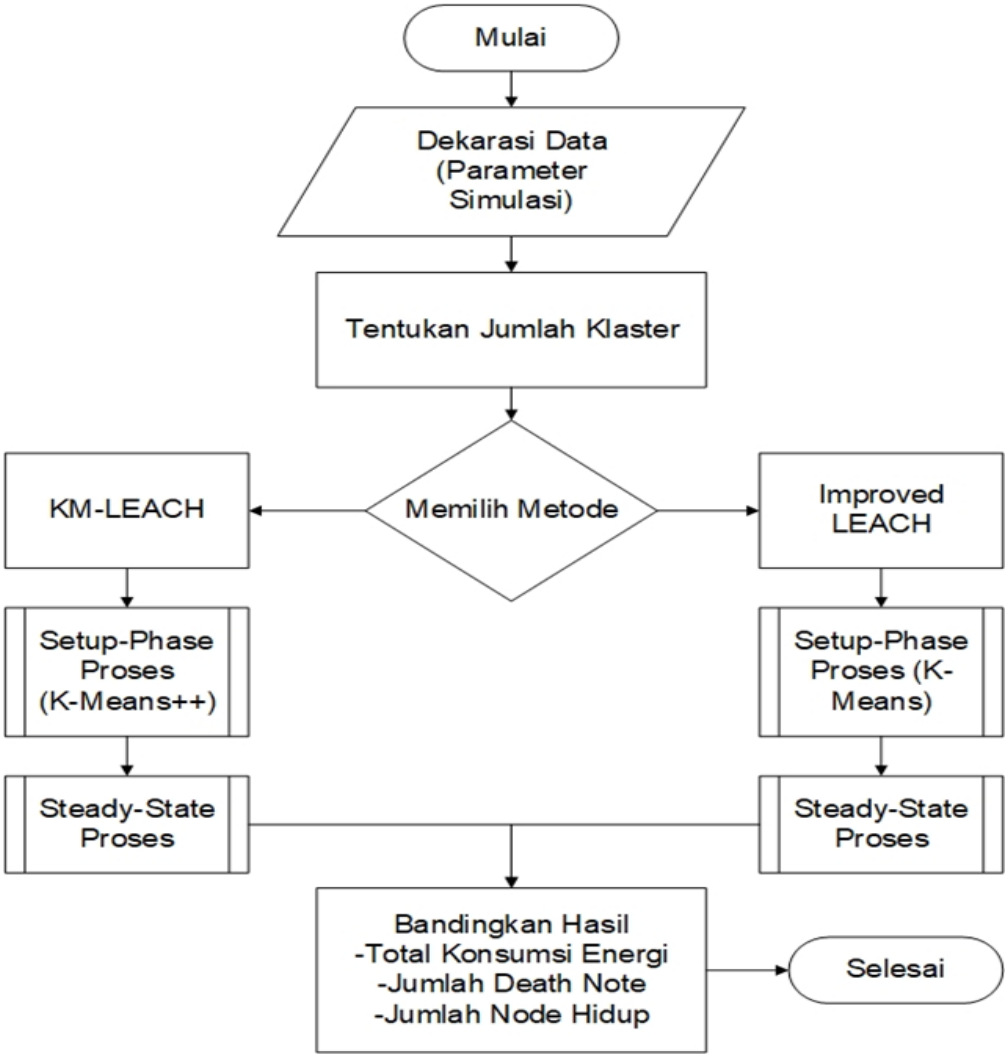

Figure 1: Program Design Flowchart

3. Leach With K-Means

In general, the LEACH algorithm has 2 main processes, namely the Setup Phase and Steady State. In this Setup Phase process, it will be modified using the K-Means algorithm. The K-Means algorithm plays a role in the cluster formation process where with the help of the K-Means algorithm it can help form clusters that are better than the usual clustering algorithm in the LEACH algorithm. The following is the LEACH protocol flowchart that has been optimized with K-Means in the following image:

ISSN 1533-9211
DOI: 10.5281/zenodo.8251884

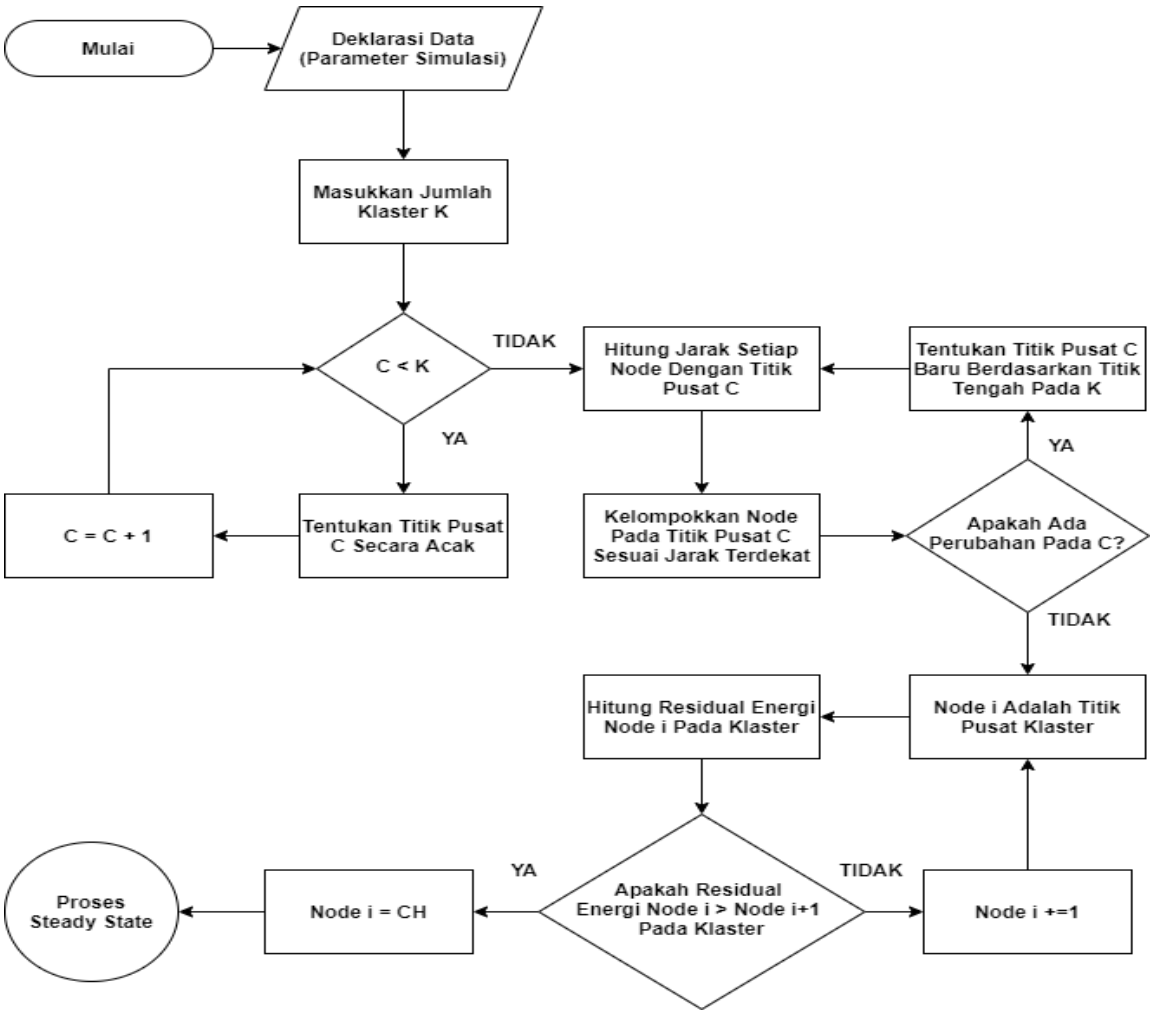

Figure 2: Setup Phase LEACH Flowchart With K-Means

Flowchart Explanation

1. The first stage is to form a network topology based on predetermined simulation parameters
2. Second, enter the number of clusters
3. Then do iteration to determine the centroid / cluster center point of C = the number of clusters specified
4. Place the cluster center point / centroid C randomly
5. Calculate the distance of each node with centroid C using the formula

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.8251884

$$J = \sum_{i=1}^{k} \sum_{j=1}^{n} \left\| X_j - C_i \right\|^2 \qquad (1)$$

Where:

J = the distance between the data point node and the centroid

$\mathbf{X_j^i}$ = the j-th node of the i-th cluster

$\mathbf{C_i}$ = the i-th centroid in the i-th cluster

6. Group of nodes at the center of the cluster/centroid C according to the closest distance
7. Then check if there is a change in the center of the cluster/centroid C
8. If yes, then determine the center point of the cluster/centroid C based on the midpoint of k with the following equation, then repeat steps 5-7 until there is no change in cluster center / centroid C

$$Ci = \frac{1}{|ki|} \sum_{X_{j \in k}}^{k} Xj \qquad (2)$$

9. If not then node i = cluster center point/centroid C
10. Then enter into determining the cluster head where in the selection pay attention to the energy residual parameters at the node for each cluster which is calculated by the following equation

$$Er(n) = E - Te \qquad (3)$$

Where:

Er = Residual energy of each node during the round

Ei = Energy of each node during the round

T_e = Amount Total node energy consumption during the round

11. Compare whether the residual energy of node i>node i+1, if yes then select node i as CH based on the largest energy residual
12. If not then node i = node i + 1 and repeat steps 9 through 11
13. If so, then node i is the cluster head and continues to the steady state process

4. Leach With K-Means++

The K-Means ++ Clustering algorithm is one of the algorithms used for data classification or grouping, this algorithm is a development of the K-Means algorithm. In the LEACH protocol

ISSN 1533-9211

DOI: 10.5281/zenodo.8251884

with the help of the K-Means ++ algorithm, it will modify the setup phase of the LEACH protocol where in the setup phase the cluster head selection uses the K-Means ++ algorithm. The following is a flowchart setup phase with the K-Means ++ algorithm which is represented in the image below.

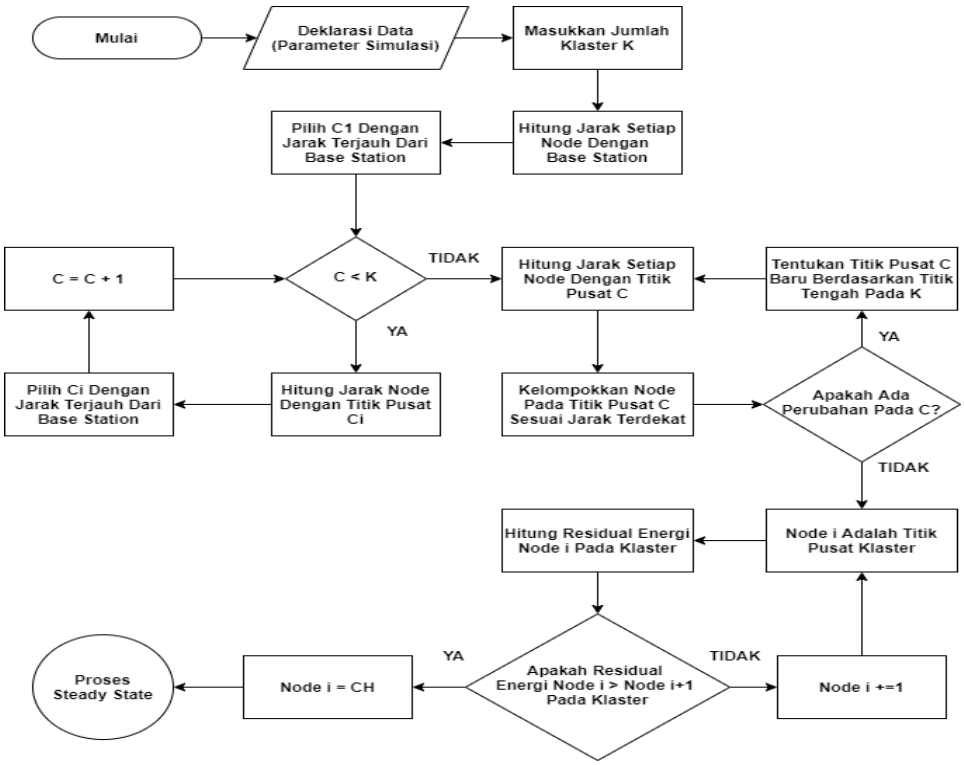

Figure 3: Setup Phase LEACH Flowchart With K-Means++

Flowchart Explanation

1. The first stage is to create a node shape according to the specified parameters
2. Input the number of clusters
3. Calculate the distance of all nodes to the base station with the formula from the equation
4. Determine the center of C1 cluster based on the furthest distance between the node and the base station.
5. Selection of Ci then calculate the distance of all nodes to the center of cluster C1 with the formula from the equation (1)
6. Determine the next Ci based on the farthest distance between the center point Ci and the base station.
7. Repeat steps 5 and 6 until all clusters are formed

ISSN 1533-9211

DOI: 10.5281/zenodo.8251884

8. Calculate the distance of each node to the center of the cluster/centroid C using the equation (1)
9. Group of nodes at the center of the cluster/centroid C according to the closest distance
10. Then check if there is a change in the center of the cluster/centroid C
11. If yes, then determine the center point of the cluster/centroid C based on the center point on k with the equation (2), then repeat steps 8-10 until there is no change in cluster center/centroid C
12. If not then node i = cluster center point/centroid C
13. Then enter into determining the cluster head where in the selection pay attention to the energy residual parameters at the node for each cluster which is calculated by the equation (3).
14. Compare whether the residual energy of node i>node i+1
15. If not then node i=node i+1 and repeat steps 13 to 14, if yes then node i is the cluster head and continue to the steady state process

5. Steady State Leach

The steady state phase is the phase of data transmission between nodes to the cluster head and base station. In this phase, it consists of 2 types of transmission processes where between the nodes to the cluster head, then each cluster head in each cluster will combine data and send it to a base station on a WSN network. The following is a steady state flowchart on LEACH which is represented in the image below.

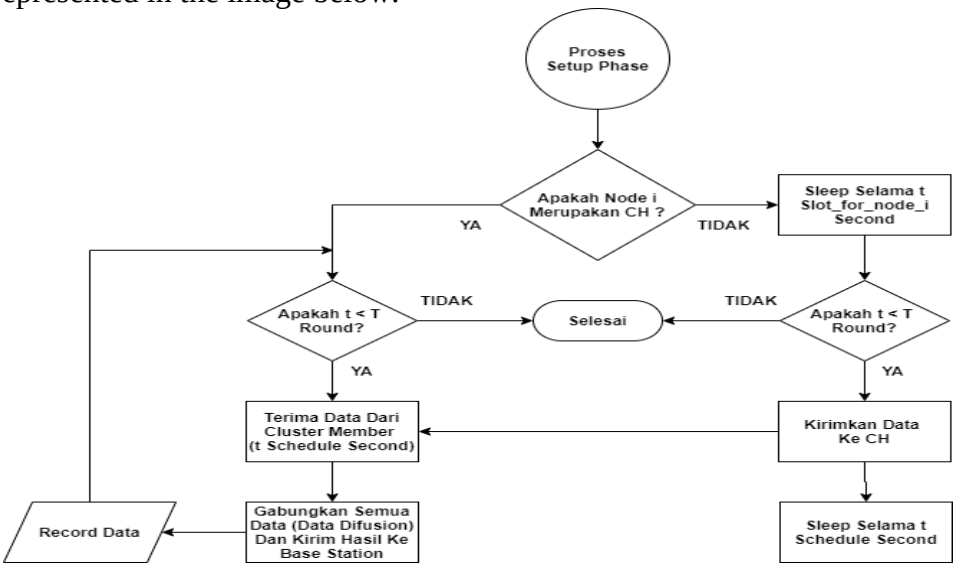

Figure 4: Flowchart Steady State LEACH [10]

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.8251884

6. Testing Scenarios

In this study, the LEACH protocol with K-Means optimization and KM-LEACH with K-Means ++ optimization was simulated with MATLAB 2015a software with several parameters adjusted based on research results [8] to create predefined simulation scenarios, including:

Table 1: Simulation Parameter [8]

Parameter	Nilai
Number of *nodes*	100
Simulation area	100 x 100 m^2
Number of rounds	400
Base station position	(50,50)
Probability cluster head %	10
The initial energy of the node	0.5 Joule
Data Length	4000 bits
Eelec	50nJ/bit
εmp	13nJ/bit/m^4
εfs	10nJ/bit/m^2
Data Aggregation (EDA)	50nJ/bit/signal

7. Measurement and Simulation Results

The performance quality of the KM-LEACH protocol with the LEACH protocol will be measured by several parameters such as the average energy consumption, the number of dead nodes and the number of nodes that are alive for 100 rounds. Where to calculate the total energy consumption per-bit expended is obtained using the same radio communication energy consumption shown as used as the LEACH protocol. This model consists of two parts: the transmission energy consumption model and the receiver energy consumption model. In accordance with the radio communication energy dissipation model, a formula is obtained to calculate energy consumption [8]:

$$\boldsymbol{E_{TX}(k,d) = (E_{elec} \times k) + \varepsilon_{fs} + k \times d^2} \qquad (4)$$

Where is the explanation of the formula variable above, namely:

$\boldsymbol{d}$ = *Threshold Distance,*

$\boldsymbol{E_{elec}}$ = *Electronic Energy*

$\boldsymbol{\varepsilon_{fs}}$ = *Parameter for Free Space Model*

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.8251884

ε_{mp} = *Parameter for Multi Path Model*

The energy consumption for receiving k-bit messages can be calculated by the equation:

$$E_{RX} = E_{elec} * k \quad (5)$$

Meanwhile, to find out the number of dead and live nodes, it is calculated by:

Dead node = total *node* - *alive node* (6)

Node hidup = total *node* – *death node* (7)

Where after the data is obtained from 2 scenarios on the LEACH protocol with the K-Means method and the KM-LEACH protocol with the K-Means ++ method, then the data is slammed with the following table to get the final result to find out which protocol is more optimal between the two protocols. The following is a table of the measurement parameters of the two protocols:

Table 2: Protocol Test Parameters

Protocol	Number of Dead Nodes/100 Round	Number of Alive Nodes/100 Round	Amount of Energy Consumption /100 Round
LEACH dengan metode *K-Means*			
KM-LEACH dengan metode *K-Means++*			

3. RESULT AND DISCUSSION

1. Testing Results of Amount Energy Consumption

The total energy consumption is the amount of energy released by each node during the data transmission process in the WSN network. In this study, each total energy consumption released by 2 network protocols, namely LEACH which is optimized with the K-Means algorithm with LEACH which is optimized with the K-Means ++ algorithm has been stored in 2 txt files during the simulation process. The LEACH protocol with the K-Means algorithm is stored in the TE-kmeans.txt file and the LEACH protocol with the K-Means ++ algorithm. Based on research conducted over 400 rounds, the results of a comparison of total energy consumption are obtained as shown in the following figure.

ISSN 1533-9211

DOI: 10.5281/zenodo.8251884

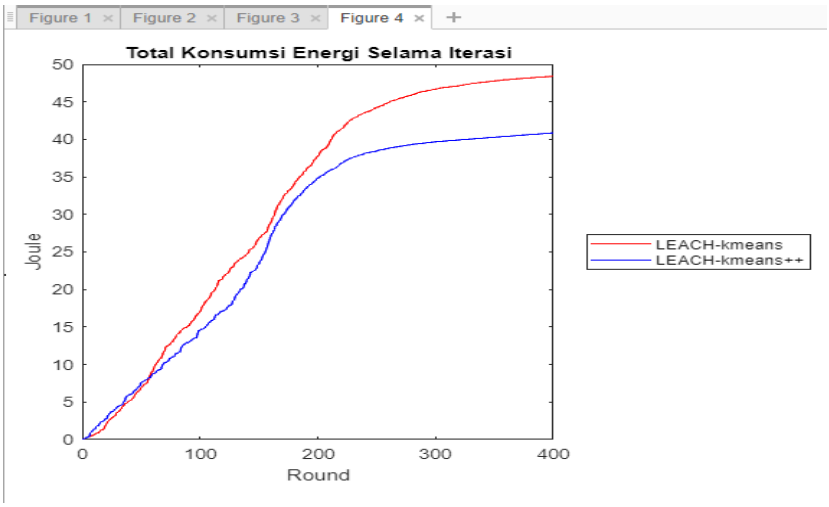

Figure 5: Comparison Results of Amount Energy Consumption

And the following is a table of the results of a comparison of total energy consumption in accordance with Figure 5 above.

Table 3

Metode	Number of Iteration (round)	Total Energy Consumption (Joule)
LEACH *K-Means*	400	48.40
LEACH *K-Means++*	400	40.86

Based on the results obtained, it shows that the LEACH protocol modified with the K-Means ++ or KM-LEACH algorithm is more energy efficient than the LEACH optimized with the K-Means algorithm, where KM-LEACH consumes a total energy consumption of 40.86 Joules. It is more efficient than LEACH with the K-Means algorithm which consumes 48.40 Joules of energy.

2. Testing the results of the number of death nodes

Dead node is a condition where the energy owned by the node is less than equal to 0 during the simulation process in the WSN topology. In this study, a comparison was made between the two simulated protocols, namely LEACH with the K-Means ++ algorithm or KM-LEACH and LEACH modified with the K-Mean algorithm, where the two simulation results were stored in 2 txt files, namely TDN-kmeans.txt and TDN- kmeansplus.txt. Based on research conducted for 400 rounds, the results of the comparison of total dead nodes are obtained as shown below

ISSN 1533-9211

DOI: 10.5281/zenodo.8251884

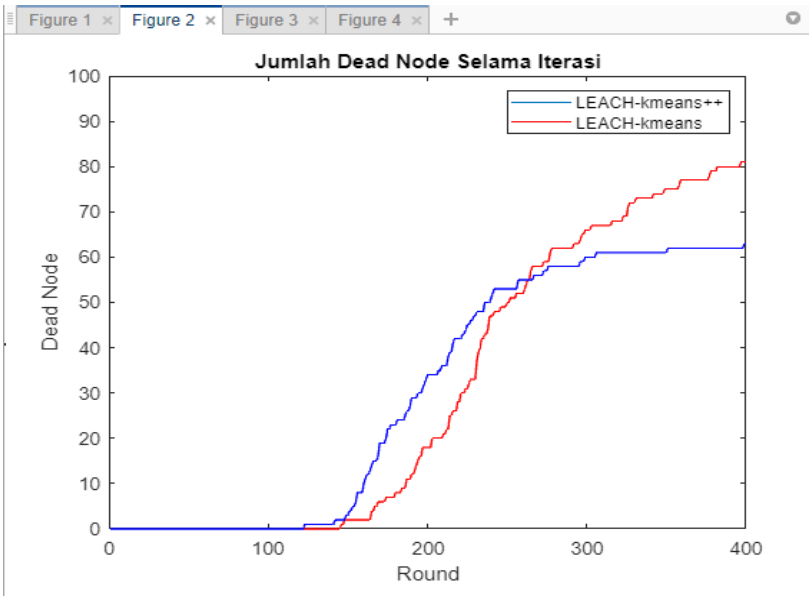

Figure 6: Hasil Perbandingan Jumlah Death Node

And the following is a table of the results of the comparison of the number of Dead Nodes according to Figure 6 above.

Table 4: Hasil Perbandingan Jumlah Dead Node

Metode	Jumlah Iterasi (round)	Jumlah *Dead Node* (node)
LEACH *K-Means*	400	81
LEACH *K-Means++*	400	63

Based on the results obtained, it shows that the LEACH protocol modified with the K-Means ++ or KM-LEACH algorithm is more stable than the LEACH optimized with the K-Means algorithm, where KM-LEACH has 63 dead nodes. Better than LEACH with the K-Means algorithm which has 81 dead nodes.

3. Testing the Results of the Number of Alive Nodes

Node alive is the number of live nodes remaining during the energy transmission process in the WSN topology. In this study, a comparison was made between the two simulated protocols, namely LEACH with the K-Means ++ algorithm or KM-LEACH and LEACH modified with the K-Mean algorithm, where the two simulation results were stored in 2 txt files, namely TNA-kmeans.txt and TNA- kmeansplus.txt. Based on research conducted for 400 rounds, the results of comparisons of total nodes alive are as shown in the following figure.

ISSN 1533-9211

DOI: 10.5281/zenodo.8251884

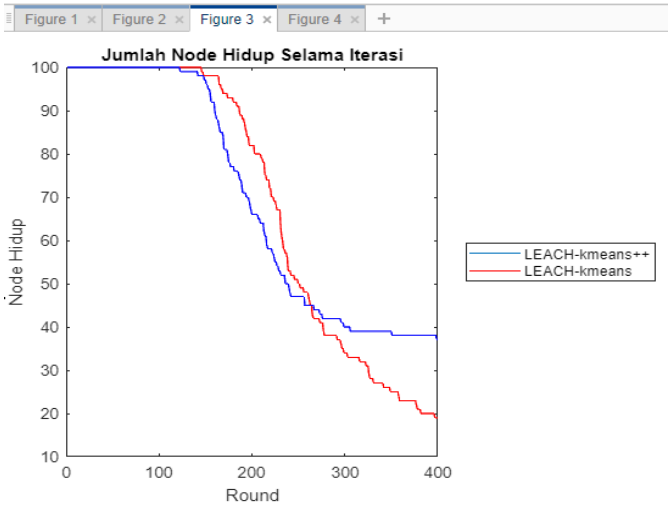

Figure 7: Comparison Result of Number of Alive Nodes

And the following is a table of the results of the comparison of the number of Alive Nodes according to Figure 7 above.

Table 5: Perbandingan Jumlah Node Alive

Metode	Jumlah Iterasi (round)	Jumlah *Node Alive* (node)
LEACH *K-Means*	400	19
LEACH *K-Means++*	400	37

Based on the results obtained, it shows that the LEACH protocol modified with the K-Means ++ or KM-LEACH algorithm has a better life span compared to LEACH which is optimized with the K-Means algorithm, where KM-LEACH has 37 nodes alive. Better than LEACH with K-Means algorithm which has 19 nodes alive.

4. CONCLUSION

Based on the implementation of the research that has been done and the results that have been obtained, it can be concluded that from the results obtained and the tests carried out in terms of the total energy consumption produced, it is proven that KM-LEACH with the K-Means ++ algorithm is more efficient and energy efficient than LEACH with the algorithm. K-Means is 40.86 Joules versus 48.40 Joules during the simulation process. And the results of the number of dead nodes obtained also show that KM-LEACH with the ¬¬K-Means ++ algorithm which has a more efficient total energy consumption can reduce the number of dead nodes generated in the network compared to LEACH with the K-Means algorithm, which is 63 nodes. Versus 81 nodes during the simulation process. The energy efficiency obtained also has an impact on the number of nodes alive, where KM-LEACH with the K-Means ++ algorithm can increase

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.8251884

the lifetime of a node better than LEACH with the K-Means algorithm, which is 37 nodes versus 19 nodes during the simulation process.

The suggestion that can be used for further research is that in this study the K-Means ++ algorithm has a more complex number of iterations than the K-Means algorithm so that if it is actually implemented on nodes it requires a larger node memory specification. In further research, other algorithms that have lower complexity than the applied method can be used, and in the process of generating random numbers the K-Means algorithm is still done manually via the read file not automatically in the program because in Matlab it is not possible to generate random numbers. unique in contrast to many iterations. So that in the next research it is expected that random, number generation can be done automatically through the program

References

1. Arthur, D., and Sergei Vassilvitskii. (2007). *K-Means++*: The Advantages of Careful Seeding. *Proc. of the annu. ACM-SIAM Symp. On Discrete Algorithm, Vol.8* hal.1027-1035. doi:10.1145/1283383.1283494
2. Heinzelman, W. B., Anantha P. Chandrakasan and Hari Balakrishnan. (2002). an Application-Specific Protocol Architecture for Wireless Microsensor Networks. IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, VOL. 1, NO. 4, OCTOBER 2002, hal.660-670.
3. Hibrian, M., Wardi and Agussalim. (2020). Peningkatan Network Lifetime Menggunakan Cluster Based Pada Wireless Sensor Network. JITCE (Journal of Information Technology and Computer Engineering), Vol. 04 No. 01 (2020), hal.16-21.
4. Han, J., Micheline Kamber and Jian Pei (2011). Data Mining Concept and Techniques. USA, Elsevier Inc.
5. Novid, I., Delsina Faiza, Thamrin and Winda Agustiarmi. (2019). Analisa Perbandingan Routing Protokol Pada Wireless Sensor Networks (WSNs). Jurnal Vokasional Teknik Elektronika dan Informatika, Vol. 7, No. 2, Juni 2019.
6. Permana, M. A. (2015). Analisa Algoritma LEACH Pada Jaringan Sensor Nirkabel. Proceeding Seminar Tugas Akhir Jurusan Teknik Elektro FTI-ITS.
7. Ridwan, A., Rian Ferdian and Rahmadia Kurnia. (2020). Optimasi Protokol LEACH Untuk Meningkatkan Stabilitas Pada Wireless Sensor Network. RESTI (Rekayasa Sistem dan Teknologi Informasi, Vol.4 No.1, hal.193 -200.
8. Saheb, P., and Kanika Sharma. (2017). Improved LEACH Protocol Based on K-Means Clustering Algorithm for Wireless Sensor Network. International Journal of Electronics & Communication Technology, 8(4).
9. Salem, A. O. A., and Noor Shudifat. (2019). Enhanced LEACH protocol for increasing a lifetime of WSNs. Springer (23), hal.901–907.
10. Sohraby, K., Daniel Minoli and Taieb Znati (2007). Wireless Sensor Networks Technology, Protocols, and Applications. New Jersey, John Wiley & Sons, Inc.
11. Susanto, E. B. (2016). Evaluasi Hasil Klaster Pada Dataset Iris, Soybean-small, Wine Menggunakan Algoritma Fuzzy C-Means dan Kmeans++. SURYA INFORMATIKA, Vol. 2 No. 1 - Mei 2016.
12. Tonapa, J. R., and Indrastanti Ratna W. (2016). Analisis Ketahanan Energi Oleh Low Energy Adaptive Clustering Hierarchy (LEACH) Pada Cluster Head Wireless Sensor Network (WSN). Semantic Scholar.