Article:

Rady, M, Hafeez, M and Zaidi, SAR (2019) Computational Methods for Network-Aware and Network-Agnostic IoT Low Power Wide Area Networks (LPWANs). IEEE Internet of Things, 6 (3). pp. 5732-5744. ISSN 2327-4662

https://doi.org/10.1109/JIOT.2019.2905134

Computational Methods for Optimal Deployment of IoT Low Power Wide Area Networks

***Abstract*—In this paper, we tackle the design issue of optimal deployment of low power wide area network(LPWAN) IoT gateways(GWs). We classify GW deployment problem into two different categories, i.e., network spatial topology (NST) aware and NST-agnostic. In NST-aware GW deployment, precise location of IoT end devices (EDs) is known and thus the design questions are: (i) where to place gateways, i.e. to maximize received signal strength; and (ii) given received signal strength which GW should the ED be associated with to balance the network load. For, NST-agnostic GW deployment, same questions are answered in the absence of precise knowledge for the locations of EDs. For the NST-aware deployment we borrow tools from machine-learning such as K-means clustering for determination of optimized GW location. Subsequently, the link assignment problem is presented as an Integer Linear Programming (ILP) optimization. We prove that the NST-agnostic GW deployment principle of placement of GWs at highest altitudes, if applied automatically, may lead to poor network performance increasing the network operational costs. Consequently, we introduce the concept of network-agnostic GW placement algorithm whereby the location of GWs can be estimated without prior knowledge of specific locations of EDs and we use it as a guiding principle to design spatial algorithm for finding GW locations. We show that spatial algorithm can, in principle, provide effective gateway placement suggestions compared to a network-aware method such as K-means clustering. We show that using a computational method for GW placement like K-means or spatial algorithm, has a potential of creating competitive network performance using just the same number of GWs, thus cutting down the financial costs of the network and increasing its sustainability.**

***Index Terms*—IoT, wireless, optimization, LPWAN, clustering**

I. Introduction

Low Power Wide Area Networks (LPWAN) are being deployed in increasing magnitudes as corresponding market size is increasing at Compound Annual Growth Rate (CAGR) of more than 90% [2]. Internet of Things (IoT), which is foreseen to be the future of smart living, heavily relies on LPWAN technologies to ensure extended coverage in both outdoor urban and rural settings. The reliability of LPWAN technology due to its long range, low power, resilient frequency hopping (for instance such as Narrow Band frequency hopping for NB-IoT and Chirp-Spread Spectrum (CSS) based modulation in LoRA networks) qualifies it as ideal candidate for dense IoT deployments spanning possibly hundreds of square kilometers in complex terrain settings. LoRa [3], Sigfox [4] and NB-IoT [5] are leading LPWAN technologies for IoT networks that are still in infancy and yet to see a truly large scale commercial deployment. Currently there are seven large scale deployment studies funded through EU Horizon 2020 which are underway [6] . Similar trials are being conducted across the globe, interested reader is directed to [7] and [8]. One of the common themes across these trials is optimal deployment which maximizes the coverage.

A typical use case of such network deployment is in smart city scenarios where there is need to cover large number of end devices (EDs) randomly distributed across a city space with some level of uncertainty about their number or geographical location prior to deployment. LPWAN technologies promise a coverage of more than 24 miles in rural and up to 3 miles in urban areas under line of sight (LoS) conditions. Along with long range, LPWANs operate on low power providing more than 10 years of battery lifetime at a low cost [9], [10], [11]. The goal of this paper is to explore the design space of LPWAN networks (such as LoRA) with the aim of optimizing the coverage while attaining appropriate load balancing.

A. Research Challenges:

Network planning is often done under varying degrees of uncertainty and constraints. In the context of LPWAN, GWs can be deployed to serve already existing EDs or as backbone infrastructure which is expected to serve a growing network with EDs whose locations are yet to be determined. Therefore there can be lack of knowledge about the complete set of EDs which presents itself as a greater challenge especially when the network is subject to high utilization. Hence, maintaining the Quality of Service (QoS) while optimally balancing the traffic load across the GWs to ensure overall network efficiency becomes crucial. Therefore, deployments of LPWANs spanning tens to thousands of nodes requires computational approaches to determine efficient placement of GWs to provide minimal cost and maximal coverage to EDs. Any variation in architecture in terms of GW placement or link assignment may result in incurring extra financial costs on purchases of GWs. It may also undermine network sustainability cost in terms of expected daily battery lifetime if it offers low link budgets, stipulating for instance larger transmit powers which increases device power consumption as explained in [11]. The problem of optimal deployment is not straightforward as: (i) the goal is to optimize coverage as the GW placement is

The work was made possible through EPSRC Network Plus (EP/P001246/1) Feasibility Study Funding and Research England QR GCRF Funding under project GRIST.

The research reported here was supported and funded by the PERCCOM Erasmus Mundus Program of the European Union (PERCCOM- FPA 2013-0231). The authors would like to express their gratitude to all the partner institutions, sponsors, and researchers involved in the PERCCOM program [1]

Mina Rady is a graduate of PERCCOM M.Sc. program , University of Lorraine, France . (e-mail: minarady@gmail.com)

Maryam Hafeez is Senior Lecturer at the School of Computer Science and Engineering, University of Huddersfield, U.K. Email:{m.hafeez@hud.ac.uk}

S. A. R. Zaidi is with the Electronic and Electrical Engineering School, The University of Leeds, Leeds LS2 9JT, U.K. (e-mail: S.A.Zaidi@leeds.ac.uk)

NP-hard problem; (ii) Additionally, when load balancing or balancing of devices and optimal associations are factored into problem, it becomes even more challenging problem considering intrinsic heterogeneity of devices and propagation uncertainties (due to terrain variability). One of the fundamental assumptions which underpins the optimal deployment of GWs is that geometric two-dimensional distance between wireless nodes is always positively correlated with path loss, **this is not always true (as discussed later in this section)**. Hence optimal placement of LPWAN GWs presents itself as an important design challenge for the practical functioning such networks in terms of ensuring coverage while balancing the load across the network at a low cost. In a complex network planning scenario, such as smart city, we are interested in finding proper placements of GWs such that the **proportion of links** which can be served by each **gateway are balanced** while **maximizing** the **average received power**. In this paper therefore, our aim is to uncover design principles for how to efficiently deploy these GWs to maximize the efficiency of future networks.

B. Related Work

1) Related Literature in LPWAN Planning & Deployment: LPWAN, as an emergent technology, has been subject of research to explore its basic landscape. In [12] authors explore fundamental limits of LoRA type LPWANs. The authors investigate scalability of properties of LoRA LPWANs. A more detailed theoretical benchmarking was introduced for LPWAN technologies such as LoRa, SigFox, Weightless, and Ingenu in [13]. Quantitative research has attempted to explore the radio limitations of various LPWAN technologies through empirical field measurements in [11], [12], [14], [9], or using wireless propagation simulation in [14]. More realistic examination of LoRaWAN scalability was done in a simulated rural environment in [15].

Internet Engineering Task Force (IETF) in [16] has clearly identified that ultra-dense deployment of EDs will be key characteristic for the future LPWANs. The density of deployment expected is several order higher than small cellular networks where the dense deployment improves both coverage and capacity due to aggressive spectral reuse. While LPWAN EDs draw parallels in terms of density with small cellular network, from planning perspective LPWANs have different characteristic. Firstly, the topology of network is not cellular by design. Each LPWAN cell spans several kms and usually employs some distance dependent transmission adaptation scheme (such as different SFs in LoRA). From spatial perspective LPWANs can be visualized as clustered spatial networks rather than cellular networks. In cellular networks the location of end users is generally unknown and in past few years research community has employed spatial stochastic models to cater for this spatial uncertainty [17]. However for LPWAN deployments generally the topology of EDs can be established and GW placement needs to be optimized accordingly. A simple approach would be to apply K-mean clustering as presented in this paper. However, such dependence of GW placement implies non-optimal coverage for nodes commissioned thereafter. Also when coverage optimization is coupled with load-balancing which is not intrinsic for K-means (which maximizes SNR but does not yield equal number of nodes per GW) the optimal association becomes an open problem. To this end, this paper is geared to address this open issue. Moreover, notice that the optimal deployment of GWs in a network-agnostic setting is particularly more challenging as compared to the scenario where locations of EDs are known a priori. Consequently, to the best of our knowledge this is first study to tackle this design challenge.

2) Related Literature in WSN Clustering: Since late 1990s, extensive number of clustering algorithms for WSNs were proposed which effectively also yield optimal location for GWs or so called cluster heads. Contributions such as [18], [19], [20], [21], [22] remain some of the most highly regarded research efforts. Each paper proposes a different cluster head (CH) election algorithm coupled with an energy efficient protocol that realizes the algorithm in a conceptualization of a WSN following purely random distribution. An energy consumption model for WSN clusters is proposed in [22] and is tested on the clustering protocol proposed in [18]. Authors of [21] distinguish it as an approach for fixed location nodes, while authors of [18] assume fixed CH locations and stationary nodes and CHs all over the network. However, several research efforts such as [23], [24] propose clustering approaches with explicit assumption of mobility of nodes. Also, the CH election approach in [21] explicitly assumes that CHs locations are independent of each other. Clustering computations are simplified by assuming homogeneous node energy constraints such in [19], [18]. Finally, all propositions aim at minimizing energy consumption in the network by providing, eventually, energy efficient protocols which implement a self-clustering algorithm in the WSN where *any* node in the network can become a CH.

These studies abstract RF aspects into certain probabilistic attributes. Consequently, they do not consider impact of terrain geometry on WSN behavior and CH election. Following from this is an overall inherent assumption of positive correlation between Received Signal Strength Indicator (RSSI) and geometric distance. Consequently, after CHs are elected, nodes are assigned to geometrically closest CH indicated by the maximum CH RSSI. Relatively recent effort such as [19] explicitly assumes that "distance between nodes can be computed based on received signal strength". While this maybe true for distances on a planar surface, the geographical element is a key determinant in realistic propagation loss which is a major factor in the effective RSSI saturation in the network. For example, a node may have strong RSSI from a farther CH when both devices are separated by low altitude terrain. But the same node could have relatively less RSSI from a closer CH when both devices are separated by high altitude terrain. The clustering algorithms are also assessed in abstract simulations of theoretical diameters of up to 200 meters as in [18] which are too small compared to LPWAN ranges spanning several kilometers and where terrain geometry is expected to play a significant role in the WSN behavior.

On the other hand, the general direction is focused only towards the logical topology of the network assuming an ad-hoc WSN with rotating CHs (i.e. each node is likely to be elected as CH at any point in time) and with multi-hop routing through CHs to the base-station. Approaches in [18], [21] explicitly elect every node i in the network at a given probability P_i to be CH to achieve load balancing in the network. However, topology of LPWAN is essentially defined as star topology or star of stars where GWs (or CHs) are predefined and are configured and mounted exclusively for this function as in [25], [11], [9]. This change presents a departure from the topological assumption in the previous papers.

Even when assuming energy efficient logical topology, the energy saving on the physical layer sets the physical lower limit of energy consumption in any transmission in the network, regardless of the logical topology. This is because higher RSSI saturation allows use of lower transmit powers, which are key components in battery consumption during any transmission.

Furthermore, an explicit assumption in all mentioned sources is that nodes geometric distribution is following a uniform random Poisson distribution across the entire surface covered by the WSN. Therefore, all efforts benchmark their approaches against homogeneous random conceptualizations of WSN deployments. However, an urban deployment of an LPWAN challenges the assumption of uniform random distribution of WSNs. In the case study we will examine later in the paper, we can see that the WSN simulated based on real smart city parameters, follows a distribution across the city area that is far from uniform, thus challenging the essential premise of such approaches for LPWAN planning in urban infrastructure monitoring.

C. Novelty Compared to state-of-the-art in Green IoT Deployment

Research in [26] represents state of the art in analytical methods to design green IoT deployments in hierarchical organization using an NP-hard optimization model for planning an optimized deployment. The model is utilized in an algorithm (MECA) based on K-means clustering heuristic to determine relay (i.e. GW) locations. The optimization method is formulated as a constrained linear optimization model subject to energy consumption constraints and budget constraints. The network hierarchy is computed as *Steiner* tree (with link energy consumption as the weights of the edges) which is the basis of the NP-hardness of the problem formulation, as explained by the authors.

Our key contributions as compared to the approach proposed in [26] are as follows:

- The MECA algorithm in [26] relies in its foundation on K-means clustering to determine relay (GW) locations as a "canonical" clustering algorithm in a straightforward utilization. However, such an assumption considers only a predefined set of sensor node locations and therefore would be inapplicable in a network-agnostic mode. Moreover, it is important to take into consideration the heuristic nature of K-means algorithm, especially with the variation of its seed parameters. This means that the computed solution is non-deterministic for the same input. However, our research evaluates K-means in conjunction with our proposed GW location deterministic algorithm that is agnostic to sensor node locations. We show how our proposed approach is deterministic for any given terrain profile and we show also how its performance is competitive to that of K-means clustering but without any predefined knowledge of sensor node locations.
- The link assignment problem is formulated in the research as a *Steiner* tree problem, an NP-hard approach. However, we propose an NP-complete approach as we solve it as a pure constrained Integer Linear Programming (ILP) formulation which enables research possibilities of looking at the same problem from the point of view of any other NP-complete approach.
- The approach is experimented on a small $100 \times 100\,m^2$ deployment area which does not consider any variation in terrestrial elevation profile. However, a realistic LPWAN IoT deployment spans hundreds of square kilometers where terrain elevation profile plays a significant role in propagation loss behavior. Furthermore, in the MECA algorithm, particularly step 4, there is an implicit assumption that Euclidian distance between nodes is synonymous with path loss since it relies on Friis model. However, this assumption contradicts the reality of radio propagation in a pervasive geographical topology where LoS and shadowing play a critical role and therefore, it is quite possible and common that a larger distance with higher LoS visibility will suffer less path loss than a small distance with lower LoS visibility. In contrast, the evaluation experimental setting used in this research uses real data set of sensor node locations spanning hundreds of square kilometers in a real terrestrial topology of highly variable elevation profile. We offer more realistic evaluation as we rely on Irregular Terrain Model propagation loss model which is aware of the geographical elevation profile of the covered terrain.
- The experiments in the research is done on randomly generated sensor node locations without defining the sort of random distribution used (and therefore assumed to be uniform distribution). However, a realistic data set for sensor node locations (as used in our validation experiments) is shown to be of highly non-uniform nature. In such case, performance of the clustering algorithm can and would differ because of biased node concentration densities in the covered area. In our experimental setup, since we use such a realistic setting, we show more reliable evaluation of all suggested methods and we show how our proposed network-agnostic approach in the Spatial Algorithm shows competitive performance to K-means clustering, even in such non-uniform distribution.
- Authors presume that all nodes (i.e. EDs or GWs) have the same wireless properties, and thus simplifying the problem to enable the proposed link assignment optimization model. However, in a realistic setting, device properties could be quite heterogeneous in terms of their wireless properties. In contrast, our proposed link assignment optimization model considers heterogeneous *link*

budgets individually (where each link budget embeds various complex factors such as network interface properties and path loss behavior). This is a significant feature since not all devices would afford the same energy capacity to transmit at the same power and range. Therefore, the link assignment approach in this research offers a link assignment approach that is flexible to adapt to both homogeneous and heterogeneous deployment settings.

D. Contributions

In this paper, our aim is to provide design principles for how to efficiently deploy LPWAN GWs to maximize the coverage and load efficiency of future IoT networks. As explained above, studies on network planning addressing the deployment of LPWAN gateways are very limited. Hence, it is essential to study various possible approaches and compare their benefits in terms of coverage and cost optimization across the network. Based on the findings, it is important to propose new methodologies that overcome the shortcomings of existing solutions. This paper is geared towards finding these optimal design principles for LPWAN deployment. In summary, our research has the following novel outcomes:

- We borrow tools from machine learning such as K-mean clustering for devising LPWAN geographical planning in an NP-complete approach when locations of ED are precisely known.
- In order, to balance the load across GWs, we formalize LPWAN network link assignment procedure in an ILP form for network QoS optimization, thus presenting a generalized NP-complete form for the problem. This makes it possible to converge on global optimal solution with maximum RSSIs in the network under defined GWs load constraints.
- In a Network-agnostic deployment approach, we prove that principle of highest altitude for GW placement, if applied automatically irrespective of terrain nature or network distribution, can lead to very poor deteriorated network coverage.
- We introduce the concept of "Network-Agnostic Wireless Planning" where GWs locations can be estimated without prior knowledge of specific locations of EDs in a dense pervasive distribution. We introduce a spatial algorithm for network-agnostic planning and show that the algorithm can, in principle, provide near-optimal solutions for GW placement given only the terrain profile of the city.
- We carry out a detailed case study of an LPWAN GW deployment for the city of Leeds, operating in the ISM band 868MHz (which is employed by LoRA in EU). We show that the network-aware method (K-means) provides the best coverage measured in terms of RSSI providing up to 20% network RSSI gain as compared to other approaches. We show that network-agnostic planning method, Spatial Method, can provide solutions with competitive RSSIs to network-aware method, without prior knowledge of existing EDs locations. We also demonstrate that ILP link assignment method can provided optimal load balancing solution in traffic hot spots without any tangible deterioration to network RSSIs.

We examine application of K-means clustering, otherwise known as Lloyd's algorithm [27] for selecting GW deployment locations assuming that ED locations are known. The resulting model is non-deterministic polynomial (NP)-complete [28]. We also adopt a network agnostic (Spatial and Grid) approaches to find optimal GW placement locations assuming ED locations are not known. For these approaches we find the propagation losses and then examine application of Integer Linear Programming (ILP) based constrained optimization for link assignment between respective GWs and EDs. We employ the Irregular Terrain Model (ITM) [29] implemented by the International Telecommunications Union (ITU), to capture the geographical profile of the terrain of the deployment area. We rely on Simplex Method implementation of ILP which is a core foundation of linear optimization applications as presented in [30]. Even though Simplex Method runs in exponential time for worst cases, it is highly efficient and normally behaves as an efficient polynomial time algorithm as theoretically explained in [31]. We compare the GW deployment and link assignment approaches and demonstrate that network agnostic approach with ILP based link assignment outperforms other approaches in terms of link quality and load distribution across the network.

E. Organization

The paper is organized as follows. In section II, we present our proposed methodology. We explain the formal foundations of our experimental techniques. In section III, we present our experimental setup. In section IV, we present the aggregation of results of all experiments and we benchmark the performance of each resulting network architecture. In section V, we present the foundations for generalization of our patterns in the classic literature of Irregular Terrain Model. In section VI, we present our conclusions.

II. Methodology:

In our methodology, we split the network planning process to four core phases as depicted in the Unified Modeling Language (UML) diagram in figure 1. The input to our experiments can be either site elevation data of the terrain to be covered or locations of EDs to be covered. Then planning process proceeds as follows:

- Phase I, GWs locations computation: this phase has two subcategories of methods: 1) Network Aware, where location of network EDs are completely known and are not expected to change much. For this category we examine K-means clustering 2) Network Agnostic, where GWs locations is computed with no prior knowledge of EDs locations. For this category we examine two methods: Grid method and Spatial Method.
- Phase II, propagation loss modeling: we compute link-budgets based on point-to-point wireless propagation simulations among all EDs and GWs
- Phase III, link assignment: this phase has two subcategories of methods: 1) Unconstrained, where GWs have virtually unlimited capacity relative to the number of EDs to be covered. Here we examine two assignment

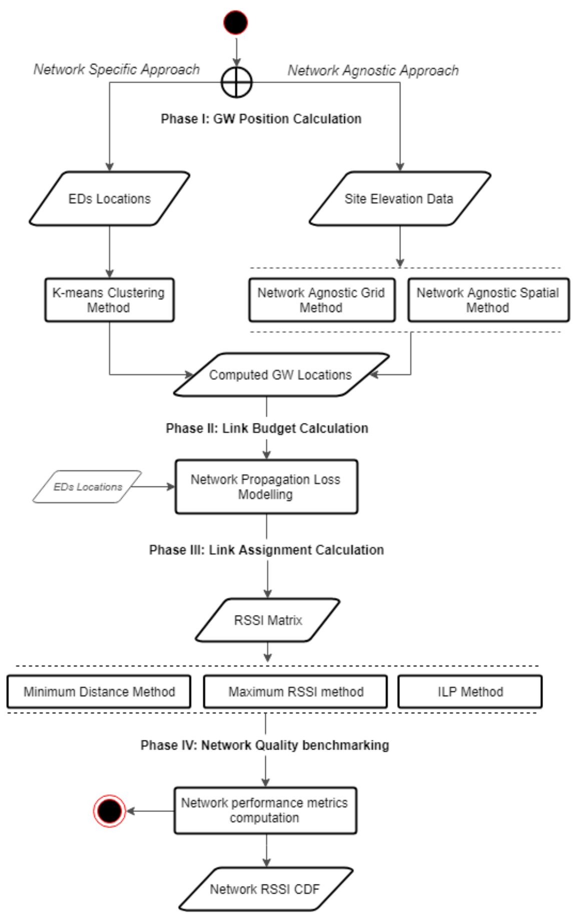


Fig. 1: Experiment Design

methods: A) Assignment to based on minimal distance and B) assignment based on maximum Received Signal Strength Indicator (RSSI). And 2) Constrained, where GWs have limited capacity relative to the number of EDs to be covered and therefore, such capacity need to be respected while assigning EDs to GWs. In this category, we present an ILP formulation and implementation

- Phase IV, the resulting network configuration is evaluated with respect to select metrics.

A. Phase I: GW Position Optimization Heuristics:

In general, we assume that at the beginning of the planning process, network engineers have at hand K GWs that are to be used to cover the terrain. We examine the following methods to compute the locations of K gateways.

1) K-means method (network-aware): In this approach, we perform clustering to compute the GW locations as K cluster centroids. This is meant to find a solution set that is appropriate to EDs with already known locations. The purpose of this approach is to optimize GW positioning to be geometrically closest to as many EDs as possible by optimizing centroid locations such that Euclidian geometric distances among EDs and cluster centroids is minimized. More specifically, it processes latitude and longitude arrays to produce geometric clusters in Lat/Long domain. As discussed in sec. I-B, the details of the algorithm can be found in [27].

2) NST-Agnostic Grid method (NAG) : We construct this method to compute the K-highest points in a terrain based on the intuition that higher altitudes for GWs placement creates better coverage and QoS. This is because of the positive correlation between elevation of GWs and their LoS coverage range as in the LoS formula [32]:

$$d \approx 3.57 \left(\sqrt{\alpha.h_1} + \sqrt{\alpha.h_2} \right) \tag{1}$$

where d is distance to horizon of a GW placed with h_1 antenna height towards an ED placed at h_2 antenna height and α is a constant representing earth's bulge. While we fix the physical antenna height for the sake of simulations, the terrain relative altitude of GWs with respect to EDs acts as an additional parameter contributing to the final antenna altitude in relation to EDs. In contrast to network-aware approach, this method is strictly dependent on the terrain altitude profile and not the EDs locations. Therefore, it can be useful in cases where EDs locations bear significant level of uncertainty in either scale or scope. The method is organized as follows:

1) Target area is split into a grid of n x n resolution. Grid resolution n should be estimated relative to the variation in terrestrial elevation such that no significant change of variation can occur within any given cell in the grid. A concrete example is given in subsection IV-A.
2) The coordinates of the geometric centroid C_{ij} of each grid $Cell_{ij}$ are computed. This results in two dimensional matrix $C_{n^2 \times 2}$. Those centroids are the candidate GW locations. Then compute and append the elevation of each centroid to its sub-array, producing matrix $C_{\ n^2 \times 3}$.
3) Sort matrix C in descending order by elevation.
4) The highest K points in elevation (regardless of their inter-distances) are selected as deployment sites.

3) NST-Agnostic Spatial method (NAS): Following the heuristics of the Grid method, we develop the Spatial method to compute the K-highest points while enforcing a minimal separation threshold σ relative to the area of the covered terrain such that GWs locations are separated by a minimal distance threshold. By definition, we compute σ in terms of the standard deviation in the set of all possible distances in the Grid. Therefore, it is constant and representative for any given grid. This is to avoid over-concentration of results on a high altitude area. This method used the four steps in the Grid method, and in addition we proceed as follows:

1) Compute the euclidean distances between each pair of centroids in a new vector $\overrightarrow{d} = [(n^2-1)^2]$, let
$$s = \text{size}\,(\overrightarrow{d}\,) = (n^2-1)^2 \tag{2}$$
1) Compute mean distance,μ in vector d as:
$$\mu = \frac{sum(\overrightarrow{d})}{s} \tag{3}$$
2) Compute standard deviation σ in d as
$$\sigma = \sqrt{\frac{\sum_{i=1}^{s}(d_i-\mu)^2}{s^2}} \tag{4}$$
3) At last, we apply the pseudo code in Algorithm 1 to

compute GW locations:

Algorithm 1 Spatial Method (ASM)

Algorithm II.1: SPATIALMETHOD($Centroids[S][3], K$)

$\begin{cases} Centroids \text{ has S Nr of centroids with lat, long, elevation} \\ K \text{is solution size} \end{cases}$

GWs ← new array
CSorted ← Centroids.Sort(desc,3)
comment: Sorted centroids by elevation in desc order

GWs.Push(CSorted.Pop())
comment: add highest point to solution

$Sigma \leftarrow$ StDev in centroids distance matrix

while GWs.Size() $< K$

- CandidateC [3] ← CSorted.pop()
- **comment:** pop next highest point
- DistanceVector ← COMPDISTVECT($GWs, CandidateC$)
- **if** DistanceVector.Min() $> Sigma$
  then GWs.push(CandidateC)

procedure COMPDISTVECT($GWs[][], C[]$)

- DistanceVector[GWs.Size()]
- **for** $i \leftarrow 1$ **to** GWs.Size()
  do DistanceVector[i] ← EuclidianDistance (C,GWs[i])
- **return** ($DistanceVector$)

B. Phase II: Propagation Loss Computation:

We rely on the Langley-Rice path loss prediction model, commonly known as Irregular Terrain Model [29]. The model, which is approved by the United States Federal Communications Commission (FCC) and implemented by ITU, offers theoretical foundations for classifying terrains into five different profiles as in table I [29]. It has been proven recently in [33] to maintain general stability and competitive accuracy for common applications. The average terrain profile according to ITM is "hills" and we rely on that to perform our experiments in a location matching this profile. This is to establish theoretical belief that our observed patterns are expected to be replicable in an average type terrain, especially since ITM is based on statistical models that are founded on empirical prediction curves. It offers estimation of propagation loss depending on empirically obtained path loss curves in different terrestrial environments. It is implemented in ITS algorithm [34] and it is also valid for a large variety of engineering problems within frequency range 20 MHz- 20 GHz including TV broadcast and mobile networks. ITM incorporates terrestrial features of the specific locations of network deployment, including terrain elevation profile, climate, land surface refractivity, and it has also been deployed for various military and land-mobile systems applications [29]. It offers two propagation loss prediction modes that are both important to our analysis:

Terrain Category	$\triangle h$ (meters)
Flat (or smooth water)	0
Plains	30
Hills	90
Mountains	200
Rugged Mountains	500
For average $\triangle h$ use value 90	

TABLE I: Terrain categories according to ITM (source: [29])

- **Point-to-point prediction mode:** In this mode, the model considers the specific physical land elevation profile between the sending and the receiving radio nodes. Therefore, it is useful for more accurate point to point link budget prediction.
- **Area prediction mode:** This mode is based on computational categorization of terrestrial features into a set of profiles. Profiling considers factors such as variability of surface elevations, climate, and ground surface refractivity. However, according to [29], most of these parameters can be set to nominal values. For example, for a distance less than fifty kilometers, the climate parameter impact is insignificant and can be simply set to the average value of "Continental Temperate Climate". The only significant variable is the Terrain Irregularity Parameter:$\triangle h$, which is used to categorize terrain elevation profiles according to intervals of $\triangle h$ as in table I.

From the Langley Rice model, we obtain an n x m matrix $\mathbf{P}$ where n is the number of EDs and m is the number of GWs in the network. $\mathbf{P}_{nm}$is the relative RSSI between ED_n and GW_m. For convenience, we follow the reporting scheme of Radio Mobile of path loss estimation: we normalize relative RSSI with respect to receiver's threshold by adding a value of fifty to the result as in equation 5

$$\mathrm{RSS}I_{norm} = \mathrm{RSS}I_{Rx} - \mathrm{R}x_{threshold} + 50 \tag{5}$$

In this way, any results above ninety-nine, is too strong and would be counted as ninety-nine anyway and any result below one is too weak and would be counted as one anyway. This allows to fix the range of the report (and normalize it) to values in $[1, 99]$ interval. For example: If the receiver (Rx) threshold is set to -130 dB and computed RSSI is -110 dB, relativeRx is: $-110-(-130) = 20$ dB and "normalized" value would be $20+50 = 70$. Similarly, if RSSI is -60 dB, the "normalized" value would be $-60-(-130)+50 = 120$ dB which would be approximated as ninety-nine.

C. Phase III: Link Assignment Procedures:

At this point, there is still room for decision making regarding how EDs are assigned to GWs. In Global System for Mobile communications (GSM) networks, choice of mobile devices to base stations, especially during handovers, is mainly governed by RSSIs of base stations at EDs. However, LPWAN enabling modulations enjoy a different capacity where their modulations allow robust long range communication with relaxed latency constraint (such as Ultra Narrow Band frequency hopping for NB-IoT and CSS modulation for LoRa). On the other hand, it is restricted by duty cycle limitations, such as 1% duty cycle in EU regulations in 868 MHz band. Therefore, it is necessary to control the allocation of EDs access to GWs for sustainable network performance. For this phase, we experiment with three different approaches.In addition to matrix $\mathbf{P}$, we define $n \times m$ matrix $\mathbf{D}$ where $\mathbf{D}_{nm}$is Euclidian distance between $ED_n$and GW_m.

1) **Maximum RSSI approach:** In this approach, we assign each ED_i to GW_j with highest RSSI measurement. That is, where $\mathrm{RSSI}_{ij} = \max\left(\overrightarrow{\mathbf{P_i}}\right)$.
2) **Minimal distance approach:** In this approach we assign each each ED_i to GW_j that is geographically closest in location. That is, where geometric distance $D_{ij} = \min(\overrightarrow{\mathbf{D_i}})$ in the Latitude/Longitude domain.
3) **Integer Linear Programming approach:** We propose an ILP model to be applied on the link budget matrix computed in phase II. In the formalization in equation 7, we minimize the cost of the network as expressed in link budget. We impose a linear constraint on number of EDs per GW $\leq N$ such that no GW will be assigned more than N EDs. N can be simply defined as the ratio of total number of EDs to the total number of GWs. In practice N can be lesser than such a ratio based on several factors such as GW RF capacity in terms of number of available receive paths and expected packet rates from connecting devices. We define

$$f\left(E,G,C\right) = \sum_{i=1}^{E}\sum_{j=1}^{G} X_{ij}.C_{ij} \tag{6}$$

where C_{ij} is assignment cost of $\mathrm{G}W_{\mathrm{j}}$ at the $\mathrm{E}D_{\mathrm{i}}$, expressed as $-(\mathrm{RelativeRSS}I_{ij} \in [1,99])$, E is the number of EDs, G is the number of GWs, and N_j is the maximum EDs capacity for $\mathrm{G}W_{\mathrm{j}}$. With such formalization, we find global optimal solution for each given network arrangement as follows.

$$\min\ f\left(E,G,C\right), \tag{7}$$

$$\text{subject to}\ \begin{array}{rcl} \sum_{i=1}^{E} x_{ij} & \leq & N_j \text{ for each } \mathrm{GW}_j \\ \sum_{j=1}^{G} x_{ij} & = & 1 \text{ for each } \mathrm{ED}_i \end{array}.$$

$$\text{where}\ X_{ij} \ = \ \begin{cases} 1 & \mathrm{E}D_{\mathrm{i}} \text{ is assigned to } \mathrm{G}W_{\mathrm{j}} \\ 0 & \text{otherwise} \end{cases}.$$

The general cost expression C_{ij} can be extended to include more complex cost metrics such as Energy per bit or device battery change cost. Since different link budget may stipulate different levels of transmission power which lead to various levels of energy consumption and, hence, expected daily battery life times. At this fundamental stage, we experiment with basic definition of C_{ij}in terms of $RSSI_{ij}$.

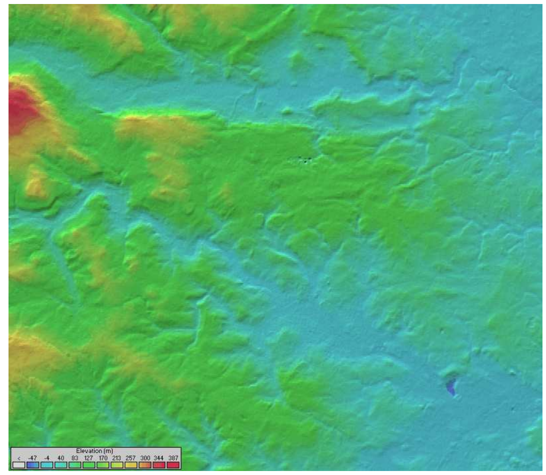


Fig. 2: Leeds Elevation Map

D. Phase IV: Evaluation Metrics:

In this phase, we benchmark our computed configuration as follows:

- Cumulative Distribution Function (CDF) of path loss l in the network as:

$$\begin{array}{rcl} P(L) & = & \int_{-\infty}^{l} p(t)\, dt \\ & = & \frac{l}{100} \qquad \text{for} \quad 0 \leq l \leq 100 \end{array} \tag{8}$$

Where $p(t)$ is the probability density function for a pathloss value less than l and $p(t) = 1/100$.

- Statistical features of path loss measurements in the final network configuration: mean, median, standard deviation. Moreover, as a nominal indicator of network quality, we calculate the percentage of EDs with relative RSSI $P \geq 50$ (i.e. above Rx threshold). We use this as a practical benchmark for EDs with bare minimum signal quality.

CDF computation is useful to compare the saturation of RSSIs in each resulting network architecture with the knowledge that all RSSIs below fifty indicate devices out of service.

III. Experimental Setup

We choose as a case study an application of monitoring operational status of lamp-posts across city of Leeds in the United Kingdom. We obtained a full listing of lamp-posts locations across Leeds from the City Council's data repository [35]. The total number of lamp-posts exceeds $106,000$. For the sake of analysis, we will perform our approaches on a random sample of almost two thousand lamp-posts. We consider the topographic nature of Leeds which has relatively high variance in altitudes as shown in figure 2. We use Radio Mobile wireless propagation simulator which incorporates ITS implementation of ITM.

For the purpose of the exploration, we fix several parameters, especially those of the ITS simulation algorithm across all experiments. For Tx/Rx devices, we fix PHY configurations as shown in table II and we fix ITS simulation parameters as in table III. For the network architecture, we assume that ten gateways are to be deployed to cover the entire city (i.e.$K = 10$). While it is of interest to load balance EDs

Parameter	Value
Frequency	868.1 MHz
Antenna gain	6 dBd
Antenna type	omnidirectional
Line Loss	0.5 dB
Antenna Height	8 m
Transmit Power	14dBm
Receiver Threshold	-130 dBm

TABLE II: PHY Configurations

Parameter	Value
Minimum frequency	868.1 MHz
Maximum frequency	868.2 MHz
Antenna polarization	vertical
Mode of variability	Spot at 70% of situations
Surface refractivity	301 N-Units
Ground conductivity	0.005 S/m
Relative ground permittivity	15
Climate	continental temperate

TABLE III: Experimental Setup

across GWs, we configure our ILP model with $N_j = 300$ EDs per GW_j. However, N_j can be configured variably for each GW capacity. And finally, for our RSSI normalization as in equation 5, any RSSI > -81 dB will be approximated as ninety-nine and any RSSI< -179 will be approximated as one.

IV. Results

A. Gateway Computation Results:

In this section we will show the results of phase I of the engineering methodology which is computing GW locations using all the four methods outlined previously.
Using K-means clustering method with longitude/latitude data of EDs locations, we compute cluster centroids as GW locations. The visualization in figure 3 illustrates computed gateway coordinates. We can see in figure 3 that the GW locations are concentrated near the city center where EDs are more concentrated. This is an expected outcome of K-means clustering since it computes centroids with reference to training data sample.

Using Grid method we run our computation with 45×45 resolution and with $K = 10$. Accordingly, the resulting GWs coordinates are visualized in figure 4. However, we can see that this approach has introduced overly concentrated solution set. Since the resolution is too fine, the algorithm returned several centroids located on top of one of the highest hills in the city.

Since this result has most solution set in adjacent grid cells, we attempt a second run with significantly less resolution of 15×15 to achieve more separation in the solution set. The resulting GW coordinates are visualized in figure 5 and we can clearly see in the figure that results are much more distributed

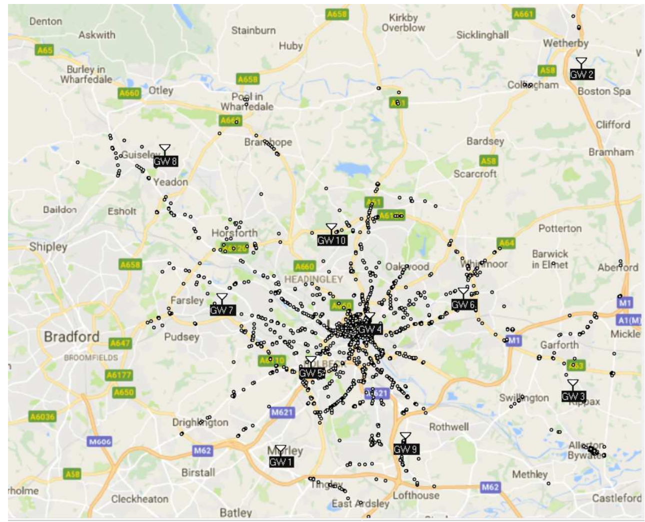


Fig. 3: Calculated gateway locations using k-means algorithms for k =10

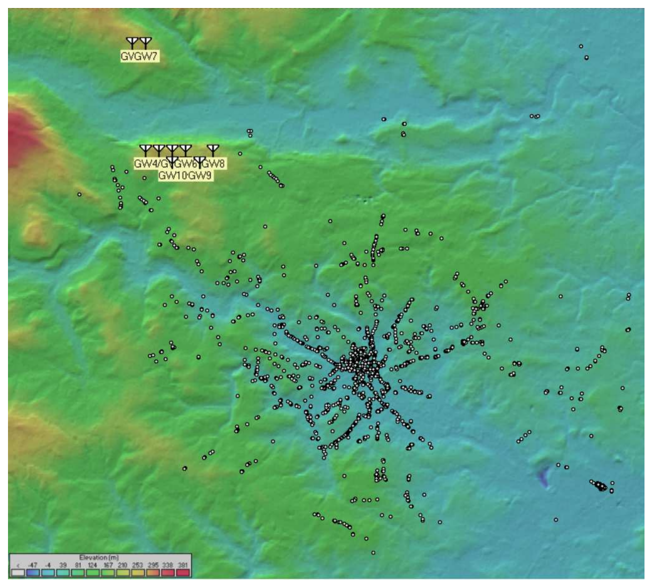


Fig. 4: Calculated gateway locations using Grid method with45×45 resolution

across the area which can offer a significant variance in the final topology.

Using the Spatial method, we define 45 x 45 grid over Leeds terrain. This fine grid resolution is convenient as the dimensions of the covered area is approximately 31 x 26 km and therefore the surface area of a grid cell is approximately $0.4\ \text{km}^2$. This figure seems appropriate as the elevation heat map of Leeds does not show very sharp shifts in elevation and it is rather smooth. The resulting gateway location coordinates are visualized in figure 6. We can observe that the GW results are almost uniformly displaced away from each other in a manner proportional to the covered area, as expected.

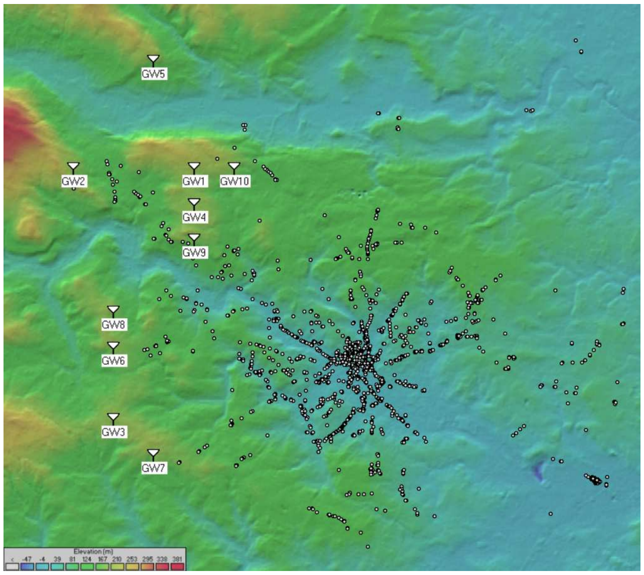


Fig. 5: Calculated gateway locations using Grid method with15×15 resolution

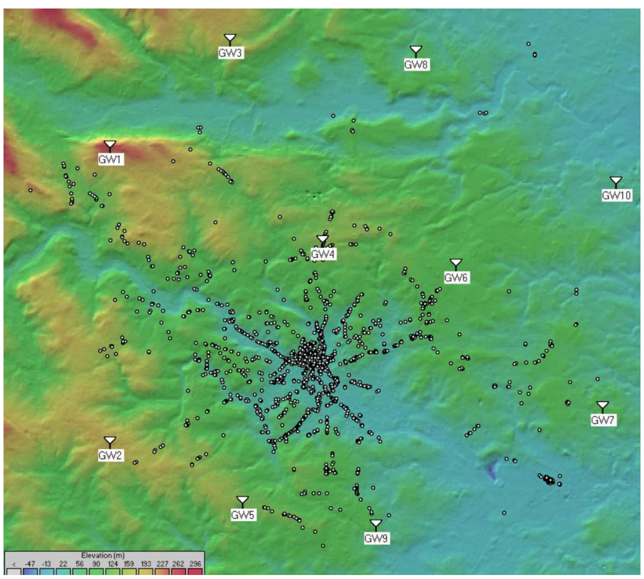


Fig. 6: Calculated gateway locations using spatial method

B. Link Assignment Computation Results:

In this section, we show the results of applying the three methods of link assignment on each of the four network configurations resulting from GW computation phase. This leaves us with a total of twelve combinations of experiments and their corresponding architectures. In table V, in the Appendix, we visualize the GWs load distribution in each architecture. For each architecture, we show the histogram of GW loads expressed as the number of assigned EDs and a scatter plot showing a geographical visualization of EDs colored according to their GWs assignment. Moreover, in table VI, in the Appendix, we visualize RSSI distribution in each network architecture. For each architecture, we show an RSSI histogram in the network and a scatter plot showing a geographical visualization of EDs colored by RSSI strength.

Moreover, we present detailed quantitative statistical description of the RSSI data for each network architecture in table IV. For each experiment, we show the traditional statistical parameters of RSSI data in a addition to the percentage of EDs with RSSI ≥ 50.

C. Comparative Analysis:

To compare the statistical behavior of network quality for the twelve architectures explored in this research, we plot the statistical box-chart for all experiments in figure 7. Moreover, we plot the CDF for the RSSI probability distribution in each network architecture in figure 8. We can deduce from the plot significant differences between the various architecture performances:

- Grid method shows comparatively the least reliable results, especially with the run of the 45×45 fine resolution. This is evidence that simply choosing the highest altitudes locations for GW deployment, regardless of their distributiveness, does not guarantee reliable architecture and may even lead to very deteriorated network performance. This is critical to note in a complex urban setting with a geographically pervasive and dense network. Therefore, the intuition that engineers commonly use to place GWs at highest points in a city may lead to deteriorated QoS.
- We can observe significant closeness in overall network quality between network-aware method (K-means) and our proposed network-agnostic Spatial method (bold lines in CDF plot). We can also observe that introduction of ILP method barely undermined network quality in terms of RSSI saturation while achieving critical load balancing on GWs. This highlights the significance of the Spatial method as a promising approach for network-agnostic planning and promotes it for future efforts of research and enhancements. It also highlights the significance of the ILP formalization for load balancing and introduces it as a promising approach worthy of further investigation.
- Given the distinctively efficient results of K-means and Spatial method we propose that network planning process must respect at least network EDs locations or minimal distance threshold among GWs to ensure reasonable acceptable network architecture quality. This is to say that it may not be the wisest practice to simply pick the highest buildings in the city as locations for GWs without performing the aforementioned calculations.

Moreover, we plot the distribution for GWs load in all assessed architectures in figure 9. We can observe significant patterns as follows:

- In overall, using ILP based link assignment proves quite significant in streamlining gaps in load distribution created by any GWs locations computation method. This is particularly useful given the highly non-uniform distribution of EDs leading to non-uniform GW load distribution.
- The general class of K-means based clustering and Grid methods (applied with lower resolutions) offer quite stable load distribution regardless of link assignment method.

	Mean	Median	Min	Max	StDev	$P(RSSI \geq 50)$
K-means with minimum distance	78.7	90	1	99	16.1	97%
K-means with maximum RSSI	87.7	82	42	99	10.6	99.8%
K-means with ILP	87.2	90	42	99	10.7	99.8%
Grid (45) with minimum distance	68.5	55	2.	99	15.1	90.6%
Grid (45) method with maximum RSSI	68.1	65	35	99	14.5	91.6%
Grid (45) with ILP	67.3	64	32	99	14.9	89.6%
Grid (15) with minimum distance	66.1	60	1	99	19.0	79.3%
Grid (15) method with maximum RSSI	73.1	73	44	99	13.8	97.4%
Grid (15) with ILP	69.7	73	41	99	14.9	91.2%
Spatial method with minimum distance	66.4	68	21	99	15.1	85.9%
Spatial method with with maximum RSSI	81.7	84	47	99	10.9	99.3%
Spatial method with ILP	79.9	81	41	99	11.3	98.2%

TABLE IV: Statistical description of EDs RSSI in resulting architectures

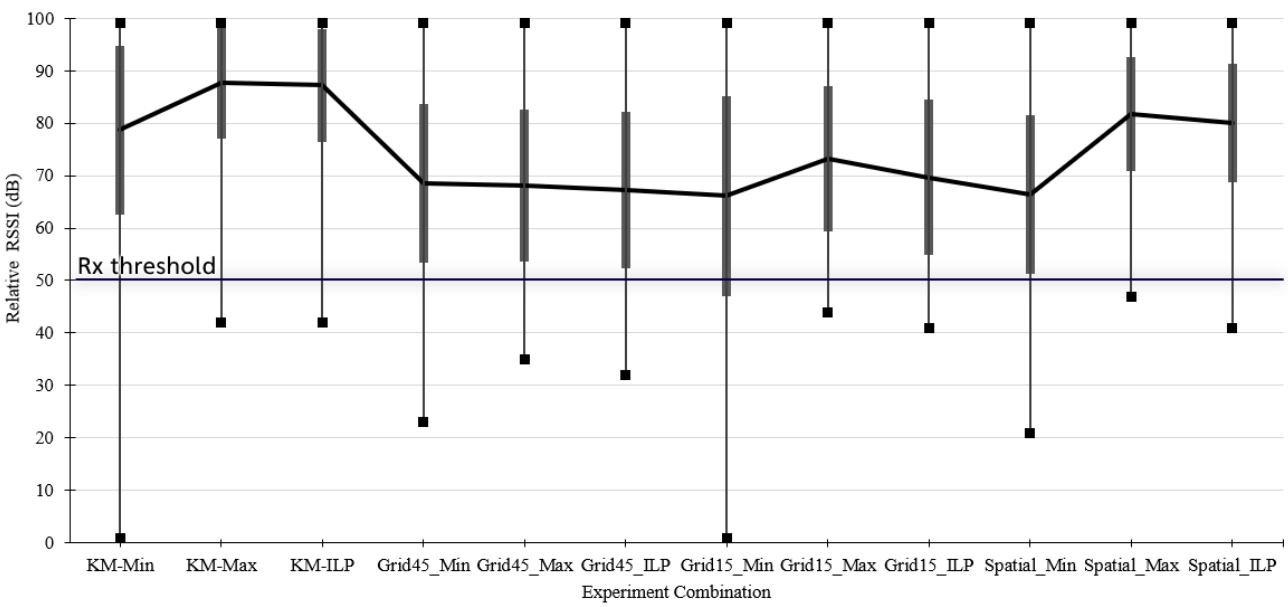


Fig. 7: Statistical overview of all results

- Spatial method offers highly variant load distribution however ILP link assignment streamlines the load distribution balance, thus offering quite competitive load distribution compared to rest of architectures.
- We can observe that unconstrained methods run the risk of leading to excessive stress on more GWs which can jeopardize overall network sustainability and quality. For instance, as in figure 9, `Grid45_min` run introduces massive load on GW8 which is the nearest GW to Leeds city center, where most nodes are concentrated. However, general class of K-means method has self contained load-distribution feature thanks to clustering approach used in determining GW locations.
- To further highlight those patterns, we plot detailed geographical visualization of GW distributions and GW assignment histograms of all network plans in the Appendix in table V. In table VI, we plot geographical visualization and histograms for resulting RSSI distribution in all network plans.

V. Generalization Scope: Leeds Terrain Profiling:

As described in section II-B, the terrain irregularity parameter $\triangle h$ was calculated for Leeds terrain. The results of the computation proved that Leeds falls under the "hills" category which sets $\triangle h = 90$m which is otherwise the default value if the terrain $\triangle h$ profile is unknown, as per the original guide to the ITS algorithm implementation. This sets a theoretical basis for generalization of the models patterns observed in this use case for any terrain within similar terrain profile. In this section, we show briefly the steps for $\triangle h$ approximation for Leeds Terrain as in the following steps:

1) We compute a grid of reasonable resolution that captures the terrain elevation profile with satisfactory inclusiveness. For our case, we use 11×14 grid.
2) We retrieve the elevation profile (in meters) for each grid line cross section. Samples of the elevation profile

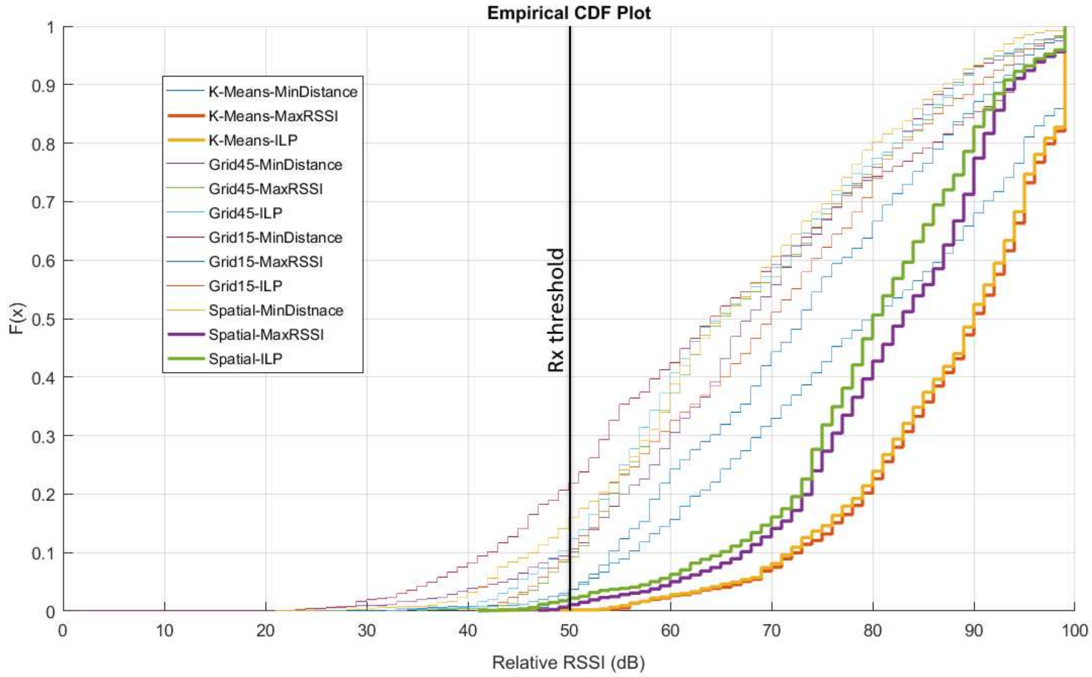


Fig. 8: Cumulative Density Function curves for network RSSI in all experiments.

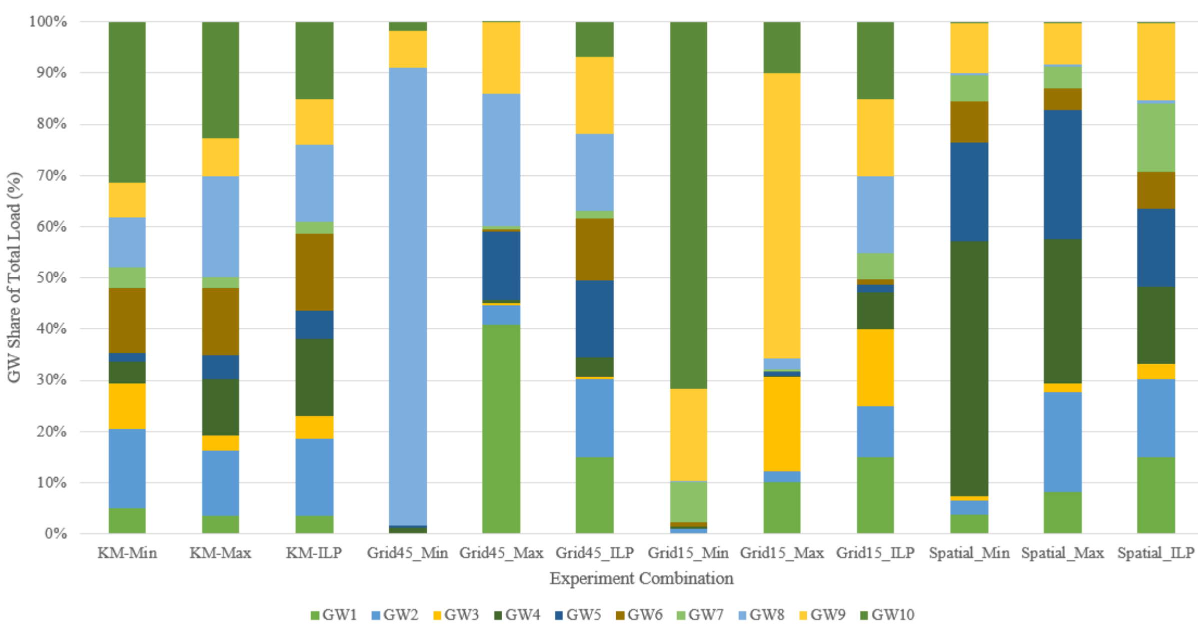


Fig. 9: GWs load distributions across all experiments

for the central vertical and horizontal cross sections (in bold) are in figure 10.

3) For each elevation profile, we compute the interdecile range (as specified by the ITS model [29]). This is to exclude occasional anomalies of exceptionally high or low altitude points.
4) The final $\triangle h$ value would be approximated as the the median value for the all interdecile ranges computed in the previous step. For our application, $\triangle h \approx 99$.

VI. Conclusions

In conclusion, our analysis examined various techniques for LPWAN wireless planning pipeline. We examined a network-aware approach using K-means clustering and a network-agnostic approach using Grid method. We also proposed an approach for network-agnostic wireless planning, the Spatial Algorithm, which showed competitive results compared to K-means method. Therefore, it is worth to continue investigation in the generalization scope of the Spatial Algorithm for various

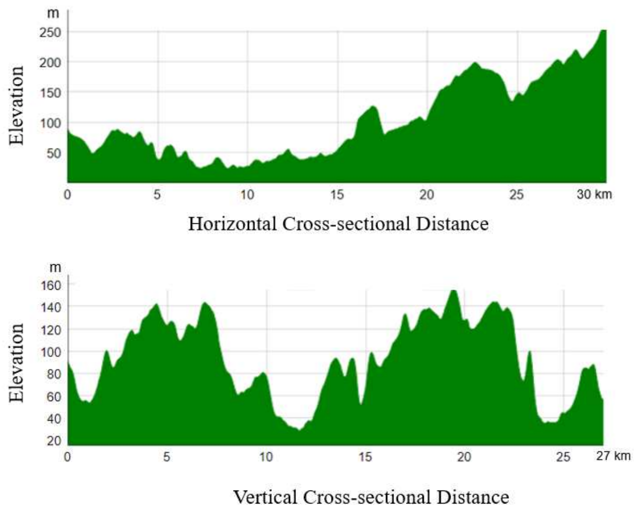


Fig. 10: Cross-sectional elevation profiles of Leeds

technologies and topographies such as indoor planning and industrial sensor networks. We show that following a principle of highest altitude for GW may undermine network performance severely and therefore may pressure network engineers to request more GWs to compensate for deteriorated QoS. However, using a computational method for GW placement like K-means or Spatial Method, has a potential of creating competitive network performance using just the same number of GWs, thus cutting down the financial costs of the network and increasing its sustainability.

REFERENCES

[1] J. Porras, A. Seffah, E. Rondeau, K. Andersson, and A. Klimova, "PERCCOM : A Master Program in Pervasive Computing and COMmunications for Sustainable Development PERCCOM : A Master Program in Pervasive Computing and COMmunications for," in *29th IEEE Conference on Software Engineering Education and Training*, no. APRIL, Dallas, United States. 2016., 2016.

[2] Research and Markets, "Low Power Wide Area Network (LPWAN) Market, Technologies, Industry Verticals, Regions, Competitive Landscape, End-User Analysis - Global Market Drivers, Opportunities, Trends, and Forecasts, 2016-2022," 2015. [Online]. Available: https://www.researchandmarkets.com/research/r6dzlv/low{_}power{_}wide

[3] LoRa-Alliance, "LoRa-Alliance Technology," 2018. [Online]. Available: https://www.lora-alliance.org/technology

[4] SigFox, "Sigfox Technology Overview," 2018. [Online]. Available: https://www.sigfox.com/en/sigfox-iot-technology-overview

[5] 3rd Generation Partnership Project (3GPP), "Standardization of NB-IOT completed," 2016. [Online]. Available: http://www.3gpp.org/news-events/3gpp-news/1785-nb{_}iot{_}complete

[6] European Large-Scale Pilots Programme, "IoT European Large-Scale Pilots Programme."

[7] G. Pasolini, C. Buratti, L. Feltrin, R. Verdone, and O. Andrisano, "Smart City Pilot Projects Using LoRa and IEEE802.15.4 Technologies," pp. 1–17, 2018.

[8] K. Mekki, E. Bajic, F. Chaxel, and F. Meyer, "A comparative study of LPWAN technologies for large-scale IoT deployment," *ICT Express*, 2018. [Online]. Available: https://doi.org/10.1016/j.icte.2017.12.005

[9] J. Petäjäjärvi, K. Mikhaylov, A. Roivainen, T. Hänninen, and M. Pettissalo, "On the coverage of LPWANs: Range evaluation and channel attenuation model for LoRa technology," *2015 14th International Conference on ITS Telecommunications, ITST 2015*, pp. 55–59, 2016.

[10] K. E. Nolan and M. Kelly, "An Evaluation Of Low Power Wide Area Network Technologies For The Internet Of Things," pp. 439–444, 2016.

[11] A. Augustin, J. Yi, T. Clausen, and W. Townsley, "A Study of LoRa: Long Range & Low Power Networks for the Internet of Things," p. 1466, 2016. [Online]. Available: http://www.mdpi.com/1424-8220/16/9/1466

[12] M. Bor, J. Vidler, and U. Roedig, "LoRa for the Internet of Things," USA, pp. 361–366, 2016. [Online]. Available: http://dl.acm.org/citation.cfm?id=2893711.2893802

[13] C. Goursad and J. M. Gorce, "Dedicated networks for IoT: PHY/MAC state of the art and challenges," *EAI Endorsed Transactions on the Internet of Things*, vol. 1, no. 1, pp. 1–12, 2015.

[14] M. Lauridsen, H. Nguyen, B. Vejlgaard, P. Mogensen, M. Sørensen, and N.-i. Lora, "Coverage comparison of GPRS , NB-IoT , LoRa , and SigFox in a 7800 km 2 area," pp. 2–6, 2017.

[15] M. Lauridsen, P. Mogensen, and S. Holst, "Coverage and Capacity Analysis of LTE-M and NB-IoT in a Rural Area," vol. 20, pp. 2–6, 2016.

[16] M. Al., "LP-WAN Problem statement." [Online]. Available: https://tools.ietf.org/html/draft-minaburo-lp-wan-gap-analysis-00

[17] M. Haenggi, "User point processes in cellular networks," *IEEE Wireless Commun. Letters*, vol. 6, no. 2, pp. 258–261, 2017. [Online]. Available: https://doi.org/10.1109/LWC.2017.2668417

[18] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," vol. 00, no. c, pp. 1–10, 2000.

[19] J.-s. Lee, S. Member, and W.-l. Cheng, "Fuzzy-Logic-Based Clustering Approach for Energy Predication," vol. 12, no. 9, 2012.

[20] H. Long, Y. Liu, X. Fan, R. P. Dick, H. Yang, and A. Arbor, "Energy-Efficient Spatially-Adaptive Clustering and Routing in Wireless Sensor Networks," 2009.

[21] S. Bandyopadhyay and E. J. Coyle, "An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks," vol. 00, no. C, pp. 1713–1723, 2003.

[22] E. J. Duarte-melo and M. Liu, "Analysis of Energy Consumption and Lifetime ˜ f Heterogeneous Wireless Sensor Networks," pp. 21–25.

[23] S. Basagni, "Distributed and Mobility-Adaptive Clustering for Multimedia Support in Multi-Hop Wireless Networks," 1999.

[24] M. Chatterjee, S. K. Das, and D. Turgut, "An On-Demand Weighted Clustering Algorithm (WCA) for Ad hoc Networks," pp. 1697–1701, 2000.

[25] K. Mikhaylov, J. Petäjäjärvi, and T. Hänninen, "Analysis of Capacity and Scalability of the LoRa Low Power Wide Area Network Technology."

[26] J. Huang, Y. Meng, X. Gong, Y. Liu, and Q. Duan, "A Novel Deployment Scheme for Green Internet of Things," *IEEE Internet of Things Journal*, vol. 1, no. 2, pp. 196–205, 2014.

[27] S. P. Lloyd, "Least Squares Quantization in PCM," pp. 129–137, 1982.

[28] H. S. W. Itsenhausen, "The Complexity of the Generalized Lloyd-Max Problem," pp. 255–256, 1982.

[29] G. A. Hufford, A. G. Longley, and W. A. Kissick, "A Guide to the Use of the ITS Irregular Terrain Model in the Area Prediction Mode," 1982.

[30] J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization Sign in Oxford Academic account Sign in via your Institution Short-term Access," pp. 1–12, 2018.

[31] S. Smale, "On the Average Number of Steps of the Simplex Method of Linear Programming," pp. 241–242, 1983.

[32] H. Rumsev, "The Various Theories on the Propagation," 1954.

[33] C. Phillips, "The Stability of the Longley-Rice Irregular Terrain Model for Typical Problems ; CU-CS-1086-11," 2011.

[34] G. Hufford, "The ITS Irregular Terrain Model, version 1.2.2 The Algorithm," 1982.

[35] Leeds City Council, "Street lights (unmetered)," 2017. [Online]. Available: https://data.gov.uk/dataset/street-lights-unmetered

APPENDIX

In this appendix, we show full visualization set of EDs GW assignment and EDs simulated RSSI for each of resulting network plan.

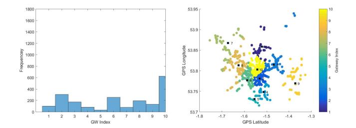
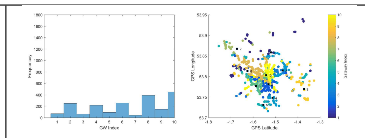
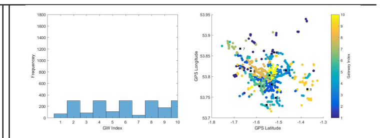
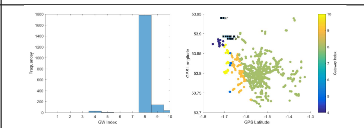
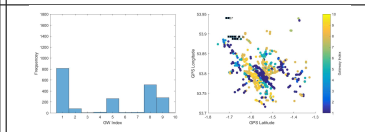
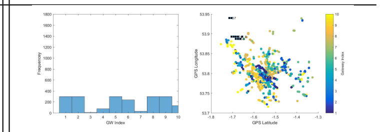
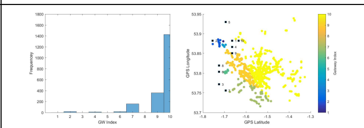
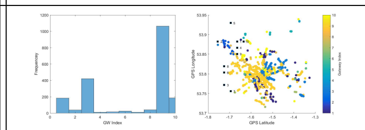
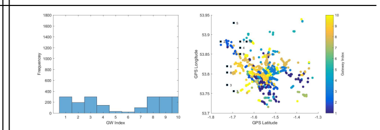
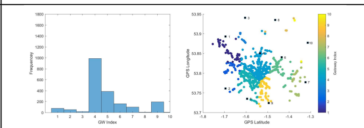
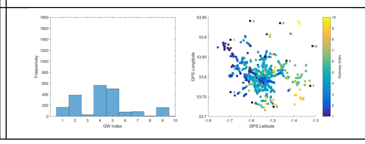
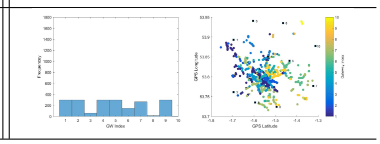
	Min Distance	Max RSSI	Constrained ILP Model
K-Means Clustering			
Grid 45			
Grid 15			
Spatial Method			

TABLE V: Visualization of gateways load distribution

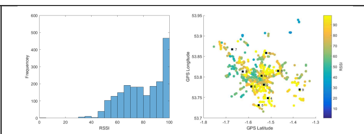
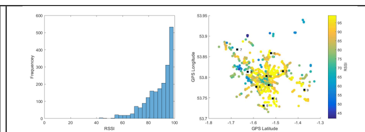
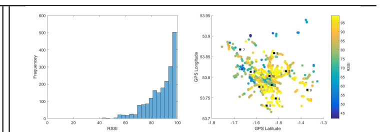
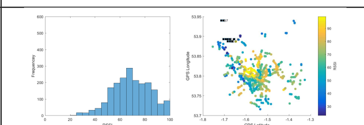
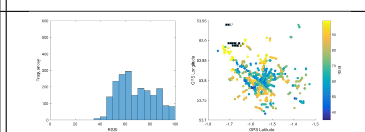
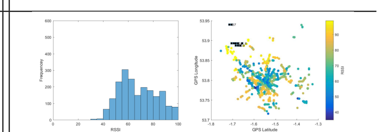
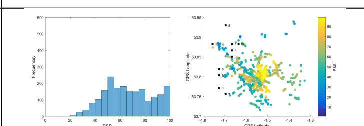
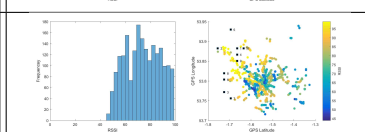
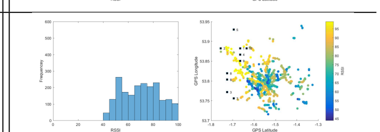
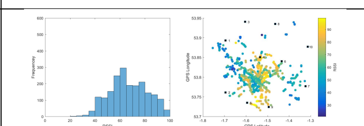
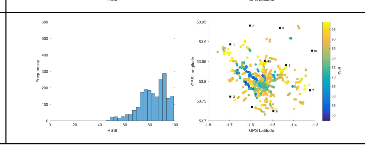
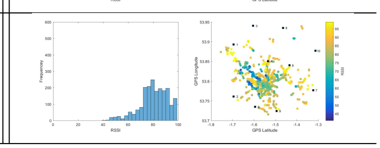
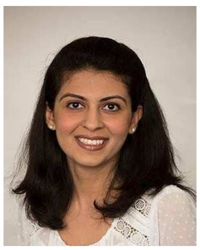
	Min Distance	Max RSSI	Constrained ILP Model
K-Means Clustering			
Grid 45			
Grid 15			
Spatial Method			

TABLE VI: Network RSSI distribution

Mina Rady is an alumni of Erasmus Mundus joint M.Sc. program in Pervasive Computing and Communications (PERCCOM) under scholarship award, at University of Lorraine (France), Lappeenranta University of Technology (Finland), and Luleå University of Technology (Sweden). He conducted this research as part of summer internship supervised by Dr. Syed. A. R. Zaidi at the University of Leeds. He completed his M.Sc. research in estimating real budget of LPWAN architectures under supervision of Professor Francis Lepage and Dr. Jean-Philippe Georges at Center for Automatic Control in Nancy (CRAN).



Maryam Hafeez received the B.Eng. degree in information and communications systems engineering from the School of Electrical Engineering and Computer Sciences (SEECS), National University of Science and Technology (NUST), Islamabad, Pakistan, and Ph.D. degree in Electrical Engineering from the University of Leeds, Leeds, U.K., in 2015. Currently, she is a Senior Lecturer at the School of Computing and Engineering, University of Huddersfield. Her research interest is towards design and analysis of protocols for next generation green intelligent wireless networks by employing tools from game theory and stochastic geometry. She has worked in the area of dynamic spectrum access in future wireless networks. She is also working towards applied research in for Industrial Internet of Things (IIoT).

Syed Ali Raza Zaidi is a University Academic Fellow (Assistant Professor) at the University of Leeds in the area of wireless communication & sensing systems. From 2013-2015, he was associated with the SPCOM research group working on the United States Army Research Lab funded project in the area of Network Science. From 2011-2013, he was associated with the International University of Rabat working as a Lecturer. He was also a visiting research scientist at Qatar Innovations and Mobility Centre from October-December 2013 working on QNRF funded project QSON. He received his Doctoral Degree at the School of Electronic and Electrical Engineering at Leeds and was awarded the G. W. and F. W. Carter Prize for best thesis and best research paper. He has published more than 90 papers in leading IEEE conferences and journals. From 2014-2015, he served as an editor for IEEE Communication Letters. He was also lead guest editor for IET Signal Processing Journal's Special Issue on Signal Processing for Large Scale 5G Wireless Networks. Currently, he is an Associate Technical Editor for IEEE Communication Magazine. Dr. Zaidi is EURASIP Local Liaison for United Kingdom and also General Secretary for IEEE Technical Subcommittee on Backhaul and Fronthaul networks. He is also an active member of EPSRC Peer Review College.