



P-LUET: A Prolong Lines of Uniformity Based Enhanced Threshold Algorithm for Heterogeneous Wireless Sensor Network Enabled Internet of Things Framework

Tanvi Sood¹ · Kanika Sharma¹

Accepted: 10 May 2021 / Published online: 9 June 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

With expanding realms of Internet of Things (IoT), researchers have started venturing into designing such algorithms for Wireless Sensor Networks (WSN) that support IoT network requirements. However, collaborating a sensor network into Internet of Things applicability faces surging data flow amidst which the sensor network is expected to provide reliable data over extended span. Amidst variant techniques like routing, data aggregation, packet scheduling etc. which can be improvised upon, clustering has been a widespread energy efficient technique that primarily shortens distances in large-scale networks while also bring down commuting packets over improved connectivity. Therefore, utilizing clustering to incorporate WSN enabled IoT (WSN-IoT) standards becomes the primary focus of this paper. Heuristic based clustering algorithm termed as Prolong—Lines of Uniformity based Energy Threshold protocol (P-LUET) has been proposed that focusses on expanding the stable operating period of WSN-IoT. This algorithm is based on certain measures of WSN-IoT's per unit, that is, residual energy, cartesian coordinate based location, shadow distance from the Sink Node, and density within the field. The parameters are employed such that they help combat hotspot issue, cluster overlapping, and network connectivity. A cluster-based conditional gridding Voronoi structure has also been consolidated that enables multi-hop communication. P-LUET algorithm is thoroughly analysed through comparisons with the existing approaches. Furthermore, analysis on P-LUET has also been carried on scenarios that are based on heterogeneity synthesis and b , c parameters.

Keywords Cluster-head · Internet of things · Stability period · Energy efficiency · Heterogeneity · Hotspot issue · Wireless sensor network

✉ Tanvi Sood
ece.tanvisood@gmail.com

¹ Department of Electronics and Communication Engineering, NITTTR, Panjab University, Chandigarh 160019, India

1 Introduction

Wireless Sensor Network—enabled Internet of Things (WSN-IoT) (Fig. 1) is an innovation in the midst of numerous advances in micro sensors, VLSI and wireless communication based technologies [1]. Apart from applications relating to environment, home, healthcare, security, infiltration detection across LoC, military, industrial process control,

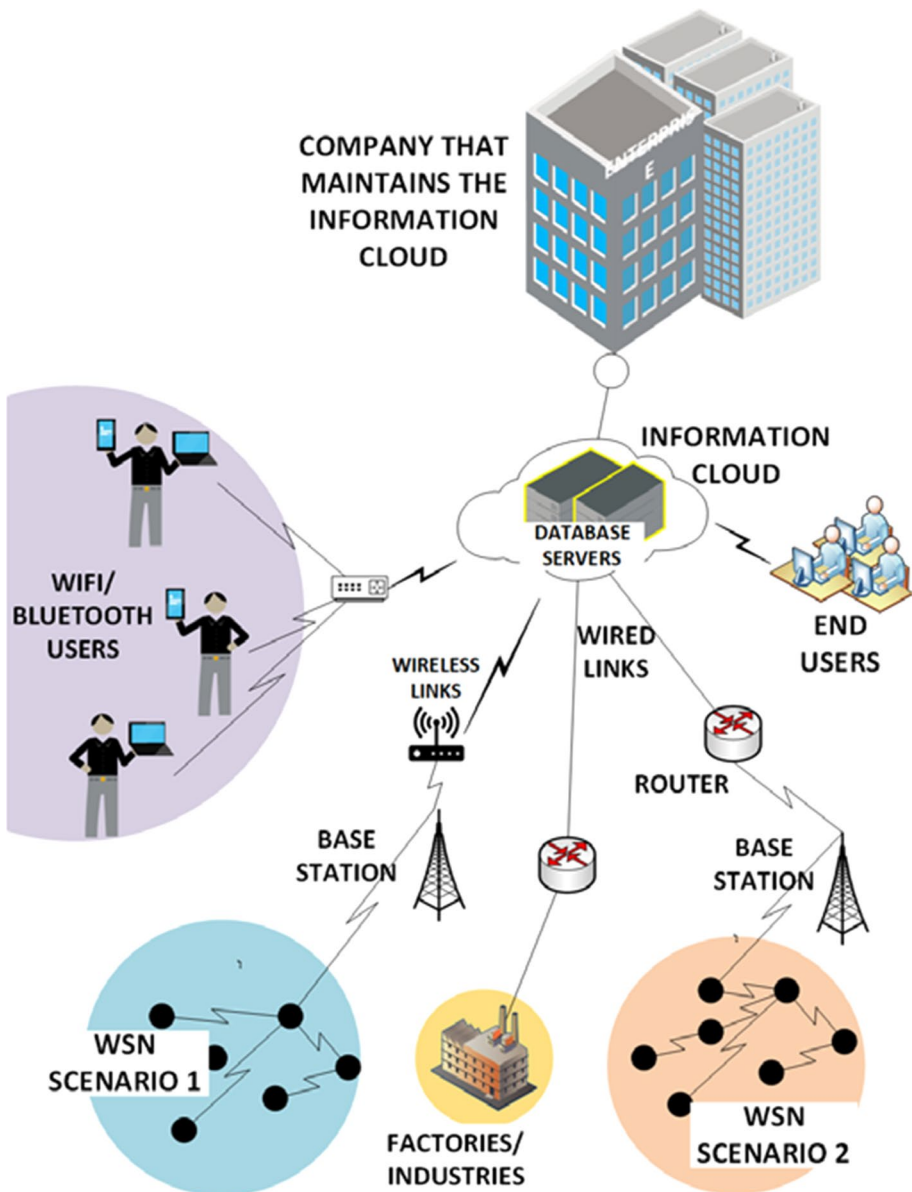


Fig. 1 Wireless Sensor Network (WSN-IoT) scenario enabled with Internet of Things (IoT)

civil, kindergartens etc. [2]–[5], WSN also adds benefit in IoT infrastructure by allowing researchers to realise a dynamic and robust paradigm [6, 7]. WSN’s autonomous energy-efficient resilient traits, dynamic network size, energy-heterogeneity, functionality of quality of service (QoS) allows it to dominate data collection task for any IoT framework [8–10]. Several real-time WSN-IoT applications are WaIoT—no flooding system based on IoT [11], Smart IoT-Connected Railways by SADEL, AAEON and Intel [12], Bosch Early Warning System [13], Intelligent Manufacturing by NEXCOM and Intel IoT Gateway [14] etc.

As WSN ventures into IoT, it efficiently uses machine-to-machine connectivity to gather field data and make it accessible anywhere [15, 16]. WSN-IoT has innovated its services to serve many applications like smart home, smart city, smart industries, smart health care, smart environment, smart transportation and many more. And therefore, with such advances, the need for efficiency infiltrated. It demanded an improvement in efficient acquiring of data monitored by IoT end devices scattered within the field area. These end devices, which hold limited energy resources and loose major battery power during a communication, have to comply with the constraints even while fulfilling the efficiency attainment demand. The WSN-IoT architectural framework follows a two-layered hierarchy of sensing layer and IoT layer as can be seen from Fig. 2. With the WSN mainly required for acquiring field data or any other environment data, it is contained in the sensing layer of the architectural framework that consist of numerous randomly deployed ID-enabled (not necessarily connected to internet and are alternatively known as sensor nodes) or IP-enabled end devices. These end-devices are equipped with one or more micro-sensors dependent of the kind of IoT application like smart homes can use IR, proximity, motion, and video sensors to detect any intrusion, smart city can use acoustic, temperature and seismic sensors to detect any natural calamity, smart industries can use gas, pollutant, and humidity sensors to prevent any industrial damage etc. [17–20]. The end-devices are connected to one or more static/mobile IoT hub in the IoT layer that performs the essential function of a Sink Node SN (or a Base Station BS) and allows to access the retrieved data from the

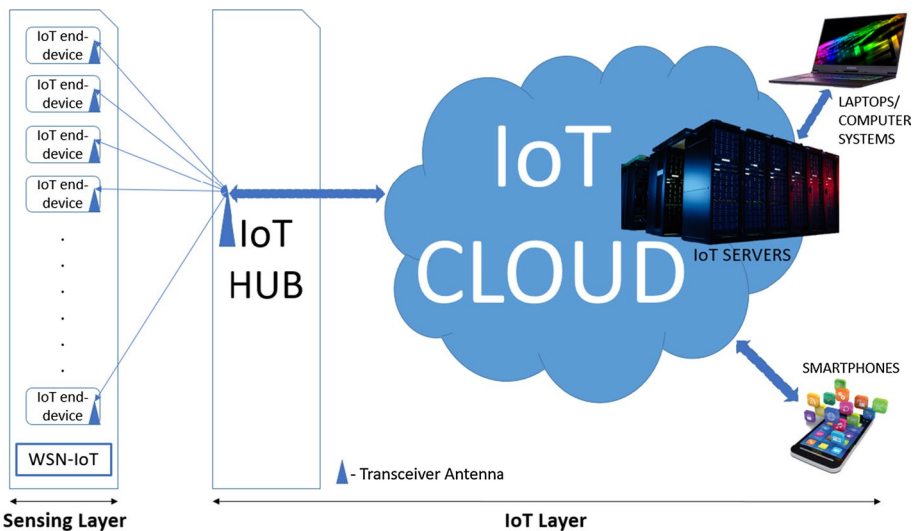


Fig. 2 WSN-IoT architectural framework

field from anywhere across the globe. The IoT hub can either lie in the field surrounding the phenomenon or may lie outside it. A flat routing arrangement between the end-devices and the IoT hub resonates a computational complexity of $O(n_{ini}^2)$ where n_{ini} is the initial arrangement count of a WSN which clearly describes an unstable operation as the population rises as can also be graphically seen from Fig. 3 [21]. Therefore, in order to enhance the scope of research in this domain, a need to prompt efficient connectivity among these IoT end-devices and with the IoT hub arises. Among the various ways of achieving an improved connectivity through minimal communication, the prominently used methods are as follows.

- Efficient grouping of the IoT end-devices.
- Strategically defining the commuting hop between these end-devices and their corresponding IoT-hub(s) such that efficiency improves over nominal computational complexity [22].

Grouping of the end-devices within a WSN either follows a Voronoi structure or a non-Voronoi structure. The most common domain of grouping has been embarked upon by employing Voronoi structure which either adopts a distributed clustering strategy or randomness-based strategy or an adaptive clustering strategy and have a basic purpose of averting transmittal of spatially correlated sensed data [23]. Clustering of IoT end devices successfully carries out the task of acquiring data from the field over reduced communication and the corresponding overhead. Clustering involves choosing one or more heads amongst the network population to other end-devices that primarily have the job of communicating with the IoT hub. These heads are popularly termed as cluster heads (CHs) which become the intermediary for the sensing layer. However, for efficient operation, there is a constant need to keep researching and innovating various CH-selection techniques that are based on certain characteristic measures of the network or of its end-devices. However, this operation is quite a challenging task as choosing an optimal set of CHs in every round of network life is an NP-hard problem where round is one cycle of sensing, data acquisition and communication to IoT hub.

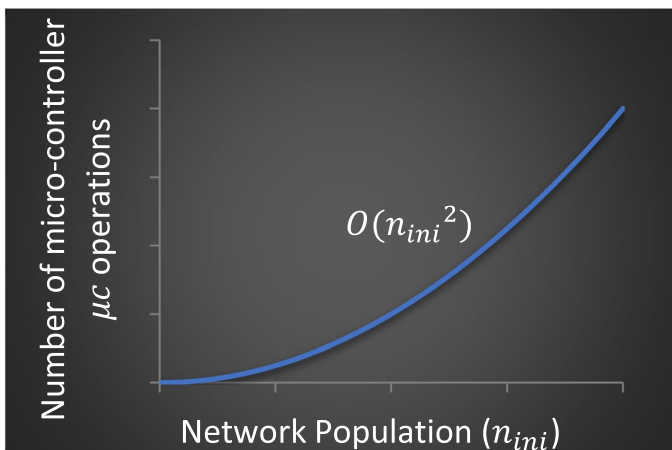


Fig. 3 Computational Complexity of flat-routed WSN-IoT

An ordinary WSN-IoT consists of hundreds or probably thousands of end-devices with fewer enacting the role of a CH. With limited energy resources and communication range, there is a possibility that a fewer section of network population (known as isolated devices) faces no connectivity with any CH of the running round causing them to commute directly with the IoT hub. As long-distance communication leads to greater energy drainage, greater isolated nodes leads to reduced network efficiency. This issue is mostly observed in environments requiring large networks. WSN-IoT may also sometimes tend to face overlapping clusters in densely populated areas that further raise the chances of facing device isolation in the other parts of the network. This may add as a disadvantage to the device isolation. Another arrogant design concern of large-scale clustered WSN-IoT is hotspot issue that hinders lifespan of the devices closer to the IoT hub that may hold a higher probability of being chosen as a CH owing to their vicinity with the IoT hub. This concern can be projected through right measures for CH selection and/or communication pattern for all the devices. An appropriate algorithm that optimally defines such set of nodes as CH that help overcome or partially overcome these issues becomes the need of the hour solution for WSN-IoT.

1.1 Major Contributions

The prominent focus of the proposed Prolong-LUET algorithm is on prolonging the stability period of the LUET (Lines of Uniformity based Enhanced Threshold) protocol [24] (which is our previous work) in a WSN-IoT environment. The main contributions addressed in this work are stated below.

- a. Apart from incorporating vicinity to the Lines of Uniformity LoU in a field, algorithm also requires the end-devices to calculate their shadow distance from their IoT hub based on this vicinity parameter to alternate the perception of obtaining the annulus ring.
- b. With the IoT hub considered within the field limits, concentric rings around it define a weight for every end-device in such a way that it helps overcome the hotspot issue.
- c. CH selection is based on a rank value that additionally depends on the device density and its residual energy apart from its shadow distance and weight value.
- d. Conditional gridding for a partial lifetime is adopted within algorithm to help overcome overlapping clusters to a certain limit.
- e. Conditional multi-hop communication has been adopted.
- f. Simulation results assess the dominance of P-LUET over the existing state-of-the-art clustering algorithms while also accessing the P-LUET performance on scenarios based on heterogeneous sensitivity and b, c parameters.

1.2 Methodology Workflow

The rest of the manuscript is coordinated as follows. Section 2 elaborates the shift in the concerned research domain of design of CH list from homogeneity to heterogeneity to improvisation in probability definitions within threshold function values to indulgence of heuristic metrics to meta-heuristic based schemes to being efficiently available for IoT incorporation and the motivation for the proposed work. Section 3 presents the preliminary notations for the opted WSN-IoT scenario. Section 4 describes the proposed methodology, and the simulation results and their analysis is given in Sect. 5. Finally, Sect. 6 concludes the paper.

2 Related Work

The main goal of the proposed algorithm is to prolong stability period over improved energy efficiency during the design of a CH-selection based clusters in sensing layer of WSN-IoT architectural framework. In practise there is a need to integrate IoT end devices with the IoT cloud for any real application. This can be achieved either via adding a few IP-enabled end-devices among the other ID-enabled end-devices or via connecting an IP-enabled IoT hub with the locally connected ID-enabled end-devices. As pointed out earlier, due to the constrained resources of an end-device (ID-/IP-enabled), there is a need to perform grouping which could either be grid-based or clustering-based. Therefore, popular network topologies have been based on a grid structure or on a clustered strategy. The approach of clustering was first used in Low Energy Adaptive Clustering Hierarchy protocol (LEACH) [25] for a homogeneous environment that divided every round of lifetime into setup phase and steady state phase. Setup phase involved CH election based on a probabilistic fitness function P_{opt} and setup of a cluster layout while a steady state phase observed uploading of data from non-CH nodes to their corresponding CH where after removal of data redundancy, aggregated data is forwarded to network's SN by these CHs. However, despite shortened commuting distances, slower power depletion, longer life, the randomness-based protocol faces reduced connectivity coverage due to non-uniform CH distribution which brings down the overall efficiency of a CH selection procedure. Thereafter, many heuristic/meta-heuristic approaches were proposed to optimize the CH-selection. Energy-homogeneity could not be relied upon for real-time applications as it contrasts with the environment dynamicity and observed unreliable stability-lifetime relationship. Later, to allow deployment of WSN in a real-time environment, energy heterogeneity was introduced through SEP protocol [26] that replaced P_{opt} with weighted election probabilities for a two-level (: 2) heterogeneity while an enhancement to a three-level (: 3) heterogeneity was introduced in EEHC protocol [27]. Both the protocols considered sub-epochs for the heterogeneous end-devices of the WSN. However, the issue of non-uniform CH distribution was still not addressed. Apart from transforming sub-epochs, the dynamicity in the environment also required nodes to adapt their epoch duration according to their energy usage pattern. This led to the discovery of DEEC protocol [28] that adhered to the shortcomings and included an energy parameter (ratio of residual and average energy) in the weighted probability function of SEP, $P_{wgt} = \{P_{nrm:2}, P_{adv:2}\}$. The same concept was applied on three-level heterogeneity in EDEEC protocol [29] on weighted probability functions for normal, advanced and super-advanced devices ($P_{nrm:3}$, $P_{adv:3}$ and $P_{sadv:3}$) as can be seen from Eqs. (1–3), respectively.

$$P_{nrm:3} = \frac{P_{opt}E_i(r)}{(1 + (\alpha + m_o\beta)m)\bar{E}(r)} \quad (1)$$

$$P_{adv:3} = \frac{P_{opt}(1 + \alpha)E_i(r)}{(1 + (\alpha + m_o\beta)m)\bar{E}(r)} \quad (2)$$

$$P_{sadv:3} = \frac{P_{opt}(1 + \beta)E_i(r)}{(1 + (\alpha + m_o\beta)m)\bar{E}(r)} \quad (3)$$

Improved epoch sure enough introduced dynamicity, however, it also observed a higher sub-epoch for the super-advanced (or advanced) end-devices leading to their early decay.

Therefore, DDEEC protocol [30] made use of a threshold residual energy Th_{res} parameter, $E_i(0.7)$ that lets both the end-devices to switch to a common probability function in a two-level heterogeneous network as shown in Eq. (4).

$$P_{wgt} = \begin{cases} P_{nrm:2}, & E_i > Th_{res} \ \&\& \ end - device \in nrm \\ P_{adv:2}, & E_i > Th_{res} \ \&\& \ end - device \in adv \\ c \frac{P_{opt(1+\alpha)E_i(r)}}{(1+\alpha)E_i(r)}, & \forall c = [0, 1], E_i \leq Th_{res} \end{cases} \quad (4)$$

A similar Th_{res} parameter was used for three-level heterogeneity network in EDDEEC protocol [31]. In order to extend dynamicity, an n -level heterogeneity based hetDEEC protocol [32] was proposed that could be applied on 1/2/3-level heterogeneous network model just by varying a Θ parameter. Along the research journey, many heuristic based CH-selection approaches have been proposed that aimed at improving certain sensor network characteristics which could be implied via network's certain performance metrics. ATEER [33] adopted three-level heterogeneous network that utilized $P_{nrm:3}$, $P_{adv:3}$ and $P_{sadv:3}$ for CH election. Along with TEEN for intra-cluster communication and CDMA for inter-cluster communication, ATEER successfully improved network's lifetime as compared to EEHC, DEEC, EDDEEC. However, ATEER mainly hybridised the existing protocols, yet despite unbalanced energy consumption due to TEEN implementation, it employed average energy estimate value for CH selection in every round leading to inappropriate results. Incompetence towards solving the isolated node value, ATEER may face an early death and hence, early energy hole issue causing the network to lose probable important data. Apart from improvising CH selection technique, protocol may also focus on working on data fusion at the CH. One such algorithm [34] outperformed SEP and was proposed for a two-level heterogeneous network with a threshold function. It comparatively enhanced results by additionally betting the cluster design upon residual energy of an end-device i , E_i , its initial energy, $E_{initial}$, its distance from the SN, d_i and distance of SN from network's farthest alive device, d_{max} . However, this operation requires the network to keep a track of the farthest alive node which adds to network's additional overhead hindering its efficiency. Therefore, distance computation dependency upon a dynamic network characteristic limits algorithm's application in a real-time scenario. PSEP protocol [35] infused the research domain with a new CH selecting policy to prolong the stability period of uniformly-distributed fog-supported WSNs. It incorporated two-tier heterogeneous settings and neighbour density-based data fusion that enabled prolonged stable time interval as compared to SEP. However, it either requires spending additional power in accumulating past energy usage of all network devices or requires keeping memory to save energy levels to avoid overhead. In both scenarios, end-devices face energy or memory constraints in real environments bring the network a disadvantage. This protocol tends to keep a higher probability rate for normal nodes due to their greater population compared to the other category nodes, however, this tends to unbalance energy consumption leading to an early death in normal nodes despite prolonging stability period compared to SEP. This unbalanced energy consumption may directly impact calculation of average energy that otherwise is responsible for $P_{adv:2}$ computation. Another disadvantage that affects its average energy computation is its applicability domain i.e., precisely critical IoT scenarios. EDCF technique [36] for two-level heterogeneous system aims at improving its energy structure while also enhancing its lifespan. Based on the total initial energy of the network and residual energy of individual network devices, this technique incorporates a new threshold function formula. However, depending solely on energy parameter for CH selection, overlooks the need of some important

network characteristics that require a researcher's eye. Some of them are a device's location, its device-vicinity degree, its SN vicinity etc. that greatly impact network's situation in terms of hotspot issue, overlapping clusters etc. Gradually, a shift in research domain of WSN was observed such that clustering algorithms supporting the IoT environment were birthed and a new term WSN-enabled IoT (WSN-IoT) was born. IoT was observed to base on continuous or query systems that intended on satisfying the requests from local/distant end-users. These requests mostly required WSN to sense and transmit proclaimed data from their deployed environment. Having to serve frequent requests meant constant data transfer over the channel, thereby, largely affecting network's efficiency. Therefore, researchers focussed on grouping algorithms that could improve efficiency. Both grid-based and cluster-based algorithms were proposed. A grid-based algorithm, namely energy efficient hierarchical clustering index tree (ECH-tree) technique [37] was developed to answer continuous queries from the end-users. This time-correlated region query technique segregated the IoT end-devices into even grids before assigning a clustering index tree to organize these grids cells. The assigned index allows clustering of sub grid regions ensuring reduced redundant data with upper-level sub-regions holding lesser dead space. However, apart from serving a network's efficacy, the researchers noticed that pertaining to real-time scenario, it was important for algorithms to serve for longer stability periods as loss of any single device may make the data comparatively less reliable. Dynamic partitioning into grid cells was proposed in virtual uneven grid-based routing protocol, VUGR [38] that directed towards honouring improved stable lifespan. The uneven gridding was based on the energy resources of end-devices which helped overcome hotspot issue of WSN-IoT. The protocol assigned three roles among the end-devices, namely, regular nodes, main cell-headers, assistant cell-headers apart from utilizing a mobile SN. The grids fulfilling the hard energy criteria is known to elect a main cell header while the grid that fails is broken down into smaller grids beheaded by assistance cell-headers. With a mobile sink, a hotspot issue is curbed, however, it adds to the delay while a query has to be reported back. Also, this strategy has been applied using homogeneous environment that makes it less feasible within the real-world dynamics. Utilizing the standard IoT standards, a pragmatic architectural framework was introduced in [22] where in two approaches, clustering based and gridding based were introduced. They primarily focussed on improving the network lifespan using either of the two approaches where heuristic based clustering approach utilizes neighbour count and residual energy of IP and ID enabled end-devices, while the grid-based approach utilizes location information of ID-enabled end-devices only. Although improved communication is achieved yet an algorithm is required that utilizes all the parameters together and yet showcase an improved stable period. A homogeneous scenario incorporates a clustering algorithm that savours initial and residual energy of an end device to reach optimal CHs [39]. However, a random CH election in the first round before the energy-based selection in the consecutive rounds, may cause a derange partition, thereby snubbing the movement and degree of wireless mobile nodes, WMNs. An evolutionary algorithm namely, memetic algorithm (memA) has been introduced into WSN (also inhabited with WMNs) clustering [21] to lower down the probability of early convergence. In order to achieve dynamic load balancing, memA supports in reaching an optimal CH set (if and when required) as early as possible. However, there are more network characteristics that need to be addressed to. A proficient bee-colony clustering protocol (PBC-CP)—based on artificial bee colony (ABC) algorithm [40], utilizes energy, degree, and distance from SN of WMNs for selection of head nodes. Despite appropriate parameter utilization, the non-involvement of mobility metric for CH selection, may lead nodes with varied mobility as compared to that of its neighbours to be chosen as CHs. No concurrence

in mobility with its neighbours may lead to frequent re-clustering as the untimely topology change in chosen CHs' vicinity causes unstable clustering.

2.1 Motivation

Over a long reign, clustering algorithms have been prioritized to enhance network's efficacy and stable lifespan, while also indulging individually into other design constraints due to clustering i.e., hotspot issue or cluster connectivity or energy hole or cluster overlapping, etc. However, a need for a technique that addresses multiple issues arises as sensors networks are being conglomerated with the Internet. IoT designs cater to real-world dynamics that rarely observe homogeneous environment, hence, WSNs deployed with energy heterogeneity are considered more feasible. Therefore, venturing completely into heterogeneous environment is a gradual yet essential requirement for any WSN-IoT. As energy heterogeneity is catered to in sensor networks, primarily energy-based CH selection strategies are focussed upon. However, for an overall optimized CH list, it is necessary to also cater to other characteristics of the end-devices actively such as, information about its vicinity, its location within the field etc. Literature also observed various grid-based and clustering based Voronoi schemes, however, either have their own disadvantages. Where clustering schemes may introduce cluster overlapping in case of unmanned CH selections when the node population is denser, static grid schemes may hamper network's efficiency as the node count decreases. Hence, a need also arises to design an algorithm that successfully captures the individual benefits of either of the schemes and strategically convenes a technique. Therefore, from the survey, it can be observed that there is a shift in research domain within design of clustering algorithms. In order to ply with one design constraint i.e., network connectivity apart from network's efficiency and lifetime, LUET [24] has been proposed in our previous work. In LUET, a novel concept of Lines of Uniformity *LoU* has been introduced that manages to lower down the average isolated node ratio within the network thereby improving its connectivity. However, it lacks in overpowering the other primary constraints of a WSN-IoT i.e., hotspot issue or cluster overlapping or energy hole due to early first device death—FDD. Therefore, this work further enhances the LUET protocol in terms of its stable region.

3 Preliminary Notations

3.1 Assumptions

The proposed network model for P-LUET algorithm follows a few basic assumptions as follows:

- After deployment, all end devices are static in position within the field and are ID-enabled.
- Every end device has alike processing—communication functionalities, however, they are energy-heterogeneous.
- The field is governed by a single static IoT hub in the centre of the field that acts as a sink end-device for the IoT end devices and does not face any memory, energy, or computation constraint.
- Energy depletion is the only potential cause considered for an end-device's failure.

- No obstacles from external environment hamper the operation of the network laid in the field.
- Every end-device holds an aggregation capability which allows it to compress redundant information into one useful packet while performing the CH duty.
- Received Signal Strength Indicator (*RSSI*) at an end-device allows it to estimate its relative distance from any other entity matching its radio standards.
- Every end-device is aware of its own location and the field dimensions through a positioning system or an energy-efficient positioning algorithm.

3.2 Network Topology

This section describes the three-grade energy-heterogeneous network model assumed for a WSN-IoT. The reason behind using a heterogeneous model is that deployment for real-time scenarios suffer dynamicity requiring a network to be capable enough to handle energy-heterogeneity. The proposed network is patterned as a square field ($\mathcal{M} \times \mathcal{M}$) of n_{ini} , ($n_{ini} \in \mathbb{N}$) end-devices dispersed haphazardly yet uniformly with fixed locations after deployment. The end-devices (alternatively termed as sensors in this work) are presump-tuously based on three-level energy heterogeneity and can be classified as normal *nrm*, advanced *adv*, and super-advanced *sadv*. Note that if E_o is considered the initial energy of a single normal device then the initial energy of advanced and super-advanced device is $E_o(1 + a)$ and $E_o(1 + s)$, respectively. The advanced and super-advanced end-devices are equipped with a and s times more energy than that of normal end-devices, respectively. Advanced end-devices are considered m_a fraction of the total end-devices and super-advanced end-devices are considered m_s fraction of advanced end-devices. The total initial energy of the network can be calculated as shown in Eq. (5).

$$E_T = (1 + m_a(a + sm_s)) \times n_{ini} \times E_o \quad (5)$$

The clustering hierarchy has been applied on this network model that leads the WSN-IoT to be segregated into several clusters, assuming that the IoT Hub is situated in the centre of the network. There are two tiers of communication in the clustered WSN-IoT, namely, sensors to their individual CHs, CHs to network's IoT Hub. The CHs execute aggregation to remove any existing correlated redundant data before forwarding any data to the IoT Hub. Removing redundancy reduces communication overhead and supports improving efficiency.

3.3 Optimal Clustering

The location of the IoT Hub (SN) is known to every sensor of the network. Therefore, average distance between any sensor i enacting the CH duty and the IoT Hub can be estimated using Eq. (6).

$$d'_{i:hub} = \frac{0.765 \times \mathcal{M}}{2} \quad (6)$$

In order to prevent non-essential formation of clusters within the network which may lead to network instability in terms of exponential increase in network energy consumption, clusters need to be created in optimal count. Communication channel incorporates a channel model dependent upon the commuting distance between any two entities. As a

common reference, a threshold distance d_0 is assumed. If the distance between the two entities is less than d_0 , it assumes a free-space channel model, fs , while otherwise multi-path channel, mp , model is assumed for communication. Therefore, computation of d_0 ($d_0 = \sqrt{\frac{E_{fs}}{E_{mp}}}$, that amounts to 87.7 meters in this case) is based on the transmit free-space amplifier energy $E_{fs} = 10pJ/bit/m^2$ and $E_{mp} = 0.0013pJ/bit/m^2$ for both the channel models. Optimum number of CH count k_{opt} and the optimum probability value P_{opt} can, therefore, be calculated utilizing Eqs. (7) and (8) dependent upon the distance of significant number of end-devices from their IoT Hub $d'_{i:hub}$ with respect to d_0 [26].

$$k_{opt} = \begin{cases} \sqrt{\frac{n_{ini}}{2\pi}} \times \frac{M}{d'_{i:hub}} & d'_{i:hub} < d_0 (\text{small - scale}) \\ \sqrt{\frac{n_{ini}}{2\pi}} \times \frac{M \times d_0}{(d'_{i:hub})^2} & d'_{i:hub} \geq d_0 (\text{large - scale}) \end{cases} \tag{7}$$

$$P_{opt} = \frac{k_{opt}}{n_{ini}} \tag{8}$$

4 Methodology of P-LUET algorithm

P-LUET algorithm has been proposed to address the needs of the current research scenario mentioned in the Motivation Sect. 2.1. It is named as P-LUET in order to justify the prolonged stability period. The basic procedure of the CH-selection in P-LUET is similar to the selection procedure in SEP which asks of every end-device to compute its threshold value, Th_{CH} , based on some pre-defined parameters so as to compare it with its generated random number. A higher Th_{CH} makes a significant criterion for the CH selection eligibility process.

4.1 End-Device Initialization

Before the onset of network’s round-time operation, every end-device uses a buffer time to compute its necessary parameters which is termed as initialisation phase. To compute its eligibility for the CH role, every end-device initializes pre-computation with using its location on a visualized 2-dimensional (2-D) field platform to calculate its proximity to network’s ‘lines of uniformity’ (LoU), its tier distance away from SN and the corresponding weight value, its initial vicinity density, and its region concurrence.

4.1.1 Proximity to LoU, $d_{\perp}(i)$.

Any geometric polygon has a circumradius and an inradius where circumradius is the smallest radius of the sphere into which the geometric polygon can fit in while inradius is the largest radius of the sphere which can completely fit inside the geometric polygon. Compared to the incircle with a radius r , polygon’s circumcircle with a radius R contains every point of the polygon within the circle as can be seen from Fig. 4. The term ‘lines of uniformity’ is given to the two-line segments individually connecting the opposite edges of the field (any polygon, in this case square) which are popularly also known as field diagonals. These line segments intersect

at the centre of the field and individually have a length of d ($d = 2R; d = 2\sqrt{2}r$). Considering the SN at the centre of the field with omni-directional connectivity, its spherical connectivity coverage should have a radiation reach of $d/2 = R$ ($= \sqrt{2}r$) to cover the entire network population. Therefore, the lines of uniformity sweeping through the entire polygon network are considered since they have a greater coverage spread around the SN.

The lines of uniformity for the field of dimensions, 300 m in length and 300 m in width, are $x - y = 0$ and $x + y - 300 = 0$. These lines intersect at the polygon field centre where the SN is placed. This polygon field has n_{ini} haphazardly scattered end-devices with coordinates as $(x_i, y_i) = [(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_{n_{ini}}, y_{n_{ini}})]$. Each end-device calculates its proximity to both the lines of uniformity $D_1(i), D_2(i)$ using Eqs. (9) and (10) where the primary proximity is known to be from the closer *LoU*. The primary proximity $d_{\perp}(i)$ is therefore obtained from $d_{\perp}(i) = \min \{D_1(i), D_2(i)\}$ while the secondary proximity is discarded by the end-device as stated in [24].

$$\text{One proximity of } (x_i, y_i) : D_1(i) = \frac{|x_i - y_i|}{\sqrt{2}} \tag{9}$$

$$\text{Second proximity of } (x_i, y_i) : D_2(i) = \frac{|x_i + y_i - 300|}{\sqrt{2}} \tag{10}$$

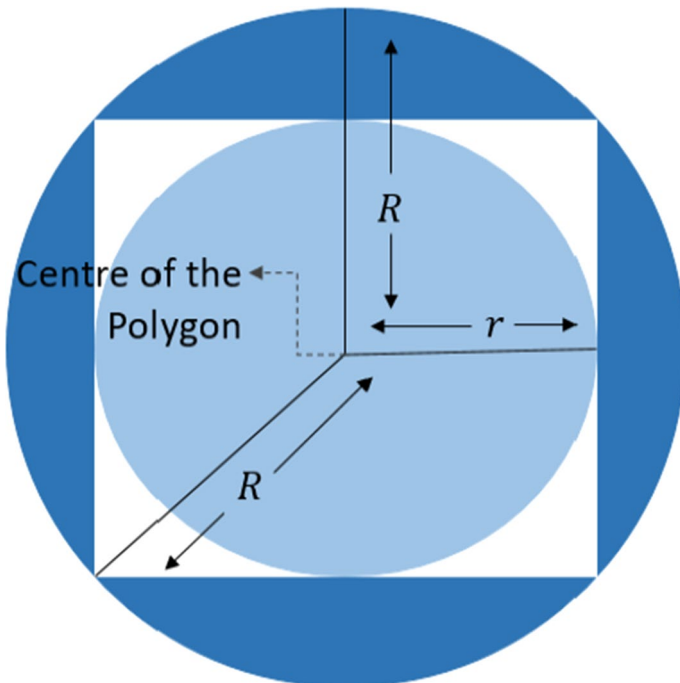


Fig. 4 Polygon with its circumcircle and incircle

4.1.2 Concentric Tiers, \mathcal{R}_b encircling IoT Hub.

P-LUET algorithm considers two concentric circles, $n = \{1, 2\}, \forall n \in n_{max}$ within the polygon field with IoT Hub as their common centre and different radii, t_n . To compute both the tier radii in the field of dimensions, 300 by 300 m (pertaining to a large-scale sensing layer), $R (= d/2)$ defining the distance of the IoT Hub from the farthest point in the field is utilized as can be seen from Fig. 5. The equation of the concentric tiers for P-LUET, can be exhibited as $(x - 150)^2 + (y - 150)^2 = t_n$. The radius of the outer tier t_2 and the inner tier t_1 can be calculated from $t_2 = R - (R/3)$ and $t_1 = t_2 - d_0$, respectively. Utilizing the tier radii distance $t_n, \forall n = \{1, 2\}$, every end-device i computes its annulus ring $\mathcal{R}_t(i), \forall t = \{1, 2, 3\}$ based on its shadow distance from the IoT Hub $d_{tier}(i)$ with respect to its primary proximity value $d_{\perp}(i)$ as can be understood from Fig. 5. Based on $\mathcal{R}_t(i), \forall t = \{1, 2, 3\}$ value, considering the global solution, a weight value ${}^iW, \forall t = \{1, 2, 3\}$ is self-assigned by the end-device or sensor i . The corresponding computations have been explained below.

Computation of sensor's (i) Annulus Ring $\mathcal{R}_t(i), \forall t = \{1, 2, 3\}$

Sensor i computes $d_{i:hub}(i) = \sqrt{(x_i - x_{hub})^2 + (y_i - y_{hub})^2}, \exists (x_i, y_i), (x_{hub}, y_{hub})$ cartesian coordinates of i , IoT Hub, respectively

Sensor i computes $d_{tier}(i) = \sqrt{d_{i:hub}(i)^2 - d_{\perp}(i)^2}$, where $d_{i:hub}(i)$ is the direct distance of i from the IoT Hub

$d_{tier}(i)$ is computed \therefore the annulus ring $\mathcal{R}_t(i), \forall t = \{1, 2, 3\}$ determination is based on the common parameter, R with respect to IoT Hub

for $(\mathcal{R}_t(i), \forall t = 1, 3)$, sensor i assumes $d_{tier}(i) = d_{i:hub}(i)$

\therefore sensor i in \mathcal{R}_1 will have a near approximate $d_{tier}(i) \cong d_{i:hub}(i)$ value and $Area(\mathcal{R}_3) < \pi(R^2 - (t_2)^2)$ which causes i in \mathcal{R}_3 to have disproportionate chances of CH selection contingency

if $(\mathcal{R}_t(i) == \mathcal{R}_1)$ then ${}^1W = 2$

if $(\mathcal{R}_t(i) == \mathcal{R}_2)$ then ${}^2W = 3$

if $(\mathcal{R}_t(i) == \mathcal{R}_3)$ then ${}^3W = 1$

${}^2W > {}^1W$ to overcome hotspot problem and ${}^2W \gg {}^3W$ to overcome high energy consumption due to direct inter-cluster communication which could result from high CH selection probability of sensors in \mathcal{R}_3

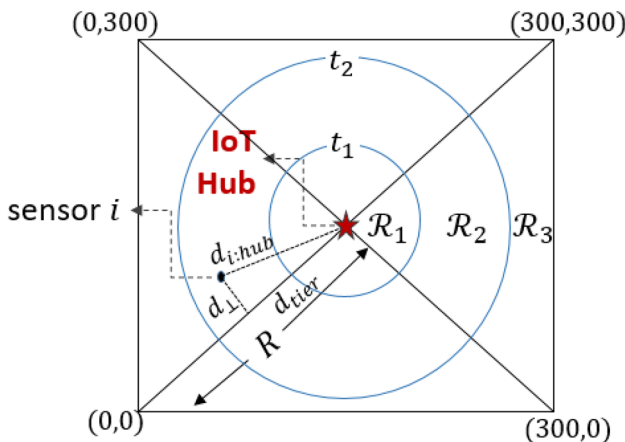


Fig. 5 Concentric tiers within the field

4.1.3 Initial Vicinity Density, ∂_i .

The initial vicinity density allows to determine an end-device i 's degree of accessibility within the network during its tenure of performing CH duty. This degree is estimated based on the effective distance of i with its neighbours within its threshold range (which could possibly be k_{nei} in count), $d_{i:j}, \forall j = \{1, 2, 3, \dots, k_{nei}\}$. As can be referred from [41], the vicinity density ∂_i for a sensor i in a network of population \aleph can be inferred from the following Eqs. (11–12). While based on d_0 value, average ∂' value can be estimated as given in Eq. (13).

$$d_{i:j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \exists (x_i, y_i), (x_j, y_j) \text{ cartesian coordinates of } i, j \text{ respectively} \tag{11}$$

$$\partial_i = 1 - \left(\frac{\frac{1}{k_{nei}} \sum_{j=1}^{k_{nei}} d_{i:j}}{r_{ini}} \right) \tag{12}$$

$$\partial' = 1 - (d_0 / r_{ini}) \tag{13}$$

4.1.4 Region Concurrence for Large-Scale WSN, \mathcal{G}

The P-LUET algorithm grids the large-scale deployment field into uniform regions of equal size. The number of virtual grids \mathcal{G} is conditionally determined by the initial count of the end-device population n_{ini} and the current alive end-devices $A_{\aleph} = [0, n_{ini}]$ as can be adhered from the Eq. (14) below. The network field is presumed to be virtually divided into k_{opt} grids as can be seen in Fig. 6 if $(0.75n_{ini} \times P_{opt} < A_{\aleph} \times P_{opt} \leq n_{ini} \times P_{opt})$ condition is fulfilled and is considered as one whole field without any grid for $(A_{\aleph} \times P_{opt} \leq 0.75n_{ini} \times P_{opt})$ condition. The CH selection does not adhere to the selection of an end-device i near the centre of these regions since the random end-device deployment may cause a shift in the centroid due to varied density composite within every region. The proposed gridding analogy for a partial lifetime allows P-LUET algorithm to overcome overlapping to a certain extent until the network has comparatively dense deployment.

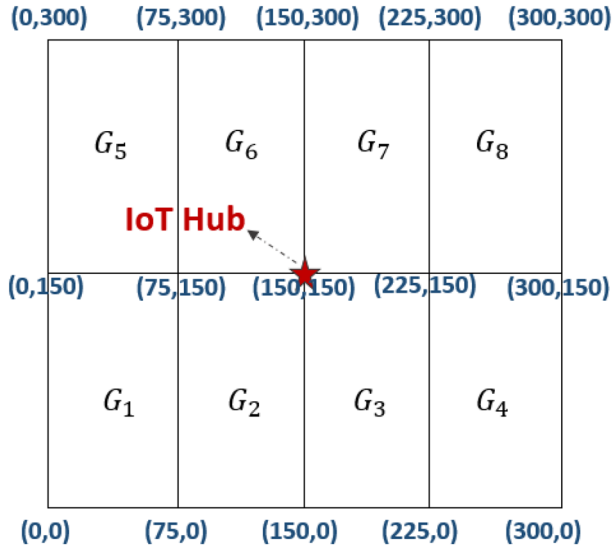
$$\mathcal{G} = \begin{cases} 1 & (A_{\aleph} \times P_{opt} \leq 0.75n_{ini} \times P_{opt}) \\ k_{opt} & (0.75n_{ini} \times P_{opt} < A_{\aleph} \times P_{opt} \leq n_{ini} \times P_{opt}) \end{cases} \tag{14}$$

After virtual grid partition, every end-device i assigns itself a grid region $G_k, \forall k = \{1, \dots, \mathcal{G}\}$ based on its cartesian coordinate based location information (x_i, y_i) .

4.2 Likelihood of CH Choice

The assumed topology adopted for configuring P-LUET algorithm follows three-level energy heterogeneity. The resultant total energy of the network amounts to $E_T = (1 + m_a(a + sm_s))n_{ini}E_o$ showcasing $m_a(a + sm_s)$ gain value in the total energy of similar network using homogeneous energy setting. Therefore, due to the additional gained

Fig. 6 Virtual grids in the large-scale network. {subjective to the scenario when $size_{WSN}$ is (300, 300), $k_{opt} = 8$, $A_{\kappa} = 85$, $n_{ini} = 100$ }



network energy, the resultant heterogeneous epoch of the network correspondingly changes to $(1 + m_a(a + sm_s)) \times (1/P_{opt})$. However, due to resource-variability amongst the end-device population in a heterogeneous sensor network, every end-device type ($nrm/adv/sadv$) has its individual sub-epoch within the heterogeneous epoch to achieve an unbiased threshold assignment. Therefore, if nrm can perform the CH duty once in the heterogeneous epoch, adv can perform the CH duty $(1 + a)$ times in the same epoch owing to the additional a times energy while $sadv$ can perform the CH duty $(1 + s)$ times in the same epoch owing to the additional s times energy. This leads the algorithm to assign an energy-based weight to their individual election probabilities (likelihood of choice in lieu of every single elected CH), giving birth to the term weighted probability P_{wgt} . In this work, the P_{wgt} term assignment for nrm is assumed as P_n , for adv is assumed as P_a , and for $sadv$ is assumed as P_s . Hence, an average count of nrm enacting as CHs per round per heterogeneous epoch is $P_n n_{ini} (1 - m_a)$, $P_a n_{ini} m_a (1 - m_s)$ is the an average count of adv enacting as CHs per round per heterogeneous epoch, and an average number of $sadv$ enacting as CHs per round per heterogeneous epoch is $P_s n_{ini} m_a m_s$ while the total average number of CHs per round per epoch stays $P_{opt} n_{ini}$. Thus, the total average CH count per round per epoch can be stated as below in Eq. (15) [42].

$$P_{opt} n_{ini} = P_n n_{ini} (1 - m_a) + P_a n_{ini} m_a (1 - m_s) + P_s n_{ini} m_a m_s \tag{15}$$

However, apart from the initial energy resources, P_{wgt} in [42] is oblivious to end-devices' other fixed/adaptive parametric measures which affect any end-device's individual capacity to be assigned the role of a CH. These measures are subjective to overcoming specified challenges in achieving the desired Quality-of-Service QoS in clustering. In this work, P-LUET assigns a rank value $rank(i)$ defined in Eq. (16) to the individual probabilities of the end-device population P_{wgt} as shown in Eq. (17). This rank value is based on a few static end-device's parametric measures like i 's primary proximity to LoU $d_1(i)$, its annulus ring association to the IoT Hub and the corresponding weight value $\mathcal{R}_r(i)_i^t W, \forall t = \{1, 2, 3\}$, its shadow distance from the IoT Hub $d_{tier}(i)$, and the degree of its

vicinity density during WSN-IoT’s stable life span ∂_i . Under dynamic environment requirements, an end-device i may have disproportionate energy consumption making it a necessary inclusion in the P_{wgt} computation. Therefore, an additional parameter defining the proportionate change in end-device’s energy resources is adopted in computing the P_{wgt} of the end-device.

Iterative CH-reaffiliation rate snubs end-devices of their battery power. With a higher rate in $adv/sadv$ of being chosen for duty as compared to nrm , P-LUET utilizes the concept of threshold energy, $E_{th} = bE_{ini} \forall b = [0, 1]$, as proposed in DDEEC [30], that lets the end-devices to switch to a common probability function $P(i)$ as can be understood from Eq. (17).

$$rank(i) = \beta \left(d_0 / d_{ier}(i) \right) + \gamma \left({}^iW / n_{max} + 1 \right) + \delta \left(\partial_i / \partial' \right), \forall (\beta + \gamma + \delta) = 1 \tag{16}$$

$$P_{wgt} \equiv \begin{cases} P_n(i) = \frac{P_{opt}}{(1+m_a(a+sm_s))} \left(E_{res} / E_{ini} \right) \times rank(i) & i \in nrm, E_{res} > E_o \\ P_a(i) = \frac{P_{opt}(1+a)}{(1+m_a(a+sm_s))} \left(E_{res} / E_{ini} \right) \times rank(i) & i \in adv, E_{res} > E_o \\ P_s(i) = \frac{P_{opt}(1+s)}{(1+m_a(a+sm_s))} \left(E_{res} / E_{ini} \right) \times rank(i) & i \in sadv, E_{res} > E_o \\ P(i) = c \frac{P_{opt}(1+s)}{(1+m_a(a+sm_s))} \left(E_{ini} / E_{avg} \right) & any\ i, E_{res} \leq E_{th}, \forall c = [0, 1] \end{cases} \tag{17}$$

4.3 Ranked Threshold computation for CH selection

The threshold value has been denoted as Th_{CH} in the current work that corresponds to $Th_{CH}(i \in nrm)$ for nrm end-devices, $Th_{CH}(i \in adv)$ for adv end-devices, $Th_{CH}(i \in sadv)$ for $sadv$ end-devices, and $Th_{CH}(i)$ for any device type (if $E_{res} \leq E_o$) to be shortlisted for the tentative CH list. Upon utilizing calculated P_{wgt} values from Eq. (17), the respective threshold value functions can be demonstrated as given in Eqs. (18–21).

$$Th_{CH}(i \in nrm) = \begin{cases} \frac{P_n(i)}{1 - (P_n(i) \times (l \times \text{mod}(1/P_n(i))))} & i \in G'_l \\ 0 & otherwise \end{cases} \tag{18}$$

The $Th_{CH}(i \in nrm)$ is applied to $n_{ini}(1 - m_a)$ normal end-devices for the current l th round. However, end-devices in nrm category have to belong to a set G'_l which have not played the CH role in the last $1/P_n(i)$ rounds of the heterogeneous epoch. Therefore, for P-LUET algorithm, the average count of $i \in nrm$ playing the CH role in a round per epoch stays $\cong P_n n_{ini}(1 - m_a)$, and $(1 + m_a(a + sm_s)) \times (1/P_{opt})$ becomes the rate of being selected for the CH role per epoch and per $[rank(i)E_{ini}] / E_{res}$.

$$Th_{CH}(i \in adv) = \begin{cases} \frac{P_a(i)}{1 - (P_a(i) \times (l \times \text{mod}(1/P_a(i))))} & i \in G''_l \\ 0 & otherwise \end{cases} \tag{19}$$

The $Th_{CH}(i \in adv)$ is applied to $n_{ini}m_a(1 - m_s)$ advanced end-devices for the current l th round. However, end-devices in adv category have to belong to a set G''_l which have

not played the CH role in the last $1/P_a(i)$ rounds of the heterogeneous epoch. Therefore, for P-LUET algorithm, the average count of $i \in adv$ playing the CH role in a round per epoch stays $\cong P_a n_{ini} m_a (1 - m_s)$, and $(1 + m_a(a + sm_s)) \times (1/(1 + a)P_{opt})$ becomes the rate of being selected for the CH role per epoch and per $[rank(i)E_{ini}]/E_{res}$.

$$Th_{CH}(i \in sadv) = \begin{cases} \frac{P_s(i)}{1 - (P_s(i) \times (l \times \text{mod}(1/P_s(i))))} & i \in G_l''' \\ 0 & otherwise \end{cases} \tag{20}$$

The $Th_{CH}(i \in sadv)$ is applied to $n_{ini} m_a m_s$ super-advanced end-devices for the current l th round. However, end-devices in $sadv$ category have to belong to a set G_l''' which have not played the CH role in the last $1/P_s(i)$ rounds of the heterogeneous epoch. Therefore, for P-LUET algorithm, the average count of $i \in sadv$ playing the CH role in a round per epoch stays $\cong P_s n_{ini} m_a m_s$, and $(1 + m_a(a + sm_s)) \times (1/(1 + s)P_{opt})$ becomes the rate of being selected for the CH role per epoch and per $[rank(i)E_{ini}]/E_{res}$.

$$Th_{CH}(i) = \begin{cases} \frac{P(i)}{1 - (P(i) \times (l \times \text{mod}(1/P(i))))} & i \in G_l \\ 0 & otherwise \end{cases} \tag{21}$$

The $Th_{CH}(i)$ is applied to $nrm/adv/sadv$ for the current l th round if they satisfy $E_{res} \leq E_{th}$ condition. However, end-devices have to belong to a set G_l which have not played the CH role in the last $1/P(i)$ rounds of the heterogeneous epoch. Therefore, for P-LUET algorithm, $(1 + m_a(a + sm_s)) \times (1/(1 + s)P_{opt})$ becomes the common rate for all end devices to be selected for the CH role per epoch and per E_{avg}/E_{ini} .

Henceforth, upon computed Th_{CH} evaluation with respect to a randomly generated number $rand()$ by the eligible end-device population in WSN-IoT, $P_{opt} n_{ini}$ (as derived from Eq. (15)) end-devices are chosen randomly from the tentative list $list_l^{CH}$ as the final CHs $list_l^{CH}$ of the network for the l th round. However, the choice of final CHs is based upon the validation of the condition for the region concurrence. If the network fulfils the condition $0.75 n_{ini} \times P_{opt} < A_N \times P_{opt} \leq n_{ini} \times P_{opt}$, the end-devices assume $g = k_{opt}$ virtual grids in the network. Based on the location information, end-devices from $list_l^{CH}$ move to $list_l^{CH}$ only if its corresponding region $G_k, \forall k = \{1, \dots, g\}$ has yet to choose a CH. Every grid region will primarily be headed by one CH. After CHs have been chosen per grid, the rest of the end-devices $i \notin list_l^{CH}$ associate themselves with the CH of their corresponding region G_k . This greatly reduces the overhead exchange during the association confirmation process in case of multiple advertisement messages received by the non-CH end-devices. Since, the regions are clearly demarcated, the cluster overlapping is also reduced, thereby improving the cluster coverage, and further enhancing the network efficiency by $\approx 6.7\%$ as can be seen from Fig. 7. The energy efficiency of P-LUET, ϵ_{p-luet} is proportional to the amount of network data T^{net} propagated by consuming a certain energy E_{used} to withstand surging data flow from the IoT Hub as given in (22) and as stated in [43].

$$\epsilon_{p-luet} = T^{net} / E_{used} \tag{22}$$

As soon as the new condition $A_N \times P_{opt} \leq 0.75 n_{ini} \times P_{opt}$ turns true, the virtual griding concept is discarded by the end-device population and thereafter, random k_{opt}

end-devices are chosen as CHs from $list_l^{iCH}$. Thereby, end-devices $i \notin list_l^{iCH}$ either associate to their closest CH and be known as member i^{CM} to that CH or be known as isolated end-devices i^{IN} that communicate directly with the IoT Hub.

4.4 \mathcal{G} -based Conditional Multi-Hop Inter-Cluster Communication

In this subsection, the inter-cluster communication model has been discussed. The inter-cluster communication model is based on the condition validation for \mathcal{G} . If the condition $(0.75r_{ini} \times P_{opt} < A_N \times P_{opt} \leq r_{ini} \times P_{opt})$ holds true for the network, the end-devices $i^{CH}, \forall i \in list_l^{iCH}$ compare their distance from their IoT Hub with their distance from CH $j^{CH}, \forall j \in list_l^{iCH}, G_k(i) \neq G_k(j), i \neq j$ in its adjoining grid region as shown in the following Eqs. (23–24).

$$d_{i;j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \exists (x_i, y_i), (x_j, y_j) \text{ cartesian coordinates of } i^{CH}, j^{CH} \text{ resp.} \tag{23}$$

$$d^{inter} = \min \{d_{i;j}, d_{i;hub}(i)\}, \forall i^{CH}, j^{CH} \in list_l^{iCH} \tag{24}$$

The proposed model focusses on minimizing the communication cost during the data uploading from the CHs overlooking farther grid regions since energy usage $E_R(size)/E_T(size, d)$ varies either proportionally to the length of the information to be imparted $size$ or also the commuting distance d as can be understood from Eqs. (25–26). In an attempt to save energy consumption, network savours prolonged aliveness NL_{p-luet} with end-devices showcasing longer lingering vitality. As can be observed from Fig. 8, a P-LUET incorporated WSN-IoT spans longest period if conditional gridding is employed with multi-hop communication while showcases the lowest span count on employing grids without multi-hop communication. In the absence of conditional grid structure, CH list is not chosen to bound to the grid coordinates, thereby, bearing higher probability of being

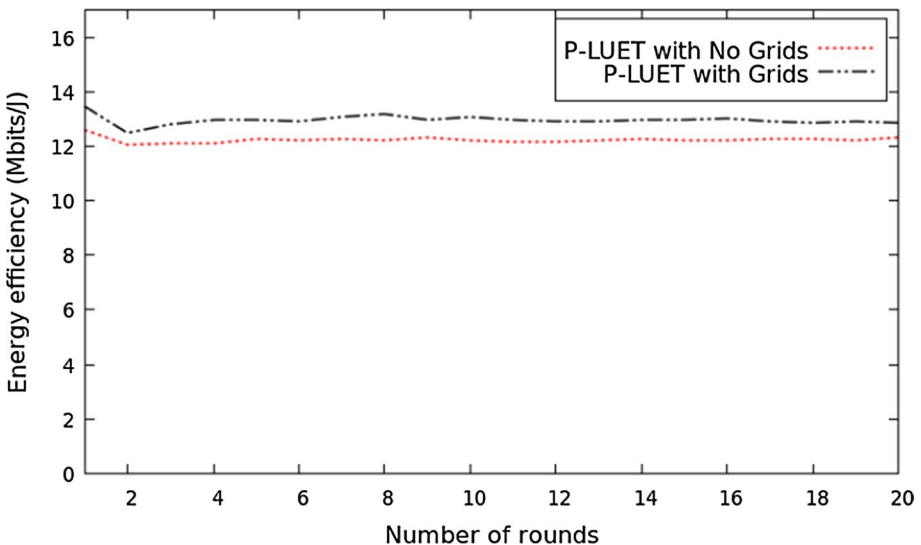


Fig. 7 ϵ_{p-luet} implying the effect of \mathcal{G}

chosen in context to the IoT Hub location. Therefore, after selection of a desired CH count that successfully qualify the algorithm specifications, as the data gets uploaded to these CHs from their respective members, the minimum distance path to the IoT Hub is computed that defines whether a multi-hop model or a single-hop model has to be presumed. However, as the previously mentioned condition turns false over the net-life's course, with a new condition turning true i.e., $A_{\text{N}} \times P_{\text{opt}} \leq 0.75r_{\text{ini}} \times P_{\text{opt}}$, the end-devices that are supposedly chosen for performing CH duty henceforth are selected randomly throughout the network. The CH count is based on the optimal count of clusters k_{opt} required for an efficient network operation. Under this condition, every end-device enacting as a CH, $i^{\text{CH}}, \forall i \in \text{list}_l^{\text{CH}}$, transmits its aggregated data directly to the IoT Hub. The energy expended in the electronic circuitry of an end-device that supports transmission or reception of information can be known as E_{elec} .

$$E_T(\text{size}, d) = \begin{cases} \text{size} * E_{\text{elec}} + \text{size} * E_{\text{fs}} * d^2, & \text{if } d < d_0 \\ \text{size} * E_{\text{elec}} + \text{size} * E_{\text{mp}} * d^4, & \text{if } d \geq d_0 \end{cases} \tag{25}$$

$$E_R(\text{size}) = E_{\text{elec}} * \text{size} \tag{26}$$

4.5 Some Explicit Remarks About P-LUET Algorithm

The quoted algorithm points at the significant steps of P-LUET. Specifically, the proposed algorithm focusses on defining the set $\text{list}_l^{\text{CH}}$ for every l -th round. *Lines*2 – 3 attempt to establish end-device deployment and thereafter, network configuration. *Line*4 starts a loop for initialization on every end-device which means every end-device i individually computes $d_{\perp}(i), \mathcal{R}_t(i)_i^t W, \forall t = \{1, 2, 3\}, d_{\text{tier}}(i),$ and ∂_i . Note that distance computations are based on the Euclidean distance computed between any two desired entities of the network.

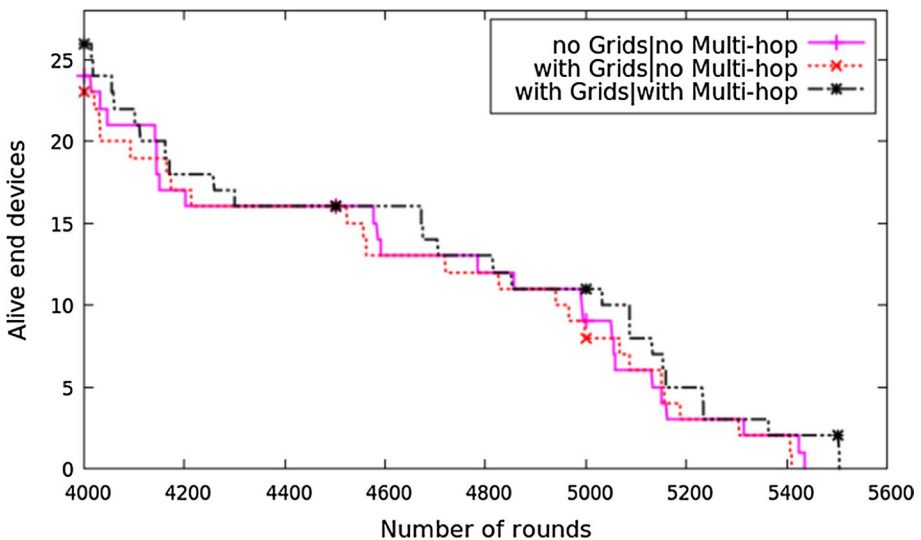


Fig. 8 NL_{p-luet} implying the effect of \mathcal{G} and multi-hop communication model

Using these static location-based parameters and an adaptive resource-based parameter E_{res} , every end-device i calculates P_{wgt} -based Th_{CH} value via *line8* for every l -th round. *Lines10 – 19* and *lines22 – 32* show routing plan from the end-devices $i \notin list_l^{CH}$ to the root centre of WSN-IoT (i.e., IoT Hub) for the g -based conditions mentioned in Eq. (14). *Line35* defines the end condition for the algorithm operational process.

4.6 MAC Scheme Adopted Within a P-LUET Cluster

P-LUET opts for Time Division Multiple Access (TDMA)/ Code Division Multiple Access (CDMA) within a cluster/ network to abridge congestion within the network during intra-cluster/ inter-cluster data transfer. For the sake, the network operation is divided into two time-stages, Setup Stage and Steady State Stage. Where Setup Stage broadly involves CH selection process within the network, Steady State Stage involves transmitting data from the source end-devices to the IoT Hub of the network either directly or via the chosen CHs.

As the network initialises, it enters the setup stage where upon obtaining $Th_{CH} > rand()$, and being chosen as a CH i^{CH} , the end-device i broadcasts its address via an advertisement packet *ADV* in its vicinity range. Every $j \notin list_l^{CH}$ may receive one or more broadcasted *ADV* from its neighbourhood. It associates itself to the nearest i^{CH} (that source of *ADV* which is the highest on the *RSSI* comparison list in j) or the only i^{CH} through an acknowledgment packet *ACK* containing its own address. After all the i^{CH} setup their clusters, the steady state stage initiates in which end-devices $j \notin list_l^{CH}$ sense and upload data to their respective $i^{CH} \in list_l^{CH}$ by opting TDMA which after aggregating the received data, opt for CDMA to transmit this aggregated data to the IoT Hub. During the setup phase, the network may contain a few $j \notin list_l^{CH}$ that do not receive *ADV* from any chosen $i^{CH} \in list_l^{CH}$ and end up getting known as isolated or disconnected end-devices. These end-devices transmit their data directly to the IoT Hub during the following steady state stage. After a time that is planned prior to the network setup, the network enters another round of CH selection i.e., the Setup Stage and its corresponding steady state stage. This repeats until the last alive end-device is drained completely of its energy resources. The opting of TDMA/CDMA allows congestion-free hierarchy via clustering on several levels within the network, leading to improved energy structure for a heterogeneous environment.

Prolong – Lines of Uniformity based Enhanced Threshold (P-LUET) Algorithm

```

1. Begin
2. Initiate Network parameters
3. Deploy end-devices in a Test Area
4. Repeat step 5 for  $i = 1$  to  $n_{ini}$ 
5.      $i$  computes  $d_{\perp}(i)$ ,  $\mathcal{R}_t(i)$ ,  ${}^iW$ ,  $\forall t = \{1,2,3\}$ ,  $d_{tier}(i)$ ,  $\partial_i$ , and  $G_k$ ,  $\forall k = \{1, \dots, g\}$ 
6. [End of step 4 loop]
7. Repeat steps 8 to 33 for  $i = 1$  to  $n_{ini}$ 
8.      $i$  computes  $P_{wgt}$  and its corresponding  $rand()$  &  $Th_{CH}$  values
9.     If  $(0.75n_{ini} \times P_{opt} < A_N \times P_{opt} \leq n_{ini} \times P_{opt})$  then
10.         $i$  checks its region concurrence and conditionally shortlists itself for  $list_i^{CH}$ 
11.        If  $(i \notin list_i^{CH})$  then
12.             $i$  associates itself to CH as  $i^{CM}$  within its corresponding  $G_k$ ,  $\forall k \in k_{opt}$ 
13.             $i$  uploads its sensed data to its corresponding CH during its assigned TDMA slot
14.        [End of If structure]
15.         $i$  receives uploaded data from its CMs
16.        If  $(d_{i,hub}(i) > d_{i,j})$  then  $\% i \in i^{CH}, j \in j^{CH}$ 
17.             $i$  forwards its aggregated data to IoT Hub via the CH,  $j^{CH}$  of adjoining  $G_k$ 
18.        Else
19.             $i$  forwards its own aggregated data and data from the CH of adjoining  $G_k$  to IoT Hub
20.        [End of If structure]
21.    Else
22.         $i$  conditionally shortlists itself for  $list_i^{CH}$ 
23.        If  $(i \notin list_i^{CH})$  then
24.             $i$  associates itself to the nearest CH as  $i^{CM}$ 
25.             $i$  uploads its sensed data to its corresponding CH during its assigned TDMA slot
26.        [End of If structure]
27.        If  $(i \notin list_i^{CH} \ \&\& \ i \neq i^{CM})$  then
28.             $i$  demarcates itself as  $i^{IN}$ 
29.             $i$  uploads its sensed data directly at IoT Hub
30.        [End of If structure]
31.         $i$  receives uploaded data from its CMs
32.         $i$  forwards its aggregated data directly to IoT Hub
33.    [End of If structure]
34. [End of step 7 loop]
35. Repeat steps 7 to 34 until  $A_N = 0$ 
36. Exit

```

4.7 Time Complexity Analysis of P-LUET

Time complexity (or computation complexity) can be described as the amount of work microcontroller μc has to perform as the input data increases (approaching to ∞) and is computed based on the Big O notation that describes the worst-case diversified computational scenario of an algorithm.

In P-LUET, it can be observed that the convergence rate to reach an optimum solution is a constant value for every end-device for every iteration (in this case it is known as a round of the lifetime). Since P-LUET follows a distributed clustering strategy, every end-device in the network is known to execute the algorithm independently to determine

its decision towards CH selection. These decisions are independent of status of the end-device's neighbours or the end-device population within the network and only depend on an end-device's current energy status apart from its static computed metrics. Therefore, the implementation complexity for CH selection can be given as $O(1)$ (Fig. 9).

However, for the initial phase of the operation when the condition $0.75n_{ini} \times P_{opt} < A_{\text{th}} \times P_{opt} \leq n_{ini} \times P_{opt}$ holds true, the network opts for static graph theoretical communication model between CHs and the IoT Hub while throughout the operation source devices follow multi-hop communication model. The diversity $g(V)$ for computational analysis of the conditionally adopted grid network has been taken as $\log(2^{n_{ini}}) = n_{ini}$ wherein network either follows single-hop communication between IoT Hub and CHs of its vicinity grids or a two-tier multi-hop communication from far away CHs and IoT Hub. Therefore, this selective multi-hop communication among CHs introduces a stronger efficient connectivity at the cost of additional $O(g(V)) = O(n_{ini})$ factor in the algorithm's implementation complexity (Fig. 10) and fulfils the property of each CH communicating successfully at least once i.e., Ψ_{min} .

5 Simulation of Results and Analysis

For the purpose of P-LUET algorithm analysis, NS2.35 has been used as the simulator that runs on a system with Windows 10 operating system. The system is i7-8750H CPU with 16 GB RAM and the processor operates at 2.20 GHz. The basic simulation environment consists of a network of haphazardly deployed 100 end-devices in a field of dimensions $300 \times 300m^2$. Each end-device has an optimal probability P_{opt} of 0.08 as computed from Eq. (7) and a maximum capacity of carrying 4000bits of data. The energy heterogeneity metrics of the basic simulation environment have been listed in Table 1.

Fig. 9 Computational Complexity of intra-cluster communication in WSN-IoT

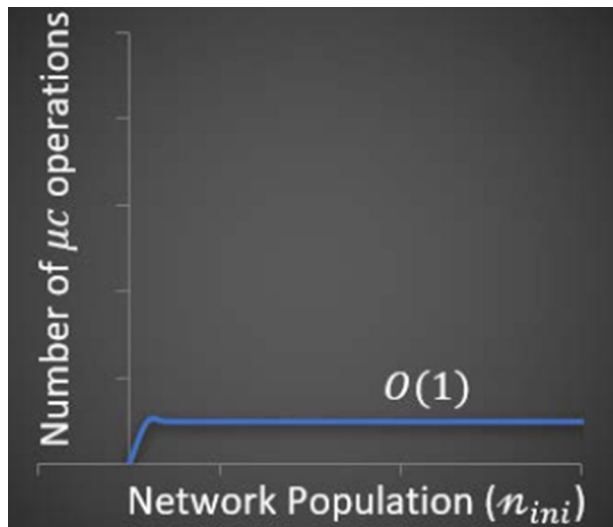


Fig. 10 Computational Complexity of inter-cluster communication in WSN-IoT

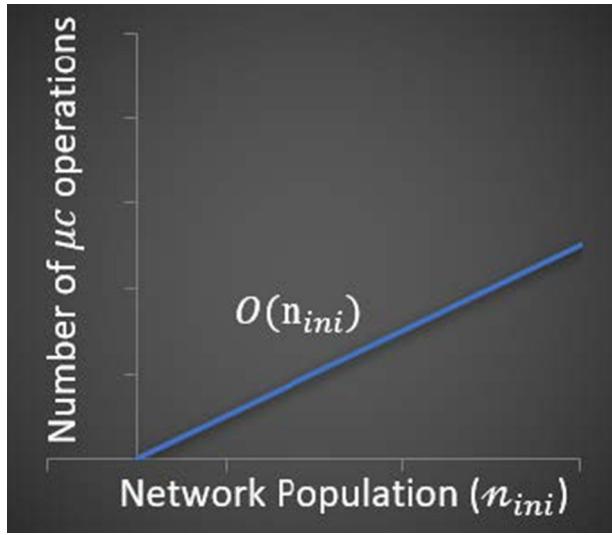


Table 1 Energy heterogeneity metrics

Heterogeneity metrics	Values
Proportion of <i>sadv</i> end-devices, m_s	0.44
Energy factor for <i>sadv</i> end-devices, s	3
Proportion of <i>adv</i> end-devices, m_a	0.41
Energy factor for <i>adv</i> end-devices, a	2
Initial energy of <i>nrm</i> end-devices (<i>joules</i>), E_o	0.5
Total energy, E_T	100

5.1 Performance Metrics

5.1.1 Stability Period, τ_{fdd}

It can be defined as the vitality period until the death of first end-device. It can alternatively be termed as reliable period despite high correlation, since field contains no energy hole which means every end-device actively participates in reading the environment. There are many critical smart applications like IoT-based health monitoring of a critical patient, or battlefield surveillance etc. that require to keep a regular check on every characteristic of the environment. Such applications may deploy end-devices with a sole purpose of updating about their respective functionality. Hence, death of any single end-device in these applications can turn detrimental owing to loss of a functionality. For example, in battlefield surveillance, loss of any end-device may result in loss of coverage of a crucial surveillance area, which may lead to minacious consequences. Hence, this metric plays a dominant role in guaranteeing acceptable performance of any protocol.

5.1.2 Unreliable Period, $\sigma_{W/I}$.

Network lifetime $NL_{W/I}$ can be defined as the period until the network's last alive device survives. Although, the last node is alive and transmitting data, the received data cannot be considered reliable as it fails to represent enough for the network. Hence, as soon as the first device dies, the network loses information about the area which was initially sensed by the dead device. Therefore, after first device death (FDD), the rest of the population may keep transmitting data, however, the data will be termed as unreliable. Hence, the corresponding unreliable period can be calculated as: $\sigma_{W/I} = NL_{W/I} - \tau_{fdd}$.

5.1.3 Energy Efficiency, $\epsilon_{W/I}$.

As stated, before in Sect. 4.3, energy efficiency, $\epsilon_{W/I}$ is proportional to the network throughput T^{net} while inversely proportional to the energy consumption E_{used} and can be computed as $\epsilon_{W/I} = T^{net}/E_{used}$ [43]. A higher $\epsilon_{W/I}$ defines algorithm's improved capability of withstanding surging data flow from the IoT Hub while balancing devices' energy-constrained resources.

5.1.4 Network Throughput, T^{net} .

Despite varying known in the literature, its inherent definition stays the same i.e., the total number of packets that can be transmitted successfully to the IoT Hub. Therefore, however differently this term may be defined in various literatures, yet the above stated remains the widely accepted definition amongst all the researchers.

5.1.5 Probability of Device Isolation, $\mathbb{P}(iso)$.

Imbalanced CH-selection could isolate certain end-devices from the chosen head list $list_l^{CH}$ in the l -th round and resultantly the count of isolated nodes per round is taken as α_{iso} . As a result, these devices resort to direct-transmission with the IoT Hub irrespective of its location. However, a direct communication with the IoT Hub leads to a computational complexity of $O(r_{ini}^2)$ that leads to exuberant energy-usage as depleted-energy increases proportionally-bifold with distance from the IoT Hub. Therefore, probability of end-device isolation $\mathbb{P}(iso)$ in every l -th round also comes across as a performance metric that can be computed as $\mathbb{P}(iso)_l = \frac{\# \alpha_{iso}}{\# A_N}$ [44].

5.2 Numerical Results

In the following, the performance of the P-LUET algorithm has been compared with the aforementioned state-of-the-art algorithms while is also tested under varied network settings.

5.2.1 Comparative Analysis with Respect to State-of-that-Art Clustering Algorithms

The proposed algorithm is run through basic simulation environment as mentioned earlier with the aforementioned algorithms ([24, 33–36]) and is evaluated based on longevity metrics ($NL_{W/I}$; τ_{fdd} ; $\sigma_{W/I}$), $\epsilon_{W/I}$, T^{net} , and $\% \mathbb{P}(iso)$ w.r.t. these algorithms. Figure 11 exhibits

Table 2 NL_{p-luet} and τ_{fd} $\tilde{\alpha}$ -values of P-LUET algorithm w.r.t. state-of-the-art clustering algorithms

P-LUET	ATEER [33]	Improved clustering algorithm [34]	P-SEP [35]	EDCF [36]	LUET [24]
$\tilde{\alpha}(NL_{p-luet})$	0.011↓	0.001↑	0.075↑	0.139↑	0.115↑
$\tilde{\alpha}(\tau_{fd})$	0.142↑	0.129↑	0.192↑	0.011↑	0.187↑

Table 3 Longevity-efficiency trade-off analysis and observations thereafter

Algorithms	% τ_{fd}	% $\sigma_{w/I}$	Average $\epsilon_{w/I}$ over $NL_{w/I}(1_\epsilon)$	Average $\epsilon_{w/I}$ over $\tau_{fd}(2_\epsilon)$	Average $\epsilon_{w/I}$ over $\sigma_{w/I}(3_\epsilon)$
ATEER[33]	19.21	80.79	11.81	12.34	11.68
Improved Clustering Algorithm [34]	19.72	80.28	11.89	12.5	11.74
P-SEP [35]	19.73	80.27	11.84	11.61	11.89
EDCF [36]	26.00	74.00	9.35	9.26	9.38
LUET [24]	20.78	79.22	11.13	11.81	10.95
P-LUET	22.63	77.37	12.07	12.75	11.87

Observations

- [33, 36] exhibit highest and lowest % τ_{fd} , resp. among simulated algorithms essaying [36] as most and [33] as least suitable for WSN-IoT (on longevity grounds only)
- The second most suitable for WSN-IoT is P-LUET on longevity grounds only
- With higher % $\sigma_{w/I}$ of $\in (79, 81)$, [24, 33–35] can be considered not suitable for being deployed in WSN-IoT as IoT framework requires network to serve high reliability more and overcome less reliable spans
- P-LUET in a WSN-IoT environment, shows an improvement of 2.2%, 1.5%, 1.9%, 22.5%, 7.8% in 1_ϵ ; 3.2%, 2%, 8.9%, 27.4%, 7.4% in 2_ϵ , 1.6%, 1.1%, $\approx 0\%$, 20.9%, 7.8% in 3_ϵ w.r.t. [24, 33–36]
- The characteristic of a typical WSN-IoT is to provide higher $\epsilon_{w/I}$ in τ_{fd} than in $\sigma_{w/I}$ oblivious of the % τ_{fd} within a field. Therefore, despite [35, 36] attach 2.4% and 1.3% more average $\epsilon_{w/I}$ to $\sigma_{w/I}$ than τ_{fd} , with P-LUET attaching 33.13% less average $\epsilon_{w/I}$ to $\sigma_{w/I}$ than τ_{fd} it is considered more efficiently suitable for WSN-IoT

the overall life pace of the algorithms involved from where it can be observed that EDCF [36] showcases a higher exponential fall rate in $A_{\mathbb{N}}$ as compared to the other algorithms, P-SEP [35] follows a gradual exponential fall rate during its major portion of the $\sigma_{w/I}$. Although, [35] may showcase an improved pace over certain period, it fails to maintain this pace until the end delivering a shorter $NL_{w/I}$. Other algorithms ATEER [33], Improved clustering algorithm [34] and LUET [24] deliver similar life pace characteristics as that of P-LUET algorithm with a constant pace throughout the journey delivering no sudden fall rate $(1 - 0.01A_{\mathbb{N}})$. However, in an IoT environment, where queries suggest data proclaimed by every end-device of the network for reliable despite redundant data. In such cases it is important to determine the stability period of the network incorporated with the designed algorithm. Under such scenario, through Fig. 12 it can be observed that P-LUET demonstrates longer τ_{fd} while P-SEP delivers shortest τ_{fd} . Therefore, despite improved life pace over $\sigma_{w/I}$, P-SEP delivers the shortest τ_{fd} compared to the aforementioned algorithms. With a strong direction towards enhancing τ_{fd} of LUET, it can be clearly observed from

Table 4 that P-LUET shows an improvement factor $\tilde{\alpha}$ of 0.187. This can be understood as a resultant of upholding CH election operation that not only works on one design constraint, however, additionally supports overcoming hotspot issue and cluster overlapping apart from poor network connectivity.

It can also be observed from Table 2 that P-LUET successfully upgrades τ_{fd} -based $\tilde{\alpha}$ value by $\in [0.011, 0.192]$ defining longer τ_{fd} making it more suitable within a query-based or continuous-monitoring-based IoT environment than the other established algorithms. The main disadvantages towards lower τ_{fd} observed in established algorithms apart from [24] can be briefed as non-consideration of device location w.r.t. the IoT Hub of the network in [33, 36], large overhead due to dependency on a dynamic location entity in [34], higher probability rate of normal devices due to comparatively large population in [35], no direction towards towing the CH selection to devices that help improve the network connectivity without enhancing its overhead in [33–36]. It can also be observed that P-LUET incorporated WSN-IoT delivers greater $NL_{W/I}$ by $\tilde{\alpha}$ factor ranging from [0.001 – 0.139] which has been clearly depicted graphically in Fig. 13. Although [33] may deliver longer $NL_{W/I}$ by $\tilde{\alpha}$ factor of 0.011, yet due to smaller τ_{fd} , it faces longer $\sigma_{W/I}$ compared to the proposed algorithm as can be evaluated from Table 3.

Observations defined in Table 3, clearly mark P-LUET as the most efficiently suitable algorithm to be considered for deployment in a WSN-IoT framework. Observations in Table 5 calibrated as per longevity parameters and throughput along with energy helps realise the energy utilization in either of the algorithms. Drawing the inferences from Table 4, the synthesized data has been utilized to reach a consensus stating that P-LUET can effectively serve an IoT-framework.

Thereafter, Tables 6 and 7 have been described that individually showcase P-LUET's energy credibility and channel utilization capability over randomly chosen round counts w.r.t. aforementioned state-of-the-art clustering algorithms. From Table 6, it can be observed that [24, 35, 36] loose energy earliest than [33, 34] and P-LUET. It can also be observed that despite holding more residual energy at the defined round (1100, 1800) than P-LUET, [35] showcases a poor channel utilization capacity during those round counts as can be confirmed from Table 7. P-LUET displays a better or similar channel utilization capacity compared to [24, 33–36] for round counts 500, 1100, 1800 and a better channel capacity compared to [24, 33, 34, 36] for round counts 2600, 3400, 4000, 4600, 5100. [35]'s channel capacity improves over latter round counts compared to P-LUET, however, over unreliable session of WSN-IoT lifetime which holds minimal importance in IoT architecture.

As an extension of LUET, P-LUET also helps in improving the network connectivity by working on keeping the $\mathbb{P}(iso)$ lower as was obtained in LUET. To further enhance this characteristic of the network, P-LUET introduces conditional gridding that allows devices to be connected to one device acting as a CH within its local grid until the condition $(0.75n_{ini} \times P_{opt} < A_N \times P_{opt} \leq n_{ini} \times P_{opt})$ is fulfilled. As can be seen from Fig. 14, unlike [35, 36] that has a few cases of higher $\% \mathbb{P}(iso)$ until round count 1700, P-LUET's $\% \mathbb{P}(iso)$ spread is on a lower front. Moving ahead in lifetime, P-LUET's $\% \mathbb{P}(iso)$ spread can be seen rising than its spread in the previous round range yet it falls on the lower front compared to the spread for [33, 35, 36] despite a fall in A_N .

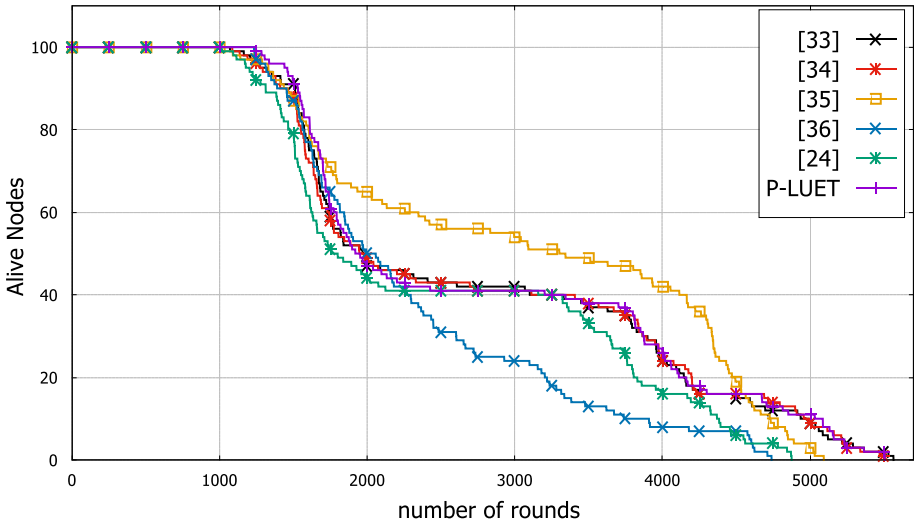


Fig. 11 Comparative life pace; ($A_N \in [0, 100]$) vs $\mathcal{M}^2 = (300)^2$

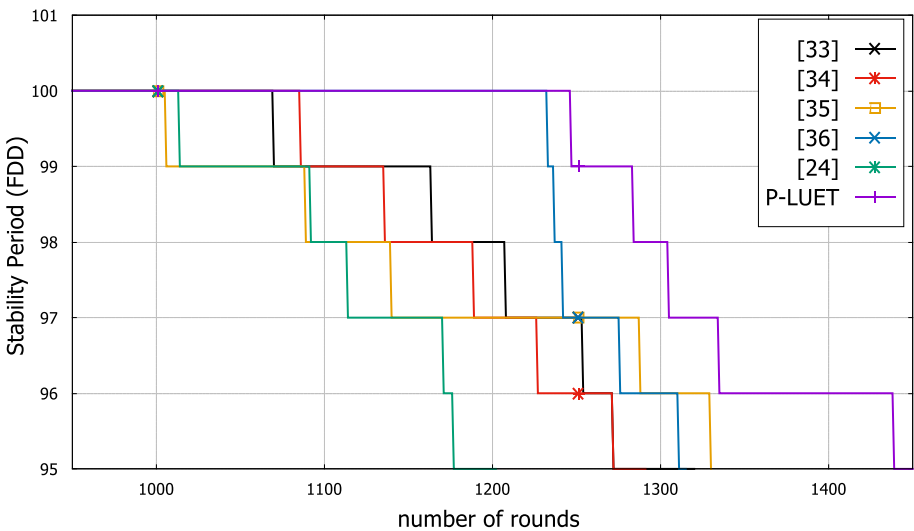


Fig. 12 τ_{fdd} - analysis; ($A_N \in [95, 100]$) vs $\mathcal{M}^2 = (300)^2$

5.2.2 Assessment based on varied assumptions on P-LUET

In this section the performance of P-LUET is tested under varying assumptions to further understand the liability of employing P-LUET. As understood from the literature, a heterogeneous network is supposedly assumed to lay devices with an additional Δ times energy to enhance network’s capability. However, varying the individual device capabilities without affecting the overall network resources can also be known as heterogeneity sensitivity.

Table 4 %Improvement in T^{net} of P-LUET w.r.t. state-of-the-art clustering algorithms

	[33]		[34]		[35]		[36]		[24]	
	l'	% T^{net}	l'	% T^{net}	l'	% T^{net}	l'	% T^{net}	l'	% T^{net}
FDD	1069	14.21↑	1085	12.92↑	1005	19.34↑	1232	1.12↑	1013	18.7↑
HDD	1988	0.33↑	1978	1.25↑	3343	44.37↓	2058	2.17↓	1800	8.92↑
LDD	5565	0.79↑	5503	0.81↑	5093	11.42↓	4739	16.4↑	4875	9.32↑

l' Round life completed

HDD Half Device Dead

Therefore, in the first assumption, while keeping the total energy of the network same $E_T = 100J$, varying scenarios have been considered on how this energy can be individually varied among devices. Under such presumption, 4 cases can be obtained. The first case is keeping the initial device count of the network $n_{ini} = 100$ and increasing the energy of every device E_{imi} from $0.5 \rightarrow 1.0J$; the second case is keeping $E_{imi} = 0.5J$ as is the energy of normal nodes in the basic pre-analysed environment and adding more devices of same E_o to the network; the third case is the two-level energy heterogeneity with two kinds of devices: normal and advanced devices where advanced devices hold $a' = 2$ times more E_o and are $m'_a = 0.5$ fraction of total nodes; the fourth case is the basic three-level heterogeneous scenario analysed in the previous section.

The observations from the Table 8 direct the analysis towards accepting Case 4 for the needful environment. Additionally, Fig. 15 graphically demonstrates the improved efficiency compared to the other cases while also showcases the FDD round graph for all the cases. Pertaining to the high E_{imi} value in Case 1, it demonstrates a much longer τ_{fdd} . However, considering it unfair for initial performance evaluation and bypassing it for valuation, Case 4 wins the race among the rest of the cases as clearly, it has the longest τ_{fdd} .

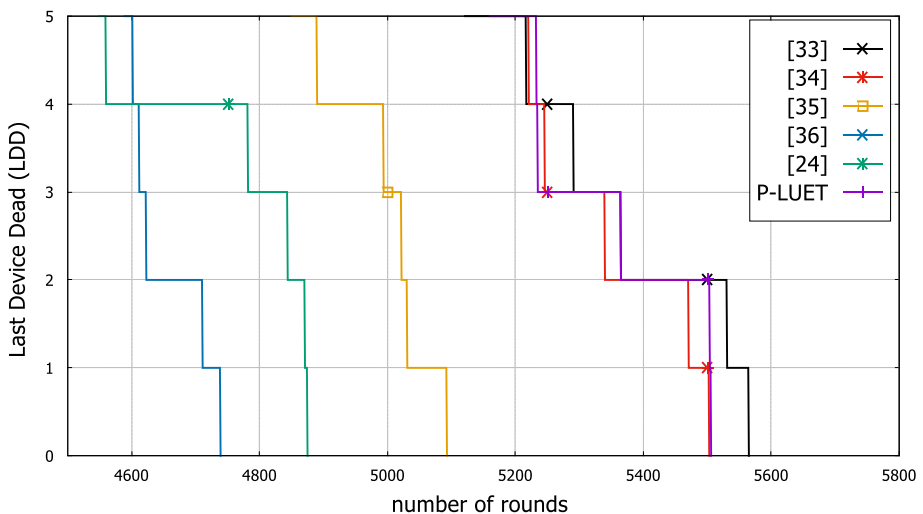


Fig. 13 $NL_{W/I}$ - analysis; ($A_R \in [0, 5]$ vs $l, M^2 = (300)^2$)

Table 5 Energy usage of P-LUET and the other established algorithms and Observations thereafter utilizing inferred results from Table 4

	[33]		[34]		[35]		[36]		[24]		P-LUET	
	l'	E_{res}	l'	E_{res}	l'	E_{res}	l'	E_{res}	l'	E_{res}	l'	E_{res}
Initial	1	99.967	1	99.969	1	99.966	1	99.953	1	99.969	1	99.970
FDD	1069	65.328	1085	65.250	1005	67.557	1232	46.824	1013	65.689	1246	60.951
HDD	1988	41.072	1978	42.177	3343	12.493	2058	21.369	1800	43.430	1952	42.829
LDD	5565	0	5503	0	5093	0	4739	0	4875	0	5506	0

Observations

1. To carry out fair performance evaluations, the initial energy environment is kept same for all the algorithms
2. While [24, 33–35] suffer a balanced fall rate ranging closely between (32,35); P-LUET suffers a higher fall rate of 39.05% and [36] suffers highest fall rate of 53.18%. A higher fall rate can be owed to higher FDD and large throughput it supports in [36] and P-LUET as can be confirmed from Table 4
3. It can also be observed through simulations that until FDD, P-LUET showcases best energy utilization while [36] shows worst energy utilization. The estimate E_{used} per l in [24, 33–36] and P-LUET have found to be 0.0324, 0.032, 0.0323, 0.0432, 0.034, 0.031, respectively. Therefore, despite higher fall rate in P-LUET, it showcases a better energy utilization until FDD
4. After reaching the state $A_N = 0.5z_{ini}$, [24, 33, 34] save about $E_{res} \in (41, 43.5)$ (in Joules) which is similar to the total E_{res} of P-LUET incorporated WSN-IoT. However, these algorithms savour high residual at the cost of lower network throughput as opposed to P-LUET algorithm that supports an approx. improvement of {0.33, 1.25, 8.92} throughput compared to [24, 33, 34]
5. [35, 36] suffer the most drop in E_{ini} at HDD, since they manage to support a large data transport of about 259,542 and 183,681 packets as compared to 179,774 packets supported by P-LUET until HDD
6. Apart from transferring a higher count of packets, P-SEP is observed to showcase best energy utilization until LDD with obtaining estimate E_{used} per l as 0.024 and 0.007 (in Joules) from FDD-HDD, and HDD-LDD, respectively. However, since the period from FDD-LDD supports transfer of unreliable data as per IoT framework, this advantage is of less importance
7. [36] suffers from worst energy utilization from FDD until HDD with obtaining a value of estimate E_{used} per l as 0.031 (Joules)
8. P-LUET faces near-average energy utilization in both FDD-HDD and HDD-LDD periods owing to value of 0.026 and 0.012 (Joules) within the range [0.024,0.031] and [0.007,0.014], respectively

Table 6 Comparison of P-LUET in terms of energy credibility over some pre-defined rounds

Protocols	Round count							
	500	1100	1800	2600	3400	4000	4600	5100
ATEER [33]	83.788	64.338	44.672	30.579	17.169	08.917	03.883	01.259
Improved clustering algorithm [34]	84.006	64.779	45.603	31.450	17.943	09.417	04.215	01.108
P-SEP [35]	85.010	67.010	48.344	32.277	17.663	08.116	01.530	00.000
EDCF [36]	78.413	52.520	26.901	13.084	05.674	02.539	00.382	00.000
LUET [24]	83.074	62.755	43.430	28.387	13.689	05.748	00.997	00.000
P-LUET	84.275	65.509	45.805	31.630	18.104	09.294	04.048	01.099

Table 7 Comparison of P-LUET in terms of T^{net} capacity over some pre-defined rounds

Protocols	Round count							
	500	1100	1800	2600	3400	4000	4600	5100
ATEER [33]	50,000	109,969	169,365	206,620	239,522	259,977	270,600	276,115
Improved clustering algorithm [34]	50,000	109,985	168,216	205,349	238,076	258,870	269,899	276,066
P-SEP [35]	50,000	109,893	170,630	219,652	262,335	290,244	308,498	311,752
EDCF [36]	50,000	110,000	169,746	204,399	222,690	229,414	233,755	234,037
LUET [24]	50,000	109,904	163,747	197,957	230,273	245,925	252,767	253,735
P-LUET	50,000	110,000	171,689	207,829	240,373	261,358	272,199	278,317

E_{th} plays an important role in defining P_{wgt} of an end-device for being chosen as the CH. The E_{th} value is greatly influenced by b parameter as $E_{th} = bE_{ini}$ while upon successful fulfilment of this condition, P_{wgt} value is greatly influenced by c parameters. Hence, it is important to understand the effect of these parameters on the performance evaluation. Therefore, in the second assumption, both the discrete parameters have been varied individually from [0.1, 1.0] while the algorithm’s FDD has been noted. Focussing on the b parameter, it can be observed from Fig. 16a that at $b = \{0.1, 0.4, 0.5, 0.7, 0.8\}$, a higher FDD has been observed depicting them as the probable optimal values of b . Upon Fig. 16b analysis for values $b = \{0.1, 0.4, 0.5, 0.7, 0.8\}$, it can be observed that $b = \{0.7, 0.8\}$ derive the network for a longer duration picking themselves as the most near optimal points if only longevity is considered. However, the proposed algorithm also focusses on improved network connectivity and thereby reduced device isolation

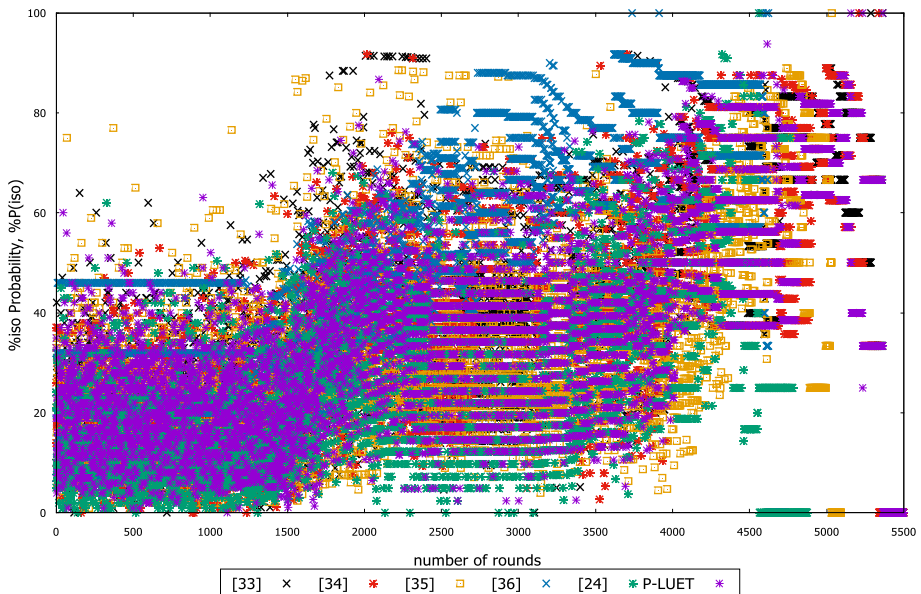


Fig. 14 %P(iso)- analysis; (%P(iso) ∈ [0, 100]) vs l, $M^2 = (300)^2$

Table 8 Longevity/efficiency performance variance w.r.t heterogeneity sensitivity

Cases	Heterogeneity scenarios in P-LUET ($E_T = 100I$)	$\% \tau_{fd}$	$\% \sigma_{w/I}$	1_ϵ	2_ϵ	3_ϵ
Case 1	One level heterogeneity $E_{ini} = 1, n_{ini} = 100$	52	48	12.634	12.648	12.618
Case 2	Two level hetero ($E_{ini} = 0.5, m_a = 0.5, a = 2, n_{ini} = 100$)	51	49	9.514	9.523	9.505
Case 3	Three level hetero ($E_{ini} = 0.5, m_a = 0.41, m_s = 0.44, a = 2, s = 3, n_{ini} = 100$)	27	73	12.355	12.723	12.221
Case 4		23	77	12.071	12.745	11.874

Observations

1. Case 1 may showcase a better $\% \tau_{fd}$ compared to the other cases, however, compared to Case 4, it suffers a drop of 30.4% in $NL_{w/I}$
2. Case 1 capitalizes high E_{ini} to sustain network for a longer τ_{fd} and obtain higher 1_ϵ . However, as mentioned earlier, 2_ϵ holds more value as compared to 1_ϵ upon considering algorithm's suitability within an IoT framework. Therefore, a higher value of 2_ϵ by 0.76% in Case 4 compared to Case 1 result, gives more advantage to Case 4 acceptance
3. Another point mentioned earlier for an algorithm's suitability for WSN-IoT is that 2_ϵ should be $>$ the 3_ϵ value as τ_{fd} is considered more purposeful as compared to algorithm's $\sigma_{w/I}$. However, it can be observed that, Case 4 shows 7.3% improvement in 2_ϵ w.r.t. 3_ϵ as opposed to 0.24%, 0.19%, 1.1% improvement in Case 1, 2, 3, respectively

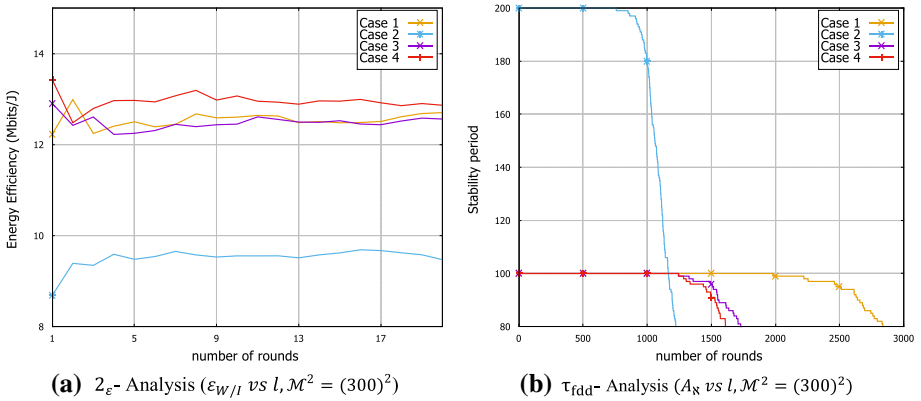


Fig. 15 Analysis based on heterogeneity sensitivity

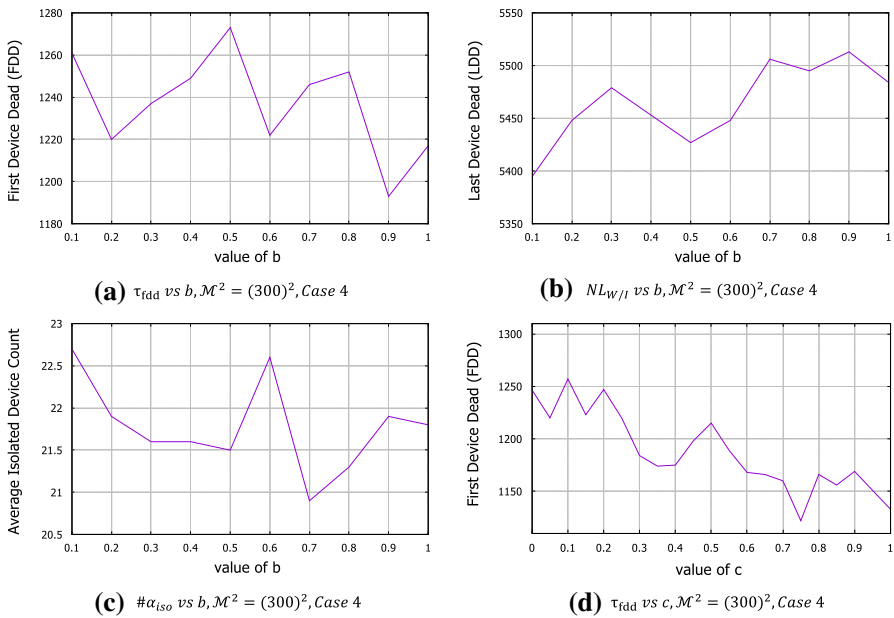


Fig. 16 Analysis based on b and c parameters

within the network. However, among the chosen probable b values, $b = 0.7$ provides the minimum isolation among the devices (as can be verified from Fig. 16c) making it the most apt value. Thereafter, stabilizing b value in the $E_{th} = bE_{ini}$ metric, c value is analysed for P_{wgt} computation. For the very purpose, c is varied from $[0.01, 1.0]$ to understand its presence. It can be observed from Fig. 16d that at smaller values of c , the network celebrates a longer τ_{fdd} while the network behaves vice-versa as c value increases. As clearly briefed in [30], c is that reel positive number that influences the CH count through the P_{wgt} computation. It is because a higher c results in comparatively a higher P_{wgt} outputting a smaller epoch for all the devices barring them to adapt as per

each residual constraints. This may result in higher CH count with most nodes communicating directly with their IoT Hub which becomes the cause of early FDD. A value of 0 is avoided for parameter c as it would result in $P_{wgt} = 0$ which again leads to a direct communication of all end-devices with the IoT Hub of the network. Therefore, from Fig. 16d, it can be inferred that $c = \{0.1, 0.2\}$ are considered the best possible options for consideration.

6 Conclusion

Exhaustive research critically analyses P-LUET algorithm w.r.t. the state-of-the-art clustering algorithms for their suitability in an IoT framework. The need of modification arises from the need to overcome hotspot issue and cluster overlapping. This algorithm has introduced the concept of annulus ring-based weight value (\mathcal{R}_b -based 3W) and conditional region concurrence (\mathcal{g}) to help overcome these issues. This algorithm successfully improves the stability period (τ_{fdd}) of LUET by 18.7%. As it is an extension of the previous work, P-LUET is expected to better the network connectivity by utilizing the lines of uniformity concept. Apart from showcasing improved efficiency of 12.8% over τ_{fdd} and a stronger connectivity, it also delivers a longer life of about 5506 rounds and a higher network throughput of 279,798 data packets. P-LUET has also been evaluated on scenarios of heterogeneous sensitivity and b, c parameters. Our work considers Constant Bit Rate (CBR) for the network. However, with the upcoming increase in the multimedia applications of IoT scenario, it is important that the network is equipped with Variable Bit Rate (VBR) that allows the sender to transmit data at a varying speed. Hence, this algorithm can be improved by considering VBR instead of CBR. The ID-enabled end devices can be replaced with IP-enabled devices to uplift the algorithm's suitability within a scenario. The impact of the algorithm can be studied on other WSN-IoT metrics like latency, packet delivery ratio and QoS in future. This algorithm can also be tested and made suitable for varying scalable network scenarios.

Funding No funding was received to assist with the preparation of this manuscript.

Availability of data and material The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Code availability The code of the algorithm has been run in Network Simulator 2.35.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

1. Ebadi, S. (2012). A multihop clustering algorithm for energy saving in wireless sensor networks. *International Scholarly Research Notices*, 2012, 1–4. <https://doi.org/10.5402/2012/817895>.

2. Yu, Y., Prasanna, V. K., Krishnamachari, B. (2006). Information processing and routing in wireless sensor networks. *World Scientific Publishing Co. Pte. Ltd.* <https://doi.org/10.1142/6288>.
3. Boujelben, M., & Youssef, H. (2008). Survey on pre shared key in WSN. In *International conference on wireless and mobile computing* (pp. 532–537).
4. Ahmad, T., Haque, M., & Khan, A. K. (2019). An energy-efficient cluster head selection using artificial bees colony optimization for wireless sensor networks. *Advances in Nature-Inspired Computing and Applications*. https://doi.org/10.1007/978-3-319-96451-5_8.
5. Al-Karaki, J. N., & Kamal, A. E. (2004). Routing techniques in wireless sensor networks: a survey. *IEEE Wireless Communications*, 11(6), 6–28. <https://doi.org/10.1109/MWC.2004.1368893>.
6. Al-Turjman, F. (2016). Impact of user's habits on smartphones' sensors: An overview. In: *Proceedings of the 2016 HONET-ICT*. <https://doi.org/10.1109/HONET.2016.7753422>.
7. Al-Turjman, F. M., Al-Fagih, A. E., Hassanein, H. S. (2012). A novel cost-effective architecture and deployment strategy for integrated RFID and WSN systems. In: *Proceedings of the 2012 international conference on computing, networking and communications (ICNC)* (pp. 835–839). <https://doi.org/10.1109/ICCNC.2012.6167542>.
8. Al-Turjman, F. (2016). Hybrid approach for mobile couriers election in smart-cities. In: *Proceedings of the 2016 IEEE 41st conference on local computer networks (LCN)* (pp. 507–510). <https://doi.org/10.1109/LCN.2016.79>.
9. Behera, T. M., Mohapatra, S. K., Samal, U. C., Khan, M. S., Daneshmand, M., & Gandomi, A. H. (2019). Residual energy based cluster-head selection in WSNs for IoT application.
10. Gubbi, J., Buyya, R., & Marusic, S. (2013). Internet of things (IoT): a vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29(7), 1645–1660. <https://doi.org/10.1016/j.future.2013.01.010>.
11. 'WaiOT Intelligent no flooding system'. [Online]. <https://www.waiot.eu/index.php>.
12. Atom, I., & Processor, E. (1991). Smart IoT-connected rail passenger infotainment solution by SADEL. <https://www.intel.in/content/www/in/en/internet-of-things/case-studies/smart-connected-railways-study.html>.
13. The Institution of Engineering and Technology, India IOT Panel. IoT Case Studies from India. <http://theiet.in/iot-case-studies-india>.
14. Summary, E. Bringing intelligence to manufacturing with the internet of things (IoT). <https://www.intel.sg/content/dam/www/program/embedded/internet-of-things/blueprints/iot-nexcom-bringingintelligence-to-manufacturing-blueprint.pdf>.
15. Kumar, J. S., & Patel, D. R. (2014). A survey on the internet of things: security and privacy issues. *International Journal of Computer Applications*, 90(11), 20–26. <https://doi.org/10.5120/15764-4454>.
16. Salem, A. O. A., & Shudifat, N. (2019). Enhanced LEACH protocol for increasing a lifetime of WSNs. *Personal and Ubiquitous Computing*, 23, 901–907. <https://doi.org/10.1007/s00779-019-01205-4>.
17. Teja, R. 'Different Types of Sensors-Electronics Hub'. [Online]. <https://www.electronicshub.org/different-types-sensors/>.
18. Heinzelman, W. R., Chandrakasan, A., & Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless microsensor networks. *Hawaii International Conference on System Sciences*. <https://doi.org/10.1109/HICSS.2000.926982>.
19. Akyildiz, I. F., Melodia, T., & Chowdhury, K. R. (2007). A survey on wireless multimedia sensor networks. *Computer Networks*, 51(4), 921–960. <https://doi.org/10.1016/j.comnet.2006.10.002>.
20. Lodhi, A. K., & Sattar, S. A. (2019). Cluster head selection by optimized ability to restrict packet drop in wireless sensor networks. *Soft Computing in Data Analytics*, 758, 453–461. https://doi.org/10.1007/978-981-13-0514-6_45.
21. Ahmad, M., Shah, B., Ullah, A., Moreira, F., Alfandi, O., Ali, G., & Hameed, A. (2021). Optimal clustering in wireless sensor networks for the internet of things based on memetic algorithm: memeWSN. *Wireless Communications and Mobile Computing*. <https://doi.org/10.1155/2021/8875950>.
22. Kumar, J. S., & Zaveri, M. A. (2018). Clustering approaches for pragmatic two-layer IoT architecture. *Wireless Communications and Mobile Computing*. <https://doi.org/10.1155/2018/8739203>.
23. Al-Humidi, N., & Chowdhary, G. V. (2019). Energy-aware approach for routing protocol by using centralized control clustering algorithm in wireless sensor networks. *Computing, Communication and Signal Processing*, 810, 261–274. https://doi.org/10.1007/978-981-13-1513-8_28.
24. Sood, T., & Sharma, K. (2020). LUET: a novel lines-of-uniformity based clustering protocol for heterogeneous-WSN for multiple-applications. *Journal of King Saud University: Computer and Information Sciences*. <https://doi.org/10.1016/j.jksuci.2020.09.016>.
25. Heinzelman, W. B., Chandrakasan, A. P., & Balakrishnan, H. (2002). An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, 1(4), 660–670. <https://doi.org/10.1109/TWC.2002.804190>.

26. Smaragdakis, G., Matta, I., & Bestavros, A. (2004). SEP: A stable election protocol for clustered heterogeneous wireless sensor networks. *Second international workshop on sensor and actor network protocols and applications (SANPA 2004)* (pp. 1–11).
27. Kumar, D., Aseri, T. C., & Patel, R. B. (2009). EEHC : energy efficient heterogeneous clustered scheme for wireless sensor networks. *Computer Communications*, 32(4), 662–667. <https://doi.org/10.1016/j.comcom.2008.11.025>.
28. Qing, L., Zhu, Q., & Wang, M. (2006). Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks. *Computer Communications*, 29(12), 2230–2237. <https://doi.org/10.1016/j.comcom.2006.02.017>.
29. Saini, P., & Sharma, A. K. (2010). E-DEEC-enhanced distributed energy efficient clustering scheme for heterogeneous WSN. In: *Proceedings of the 2010 First International Conference On Parallel, Distributed and Grid Computing (PDGC 2010)* (pp. 205–210). <https://doi.org/10.1109/PDGC.2010.5679898>.
30. Elbhiri, B., Saadane, R., Elfidhi, S., & Aboutajdine, D. (2010). Developed distributed energy-efficient clustering (DDEEC) for heterogeneous wireless sensor networks. In: *Proceedings of the 2010 5th international symposium on IV communications and mobile network* (pp. 1–4). <https://doi.org/10.1109/ISVC.2010.5656252>.
31. Javaid, N., Qureshi, T. N., Khan, A. H., Iqbal, A., Akhtar, E., & Ishfaq M. (2013). EDDEEC: Enhanced developed distributed energy-efficient clustering for heterogeneous wireless sensor networks. *International workshop on body area sensor networks (BASNet-2013) in conjunction with 4th international conference on ambient systems, networks and technologies (ANT 2013)* 19 (pp. 914–919).
32. Singh, S., Malik, A., & Kumar, R. (2017). Engineering science and technology, an international journal energy efficient heterogeneous DEEC protocol for enhancing lifetime in WSNs. *Engineering Science and Technology, an International Journal*, 20(1), 345–353. <https://doi.org/10.1016/j.jestech.2016.08.009>.
33. Singh, R., & Verma, A. K. (2017). Energy efficient cross layer based adaptive threshold routing protocol for WSN. *AEU International Journal of Electronics and Communications*, 72, 166–173. <https://doi.org/10.1016/j.aeue.2016.12.001>.
34. Wu, W., Xiong, N., & Wu, C. (2017). Improved clustering algorithm based on energy consumption in wireless sensor networks. *The Institution of Engineering and Technology, IET Networks*, 6(3), 47–68. <https://doi.org/10.1049/iet-net.2016.0115>.
35. Naranjo, P. G. V., Shojafar, M., Mostafaei, H., Pooranian, Z., & Baccarelli, E. (2017). P-SEP: a prolong stable election routing algorithm for energy-limited heterogeneous fog-supported wireless sensor networks. *The Journal of Supercomputing*, 73, 733–755. <https://doi.org/10.1007/s11227-016-1785-9>.
36. Priyadarshi, R., Rawat, P., & Nath, V. (2019). Energy dependent cluster formation in heterogeneous wireless sensor network. *Microsystem Technologies*, 25, 2313–2321. <https://doi.org/10.1007/s00542-018-4116-7>.
37. Tang, J., Zhou, Z., Niu, J., & Wang, Q. (2014). An energy efficient hierarchical clustering index tree for facilitating time-correlated region queries in the Internet of Things. *Journal of Network and Computer Applications*, 40, 1–11. <https://doi.org/10.1016/j.jnca.2013.07.009>.
38. Liu, X., & Liu, Q. (2018). A virtual uneven grid-based routing protocol for mobile sink-based WSNs in a smart home system. *Personal and Ubiquitous Computing*, 22, 111–120. <https://doi.org/10.1007/s00779-017-1093-2>.
39. Behera, T. M., Mohapatra, S. K., Samal, U. C., Khan, M. S., Daneshmand, M., & Gandomi, A. H. (2019). Residual energy-based cluster-head selection in WSNs for IoT application. *IEEE Internet of Things Journal*, 6(3), 5132–5139. <https://doi.org/10.1109/JIOT.2019.2897119>.
40. Pathak, A. (2020). A proficient bee colony-clustering protocol to prolong lifetime of wireless sensor networks. *Journal of Computer Networks and Communications*. <https://doi.org/10.1155/2020/1236187>.
41. Verma, S., Sood, N., & Sharma, A. K. (2019). Genetic algorithm-based optimized cluster head selection for single and multiple data sinks in heterogeneous wireless sensor network. *Applied Soft Computing*. <https://doi.org/10.1016/j.asoc.2019.105788>.
42. Kumar, D., Aseri, T. C., & Patel, R. B. (2009). EEHC: energy efficient heterogeneous clustered scheme for wireless sensor networks. *Computer Communications*, 32(4), 662–667. <https://doi.org/10.1016/j.comcom.2008.11.025>.
43. Iqbal, A., & Lee, T. J. (2019). GWINs: group-based medium access for large-scale wireless powered IoT networks. *IEEE Access*, 7, 172913–172927. <https://doi.org/10.1109/ACCESS.2019.2956029>.

44. Wang, Q., Dai, H. N., Zheng, Z., Imran, M., & Vasilakos, A. V. (2017). On connectivity of wireless sensor networks with directional antennas. *Sensors (Basel)*. <https://doi.org/10.3390/s17010134>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Tanvi Sood received a B.Tech degree in Electronics and Communication Engineering from Punjab Technical University, Jalandhar, Punjab, India, in 2012 and the M.E. degree in Electronics and Communication Engineering from National Institute of Technical Teachers Training and Research, Panjab University, Chandigarh, India, in 2014. She is currently pursuing a Ph.D. degree in Electronics and Communication Engineering at National Institute of Technical Teachers Training and Research, Panjab University, Chandigarh, India. Her research interest includes wireless sensor networks.



Kanika Sharma received a B.E. degree in Electronics and Communication Engineering from Maharishi Dayanand University, Rohtak, the M.E. degree in Electronics and Communication Engineering from Punjab Engineering College, Punjab University, Chandigarh and the Ph.D. degree in Electronics and Communication Engineering from Punjab Technical University, Chandigarh. This author is a member of IETE, ISTE. She is currently serving as Assistant Professor in the Department of Electronics and Communication Engineering, National Institute of Technical Teachers Training and Research, Chandigarh, India. She has more than 30 publications in International journals and conferences to her credit. Her research interests include embedded systems, digital system designing, wireless sensor networks, mobile communication.