(SEES)

A SCALABLE & ENERGY EFFICIENT SCHEME FOR HETEROGENEOUS WIRELESS NODES

A Major Project Report

submitted in partial fulfilment of the

requirements for the award of the degree

of

Bachelor of Technology

in

COMPUTER SCIENCE & ENGINEERING

BY

PAPANI SAICHARAN (15115060)

VADDINENI PAVAN KUMAR (15115083)

THADISETTY SRINIJA (15115080)

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY

RAIPUR, CG (INDIA)

APRIL, 2019

ACKNOWLEDGMENT

We would like to thank our project advisor, Dr. Sarsij Tripathi, Assistant Professor, Dept. of CSE, NIT Raipur, for his inspiring guidance, constructive criticism, valuable suggestions, and support. We also thank him for constant supervision as well as for providing necessary information regarding the project. We feel extremely grateful for the opportunity to spend a semester working on a project in a field of great personal interest under his direction.

Every project is having it's own importance in it's respective field. Every attempt is equally importance to form a chunk. We sincerely attempted to create this project by gaining the knowledge from internet and our professor. We are thanking our professor for our proper guidance.

(Our Acknowledgment)

Papani Saicharan (15115060)

Vaddineni Pavan Kumar (15115083)

Thadisetty Srinija (15115080)

ABSTRACT

Developing energy-efficient communication in IoT-based large scale network has become a major challenge since few years. Here main motive is to reduce energy usage for communication in battery-operated devices in order to increase network lifetime and to reduce data transmission cost. In this project, we propose SEES, a scalable energy-efficient scheme for heterogeneous, homogeneous wireless sensor mobile nodes. By using ambient energy-harvesting relay nodes, we study impact of energy-harvesting techniques so that it provides a larger energy conservation and guarantees a longer network life-time. SEES includes: (1) a hybrid-placement of static and mobile nodes, (2) a Multi-Stage Weighted Election heuristic (MSWE), and (3) a Minimum Cost Cross-layer Transmission model (MCCT). Our main aim here is to ensure an random placement of heterogeneous or homogeneous nodes, pre-deterministic and scalable placement of energy-harvesting nodes, fair load balance of energy among all zones and minimum energy cost for transmission of data from bottom to top most layer. SEES is a scheme which supports n-level heterogeneity, m different election parameters (static and dynamic, associated with generated m weights) and can be used for any type of IoT-based deployment. By stimulation we want to compare SEES protocol's network lifetime, conservation of energy, throughput over LEACH, SEP and ZSEP protocols respectively. We will also tune the heterogeneity parameter for better performance.

Keywords: Green IOT, Energy-efficiency, Energy harvesting, IoT-based networks, Heterogeneous objects.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 IoT (INTERNET OF THINGS)

As an innovative technology, the IoT plays a key role in various domains and promises significant advantages that enable a broad variety of applications such as building automation, inventory control, healthcare systems, transportation, smart grid, security, fire detection, asset tracking, border surveillance, smart lighting, etc [1]. IoT-based data collection also called intelligent ambient is one of major application areas. The major task here is to gather data periodically that describes our physical world components and surrounding environment so that necessary action can be taken. This includes intelligent sensing, video surveillance, remote monitoring, big data analytics, etc [2]. Figure 1 .1 shows an example scenario of a typical IoT based heterogeneous network. However, the estimation said that more than 25 billion of smart things will be in use by 2020 [3], enabling an intelligent interconnection without human intervention. To keep pace with the commercial development and help vendors to create standardized IoT products, several efforts and initiatives, driven by major industries and researchers, have been made to provide IoT reference models and architectures [4]. Some of the applications require that the wireless object, which is equipped with a limited capacity battery, has to provide a variety of independent services while some others need that it works in an uninterrupted manner for a longer period of time ranges from a month to many years [5]. As a battery-powered device, a wireless node is no longer able to fulfill its task in the network once its battery energy ran out [6]. An always-on IoT object needs a continuous power supply to extend the network lifetime so as to offer its services as long as possible. Thus, greening efforts and energy-aware designs (hardware and software) have been recently applied to the ICT sector in different optimization trends, for traditional WSN as well as IoT-based systems, and other relevant applications [7]. Energy-efficient communication is the major goal of the green IoT, reaching other goals such as scalability, stability, load balancing, fault tolerance, and optimal resource allocation.

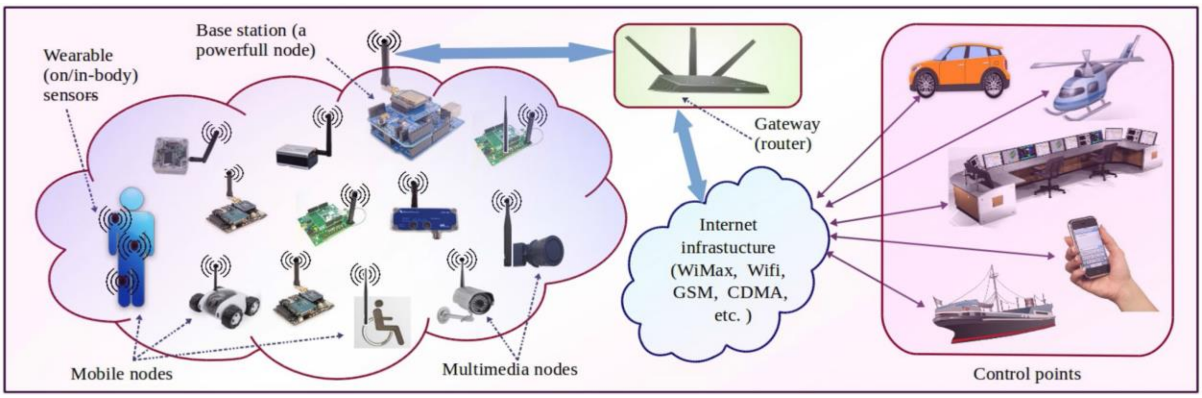


Figure 1. 1: A general architecture of IoT-based heterogenous wireless network

1.2 SEES (SCALABLE ENERGY EFFICIENT SCHEME)

In this project, we implemented a new protocol called a scalable and energy efficient scheme for wireless heterogeneous or homogeneous nodes. We study the effectiveness of cluster-based solutions when utilizing energy-harvesting relay nodes in such a way that enables a high scalability, load balancing and guarantees a longer lifetime. SEES is a heterogeneity-aware algorithm which considers n-types of various objects with n different levels of the initial energy. However, it highly reduces the data transmission cost by dividing the working area into an appropriate number of equal zones depending on the total population of nodes and by providing an energy-harvesting intermediate layer to forward packets from the lower layer to the base stations layer. Further, it elects the zone aggregator nodes using a number of parameters, static and dynamic, including the node energy, in a way that balances the energy load among all the objects in the zones and ensures that only the most robust nodes are the real candidate to perform the aggregation task. However, as there exist many other techniques to save energy (such as resource allocation, data reduction, duty cycling, sleep/wakeup schemes, radio optimization, etc.) as classified by [8], which use different methods and approaches; SEES mainly concentrates on reducing communication energy [the highest energy dissipation source in the wireless sensor nodes [9], and can be adopted in a joint optimization manner, along with such techniques, to significantly improve energy-efficiency.

The major contributions of this project is :

Proposes SEES (a scalable and energy-efficient scheme) for green IoT-based heterogeneous wireless nodes, which includes three main algorithms: (i) A hybrid-placement scheme that statically divides the working area into a number of equal-sized zones based on the node population; randomly distributes the wireless objects in all the resulted zones; and deterministically places the relay nodes at the zones' borders.

(ii) A Multi-Stage Weighted linear combination based Election heuristic (MSWE) that considers a number of static and dynamic parameters, that affect overall network performance, such as node residual energy, centrality, and distance, with a generated weight for each.

(iii) A Minimum Cost Cross-layer Transmission model (MCCT) for data dissemination from heterogeneous objects at the lowest layer to the base stations at the topmost layer.

• Considers energy-harvesting techniques by utilizing ambient energy-harvesting nodes as an intermediate relaying layer, statically distributed in a grid manner over the working area in such a way that enables a high scalability, energy conservation, and load balancing among all the zones.

• Extensive experimental simulations and analysis to compare SEES with other traditional solutions in literature using different scenarios and network conditions.

1.3 WORK FLOW

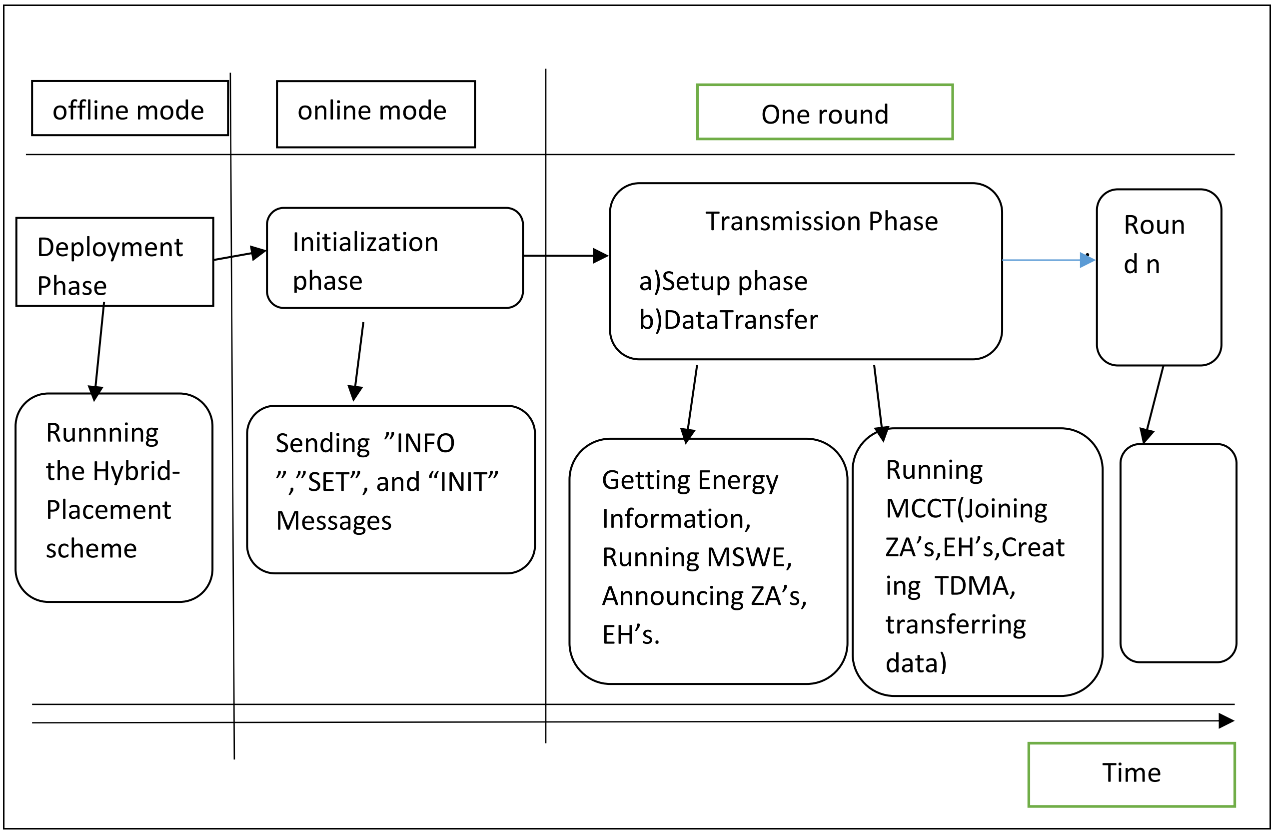


Figure 1. 2: SEES operation phases

2 LITERATURE REVIEW

2.1 RELATED WORKS

In standardization efforts, LoRa, for instance, is one of the LPWAN technologies, which defines a data-link layer with long-range, low-power, and low-bitrate M2M communications (Machine to Machine refers to direct communication without human assistance). LoRa was primarily aiming to communicate to long distance nodes by using less amount of transmission energy. NB-IoT(Narrow Band IoT) is an LTE-based LTE-based narrowband system introduced as a low-cost, low-power, wide-area (LPWA) technology that enables a robust cellular connectivity for a wide range of IoT devices and services aiming to extend the poor coverage of end-user devices, so optimizing their energy-expenditure.

cluster-based solutions have been widely utilized as an effective way to extend the battery lifetime of resource-constrained and energy-limited wireless devices Two types of hierarchical clustering approaches have been identified: dynamic and static.

- In dynamic clustering, the number and size of clusters are changed in a regular interval of time called round
- In static clustering, the clusters are formed in the initialization phase and remain the same during the whole network operation time.

LEACH [10]is one of the most important protocols for energy-efficient homogenous WSNs that introduces a distributed and dynamic cluster forming algorithm. This algorithm rotates the role of cluster-head, which consumes the highest energy, among all the network nodes and hence provides an energy load balancing and prolongs the network lifetime. LEACH-C, a centralized version of LEACH, is introduced by [10]in which the process of creating clusters is run at the base station to find *k* optimal clusters. Here, only those nodes that have a residual energy more than the average energy, are considered as candidates to become cluster heads.

Smaragdakis [11]proposed a Stable Election Protocol (**SEP**) that considers two types of sensor nodes having two different levels of the initial energy. The cluster head is elected based on weighted probability according to the node initial energy where the highest energy nodes become cluster heads more times than the normal nodes.

Z-SEP protocol is proposed by [12] faisal, where they considered both static and dynamic clustering. The working area is partitioned into three rectangular zones: a middle zone containing the normal nodes that directly communicate with the base station in the center, and two head zones containing the advanced nodes that communicate with the base station through forming clusters using techniques as in SEP. However, the way of fixed partitioning and direct communication with the base station may lead to undesirable effects in the case of largescale networks such as IoT.

Huang et al. [13] proposed an energy-efficient deployment scheme for green IoT, introducing a hierarchical framework that uses relay nodes as an intermediate layer depending on the communication radius. The number of relay nodes increases as the transmission range of sensor nodes decreases and vise versa. However, the improvement in the energy saving and network lifetime indicate its flexibility and thus applicability for the IoT-based systems.

Recently, integrating multiple techniques has been applied for better energy savings as well as keeping a desired system performance in terms of other major parameters (such as security, data accuracy, reliability, etc.), depending on the nature of applications. For instance, [9] investigated how can we use cubic B-splines along with F-transform to achieve a low distortion and low computational cost data compression. They applied these techniques to provide a compression encryption scheme that improves the compressed data security.

Daniello proposed a new quality-aware data management framework that integrates the use of virtual sensors and a data estimation technique using rule mining. The virtualized quality-aware sensor network provides different users with sensor data having different levels of quality according to their individual requirements; While the association rule mining is used to estimate the missing values in case of low perception quality. However, due to the lack of space, we have not included further protocols like those proposed by [14], [15], [16].

The complexity and communication overhead, in the dynamic solutions, grow up as the number of nodes increases, which requires more time to form clusters in every round. The zone-based static clustering schemes, which gain the advantages in large-scale networks, do not consider many important factors at the time of zone formation such as the total number of nodes. Further, they do not consider higher levels of node heterogeneity and larger numbers of election parameters or utilize efficient schemes for relay node placement. Hence, they are not flexible enough and become ineffective in the case of IoT-based systems. To address such challenges, and to give a better solution, with the requirements of green IoT as our major objectives, we are motivated to develop a new zone based approach for green IoT-based battery-operated nodes called SEES. We present our solution, in details, in the next section.

3 METHODOLOGY

3.1 SYSTEM MODEL

The network model considered in SEES is supposed to consist of N heterogeneous nodes (HNs), R ambient energy harvesting relay nodes (EHs), B local base stations (LBSs), and one master base station (MBS). Figure 3.1 shows the block diagram of SEES which includes four primary logical components: hybrid-placement, multi-stage weighted election, minimum cost cross-layer data dissemination, and energy-harvesting modules. The hybrid-placement scheme defines the general deployment architecture of the heterogeneous nodes, relay nodes, and base stations in the network field. The multi-stage weighted election heuristic selects a number of nodes as zone aggregators using a new hybrid sorting algorithm . The minimum-cost cross-layer data transmission model defines a multi-layer architecture for policy-based data transmission from working field to the base stations. Table 3.1 presents the notations and quantities used in this paper. For the development of SEES, we have made the following necessary assumptions of the considered network model:

• Here network consists of n level heterogeneity and n different levels of energy

• All HNs can determine their locations, at the setup phase (either by GPS or through any localization algorithm).

• All HNs, EHs, LBSs, and MBS are fixed once deployed and we can also make the nodes mobile to check performance .

• All EHs are homogeneous having the same capability and harvesting rate, and derive their energy from ambient power source.

• All HNs have a power control by which they can vary the transmission energy according to the communication distince between the source and destination nodes.

• All HNs, periodically, has some amount of data to be sent, that is time-driven network model, and these data are correlated.

• Energy dissipation is costed only for data transmission, reception, and aggregation; other energy consumers are neglected.

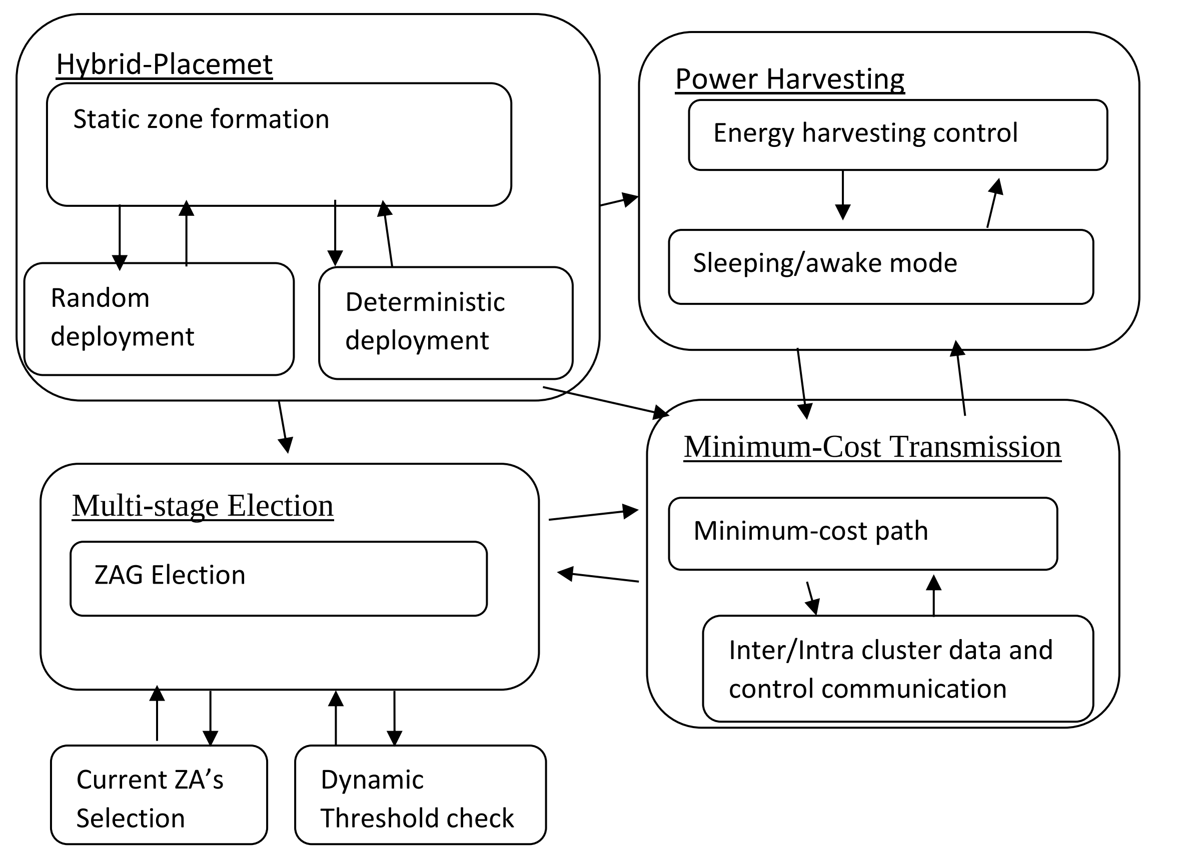


Figure 3. 1:SEES block diagram

Notation	Description	Notation	Description
HNs	Set of heterogenous nodes	ZAs	Set of current zone aggregators
EHs	Set of energy harvestin nodes	N_{zg},N_{za}	Cardinalities of ZAG and ZAs
LBSs	Set of local base stations	th_{zg}	ZAG energy threshold
MBS	Master base stations	β, γ, θ, ά	Heterogeneity parameters and constants
N, R, B	Cardinalities of HNs, EHs, LBSs.	μ, ϑ, κ, Ψ	Election constants
n	Number of heteroginity levels.	E_r^j,C_{nt}^j	Residual energy and centrality of node j
N_t^i,E_t^i	Total number and energy level of cate-i nodes	$d_{ln}^{j\to l}$,$d_{rn}^{j\to c}$	Distance of node j to node c and relay l
R_h	Energy harvesting rate	E_{Tx},E_{Rx}	Energy cost of data transmission and reception
E_{th}^{up},E_{th}^{down}	Upper and lower energy threshold of EHs	$E_{Tx-elec}$	Energy dissipation in transmitter electronics
E_h	Harvested energy at EH node	$E_{Rx-elec}$	Energy dissipation in receiver electronics
F_s	Scalability factor	E_{mp},E_{fs}	Energy of multipath and free space amplifiers
Z	Number of zones	E_{na}^j	Energy consumption of a non-ZA node j
N^z	Number of HN nodes in the zone z	E_{za}^k	Energy consumption of a ZA node k
HNs^{alv}	Set of alive nodes in a particular zone	E_{rn}^c	
N^{alv}		d_{za}^j	
m			
p,w,g			

Sc^j P_i^{hn} N_i^{st} ZAG	Cardinality of HNsalv Number of election parameters Set of params, weights and levels of importance Total WLC score of node j Scaled value of parameter i of node hn. Number of nodes input to election stage-i Set of elected zone aggregator group	d_{rn}^k d_{bs}^c E_{da} $N_{na\to za}^k$ $N_{na\to rn}^c$	Energy consumption of a relay node c Distance of a non-ZA node j to the nearest ZA Distance of a ZA node k to the nearest relay Distance of a EH relay c to the nearest LBS Energy cost of data aggregation Number of non-AZs joined to zone aggregator k Number of non-AZs joined to the relay node c

Table 3. 1: Notation used in this project

3.2 HETEROGENITY MODEL

SEES considers a multilevel model of network node heterogeneity formulas that is given recently by Singh [17].Here generally we have considered n level of heterogeneity which means n types of nodes with n different energy levels. If N is the total number of nodes and n is the number of heterogeneity levels, then the network has n categories (types) of nodes: cate-1, cate-2, cate-3, ..., cate-n nodes with their cardinalities: $N_{t,}^1\ N_t^2\ , N_t^3\ \dots\dots\dots N_t^n$ and energy levels: $E_t^1\ , E_t^2,\ E_t^3, \dots\dots\ E_t^n$ respectively, where:

If a zone consists of more number of nodes then due to load balance of energy property each node in that zone has energy less compared to energy of node belonging to zone with less number of nodes.

$$N_t^1 > N_t^2 > N_t^3 > \cdots > N_t^n \tag{1}$$

$$E_t^1 < E_t^2 < E_t^3 < \cdots < E_t^n \tag{2}$$

The different levels of node energies are related according to the following formula, where the initial energy of cate-i nodes is α times more than that of cate-(i − 1):

$$E_t^I = E_t^1 \times (1 + (I - 1) \times \alpha) \tag{3}$$

where α is constant. The number of cate-i nodes in the network is given by:

$$N_t^i = N \times (\beta - \gamma_1) \times (\beta - \gamma_2) \times (\beta - \gamma_3) \times \dots \times (\beta - \gamma_i) \tag{4}$$

where :

$$\sum_{i=1}^{n} N_t^i = N \tag{5}$$

β and γ_i are the primary and secondary parameters of heterogeneity in the Singh model respectively [17], and are related according to the following formula:

$$(\beta - \gamma_1) \times \left(1 + (\beta - \gamma_2) \times \left(1 + (\beta - \gamma_3) \times \dots \times \left(1 + (\beta - \gamma_n)\right)\right)\right) = 1 \quad (6)$$

and γ_i is given by:

$$\gamma_i = \gamma_{i-1} - 2 \times \theta \quad (7)$$

where θ is constant. The value of first secondary parameter γ_1 is predefined and must satisfy the following inequality:

$$\frac{\gamma_1}{2 \times (n-1)} > \theta \quad (8)$$

3.3 HYBRID PLACEMENT SCHEME

We propose a scalable hybrid-placement scheme that combines three main functions:

(1)Partition of given area into equal sized zones

(2)Random deployment of heterogeneous nodes, and

(3) Efficient placement of relay nodes so as determined at corners of each zone.

We can define the problem as follows: given n different types of heterogeneous nodes (HNs) with corresponding cardinalities $N_{t,}^1\ N_t^2\ , N_t^3\ \dots \dots .. N_t^n$ and energies $E_t^1\ , E_t^2,\ E_t^3, \dots \dots\ E_t^n$, where $N_t^1 + N_t^2 + N_t^3 + \dots \dots .. + N_t^n$ = N, here we need to deploy all the available N nodes to their respective cates whose total size is L*L square meters along with usage of relay nodes and conservation of energy.

First, it determines the number of zones Z, into which A will be divided based on two main parameters: the total number of HN nodes (N) in the network, and the scalability factor (F_s) which we define as the maximum number of nodes to be accommodated in a single zone. The value of Z is given by:

$$Z = \left(\left\lceil \frac{\sqrt{N}}{\sqrt{F_s}} \right\rceil\right)^2 \quad (9)$$

Second, it decides the number of HN nodes that will be deployed in each single zone. according to the following formula:

$$N^z = \begin{cases} \left\lceil \frac{N}{Z} \right\rceil, & if\ z = 1 \\ \left\lceil \frac{N - \sum_{i=1}^{z-1} N^i}{Z - z + 1} \right\rceil & if\ z > 1 \end{cases} \quad for\ z = 1,2,3,4, \dots Z \quad (9)$$

Third, it defines the value of R, the total number of EHs that should be deployed as relay nodes over all the zones as given by:

$$R = (\sqrt{Z}+1)^2 \quad (11)$$

The working area A is divided into Z equal zones each $L/\sqrt{Z} * L/\sqrt{Z}\ \ m^2$. Then each zone is filled with random placement of N^z nodes. Here, we assume that the node sensing and communication ranges r_s and r_c are set to $L/\sqrt{Z}*\sqrt{2}$ and $2* L/\sqrt{Z}*\sqrt{2}$ respectively. Relay nodes at predeterministic location such as corners of each zone i.e, 4 per each zone as shown in figure 3.2. This helps in maintaining high network connectivity and fault tolerance. The EHs are wirelessly connected to LBSs which are, in turn, connected to a powerful base station called the master base station (MBS) which is located at, relatively, a large distance from the working area. The number of LBSs and their exact locations are determined by the network administrator according to the network density and environmental conditions. Algorithm 1, indicates the pseudo code of the zone-based hybrid-placement scheme, and Figure 3. 2 shows an example of a typical deployment architecture for the proposed scheme indicating the distribution of HNs, EHs, LBSs, and MBS.

The number of zones (Z) is directly proportional to the total number of nodes (N) in the network and the number of energy harvesting relay nodes (R) is, in turn, directly proportional to the number of zones (Z). Thus, if there is any increase in N, there will be an increase in Z and consequently in R according to [Eq. 9,10,11]; Hence, preserving network scalability

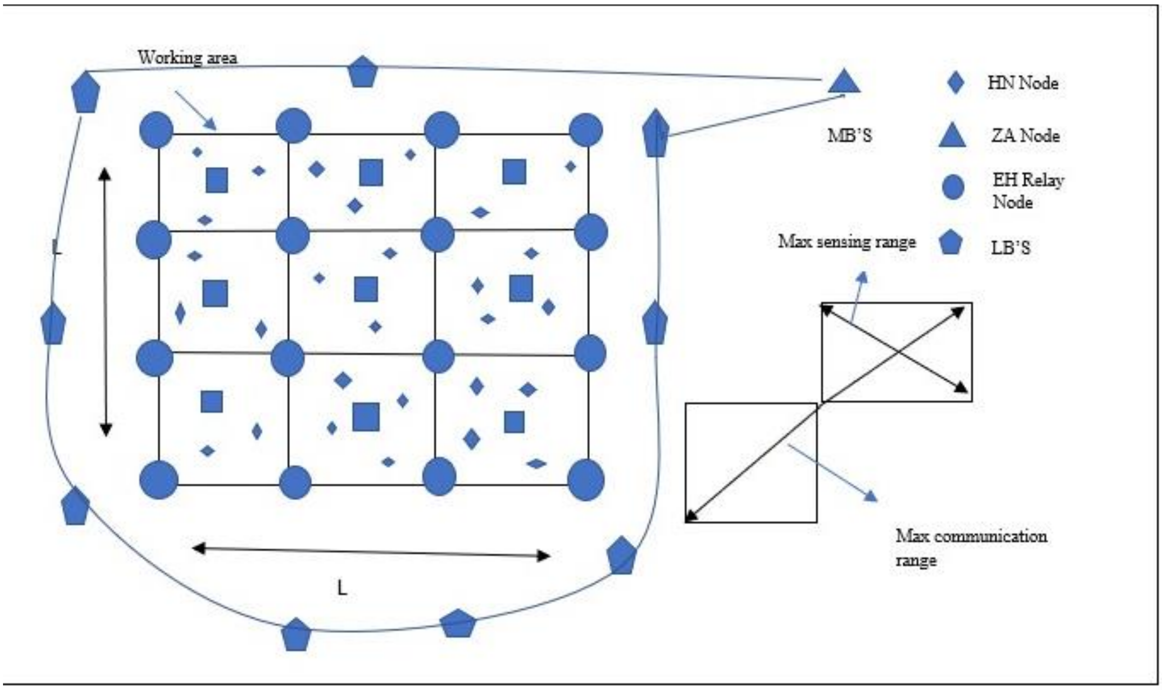


Figure 3. 2:Typical structure of hybrid placement

Algorithm 1: HYBRID PLACEMENT ALGORITHM

Inputs : N, Fs, L, B, n, α, β, Θ, γ

Parameters Setting:

1. Z = $\left(\left[\frac{\sqrt{N}}{\sqrt{F_s}}\right]\right)^2$
2. For each category of nodes cate-I do
3. Compute N_t^i using Equation 4
4. Compute E_t^i using Equation 3
5. End for
6. For each zone z do
7. Find N^z using Equation 12
8. End for
9. For each hn € HNs do
10. Set $r_s^{max} = \frac{L}{\sqrt{Z}} \times \sqrt{2}$
11. Set $r_c^{max} = \frac{L}{\sqrt{Z}} \times \sqrt{2} \times 2$
12. End for
13. $R = \left(\sqrt{Z} + 1\right)^2$

Network Formation:

14. $D = \frac{L}{\sqrt{Z}}$
15. Divide A into Z (D X D) square zones
16. For each zone z do
17. Randomly deploy N^z HN nodes in the zones z
18. Place one EH node at each corner of the zone z
19. End for
20. For each bs € LBSs do
21. Place bs at its predefined location
22. End for

3.4 MUTLI STAGE ELECTION HEURISTIC

SEES provides election heuristic which selects zone aggregation group(ZAG),here ZAG task is to aggregate data from their respective zone nodes and provide this aggregated information to relay nodes. This process is called Multi staged weighted election heuristic different from existing solution [15] etc. MSWE supports up to m different parameters (p = $p_1, p_2, p_3, \ldots, p_m$) arranged in m different levels of importance (g = $g_1, g_2, g_3, \ldots \ldots g_m$). However, it selects ZAG in m consecutive stages (stage-1, stage-2, stage-3, ..., stage-m) based on a set of given weights of importance (w =$w_1, w_2, w_3, \ldots \ldots w_m$). In each stage, one of the given parameters is taken as the dominant parameter (most important parameter) and is given the highest weight (in a stage-i, the dominant parameter is pi and is given w1).

MSWE consists of two main procedures:

(1) ZAG member nodes are elected, and

(2) Select zone aggregators (ZAs) for the current transmission round from ZAG nodes.

SEES provides a general model to generate the set of weights w used in MSWE. The weight w_i assigned to the parameter in the importance level g_i is given according to the following formula:

$$w_i = \mu \times w_{i-1} \; for \; i = 1,2,3, \ldots, m \; and \; 0 < \mu < 1 \tag{12}$$

where:

$$\sum_{i=1}^{m} w_i = 1 \tag{13}$$

and µ is the weighting factor defined based on some considerations (e.g., application nature and election parameters used). µ value can be used to tune weights and obtain balance amond different parameters. This helps giving weightage to static(distance) ,dynamic(residual energy) or some other parameters From [Eq. 12], we have:

$$w_2 = \mu \times w_1 = \mu^1 \times w_1$$

$$w_3 = \mu \times \mu \times w_1 = \mu^2 \times w_1$$

$$w_4 = \mu \times \mu \times \mu \times w_1 = \mu^3 \times w_1$$

Thus we get:

$$w_i = \mu^{i-1} \times w_1 \tag{14}$$

From [Eq 13] we have :

$$w_1 + w_2 + w_3 + \cdots + w_m = 1$$

we get :

$$w_1 + \mu \times w_1 + \mu^2 \times w_1 + \cdots + \mu^{m-1} w_1 = 1$$

Therefore, the value of the initial weight w_1(the weight of the most significant/dominant parameter in every stage) is given by:

$$w_1 = \frac{1}{(1 + \mu + \mu^2 + \mu^3 + \cdots + \mu^{m-1})} \tag{15}$$

Using [Eq 16,17], we can get all the importance weights (w = $w_1, w_2, w_3, \ldots\ldots w_m$). These weights are used in every stage of ZAG election without any change in their values.

In the first procedure of MSWE, ZAG is selected in m consecutive stages according to weighted linear combination. The final score (Sc^j) of an element j is attained by multiplying the standardized value of each parameter/criteria of that element (P_i^j) by the corresponding weight of importance (w_i) assigned to the parameter; and then adding the product of all these values as follows [18]:

$$Sc^j = \sum_{i=1}^{m} P_i^j \times w_i \tag{16}$$

In the first stage of election (stage-1), the set of parameters (p = $p_1, p_2, p_3, \ldots, p_m$) are arranged, initially, based on their levels of significance (g = $g_1, g_2, g_3, \ldots\ldots g_m$) in the network. The most important parameter is the one that has the highest impact on the network performance. Initially, p and g should be arranged as follows:

having :

p1 $\longrightarrow$ g1 (Most Important)

p2 $\longrightarrow$ g2

p3 $\longrightarrow$ g3 $\downarrow$

. .

pn $\longrightarrow$ gn (Least Important)

where $g_1 > g_2 > g_3 > \cdots .. > g_m$.

Then, the total scores of all the alive HN nodes in the network (N_{st}^{1} = N^{alv}) are calculated using [Eq 16].Nodes then are sorted in descending order according to their computed scores. Only the top ϑ% nodes are taken as inputs to the second stage. In stage-i, the pi parameter is taken as the most important parameter and given the highest weight (w_1). The parameters above the pi are shifted one step to the lower importance level, whereas those below the pi remain in the same level of their importance. This procedure continues m times. The input nodes of stage-i (equal to N_{st}^{i}) then are scored using [Eq 16], and then sorted descendingly. Only the top ϑ% nodes out of N_{st}^{i} nodes are considered as input nodes of stage-i + 1 as given by:

$$N_{st}^{i+1} = \vartheta\% \times N_{st}^{i} \qquad (17)$$

where ϑ% is predefined. In the last stage (stage-m), the top ϑ% of the scored nodes are selected as the final ZAG member nodes; thus we have:

$$N_{zg} = \vartheta\% \times N_{st}^{m} \qquad (18)$$

The election process is triggered again at the beginning of a round only if N_{zg} falls below N_{zg}^{min} which is given by:

$$N_{zg}^{min} = \varphi\% \times N^{alv} \qquad (19)$$

where φ% is predefined. This is shown in Algorithm 2 (lines 1–9).

The second procedure in MSWE is to select a number of ZAG member nodes as the zone aggregators (ZAs) for the current round. This runs in every new transmission round, where N_{za} nodes are selected from ZAG group alternately, according to their residual energy, which is given by:

$$N_{za} = \kappa\% \times N^{alv} \qquad (100)$$

where κ% ≤ φ% is chosen in advance and N^{alv} is the number of alive nodes in the current zone.

Algorithm 2: MSWE PROCEDURE
Input : N^{alv},HNs^{alv}, m, p, g, w, ϑ%, κ%, ϕ%,th_{zg}
Checking current ZAG 1. For each zg € ZAG do 2. If zg.energy ≤ th_{zg} then 3. Discard zg 4. $N_{zg} = N_{zg} - 1$

5. End if 6. End for 7. If $N_{zg} = (\phi\% \times N^{alv})$ then 8. Goto ZAs Selection 9. End if New ZAG election: 10. $N_{st}^{1} = N^{alv}$ 11. $HNs^{1} \leftarrow HNs^{alv}$ 12. For each j = 1,2,3…,m do 13. Re-arrange p where $P_j \xleftarrow{takes} g_1, w_1$ 14. For each hn € HNs^{j} do 15. Find total score sc^{j} according to Equation 18 16. End for 17. $HNs \leftarrow sort\ HNs^{j}\ descendingly\ according\ to\ Sc$ 18. $N_{st}^{j+1} = \vartheta\% \times N_{st}^{j},$ 19. $HNs^{j+1} = HNs[1{:}N_{st}^{j+1}]$ 20. End for 21. $N_{zg} = N_{st}^{m+1}$ 22. $ZAS \leftarrow HNs^{m+1}$ 23. End if ZAs Selection:: 24. $ZAS \leftarrow sort\ ZAG\ descendingly\ according\ to\ E_r$ 25. $N_{za} = \kappa\% \times N^{alv}$ 26. $ZAs \leftarrow ZAG[1{:}N_{za}]$
Output : Zag, ZAs

To study the effectiveness of our solution, we consider three election parameters (m = 3): the residual energy (E_r), the centrality (C_{nt}), and the distance to the nearest relay node (d_{rn}). Thus, we have p1 =E_r, p2 = Cnt, and p3 = 1⁄drn arranged in three levels of importance (g1 = 3, g2 = 2, and g3 = 1 respectively); and accordingly, we have three importance weights (w1, w2, and w3) defined as by [Eq 14,15]. However, the ZAG election process is performed in three stages (stage-1, stage-2, and stage-3). According to [Eq. 16], the total score of node j in stage-1 is given by

$$Sc^{j} = w_1 \times f\left(E_r^{j}\right) + w_2 \times f\left(C_{nt}^{j}\right) + w_3 \times f(d_{rn}^{j}) \tag{21}$$

The node centrality C_{nt}^{j} is taken as the inverse of the average distance to all other nodes :

$$C_{nt}^{j} = \frac{N^{z} - 1}{\sum_{c=1}^{N^{z}} d_{rn}^{j \to c}} \tag{112}$$

and the distance to the relay d_{rn} is the minimum value of all distances to the neighbouring EH relay nodes as given by:

$$d_{rn}^{j} = \min_{\forall l \in [1,R]} d_{rn}^{j \to l} \tag{123}$$

where $d_{ln}^{j \to c}$and $d_{rn}^{j \to l}$ are the distances from the HN node j to the HN node c and the EH relay node l respectively. In stage-2, the levels of parameters are changed. C_{nt} becomes the most important one having the highest level (g_1) and the weight w_1. Er is shifted one level down as the second important one with the level of g_2 and weight w_2, while drn remains at the same level (g_3) and having w_3,. Hence, the total score of a node in stage-2 is given as:

$$Sc^{j} = w_2 \times f\left(E_r^{j}\right) + w_1 \times f\left(C_{nt}^{j}\right) + w_3 \times f(d_{rn}^{j})$$

Similarly, the total score of a node in stage-3 is given as:

$$Sc^{j} = w_3 \times f\left(E_r^{j}\right) + w_2 \times f\left(C_{nt}^{j}\right) + w_1 \times f(d_{rn}^{j})$$

Other parameters like N_{st}^{i} , N_{zg} , N_{za} , etc., are determined as described above. In Sect. 5, we show the simulation results of our solution compared to the other protocols.

However, even MSWE does not guarantee that all the nodes in the zone will be a ZA the same number of times. It, instead, ensures that the most robust nodes which have a better balance of all the election parameters (for example, a large enough energy, a higher centrality, and a minimum distance to a relay node) are the best candidates to be selected as ZA nodes. Since we consider n levels of heterogeneity (as described in Sect. 3.2), and that the residual energy is taken as the most important parameter in the initial stage of MSWE, thus, the nodes with the highest levels of energy have the greater opportunity to become ZAs more than the other nodes. In Sect. 3.7, we present SEES operation and explain how it runs (including MSWE) and how the parameter values are acquired. MSWE heuristic is shown in Algorithm 2.

3.5 MINIMUM COST CROSS-LAYER DATA TRASMISSION

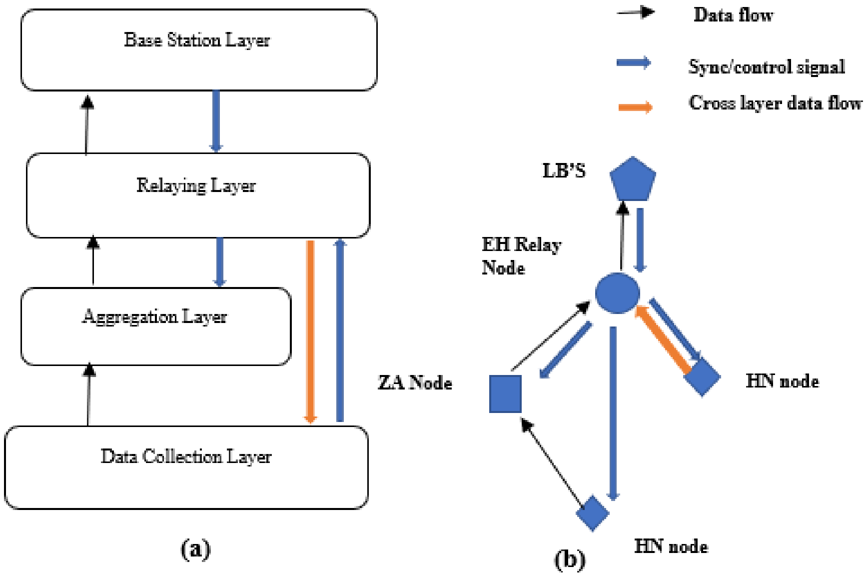


Figure 3. 3:cross-layer data transmission:(a)model architecture(b)transmission paths

SEES provides a Minimum Cost Cross-layer Transmission (MCCT) architecture for energy-efficient data transfer. The collected data can be transmitted to the relying layer either through a singlehop path (cross-layer communication) or a multi-hops path (through *ZA* nodes), whichever has the least energy cost. the transmission framework model consists of four layers: (1) sensing/working layer (the first/lowest layer) comprising all the non-*ZAs* heterogeneous nodes, (2) the aggregation layer (the second layer) consisting of all the zone aggregator nodes (*ZAs*), (3) the intermediate relaying layer (the third layer) including all the *EH* nodes, and (4) the base stations layer (the highest/convergence layer) comprising all the *LBSs* nodes and *MBS* as well. The key function of an *LBS* is to serve a number of nearby zones for further processing/forwarding of received data, coordinating/synchronizing network operation, and running the significant protocols and algorithms. *LBSs* and *MBS* are assumed to have a continuous power supply, and so they are non-constrained in terms of energy.

Algorithm 3 presents MCCT pseudocode.

Algorithm 3: MCCT PROCEDURE
Input : ZAs, EHs
Choosing Next-hop: 1. $nxt \leftarrow Find\ nearest\_relay$ 2. $d_{min} = nxt_{hop}.distance$ 3. **If** (i_am_ZA() == True) **then** 4. Goto **ZA-node**

5. **End if**
6. $my_{za} \leftarrow Find\ nearest\_ZA$
7. $za_{relay} \leftarrow\ Find\ my_{za}.nearest_{relay}$
8. $d_{za} = my_{za}.distance$
9. $d_{rn} = za_{relay}.distance$
10. If $((d_{za}^2 +\ d_{rn}^2) <\ {d_{min}}^2)$ then
11. $nxt_{hop} = my_{za}$
12. $d_{min} = d_{za}$
13. End if

Joining Next-hop:

14. Send “JOIN” message to nxt_{hop}
15. Wait unit “TDMA” scheduling
16. $\mathrm{New\_data}_{\mathrm{packet}}\ \leftarrow\ current\ data$
17. Goto **Data-send:**

ZA-node:

18. $send$ JOIN $message\ to\ nxt_{hop}$
19. Wait unit receiving “TDMA” schedule
20. If (received “JOIN” message from non-ZAs) then
21. Send “TDMA” schedule to non-ZAs
22. End if
23. Received data packets from all non-ZAs in the zone
24. $\mathrm{New\_data}_{\mathrm{packet}}\ \leftarrow\ aggregate\ recieved\ data\ packets$

Data-send:

25. Send $new_{datapacket}\ to\ nxt_{hop}$

3.6 SEES OPERATION

ALGORITHM 4: **SEES MAIN PROCEDURE**
Input: N,F_s,L,B,p,m,n,E_{th}^{down}, E_{th}^{up}, R_h^{max}, th_{zg},α,β,δ,μ,k,Ø
Network Formation**:* 1. Call Hybrid-placement(N,F_s,L,n,α,β,ϒ,Θ) ***Intilisation:

for each hn € HN’s do
    hn broadcasts “INFO” message upward
end for
for each rn € EHs received “INFO” message do
    rn adds its own information from the message
    rn re-broadcasts the “INFO” message upwars
end for
for each bs € LBs received “INFO” message do
    bs extracts information from the message
    bs forwards the message to the MBS
end for
MBS sends SET message to all LBSs
for each bs € LBSs received SET message do
    bs sends “INIT” message downward to its zones.
end for

Data Transmission:

Generate w set according to Equations 16 and 17
for each zone z served by bs € LBSs do
    $N^{alv} = N^{z}$
    while ($N^{alv} > 0$) do
        [N^{alv},HNs^{alv}]=zone_is_alive()
        for each hn € HNs^{alv} do
            hn sends E_r info. to bs
        end for
        for each rn € EHSs do
            rn -> sends E_r info.to bs
        end for
        bs runs MSWWE(N^{alv}, HNs^{alv},m,p,g,w,v,k,Ø,th_{zg}) for zone z
        bs announce ZAs , and EHs info.to zone z
        for each hn € HNs^{alv} ZAs/EHs info do
            hn runs MCCT(ZAs,EHs)
        end for
    end while

34. end for

SEES operation consists of two procedures: (1)deployment (offline) (2)execution (online). It starts by running the hybrid-placement scheme. This is performed in offline mode in which the IoT network is formed by deploying HN nodes either static or mobile, EH relay nodes, and LBSs. SEES then runs the online procedure which consist of two main phases: one-time initialization phase and a repeated transmission phase, as indicated in Figure 1.1.

Intialization Phase : - The nodes in all the zones broadcast an "INFO" messages containing its ID, its type, and its location, to the relaying layer. EH nodes nearby the zones receive these messages, add their own information (i.e. ID, and location), and then re-broadcast to the LBSs layer. The LBSs receive these messages, keep the information for the future use, and forward the messages to the MBS. Once MBS receives "INFO" messages, it sends a "SET" message, to the deployed LBSs, that contains information on which zones are served by which LBSs. Each LBS then send "INIT" message, through the EHs, to the nodes in the concerned zones, which provides the nodes with all necessary information that describes the main functionality of a node, and transmission policies.

Transmission phase:- MSWE and MCCT algorithms are run in this phase in two different distributed manner: MSWE is run at the LBSs level while MCCT scheme is run at the node level as we explained previously. This phase is divided into several repeated rounds similar to what has been in literature [10]. In each round, all the HN nodes at the lowest layer are assumed to have data to be sent to LBSs at the topmost layer. Each round is further divided into a setup phase and a data transfer phase

1)Setup Phase : All the nodes in the lower layers, as well as EH relay layer send the up-to-date information on their current status and residual energy to the serving LBS. Based on this information as well as other already available information, the LBS runs MSWE to perform ZAG⁄ZAs election/removal as described in Sect. 3.4. As aforementioned, ZAG election does not run in every round. Instead, it runs every several rounds when needed. Based on the outcome of MSWE, LBSs inform all the nodes in the zones via broadcasting a small message containing the IDs and locations of nodes selected as ZAs as well as of EH nodes that are in the working mode.

2)Transfer Phase: This includes calculating distances to the announced ZAs, EHs nodes, sending joining messages, creating TDMA schedules and starting data transfer. All these procedures are described in details by MCCT scheme in Sect. 3.5. The control/synchronization messages, here, are also forwarded from LBS downward to non-ZA/ZA/EH nodes in each round according to the network conditions and requirements. The time interval for each round is set such that we ensure that all the HN nodes at the lowest layer can send their data to LBSs at the upper layer through ZA/EH nodes.

3.7 ENERGY DISSIPATION MODEL

We employ the same first order radio energy dissipation model used in literature [10]. We consider only energy consumed during data transmission, reception and dissipations such as static energy and data sensing energy can be neglected [10]. According to this model, the energy consumed at the transmitter is due to running both the radio electronics and power amplifier depending on the length of transmitted data (l) and the distance (d) between the source and destination nodes.

If d is equal to/or less than a given threshold (d0), the free space model (E_{fs}) is used, otherwise, the multipath model (E_{mp}) is used [10]. Thus, the energy consumption at the transmitter side defined as follows:

$$E_{Tx} = \begin{cases} l * \left(E_{Tx-elec} + E_{fs} * d^2\right), if\ d \le d_0 \\ l * \left(E_{Tx-elec} + E_{mp} * d^4\right), otherwise \end{cases} \tag{24}$$

and the energy consumption at the receiver is given by:

$$E_{Rx} = l * E_{Rx-elec} \tag{25}$$

where d_o is given by:

$$d_0 = \sqrt{\frac{E_{fs}}{Emp}} \tag{26}$$

and $E_{Tx-elec}$, $E_{Rx-elec}$, are the energies dissipated to run the transmitter and receiver electronics respectively. However, as for the lowest layer in SEES, the power consumed per round by a non-ZA node j when its distance $d_{za}^{j}$to the nearest ZA or EH is equal to/or less than d_o is given by:

$$E_{na}^{j} = l * (E_{Tx-elec} + E_{fs} * {d_{za}^{j}}^{2}) \tag{27}$$

For the second layer, the energy dissipated per round by a ZA node k when its distance d_{rn}^{k} to the nearest EH node is equal to/or less than < d_o, is given by:

$$E_{za}^{k} = l * \left[E_{Rx-elec} * N_{na\to za}^{k} + E_{da} * \left(N_{na\to za}^{k} + 1\right) + E_{Tx-elec} + E_{fs} * {d_{rn}^{k}}^{2}\right] \qquad (28)$$

where E_{da} is the energy cost of data aggregation(processing) and $N_{na\to za}^{k}$is the number of non-ZA nodes that joined the ZA node k. Similarly, for the relay layer, the energy dissipated by a EH node c per round when the distance dc bs to an LBS is < d_o, is given by:

$$E_{nc}^{r} = l * [E_{rx-elec} * (N_{za\to rn}^{c} + N_{na\to rn}^{c}) + E_{da} * N_{na\to rn}^{c} + E_{Tx-elec} + E_{fs} * {d_{bs}^{c}}^{2}] \qquad (29)$$

where $N_{za\to rn}^{c}$and $N_{na\to rn}^{c}$are the number of ZAs and non-ZAs nodes respectively that joined the relay node c.

4 EXPERIMENTATION

4.1 SIMULATION PARAMETERS

CLASS	PARAMETER	VALUE
NETWORK	N,l,E_{tot} F_s B n α ϒ ϴ E_{int} EH_{cap} l	As per Scenario (Table3) 20 12 2,3,4,5 2 0.4 0.025 0.5J 0.5J 4000bits
RADIO	$E_{Tx-elec}$ $E_{Rc-elec}$ E_{da} E_{fs} E_{mp}	50nJ/bit 50nJ/bit 5nJ/bit/signal 10pJ/bit/m^2 0.0013pJ/bit/m^4
PROTOCOL	µ v k Ø M R_h^{max} R_h^{min} E_{th}^{up} E_{th}^{down}	0.5 0.7 0.1 0.2 3 0.027 0.0036 3/10EH_{cap} 1/10EH_{cap}

Table 4. 1: simulation parameters

The heterogeneity settings are taken such that the total number of nodes and energy value for each level are the same in all the protocols. We set the locations of *LBSs* in SEES to be in the range between 20 and 50 m far beyond the four sides of working area, uniformly distributed at random points. However, the exact locations of *LBSs* do not affect the overall energy consumption of the sensor nodes in the first layer since the nodes communicate their data only to the energy harvesting nodes in the intermediate layer and have no direct communication with *LBSs*.

In all the other compared solutions, the base station is positioned in the middle of the area. Even though this reduces the communication cost in these protocols; nevertheless, it is not the case in the real IoT applications. The *MBS* location is assumed to be at 150 m far away from the working area.

4.2 EVALUATION METRICS

The set of measures and parameters we used for evaluating our SEES and comparing its performance with that of existing protocols are defined as follows:

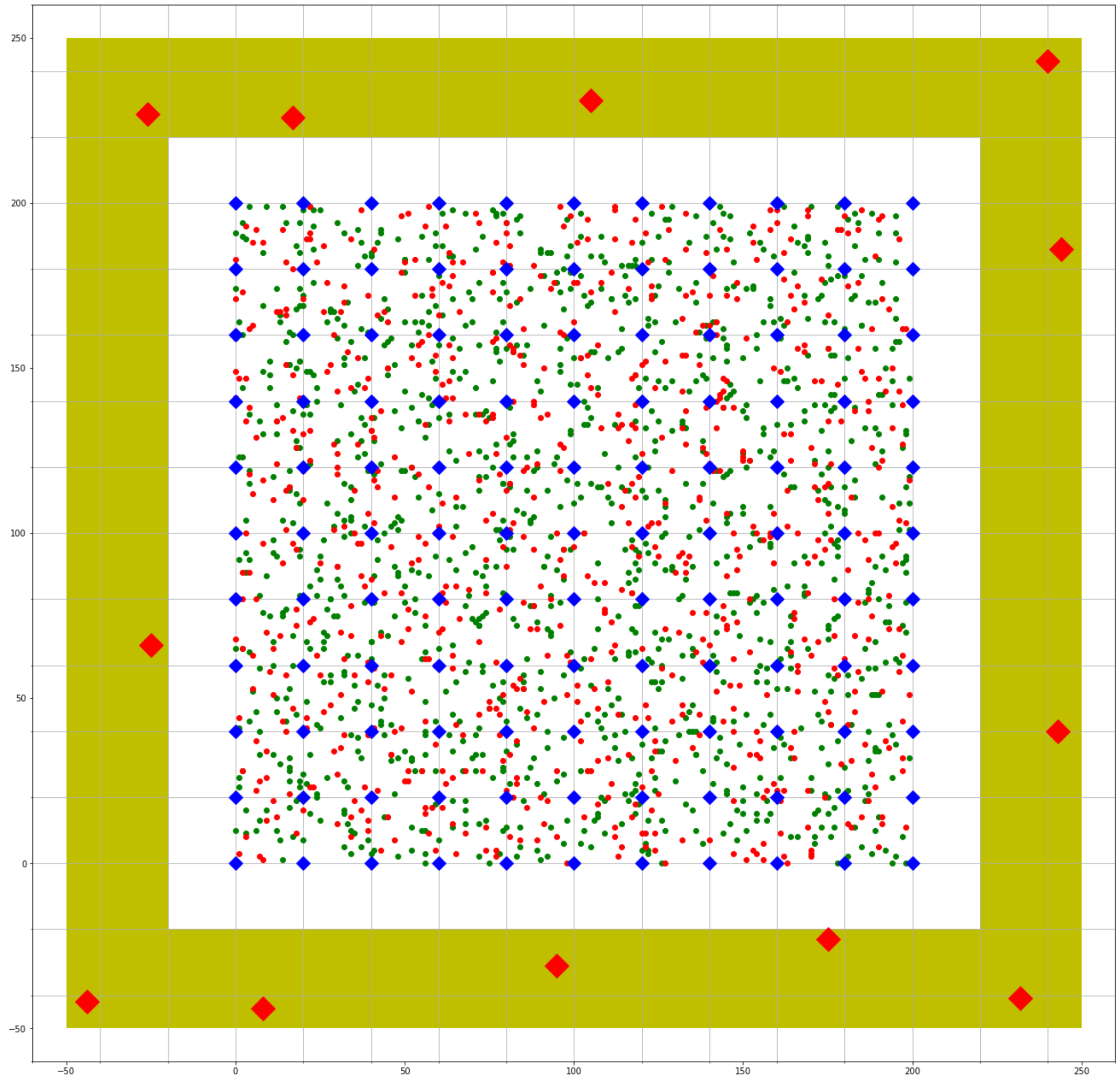
- FND (First Node Dead): measures the length of the stable region which is defined as the number of rounds from the beginning of network operation until the first node dies.
- LND (Last Node Dead): represents the length of the network lifetime, which is defined as the number of rounds from the beginning of network operation until the death of the last alive node.
- HND (Half Nodes Dead): indicates the number of rounds from the beginning of network operation until the death of 50% of the network nodes.
- Number of alive/(dead) nodes: gives the total number of nodes still active/(have expired) with respect to rounds.
- Lifetime gain: measures the percentage improvement in the network lifetime gained by SEES over the other protocols after the death of P% of the network nodes.
- Energy gain: measures the percentage of energy saved by SEES when it runs for the same number of rounds as the other protocols.
- Residual/(consumed) energy: shows the total amount of remaining/(dissipated) energy for all the network heterogeneous nodes with respect to rounds.
- Node energy-cost: measures the average amount of energy dissipated by a single node per round. This is computed by first taking the average energy for a single node per each round and then taking the average of these values with respect to the total number of rounds.
- Network throughput: represents the total amount of data (in gigabits or number of packets) sent from the lower layer to base stations in the upper layer.
- Data energy-cost: gives the amount of energy required to get a particular amount of data to the base stations.
- Throughput gain: measures the percentage of improvement in throughput achieved by SEES when it runs for the same portion of the total network lifetime as the other protocols.
- Scalability metrics: measure the ability to function well in the case of any increase in the number of nodes or area size without affecting the network performance. Here, we present some of the measures defined above with different values of the total number of nodes.

5 RESULT

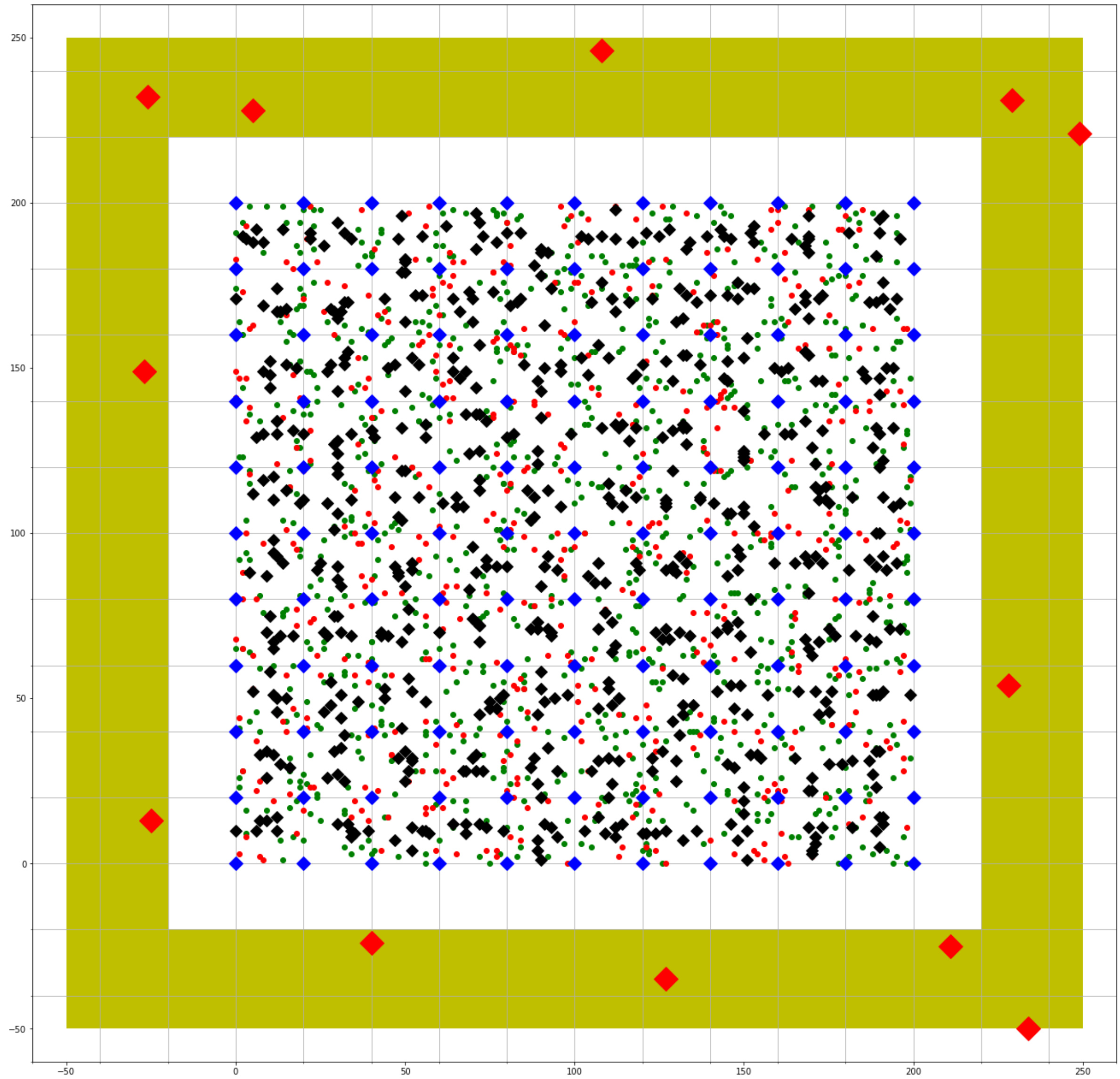
Hybrid Placement Algo:

Input:

1. Enter N(the total number of HN nodes)2000
2. Enter Fs()20
3. Enter L(length of area)200
4. Enter B(LBS)12
5. Enter n(number of heterogeneity level)2
6. Enter alpha2
7. Enter gamma0.4
8. Enter theta0.025
9. Enter E1(initial energy)0.5



MSWE Algo:
Input: Enter the μ value0.8 Enter m value3 Enter the cut of Zag = 0.7



6 FUTURE WORK

ws

7 Hkhj

8 jn

9 ff

References

[1] Maksimovic, “Greening the future: green internet of things as a key Technology enabler of sustainable developemnet,” Springer, 2018, pp. 283-313.

[2] S. Q. F. N. D. S. W. H. V. A. Qin Y, “ When things matter: a survey on data-centric IoT,” 2016, pp. 137-153.

[3] R. M. B. M. A. P. B. P. Mahdavinejad MS, “ Digital Communications and Networks,” 2017.

[4] E. C. Weyrich M, in *Reference architectures for IoT*, 2016, pp. 112-116.

[5] C. L. H. W. (. Soyata T, in *Rf energy harvesting for embedded systems*, 2016, pp. 22-57.

[6] M. D. Z. T. Sohraby K, Wireless sensor networks, 2007.

[7] J. R. Gandotra P, A survey on green community and security challanges 5g, 2017.

[8] B. A. C. Y. Rault T, “Energy efficiency in wireless sensor networks,” 2014, pp. 104-122.

[9] L. V. T. S. Gaeta M, in *Cubic bspline fuzzy transforms*, 2016, pp. 18-30.

[10] C. A. B. H. Heinzelman WB, in *An application-specific protocol architecture for wireless microsensor networks.*, 2002, pp. 660-670.

[11] M. I. B. A. Smaragdakis G, “Sep: a stable election protocol for clustered heterogeneous wireless sensor network,” International Workshop on SANPA, 2004, p. 251–261.

[12] J. N. J. A. K. M. B. S. K. Z. Faisal S, “ Z-SEP: zonal-stable election protocol for wireless sensor networks.,” J Basic Appl Sci Res (JBASR), (2013), p. 132–139.

[13] M. Y. G. X. L. Y. D. Q. Huang J, in *A novel deployment scheme for green internet of things*, (2014), p. 196–205.

[14] R. C. S. K. Lindsey S, in *Data gathering algorithms in sensor network using enrgy metrics*, 2002, pp. 924-935.

[15] P. R. Gupta V, in *An improved energy aware distributed unequal clustering protocol for heterogeneous wireless nodes*, 2016, pp. 1015-1058.

[16] R. S. C. V. M. R. S. H. A. S. Brar GS, in *Energy efficient direction-based pdorp routing protocol for wsn.*, 2016, pp. 3182-3194.

[17] S. S, in *Energy efficient multilevel network model for heterogeneous WSNs*, 2017, pp. 105-115.

[18] A. L. Samo D, Multi-attribute decision analysis in gis:weighted linear combination and ordered weighted averaging, 2009.

[19] W. P. N. D. K. D. H. Z. Lu X, in *Wireless networks with rf energy harvesting*, 2015, pp. 757-789.

[20] Z. X. D. Y. L. X. Meng J, in *Adaptive energy-harvesting aware clustering routing protocol for wireless sensor networks*, 2012, pp. 742-747.