An Energy-Efficient Dynamic Clustering In WSN Using Hybrid K-Medoid Simulated Annealing Approach

Arvind Kumar Shrivas[1], Dr. Nirupma Tiwari[2]
[1]Dept of computer Science & Engineering
[2]Asst. Professor, Dept of computer Science & Engineering
[1, 2] ShriRam College of Engineering & Management,Banmore

Abstract- *In recent decades, both academic and industrial communities have seen the advancement of Wireless Sensor Networks (WSNs). Clustering is the most common way of increasing the longevity of WSNs. In clustering techniques, the maximum no. of CHs (Cluster Heads) and how to arrange clusters are the most critical problems. The paper-based on effective CH selection method, which revolves between nodes with greater energy levels than others, and rotates the cluster head position. Residual energy, Initial energy & optimum value of CHs are considered by an algorithm for the election of next CHs for the network that suits applications of IoT like environmental monitoring, smart cities & systems. To avoid premature sensor death & prevent hot spot issues, a new dynamic clustering solution is presented. Also, Simulated Annealing is implemented as the shortest path tree for mobile nodes which is constructed to establish the connection between the nodes for finding the shortest and secure path for data transmission hence resulting in faster data sending and receiving process.*

Keywords- WSN, IoT, Residual energy, CH selection, Energy-efficient, Network Lifetime.

I. INTRODUCTION

A WSN is described as a collection of sensors situated in the environmental sensor field. Sensors can collect various data types based on their defined functions. Gathered data is transmitted to BS(Base Station) or sink after such processing operations. This procedure is carried out using multiple routing techniques. A power source or battery which is generally restricted is a component of the sensor. Moreover, these sensors are commonly placed in areas that are not humanly available and thus these components cannot be reloaded or exchanged. For this purpose, a power supply of sensors can be used to the maximum possible extent, defined as energy efficiency that will boost network longevity. By reducing or balancing node energy consumption energy efficiency can be achieved. Clustering is an energy efficiency approach. There are multiple ways to implement clustering strategies to wireless sensor networks including lower load, increasing scalability, lower energy consumption, reducing latency, preventing collisions, ensuring connectivity, load balancing, fault tolerance, avoiding energy hole &enlarging network longevity[1]. Also, separating the area of the network into subareas supports to track coverage hole issue [2], provided that any network area is not covered by sensor nodes (SNs).

Using the wireless sensor network, a remote area where human intervention is not feasible will be mounted. In this challenging environment, the network is built on an ad hoc basis and the nodes will sense events& communicate to a central node named sink node or BS for the next level of analysis& processing. WSN has an energy hole problem during static routing, where multiple communication is performed while transmitting sensed data to the static sink. Sink mobility is valuable in energy protection, balancing load-dependent on residual energy, sparse network access, and efficient data transmission. Some applications need decreasing mobility in the sensor area, including healthcare services, disaster response, intrusion detection, battlefield, as well as, sliding area, etc. Discovering sink mobility aims to boost the existence of the WSN by decreasing energy conservation and route reconfiguration procedure[3].

Figure 1 illustrates sensor nodes deployment in a wireless network [3].

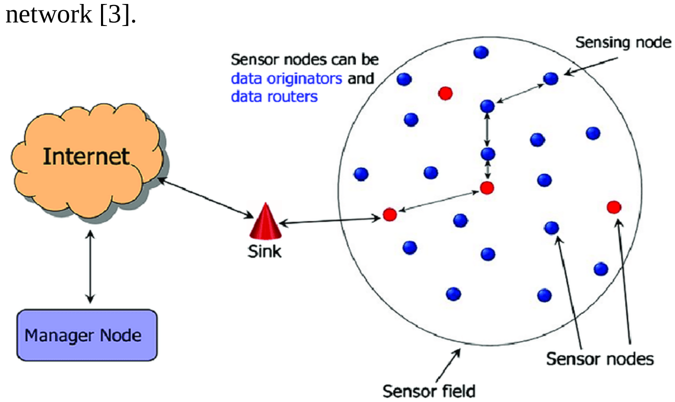

Fig. 1. Architecture of WSN

Clustering is one of the energy-efficient methods the research community has suggested to collect information from a WSN. This creates a set of clusters. There are a set of members and CH for every cluster. It receives information from its members (intra-cluster communication). CHs cooperate with a centralizedBS to transfer data (inter-cluster communication).

The remaining paper is structured accordingly: Section II outlines a related paper on literature's approach to the collection of dataand routing to mobile sink in WSN. Section III includes a detailed overview of the proposed scheme. In Section IV, performance assessments for the proposed scheme of simulation set-up and outcomes are explained in detail.Section V would also explain the conclusion and future work.

II. LITERATURE REVIEW

Shah et al. [4] suggested the fundamental concept of mobile sinks, which mobile sinks are named "data mules." Mule performs a random walk through them to aggregate data packets then data is often lost in separate access points. Although the node's transmission range is short, energy utilization may be dramatically minimized.

Wang et al. [5] Mobility-based Data Collection Algorithm suggested maximizing network life. The network area has a circular structure by MSs running in back and forth motion in the circle periphery.

Wang et al. [6] implemented mobility-based technology to refine the smart home network concept clustering algorithm &BS deployment approach. They considered various no. of mobile sinks under 2 different conditions. Firstly, they utilized a single sink that has circular motion in various radius. Formerly they utilized multi-mobile sinks in a circular area to search best no. of mobile sinks. They assumed that mobile sinks were moving at a constant velocity. The velocity of mobile agents is a significant limitation of mobility-based approaches.

In [7], the mobile sink's velocity is decreased to ensure the delivery of messages. This methodcan increase network longevity by using numerous, fixed-speed mobile sinks and achieving a high packet delivery ratio.

Though, lower sink velocity raises data delivery latency [8], when using all kinds of sinks issue of data delay can also be overcome and the packet delivery ratio is improved. This document tries to resolve the problem of the energy hole by involving unequal velocities among MSs. Furthermore, the velocity of MSs can be changed at various angles separately.

Nodes' energy utilization is assumed to define optimum cluster size in this work [9], before clustering. A two-stage Genetic Algorithm (GA) is used to evaluate optimum cluster size interval & to extract precise value from the interval. Energy hole is also an underlying problem that causes a major reduction in network lifespan. Asynchronous energy depletion of nodes in several networking layers is due to this issue. To balance the energy consumption ratio of CH, they suggest the CM2SV2(circular motion of mobile-Sink with varied velocity) algorithm.

In this context [10] novel cluster-based method with controlled flooding is introduced by the various mobile sink to extend the lifetime of WSN by including residual energy of every node &consistent data transmission over WSN. [10] Regardless of medium, this method utilizes a predetermined & controlled mobility model to determine movement direction for the mobile sink. Simulations demonstrate that the multiple mobile sinks have enhanced performance of energy and expanded life for the WSN with decreased route reconstruction.

- For effective routing, a new CH selection algorithm is introduced.
- Proposed selection method for cluster head increases the life of the network sensor by dynamic load balancing method.
- Simulated Annealing technology is built into mobile sinks to improve the longevity of the network of WSNs.
- Numerical simulation illustrates the benefit of prior optimization techniques of the proposed algorithm.

III. SYSTEM MODEL

For implementation of sensor network, we take following common assumptions in [11]:

- Sensor nodes are homogeneous.
- At first, the same residual energy was found in all sensor nodes.
- All sensors are homogeneous and operate equally well, including data collection, processing & communication.
- Bidirectional and symmetric communication links between sensor nodes.
- Mobile sink node has unrestricted energy also it is up to traveling anywhere in provided region R.
- Between sensor nodes, there are no problems.

Sensors are randomly implemented in an area& several clusters are separated in the whole region. CH for the region is chosen dependent upon the energy level of sensors. CH usually selects a sensor with maximum residual energy. Other members then give data to the CH. The CH shall transmit data to the mobile sink.

3.1 ENERGY MODEL

Two models have been utilized for examination of energy utilization, multipath fading model $\varepsilon_m d^4$. & free space model $\varepsilon_f d^2$. The distance between receiver & transmitter is dependent on both models. Fig. 2 demonstrates the radio energy model. 2. The radio therefore uses the following to relay a k- packet at distance d:

$$\mathbf{E_{TX(k,d)} = E_{TX-elec}(K) + E_{TX-amp}(k,d)}$$
$$\begin{cases} \mathbf{k * E_{elec} + k * \varepsilon_f * d^2} \\ \mathbf{k * E_{elec} + k * \varepsilon_m * d^4} \end{cases}$$

$$E_{RX}(k) = E_{RX-elec}(K) = k * E_{elec}$$

Here**,** $\mathbf{E_{TX}}$ needs energy usage for packet transmission, $\mathbf{E_{elec}}$is electronic energy which counts on filtering, spreading of signal and modulation digital coding, $\mathbf{E_{RX}}$ needs energy usage for receiving of the packet.$\mathbf{d_0}$is identical to the square root of EDA's multipath fading model for free space.

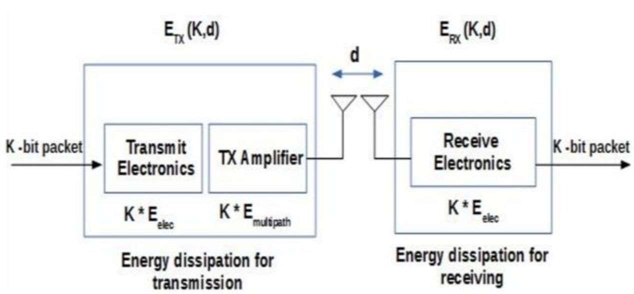

Fig. 2. Radio Model

3.2 NETWORK MODEL

The general model of the WSN is utilized in this article. It has the following functions and characteristics. I n the sensing area each sensor node is used randomly and can measure the distance to its neighboring nodes [12]. All sensor nodes (SNs) become stationary until the random deployment completes and engage in the cluster selection procedure. Every sensor may act as a normal node or as CH in the sensing field based on its properties, comprising distance among sensors, the distance among clusters, and residual energy. Each member of the cluster detects the environment & transmits data to its CH. The number of CHs remains often smaller than no. of sensors in the region. Each sensor typically works at various sensors& power levels. The communication medium is wireless, also when they are in a communications radius communication will occur between the cluster heads (CHs) & cluster members (CMs).

IV. RESEARCH METHODOLOGY

4.1 PROBLEM FORMULATION

In three fundamental problems, the ACO protocol disadvantages may be briefed.

1) 1st problem deals with inefficiency in dealing with large-scale combinatorial issues. Since the time complexity of ACO is O (n*(n-1)*m*T/2), it takes a longer time to procedure large-scale issues.
2) Since several parameters in ACO are involved, the inappropriate election of original values can lead to local optimal values, stagnation within an optimal local range, an untimely convergence which eventually leads to convergence of whole antcolony solution & failure to find an optimum solution.
3) As the situation shows that when a problem is so high, individual ants are moved separately, which takes extensive time to execute a complete algorithm, which limits the effectiveness of finding the best.

4.2 PROPOSED APPROACH

In the proposed methodology, we have used the K-medoid technique for clustering purposes and Simulated Annealing for routing purposes. CH selection is selected by minimum energy consumption of nodes, distance among cluster members &CH, and range of communication.

4.2.1 K-MEDOID

This algorithm is a clustering method related to k-means clustering of partitioning datasets in k clusters or groups. Each cluster is represented in k-medoid clusters by data points within a cluster. These points are called clusters medoids.

The term medoid refers to objects in the cluster, the average distance among them and all other CMs is negligible. It affects the most central point in the cluster. These (one per cluster) objects may be taken as representations of cluster members useful in some cases. Note that a particular cluster center is measured as a mean value for all data points in the cluster in k-means clustering.

ALGORITHM

1. Initialize: Choose k random points out of n medoid data points.
2. Allocate each data point with all traditional distance metric methods to the nearest medoid.
3. Although cost is declining:
For every medoid m, for every data or point not medoid:
1. Swap o and m to the nearest medoid for each data point, calculate the cost.
2. If the cumulative cost reaches the previous step, the swap is revoked.

K-medoid is a clear alternative to clustering of k-means. An algorithm is also less vulnerable to noise & outliers relatedto k-means since medoids are used instead of means as cluster centers (utilized in k-means).

The idea behind this approach is that it is good to spread the k initial Cluster Center (CC): the first CC is selected by random from clustered data points, after which the remaining data points are selected in a probability proportional (PP) to its square distance from most similar cluster centers.

4.2.2 CLUSTER HEAD SELECTION:

An effective method of cluster head selection, such as the minimal energy consumption of nodes, the distance amongCMs&CH, as well as communication range should be considered to boost the lifespan of the network. A mobile sink can also move through region R to efficiently protect all CHs. Thus, the best CH for a cluster is chosen &the best route for MDC is determined based onthe suggested approach.

We mention that the Boolean value x_{ij} is the membership of SNs to CHs:

$$\begin{cases} 1, if\ s_i\ joins\ to\ c_j \\ 0, \quad otherwise \end{cases}$$

for 1 ≤I ≤ n, 1 ≤ j ≤ m. There are several clusters separated into the whole network. A cluster node serves as CH&the remaining nodes work as cluster members. CH is chosen as follows: firstly, measure network's total residual energy:

$$E_r = \sum_{i=1}^{n} E_{res}(i) \quad (1)$$

Find average residual energy of sensors, then calculate the threshold value Th:

$$T_h=1/n.E_r: \quad (2)$$

For sensor si to be selected as CH, residual energy of si can be greater than a threshold value, also it can be maximum related to that of other sensors. Further, the total distance among si also sensors in that sub-region must be least. Distance among sensors si(x1; y1)&sj(x2; y2) within sub-region where (x1; y1) are coordinates of sensor si& (x2; y2) are coordinates of sensor sj. CHcj is chosen by the above-mentioned constraints, also its information is spread to whole CMs.

4.2.3 SIMULATED ANNEALING

It is a metaheuristic method, influenced by the annealing procedure in metallurgy. This is a basic optimization process consisting of a cooling &heating controller of material to strengthen its crystal size. Energy is minimized to eliminate defects in metallic structures, as per room temperature. The technique of simulated annealing utilizes its progress in temperature and internal energy as its objective feature. This process continues by key solution S & modified solution formed as S′. Solution for the process is created if fitness function F(S*) value is lower than F(S).

$$Pb= \exp\left(\frac{-(f(s^*)-f(s))}{T_m}\right) \quad (3)$$

The greater fitness value of S* is accepted by stated probability in Equation (3). To prevent interference in local optima, this particular policy facilitates the search process. Where F(S*) is the fitness function of a neighbor solution, and F(S) is the fitness function of the present method. Tm (Temperature) determines the control parameter. An equilibrium state is reached by series of moves &the temperature control parameter is calculated to depend upon the cooling rate. Tm control parameter impacts global search efficacy. A Simulated anneal process would have a better probability if the temperature exceeds a high initial value. The SA protocol would be terminated if there are no changes after the succession of a decrease in temperature. If the original temperature is low & calculation time is shorter potential is further reduced for seeking a global solution.

$$Tm = \delta k + To + Tfn \quad (4)$$

Where as δk is a descendent rate of Tm, 0<δ < 1, k is no. of stints, neighbor solutions formed; To is the original temperature value, and Tfn is the final temperature value. The below algorithm determines the process for SA.

Simulated Annealing Algorithm

Create primary solution s0
Do
Create present solution s0 with neighbor of s0*
Compute probability Pb depends upon Equation (3)
New solutions accept or reject as per Pb
Modify best solution amongst previous one
Reduce Temperature
Do not exceed the stop conditions.

Pseudocode: SA based K-Medoid

{si: Set of sensors in provided region R}
{cj: Set of CHs in si}
{Eres: Sensors residual energy}
{Th: Threshold value for si to be chosen as CH}
{CRi: Max. communication range of si }
{dis(si, cj): Distance amongcj&si}
{Dcj: Degree of CH}
{Xicj: Constraint value for si to join CHcj }
{n: No. of sensor nodes}
{Nd: No. of dead nodes}
Deploy sensors randomly
Partition region R into no. of sub-regions
while T RUE do
loop i:= 1 to n
Compute residual energy of si
end loop
Compute Th by Eq. (5)
Findcj residual energy
Compute distance amongcj&si
Find degree of cj
ComputeXic j by Eq. (10)
Sensor si joins cj if Xicj is maximum
Call ACO with coordinates of CH
Traverse MDC to gather data
Discover Nd
if (Nd = n) then
FALSE
end if
end while

V. RESULTS ILLUSTRATIONS

The efficiency of the proposed approach is contrasted with current research work into the various sensor nodes in this section. In the MATLAB environment, a simulation test was done. The parameters of the simulation are demonstrated in Table 1. Figure 3 indicates the deployment and sink position of fundamental sensor nodes.

Table 1. Parameters detail

Value of Parameters

Network size	200.200
no. of sensors	100–500
no. of rounds	3000–5000
Sensor energy	0.5 J
BS position	(100, 100)
Clustering	Dynamic
CH probability	0.05
collection of data	MDC

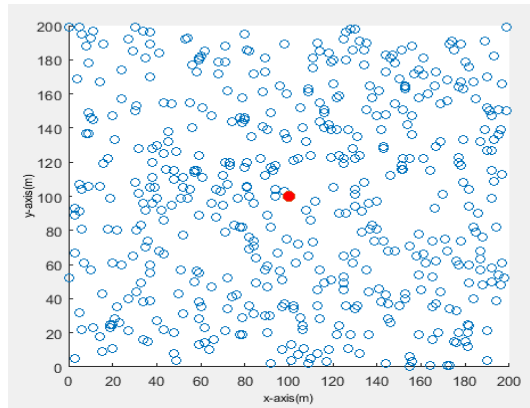

Fig. 3. Nodes Deployment

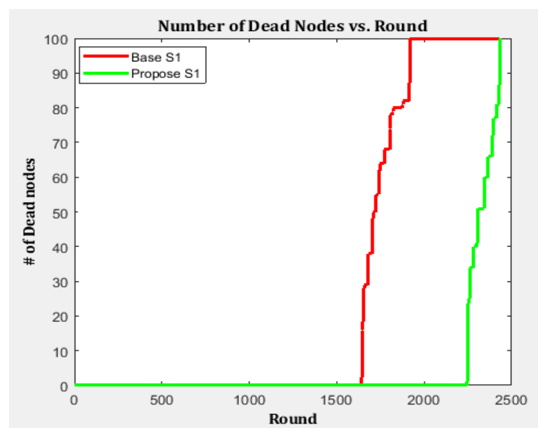

Fig. 4. Network lifetime of 100 nodes

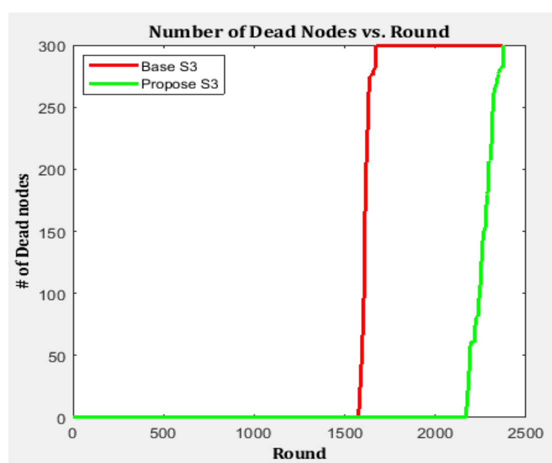

Fig. 5. Network lifetime of 300 nodes

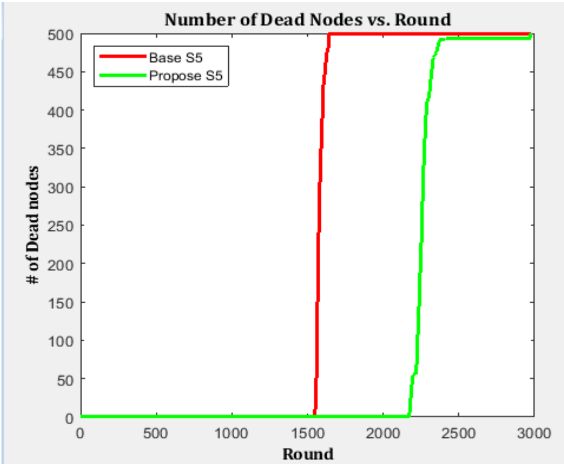

Fig. 6. Network lifetime of 500 nodes

When the first node dies andthe last node dies correspondingly, the network lifespan of previous algorithms, as well as the proposed algorithm, are compared in Figures 4, 5, and 6. Initially, the Algorithm is tested with dead conditions for the first and last nodes in terms of network existence. The number of SNs for this assessment ranges between 100 and 500.

The above 3 figures indicate that the proposed method completes an extended network's lifespan than the previous algorithm. Also, the proposed algorithm guarantees effective selection & gateway avoidance. The proposed algorithm in every iteration selects cluster head depending on the node threshold value. This strategy guarantees dynamic clustering by load balancing & avoids premature sensor death.

In Figures 7, 8, and 9 below, the existing energy usage is compared to the ***proposed*** algorithm's energy consumption. The number of sensors is varied from 100 to 500. All figures show that with different sensor nodes the proposed algorithm requires less energy than existing approaches.

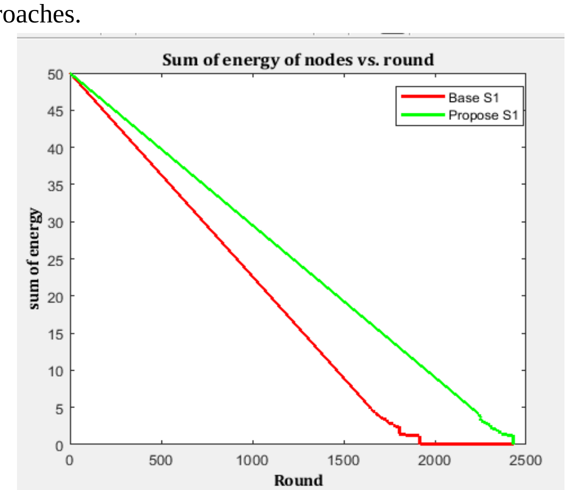

Fig. 7. Remaining energy of 100 nodes

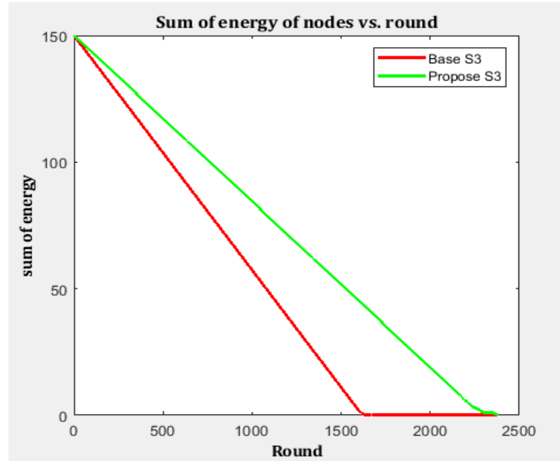

Fig. 8. Remaining energy of 300 nodes

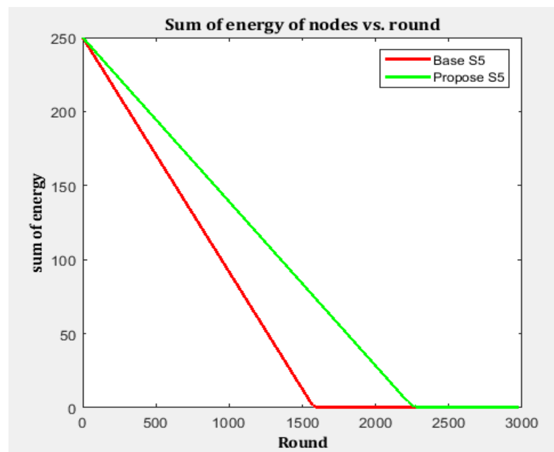

Fig. 9. Remaining energy of 500 nodes

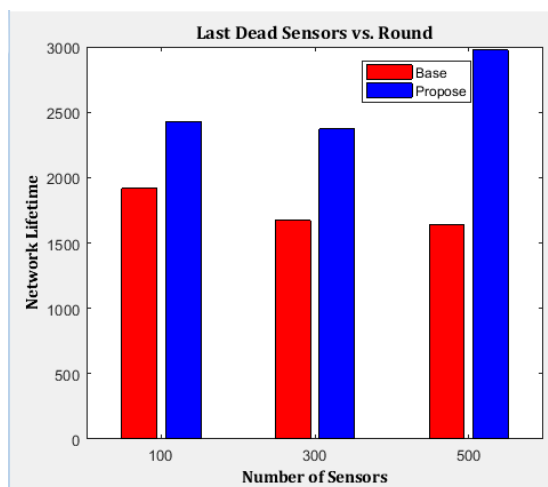

Fig. 10. Comparison graph of the existing and propose work for last node dead round

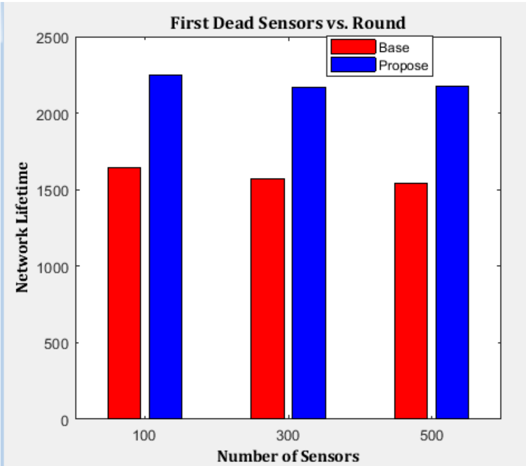

Fig. 11. Comparison graph of the existing and propose work for first node dead round

VI. CONCLUSION

Since the energy & lifetime of any WSN routing protocol are two key constraints, a lot of research work has been conducted to accomplish the goal. It is a complicated method to select an energy-efficient routing-algorithm that equally allocates load throughout the network.

This paper suggested an effective routing algorithm for path selection for mobile sinks based on SA. Proposed method comprises 2 steps: energy-efficient load-balanced clustering and an effective dataset by SA mobile sinks. Findings demonstrate that the presented clustering algorithm balances energy by the selection of CHin every round as well as minimizes CMs' energy utilization during transmission. Due to their complex CH selection at every round, the CH selection algorithm is implemented by an appropriate load balancing method. The findings show that the standard SA activity for mobile sink routing increases the total network life of the sensor thus ensuring maximum functionality. The experimental findings also reveal that the efficiency of the proposed CH selection algorithm with SA mobile sink routing is improved than that of the previous algorithm.

In the future, MDC (Mobile Data Collector)can be deployed with additional optimization algorithms based on the needs of the application with the self-healing method.

REFERENCES

[1] Liu X. A survey on clustering routing protocols in wireless sensor networks. Sensors. 2012;12:11113–11153.

[2] Latif K., Javaid N., Saqib M.N., Khan Z.A., Alrajeh N. Energy consumption model for density controlled divide-and-rule scheme for energy-efficient routing in wireless sensor networks. Int. J. Ad Hoc Ubiquitous Comput. 2016;21:130–139.

[3] S. Chinnappen-Rimer and G.P. Hancke, "Calculation of an optimum Mobile Sink path in a Wireless Sensor Network", pp.49-70, Chapter3, INTECH,

[4] Shah R.C., Roy S., Jain S., Brunette W. Data MULEs: Modeling a three-tier architecture for sparse sensor networks. Ad Hoc Netw. 2003;1:215–233.

[5] Wang J., Zuo L., Shen J., Li B., Lee S. Multiple mobile sink-based routing algorithm for data dissemination in wireless sensor networks. Concurr. Comput. 2015;27:2656–2667.

[6] Wang J., Yin Y., Zhang J., Lee S., Sherratt R.S. Mobility-based energy-efficient and multi-sink algorithms for consumer home networks. IEEE Trans. Consum. Electron. 2013;59:77–84.

[7] Ma J., Chen C., Salomaa J.P. mWSN for large-scale mobile sensing. J. Signal Process. Syst. 2008;51:195–206.

[8] Khan A.W., Abdullah A.H., Anisi M.H., Bangash J.I. A comprehensive study of data collection schemes using mobile sinks in wireless sensor networks. Sensors. 2014;14:2510–2548.

[9] Niayesh Gharaei, Kamalrulnizam Abu Bakar, Siti Zaiton Mohd Hashim Ali Hosseingholi Pourasl ID, Mohammad Siraj and Tasneem Darwish "An Energy-Efficient Mobile Sink-Based Unequal Clustering Mechanism for WSNs" Sensors 2017, 17, 1858

[10] Nagamalar, T., & Rangaswamy, T. R. (2015). Energy-efficient cluster-based approach for data collection in wireless sensor networks with multiple mobile sinks. 2015 International Conference on Industrial Instrumentation and Control (ICIC).

[11] Chen, T. S., Tsai, H. W., Chang, Y. H., & Chen, C. T. (2013).
Geographic convergecast using mobile sink in wireless sensor
networks. Computer communications, 36(4), 445–458.

[12] Xu, J., Liu, W., Lang, F., Zhang, Y., & Wang, C. (2010). Distance measurement model based on RSSI in WSN. Wireless Sensor Networks, 2(8), 606–611.