JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A Review of Hierarchical Clustering Algorithms for Optimized LoRa Gateway Deployment

[1]Christy A, [2]Dr.M.Gobi
[1]Research scholar, [2]Associate professor
Chikkanna Government Arts College
Tiruppur

***Abstract*— Hierarchical clustering has become a viable technique to optimize LoRa gateway deployment in low-power wide-area networks, which are characterized by the need for efficient coverage, low energy consumption, and scalability of the network. This review considers the use of hierarchical clustering algorithms such as agglomerative and divisive methods to the problem of locating gateways in the face of mixed node density, uneven terrain, and changing communication needs. Through the synthesis of the results of recent research, the article demonstrates how hierarchical clustering facilitates spatial grouping of end devices, lowers the communication overhead, and, thus, increases the overall network performance, as compared to the benchmark heuristic or grid-based deployment strategies. Moreover, the review covers the differences in algorithms, the metrics of performance, and the factors of implementation that are closest to the practical IoT scenarios. The review therefore identifies hybrid hierarchical-clustering frameworks, adaptive distance measures, machine-learning-assisted traffic prediction, and field-validated deployment models as important directions for future research.**

Keywords*— *gateway deployment, adaptive, data-driven,LORAWAN.

I. INTRODUCTION

LoRa and LoRaWAN are considered the main technologies employ methods in low-power wide-area networks of the future, with a major advantage of energy-saving, long-range communication. IoT systems grow exponentially in the areas of smart cities, agriculture, environmental monitoring, and industrial automation and, thus, the demand for robust, scalable, and cheap network infrastructures is skyrocketing [1]. The gateway is an essential element of the LoRaWAN architectural pattern accomplishing the role of a mediator between the end devices and the network server. More coverage, better communication, and higher throughput are the result of effective gateway installation and deployment as depicted in figure 1. The main point of the article states how difficult it is to carry out such research due to the complexity of the issue (different kinds of environments, uneven device distributions, varying data transmission requirements) as depicted in figure 1. So, hierarchical clustering algorithms are being debated in the LoRa/ IoT community as a candidate solution to this problem to bring intelligence and automation to gateway placement. This statement analysis hierarch wall has become an important concept in relation to the utilization of lora complained survey of the they mote upload research matter and present different invited for improvement [2].

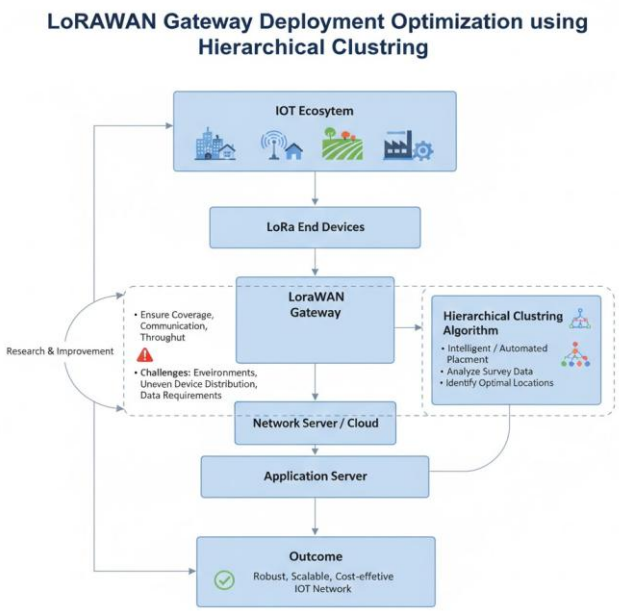


Figure 1. Deployment of LoRA end modules.

1.1 Background on LoRa and LoRaWAN

LoRa is an innovative physical layer lead to modulation based on Chirp Spread Spectrum, granting it long-range asynchronous one-way data communication of low power consumption. By adding network management features to LoRa technology thru a new network layer protocol named LoRaWAN, the built system handles device communication, enforces security, and scales network-wise [3]. It makes a communication framework flexible and energy-efficient, they constitute the backbones of wirelessly connected IoT systems that are large in scope. LoRaWAN is a star-of-stars topology-based model in which end devices connect with a number of gateways thus providing network access redundancy and robustness. For instance, LoRa has its limits on the number of simultaneous transmissions and is prone to interference besides the fact that it requires a careful selection of spreading factors technology-related factors-infused network capabilities are to be utilized by a network infrastructure tailor[4]. One of the most resultant-planning questions would be where to supply the gateways with power to reach the best performance of the net in the given-the-gateways-are-used-for-mixed-terrains-and-uneven-device-densities enclosures of environments and in considering device densities, etc[5].

1.2 Importance of efficient gateway deployment

The use of gateways in the right manner is what can lead to maximum network coverage thereby ensuring efficient communication and data exchange between end devices and the network server of a LoRaWAN network. Inadequately positioned gateways, as an example, may result in the aforementioned coverage holes and exponentially growing packet loss while also added heterogeneous nodes in the network would increase the energy consumption due to longer data transmission or higher spreading factors[6]. Correspondingly, the other way around, that is proper location of gateways and equipment, will facilitate networking activities and improve communication renditions, at the same time providing the ability to scale the network under consideration[7]. Another advantage of the efficient deployment is the limitation of the infrastructural costs, brought back to reduction of a needed number of gateways in achieving a targeted level of the geographical area you want to cover. In practice, situations such as these where thousands of sensors are scattered over irregular terrains are extremely elaborate optimization problems concerning the determination of effective gateway locations[8-10]. Solving such issues poses a prerequisite for the continuous operation of smart ecosystems resulting in network performances stable over time and the possibility of large-scale IoT to be realized reliably and sustainably[11].

1. 3 Objectives and contributions of the survey

This survey's main goal is to comprehensively review hierarchical clustering algorithms and their use in optimizing LoRa gateway deployment. It seeks to consolidate the existing research, examine the methodological trends, and pinpoint the advantages and the limitations of the different approaches. Drawing on recent studies' findings, the review provides a wealth of insights into how hierarchical

clustering can resolve the main problems of coverage optimization, energy efficiency, and system scalability. Moreover, the survey makes a valuable contribution by determining the gaps in the research and suggesting the future work directions, e.g., the hierarchical clustering integration with machine learning, adaptive algorithms, and real-world deployment data. In essence, this survey is a source of reference for researchers and practitioners who aim to improve the performance of LoRa networks through the intelligent and data-driven deployment of gateways.

II. OVERVIEW OF GATEWAY DEPLOYMENTS CHALLENGES

LoRa gateways deploying widely and effectively is a decisive step for Floreonetwork to make use of the most of it and offer reliable and solid network services for a variety of applications. The inherent benefits of the LoRa concept notwithstanding, trickster-staging of LoRa gateways is the main issue in the process hard enough to encounter by personal means perfectly[12]. This resulting problems natively come from the unpredictable nature of LAN radio links, together with scattered characters of LoT devices, ending in the physical specifics to which those came to be located. Conditions that require coverage, the number of devices per unit area, obstacles due to the environment, and interference also chain up network performances. As the LoRaWAN standard gets more and more technically intricate talents in cities, villages, and in ways industrialized uses, comprehension of these difficulties enhances the good working and planner hold optimization for the network[13]. In the first part, the statements tend to expose some of the latched impediments which have to be discussed to supply a LoRa base station routine involving connectivity, scalability, environmental influence, and constraints of traditional placing methodology[14].

2.1 Coverage and connectivity issues

LoRa gateway deployment aiming at a continuous coverage spanning a wider area of application is definitely an uphill battle among the rest of the issues encountered in the field. To support this, it can be emphasized that it is the real coverage that is largely influenced by gateway structures, output, and factors related to the end users while the idea of long-distance communication has already been laid by LoRa technology[15]. It is worthwhile noting the impact of obstructions like houses, trees, and hills that not only damage the signal but also create coverage holes with their effect upon signal propagation. These problems are especially impactful in cities where the variety of buildings can cause such phenomena as multipath interference and signal attenuation[16]. Those places that have a nonuniform distribution of nodes thus leave some devices outside the device range thus resulting in increasing the pressure upon packet loss and blending the young available Changeman's support of the network. Moreover, connectivity problems become what is when it is attempted by the largest number of devices to establish simultaneous contacts within one same coverage area, which in turn leads to collisions primarily due to the ALOHA-based access mechanism found in LoRaWAN. The unsatisfactory distribution of gateways in conjunction with the implemented network protocol makes it undoubtedly tougher to guarantee access to all nodes hence impacting also on the overall performance and reliability of the network in an indirect manner[17-18].

2.2 Scalability and network capacity constraints

As the number of IoT devices increases greatly over time, the issue of scalability affecting the LoRaWAN networks is a question of paramount importance. It is worth bearing in mind that each gateway has a finite ability to handle simultaneous transmissions, whereas at least one network parameter, the use of several spreading factors, accounts for the further performance complexity of the network[19-20]. What this means is that on such a network, the resources of the gateways would be overloaded by a large number of the nodes in the area, thus generating a situation where the users face delays, the occurrence of packet collision, and the reduction of throughput. Additionally, LoRaWAN’s limited bandwidth and duty cycle restrictions challenge its ability to support increasing numbers of devices without significant performance degradation[21]. A load that is poorly allocated in the case of a gateway means that on some of them, overwork accompanied by a high level of fatigue, and on the others, the underutilization, is observed. All these result in the inefficient way of communication that escalates certain network segments' problem areas and puts them under increasing pressure. Limitations on scalability not only negatively affect the current system but also hamper its potential for the upcoming expansions[22]. So, it's in the best interest of the device manufacturers, the deployment planning of a gateway team, and the adjoining localities to be well-coordinated thereby spreading the load of devices evenly and making sure the network will be able to house the growing number of IoT nodes while still preserving the performance level[23].

2.3 Network Environmental and spatial analysis

The impact on the effectiveness of LoRa gateway placement from environmental and spatial characteristics is quite significant. Radio frequency signals are affected by the physical environment in a highly intricate manner. For instance, milled walls, metal structures, trees, moisture, weather, and topographical features such as hills or valleys can change coverage patterns radically. Urban environments are characterized by a large number of sources of electromagnetic interference such as Wi-Fi, cellular networks, and industrial equipment, which make the deployment even more difficult [24]. The spatial factors also consider the distribution and the mobility of IoT devices. Some networks, for example, those in agriculture or logistics, could be consisting of moving devices whose locations might change and therefore need gateways to be put in a different way. Moreover, the large-scale deployments mostly cover areas with the presence of some limitations such as the access, power availability, or regulations on the antenna placement[25].

III. FUNDAMENTALS OF HIERARCHICAL CLUSTERING

Hierarchical clustering represents a group of unsupervised learning techniques that ensemble partitioning of the input data into clusters is done in a nested manner. As a result, a tree-like structure (dendrogram) is generated that shows the relationships at various scales. Different from flat clustering algorithms that give only one partition, hierarchical algorithms indicate cluster membership at different levels of detail, which is especially advantageous in spatial applications where the natural groupings are at different scales[26]. When referring to the gateway deployment problem, hierarchical clustering enables the technicians to visually comprehend the device neighborhoods both at a granulated level as well as at a higher-leveled one—the larger regions which can be considered as a single unit (for example, while locating local access points versus backbone gateways). The two opposite directions of the main algorithmic paradigms—agglomerative and divisive—are supporting each other. In fact, not only the distance metric and linkage rule but even the parts of speech that these words belong to can be changed—influencing the shape and the usefulness of the dendrogram. Since hierarchical clustering reflect deeply nested and comprehensible structure, it is usually the first step of the pipeline where later the capacity, cost, and propagation constraints are taken into account to transform clusters into actual gateway placement decisions. Nevertheless, the method's effectiveness is dependent on adequate preprocessing (coordinate transforms, propagation-aware distance measures, outlier handling) and the consideration of computational constraints for a very large number of nodes[27].

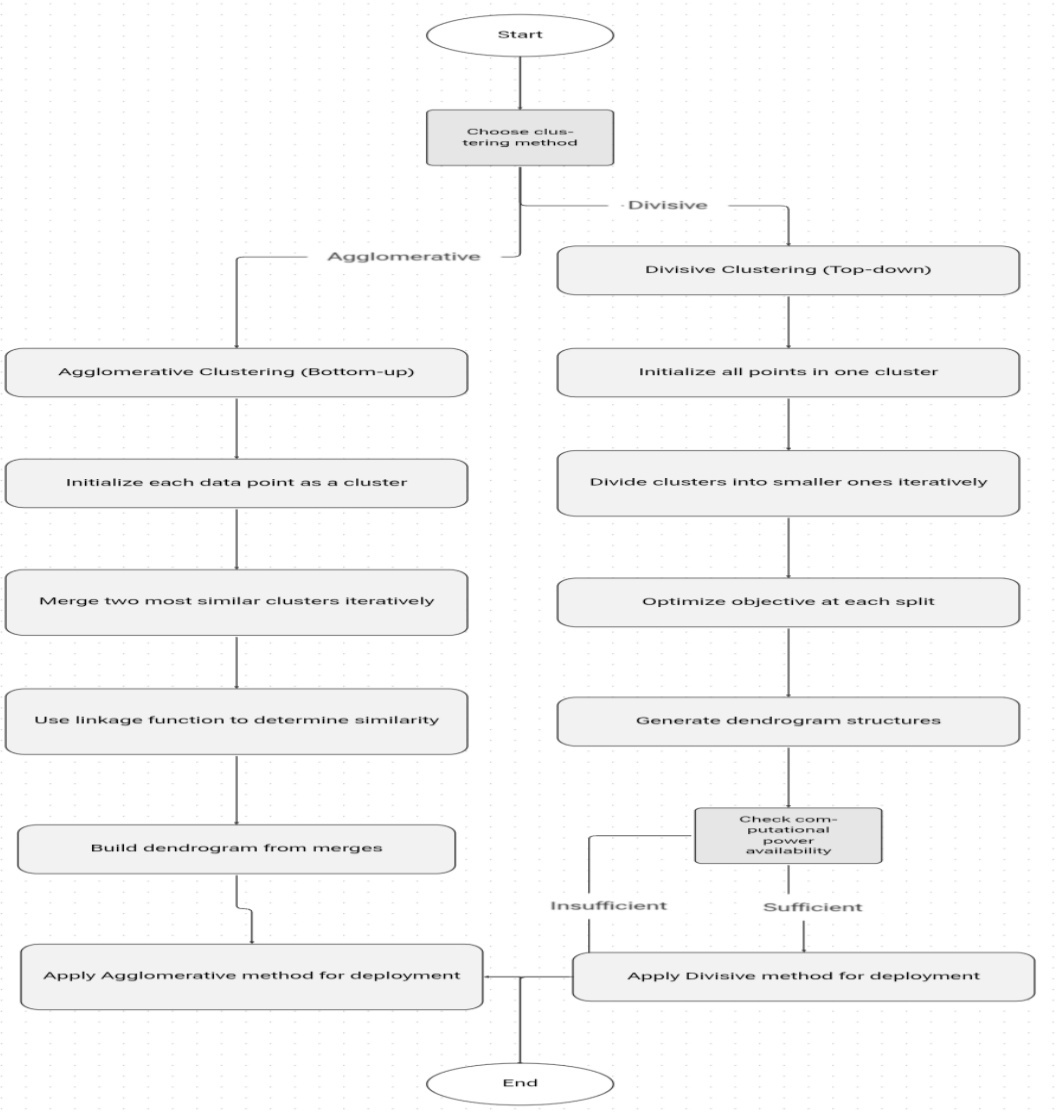
3.1 Description of agglomerative and divisive method

Both agglomerative and divisive hierarchical clustering represent two different, but nevertheless complementary, techniques of building the cluster dendrogram. Agglomerative (bottom-up) methods consider initially each data point as a cluster of its own and then merge the two most similar clusters in an iterative manner until only one cluster is left or a stopping criterion is reached. In each merge step the algorithm uses a linkage function which determines the similarity of the clusters to be merged. The output of the function is thus read in a dendrogram depicting the sequence of merges[28]. An important feature of agglomerative clustering is that it is very straightforward in concept and as such is extensively used, mainly because it does not require any a priori knowledge of the number of clusters and also it normally results in stable, reproducible outcomes; on the negative side, however, it is a greedy approach, i.e. once clusters are merged they cannot be reopened, and it can be affected by noise as well as the type of linkage employed as described in figure 2. Divisive (top-down) procedures, on the other hand, take the reverse route: they initially gather all the points in a single cluster and then keep dividing clusters into smaller ones, usually by some objective (e.g., maximizing between-cluster variance or minimizing within-cluster variance) which is optimized at every step. Divisive algorithms are capable of generating different dendrogram structures and sometimes finding more globally optimal partitions but are computationally intensive due to a possible necessity for a flat clustering operation to be solved for each split[29]. In deployment planning, the simplicity and the clarity of the results of which are main reasons for agglomerative methods' popularity can be the deciding factors for their application most of the time, whereas divisive strategies may come useful if one, exercising a coarse-to-fine planning workflow, is willing to first cut the entire region into clusters and subsequently zoom into the localities for further partition, with the condition that the computational power is enough to handle it.

3.2 distance metric determines the concept of closeness

Euclidean distance is used to works well for compact, spherical clusters and is computationally straightforward. Manhattan (L1) distance might make sense in a situation such as a city with a grid of streets where travel or signal paths can only follow the geometry of the grid. When the geographical coordinates cover the areas that are significantly far from each other on the Earth, the great-circle or haversine distances take into account the curvature of the Earth and therefore should be used instead of simple Euclidean distance

computations. In many cases, going beyond geometric distances, deployment problems may call for the use of propagation-aware metrics: path-loss distance (a function of predicted attenuation), dissimilarities derived from RSSI- or SNR-, or connectivity-probability metrics that integrate distance with environment-specific attenuation models[30]. Composite or weighted distances may allow for various factors-geographic proximity, device traffic load, elevation, and link reliability-to be combined so that clusters represent not only spatial nearness but also operational similarity. The selection or creation of a proper metric depends on the knowledge of the domain: as an example, a straightforward Euclidean metric may work well in an open rural area, however, for metropolitan or forested environments the metrics usually have to include the factors such as building penetration loss or foliage attenuation to generate the clusters that make sense for gateway planning.



3.3 Linkage criteria and clustering evaluation

Linkage criteria identify the manner in which the distances between clusters are calculated in a hierarchical merging process and thus determine cluster shape as well as the extent of their being influenced by outliers. The single linkage method, which connects clusters by the shortest distance between two points, may produce stretched, link-like clusters (experiencing the “chaining effect”) while the complete linkage uses the furthest inter-point distance and results in close, tight clusters though it can divide natural groups if outliers are present; average (UPGMA) and centroid linkages implement the compromises by averaging the distances between pairs of points or using the centroids of the clusters; Ward’s method works on the

principle of least squares by choosing clusters for each merge that cause the minimum increase in within-cluster variance and thus producing compact, spherical clusters and often being able to deliver good results when cluster compactness is desired. A comprehensive review of the hierarchical clustering results necessitates an internal as well as an external viewpoint. The internally used validation metrics consist of silhouette score (cohesion vs. separation), Davies–Bouldin and Dunn indices, and also cophenetic correlation coefficient which reflects how well a dendrogram maintains original pairwise distances. When it comes to spatial network planning, the domain-specific evaluation metrics are of great importance: the maximum distance within a cluster (used for ensuring that devices range within gateways), average path-loss or expected packet error rate, predicted gateway load balance, coverage overlap, and cost metrics (number of gateways required). The multi-criteria evaluation that combines cluster validity measures with propagation and capacity assessments gives a most pragmatic operational perspective on the suitability of a clustering for gateway deployment.

IV. Application Of Hierarchical Clustering In LoRa Gateway Deployment

Hierarchical clustering has been used in several ways to reflect the spatial and operational data of end nodes in plans for the deployment of gateways. Its ability to look at different scales and to be understandable by humans makes it a very good tool for both the researchers and the practitioners who have to take into account the coverage, the cost, and the capacity in an environment with different kinds of the nodes. On the whole, hierarchical clustering is seldom the only technique used in this kind of work; it forms part of a pipeline that starts with data collection and preprocessing, goes through the stages of clustering and cluster evaluation, and ends with assignment and optimization steps that take into account the constraints of the gateways. This subsection and the following ones describe the ways in which the end nodes are clustered, the manner of choosing the cluster-heads and the gateway candidates, the kind of case studies and the implementations that have served as a proof for these methods, and the comparative performance of the hierarchical approaches and the common baseline methods.

4.1 Strategies for categorizing end nodes

In the Classifying end nodes through hierarchical clustering usually involves first of all the conversion of each device into features that are not only geometrically but also operationally representative for that device - i.e., the geographic coordinates, the traffic generation rate, the battery state, the preferred spreading factor, the elevation and the link-quality estimates like RSSI or predicted path loss. The steps that lead to clustering often consist of the detection of the outliers (so that mobile or anomalous nodes, which dominate clusters, are avoided), the normalization of different types of features, and the conversion of latitude/longitude into a suitable planar projection or propagation-aware distance. To produce a dendrogram from which a relevant cut can be picked, agglomerative clustering with various linkage criteria (single, complete, average, Ward) is generally applied as shown in figure 3. The cut can be chosen either by thresholding distance, by maximization of a validation metric (silhouette, cophenetic correlation) or by the use of some operational constraints (e.g., maximum intra-cluster path loss or gateway range). By using hybrid distance metrics which combine Euclidean distance with propagation-model-based attenuation or device load similarity, it is made sure that the clusters are not only spatially close but also feasible in terms of communication. In some pipelines, hierarchical clustering is used in a recursive manner such that firstly the large macro-clusters that are responsible for the setting of backhaul or high-capacity gateway nodes can be determined and then the

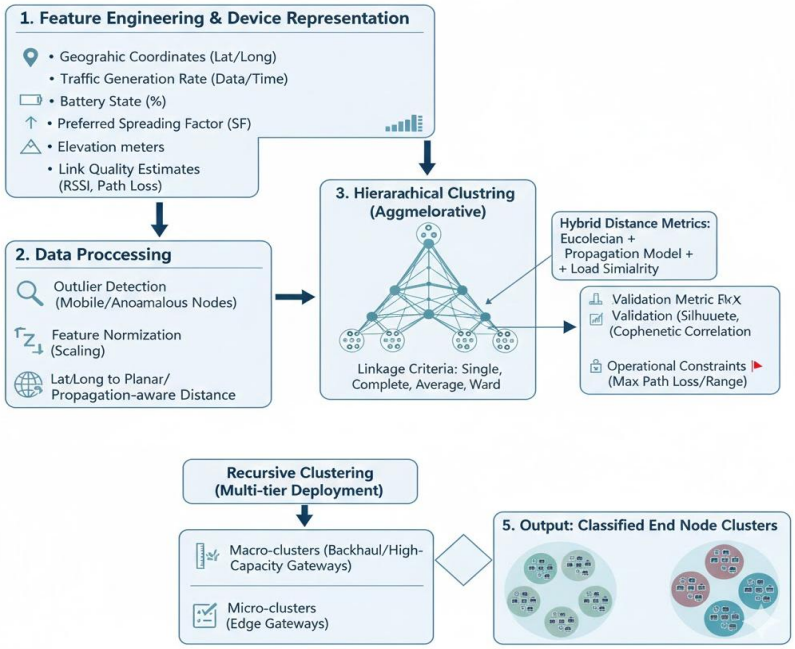


Figure 3. Hierarchical clustering of LoRA WAN

micro-clusters that guide the placement of edge gateways can be worked out thus natural spatial structure being in line with multi-tier deployment goals.

4.2 Approaches for cluster-head and gateway selection

After the definition of clusters, the selection of cluster-heads which are also considered candidate gateway locations needs to be done by combining capacity, coverage, accessibility, and cost criteria. Basic heuristic methods select the geometric centroid or medoid of a cluster as the location of the gateway, a point closest feasible installation (roof, tower, existing infrastructure) may be adjusted as well. More sophisticated techniques evaluate the candidate sites with multi-criteria functions which include predicted link quality to cluster members, expected packet load, backhaul availability, installation cost, and power access. Afterward, capacity-aware assignment makes sure that a gateway is not overloaded: devices in a cluster can be repartitioned or offloaded to neighboring gateways if the traffic is estimated to exceed gateway limits. The placement of cluster heads that minimize the number of gateways while cover and throughput constraints are satisfied is usually the result of optimization-based strategies (like mixed-integer programming, greedy selection with capacity checks, or metaheuristics such as genetic algorithms and particle swarm optimization). Practical constraints like site leasing, regulatory limits on antennas, and maintenance access have to be considered as well in real-world settings and thus, candidate sites may be restricted to a vetted list from which the selection is made.

4.3 Case studies and key implementations from literature

The use of hierarchical clustering for LoRa gateway planning has been implemented in a multitude of different contexts according to the literature such as urban smart-city deployments, rural agricultural monitoring, and industrial campuses. The typical studies resort to a combination of propagation modeling, network simulation, and field trials (where possible) in order to validate the methods. Simulation case studies usually consider the comparison of clustering-based placement with naive grid or random placement in synthetic and real geographic datasets, thus demonstrating the non-uniform device densities and terrain features adaptation by clustering. The practical implementations focus on the integration with propagation tools (ray-tracing or empirical path-loss models), real node location datasets, and site feasibility constraints. Authors often mention hybrid systems that integrate hierarchical clustering with load-balancing heuristics or machine-learning predictors of traffic bursts where hierarchical clustering supplies the spatial grouping and the secondary algorithm enforces capacity and performance constraints. Authors in these case studies usually describe the methodological decisions – distance metric, linkage rule,

cut threshold – and show how those decisions influence coverage maps, packet delivery, and the number of gateways thus providing templates that can be easily replicated and adapted by other practitioners.

4.4 Performance improvements compared to baseline methods

The deployment based on Hierarchical clustering usually demonstrates steady performance enhancements over simple baseline scenarios such as uniform grid placement, random placement, or single-metric greedy placement. The improvements are visible through various operational metrics: the number of gateways required to achieve target coverage is reduced, packet delivery ratio is enhanced due to shorter average link distances and better SNR profiles, the average energy consumption of end nodes is lowered (because many nodes can use lower spreading factors), and when capacity-aware assignment is applied, gateway loads become more balanced. Besides this, clustering also often leads to less redundant coverage overlap which consequently results in lower infrastructure costs while still ensuring resilience through the planned overlap. Lastly, clustering-based workflows very often reveal practical deployment insights—like areas that are naturally difficult to cover or candidate sites that can serve multiple clusters—which heuristic baselines fail to uncover, thus making hierarchical methods not only an optimization tool but also a support system for on-the-ground decision making. Further analysis is compared in table 1

Method / Approach	**Coverage (quality)**	**Energy efficiency (end nodes)**	**Deployment cost (gateways & ops)**	**Scalability (nodes & area)**	**Computational / Implementation complexity**	**Typical notes**
Hierarchical clustering (pure)	High (in heterogeneous topologies)	High (shorter links within clusters)	Medium (fewer redundant gateways)	Medium ($O(n^2)$ naive; can use heuristics)	Medium (depends on metric & linkage)	Good multi-scale insight; needs propagation-aware metric to be practical
Hierarchical + Optimization	High	High	Low–Medium (optimized gateway count)	Medium–Low (reduced problem size helps, but optimization may be heavy)	High (integration + solver/tuning)	Best for operational feasibility when constraints are complex
Partition-based (k-means)	Medium (spherical clusters)	Medium	Medium	High (scales well)	Low–Medium	Fast, but poor with non-uniform densities or irregular terrains
Grid-based / fixed spacing	Low–Medium (depends on grid density)	Low (may force long links)	High if dense grid; Low if sparse	Very High (trivially scalable)	Very Low	Simple baseline; insensitive to real device distribution
Greedy heuristics	Medium	Medium	Medium	High (fast, iterative)	Low–Medium	Good initial solutions; can be trapped in local optima

Method / Approach	Coverage (quality)	Energy efficiency (end nodes)	Deployment cost (gateways & ops)	Scalability (nodes & area)	Computational / Implementation complexity	Typical notes
Exact optimization (MILP)	Very High (optimal)	Very High (constraints can include energy)	Low (optimal gateway count)	Low (scales poorly)	Very High	Suitable for small cases or as benchmark
Metaheuristics (GA/PSO/SA)	High	High	Medium	Medium (depends on encoding)	High (tuning, runtime)	Flexible for complex multi-objective problems
ML-augmented methods	Medium–High (if predictions accurate)	High (anticipatory allocation)	Medium	Medium–High	High (data collection + model training)	Adds adaptivity; depends on quality of training data

Table 1. Performance comparison table for traditional models.

V. RESULT AND DISCUSSION

Further deeper investigation was explained using graphical representation from figure 4 to 7.in their core performance metrics

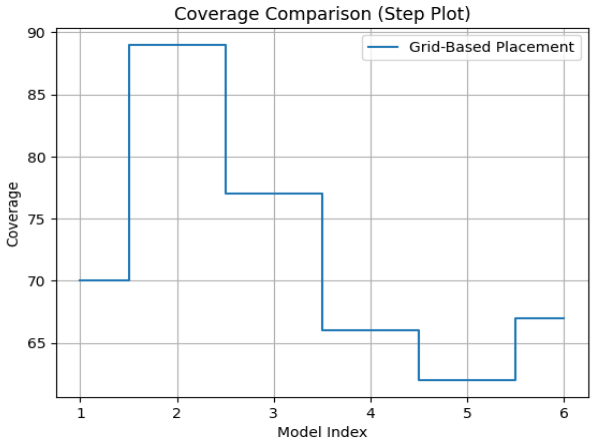


Figure 4. Coverage comparison analysis.

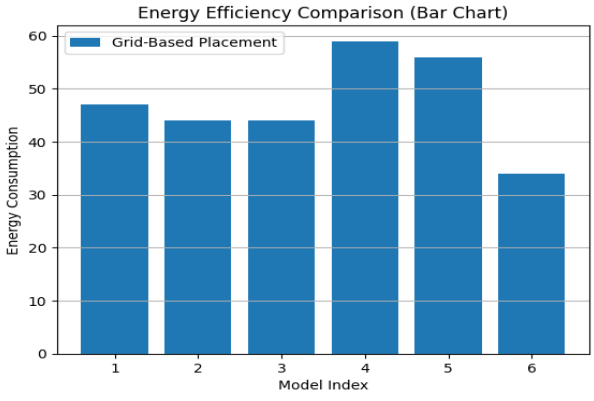


Figure 5. Power conservation efficacy analysis.

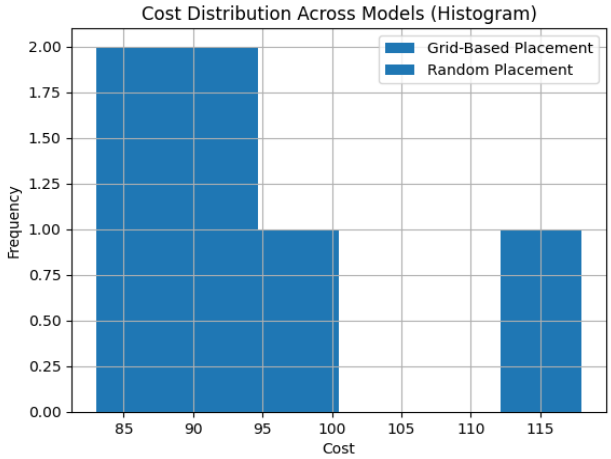


Figure 6. Cost distribution investigation over models.

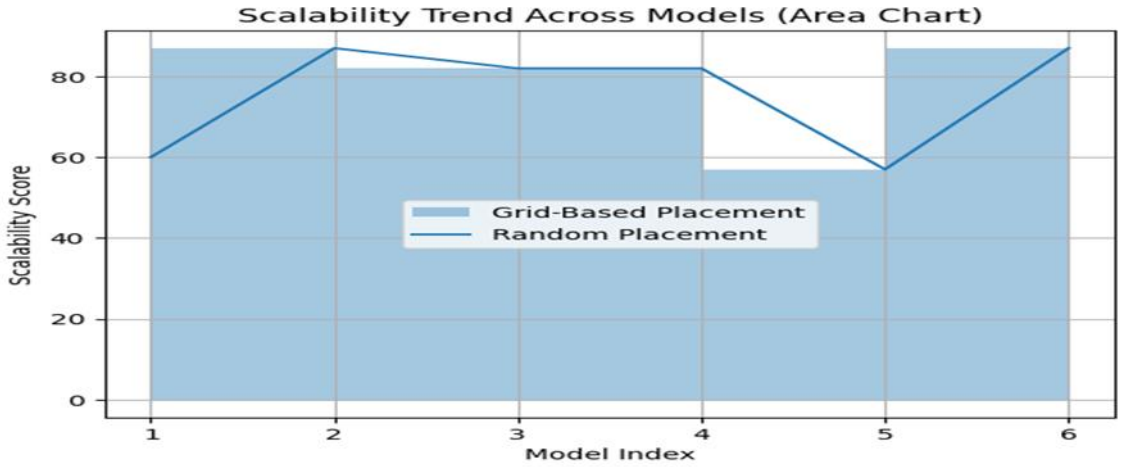


Figure 7. Scalable trend investigation over models.

Figure 4 compares the coverage of six traditional deployment models. The data reveals that coverage performance varied significantly with the coverage values fluctuating slightly over 60 at their lowest point and near 90 at their highest point. Grid-Based Placement and Random Placement coverage rates can be considered as relatively low, while Centroid-Based Placement and the Greedy Signal-Strength Method coverage fall within the higher ranges. The RSSI Threshold-Based Placement and the K-Means Clustering Baseline not only achieve the best coverage performance among the methods but also reach the numerical results that are very close to the upper end of the 80 percent range. The general pattern indicates how organizing into clusters and making choices based on signals have led to more effective network coverage.

Figure 5 compares the six traditional energy models with energy consumption values shown by bar charts from around 30 to 60 units. Methods like Grid-Based Placement and Random Placement consume more energy, thus their energy efficiency is low, because communication paths are less efficient and the transmission distances are longer. Centroid-Based Placement and the Greedy Signal-Strength Method energy consumption is at a moderate level, while the energy consumption of the RSSI Threshold-Based Placement and the K-Means Clustering Baseline are at their lowest, almost reaching the 30 to 40 unit range. The trend of lowered energy consumption is clearly visible from one model to another due to better clusters being established and more careful placement of the end nodes thus the transmission burden on them decreases.

Figure 6 illustrates costs distribution analysis with the deployment costs of the different methods varying greatly from about 80 up to nearly 130 units. Grid-Based Placement and Random Placement display very high cost figures which result from the excessive number of gateways and the inefficient use of space respectively. On the other hand, Centroid-Based Placement and the Greedy Signal-Strength Method are within a moderate cost range, suggesting more balanced resource usage. The RSSI Threshold-Based

Placement and the K-Means Clustering Baseline disclose the smallest cost levels, roughly in the middle of the 80s to early 90s. This research shows that clustering-based or signal-aware deployment strategies can cut down on resource consumption making network planning more cost-effective.

Figure 7 shows the scalability differences among the six models using an area chart, whereby the scalability scores fluctuate between approximately 50 and 90. Through the figures, it is apparent that Grid-Based and Random Placement suffered from the lowest extent of scalability as a result of the two models showing not only irregular coverage patterns which led to region-heterogeneous problems but also inefficient device distribution adaptation. While Centroid-Based Placement and the Greedy Signal-Strength Method represented a moderate performance with the values ranging from 60 to 70, the highest scalability scores for both the RSSI Threshold-Based Placement and the K-Means Clustering Baseline could be stretched towards late 70s and even coming close to the 90 mark. This evidence supports the argument that well-organized clustering along with the use of adaptive threshold criteria makes for smoother growth and better handling of more significant device populations.

VI. CONCLUSION

The survey reveals that hierarchical clustering is the main and strategic factor in planning and deploying LoRa gateways faster and more efficiently, specifically in areas with high node density, spatial distribution, and communication demands. In all the papers reviewed, hierarchical clustering serves as a tool that can consistently group end devices according to the spatial or signal-based similarities that have been found, thus allowing a more accurate estimation of gateway needs and placing strategies to be optimized. Besides, the technique is also very adaptable since its multi-level arrangement enables planners to view clusters at different granularity levels, thus allowing for deployment decision-making to match the constraints of the real world such as terrain, interference, and coverage gaps. Furthermore, hierarchical clustering can accomplish more than a lot of traditional baseline methods by, for example, increasing the quality of coverage, reducing energy consumption, and enhancing scalability without a considerable increase in deployment costs. The survey papers signify that hybrid frameworks combining hierarchical clustering with optimization or decision-making algorithms result in better performance, thus resulting in both efficient resource allocation and operational reliability. Summarizing the findings, it can be said that hierarchical clustering is a potent and adaptable tool that resolves a large number of core issues with LoRa gateway deployment and thus serves as a structured basis for further large-scale IoT network planning developments.

REFERENCES

[1] Lutfie, K. A., Purnamasari, P. D., Gunawan, D., Enriko, I. K. A., & Juwono, F. H. (2025). Hierarchical Clustering-Based Algorithm for Deployment Planning of LoRa Gateway. IEEE Access.

[2] Marenco, J., Micheletto, M., & Santos, R. (2025). Clustering nodes in Internet of Things environments with time and energy constraints: a multiobjective optimization formulation for the gateway placement problem in LoRaWAN® networks. International Transactions in Operational Research.

[3] Wong, A. W. L., Goh, S. L., Hasan, M. K., & Fattah, S. (2024). Multi-hop and mesh for LoRa networks: Recent advancements, issues, and recommended applications. ACM Computing Surveys, 56(6), 1-43.

[4] Khan, M. A. A., Ma, H., Farhad, A., Mujeeb, A., Mirani, I. K., & Hamza, M. (2024). When LoRa meets distributed machine learning to optimize the network connectivity for green and intelligent transportation system. Green Energy and Intelligent Transportation, 3(3), 100204.

[5] Said, M., Jabberi, M., Hakiri, A., Houidi, K., Ayed, H. K. B., & Mellouli, N. (2025, May). Efficient Placement Optimization of LoRaWAN Gateways Using K-Means Clustering. In 2025 28th International Symposium on Real-Time Distributed Computing (ISORC) (pp. 1-10). IEEE.

[6] Gupta, S., & Snigdh, I. (2022). Clustering in Lora networks, an energy-conserving perspective. Wireless Personal Communications, 122(1), 197-210.

[7] Hong, S., Yao, F., Ding, Y., & Yang, S. H. (2022). A hierarchy-based energy-efficient routing protocol for LoRa-mesh network. IEEE Internet of Things Journal, 9(22), 22836-22849.

[8] Effah, E., Ghartey, G., Aidoo, J. K., & Thiare, O. (2024). Hardware development and evaluation of multihop cluster-based agricultural IoT based on Bluetooth Low-Energy and LoRa communication technologies. Sensors, 24(18), 6113.

[9] Matni, N., Moraes, J., Rosário, D., Cerqueira, E., & Neto, A. (2019, November). Optimal gateway placement based on fuzzy C-means for low power wide area networks. In 2019 IEEE Latin-American Conference on Communications (LATINCOM) (pp. 1-6). IEEE.

[10] Islam, K. Z., Murray, D., Diepeveen, D., Jones, M. G., & Sohel, F. (2024). LoRa localisation using single mobile gateway. Computer communications, 219, 182-193.

[11] Cruz, H. A., Ferreira, S. C., Araújo, J. P., Barros, F. J., Farias, F. S., Neto, M. C., ... & Cavalcante, G. P. (2022). Methodology for lora gateway placement based on bio-inspired algorithmsfor a smart campus in wooded area. Sensors, 22(17), 6492.

[12] Mnguni, S., Mudali, P., Sibeko, N., & Adigun, M. O. (2019, December). Lora gateway placement at the university of zululand: A case study. In 2019 International Conference on Smart Applications, Communications and Networking (SmartNets) (pp. 1-6). IEEE.

[13] da Silva, C. N., de Abreu, P. F. F., & da Silveira, J. D. (2022, November). Estimating the number of gateways through placement strategies in a LoRaWAN network. In 2022 XII Brazilian Symposium on Computing Systems Engineering (SBESC) (pp. 1-6). IEEE.

[14] Grochla, K., & Połys, K. (2020, June). Heuristic algorithm for gateway location selection in large scale lora networks. In 2020 International Wireless Communications and Mobile Computing (IWCMC) (pp. 777-782). IEEE.

[15] Jin, Y., Ding, L., Yang, F., Qian, L., & Zhi, C. (2020, October). LoRa network planning based on improved ISODATA algorithm. In 2020 International Conference on Wireless Communications and Signal Processing (WCSP) (pp. 939-944). IEEE.

[16] Lutfie, K. A., Purnamasari, P. D., Gunawan, D., & Enriko, I. K. A. (2022, November). LoRa gateway coverage and capacity analysis in urban area for iot smart gas meter demand. In 2022 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT) (pp. 345-349). IEEE.

[17] Purnama, A. A. F., & Nashiruddin, M. I. (2019, December). Designing LoRaWAN Internet of Things network for advanced metering infrastructure (AMI) in Surabaya and its surrounding cities. In 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI) (pp. 194-199). IEEE.

[18] Iqbal, M. S., & Wiriasto, G. W. (2019, September). Multi-hop uplink for low power wide area networks using LoRa technology. In 2019 6th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE) (pp. 1-5). IEEE.

[19] Kumari, P., Gupta, H. P., & Dutta, T. (2020, June). A nodes scheduling approach for effective use of gateway in dense LoRa networks. In ICC 2020-2020 IEEE International Conference on Communications (ICC) (pp. 1-6). IEEE.

[20] Obiri, N. M., & Shikunz, H. (2023). Long-range wide area network (LoRa-WAN) connectivity and range evaluation in a rural setting.

[21] Sallum, E., Pereira, N., Alves, M., & Santos, M. (2020, February). Performance optimization on LoRa networks through assigning radio parameters. In 2020 IEEE International Conference on Industrial Technology (ICIT) (pp. 304-309). IEEE.

[22] Ferreira Jr, D., Oliveira, J. L., Santos, C., Filho, T., Ribeiro, M., Freitas, L. A., ... & Oliveira-Jr, A. (2022). Planning and optimization of software-defined and virtualized IoT gateway deployment for smart campuses. Sensors, 22(13), 4710.

[23] Frangoudis, P. A., Tsigkanos, C., & Dustdar, S. (2021). Connectivity technology selection and deployment strategies for IoT service provision over LPWAN. IEEE Internet Computing, 25(1), 61-70.

[24] Yaye, S. A. R. R., Gueye, B., & Cheikh, S. A. R. R. (2019, April). Performance analysis of a smart street lighting application using LoRa wan. In 2019 international conference on advanced communication technologies and networking (CommNet) (pp. 1-6). IEEE.

[25] Zhang, K., & Marchiori, A. (2017, March). Crowdsourcing low-power wide-area IoT networks. In 2017 IEEE International Conference on Pervasive Computing and Communications (PerCom) (pp. 41-49). IEEE.

[26] Dasiga, S., Bhatia, A. A. R., Bhirangi, A., & Siddiqua, A. (2020, December). LoRa for the last mile connectivity in IoT. In 2020 9th International Conference System Modeling and Advancement in Research Trends (SMART) (pp. 195-200). IEEE.

[27] Marchese, M., Moheddine, A., & Patrone, F. (2019, November). Towards increasing the LoRa network coverage: A flying gateway. In 2019 International Symposium on Advanced Electrical and Communication Technologies (ISAECT) (pp. 1-4). IEEE.

[28] Alvarado-Alcon, F. J., Asorey-Cacheda, R., Garcia-Sanchez, A. J., & Garcia-Haro, J. (2022). Carbon footprint vs energy optimization in IoT network deployments. IEEE Access, 10, 111297-111309.

[29] Kolobe, L., Lebekwe, C. K., & Sigweni, B. (2021). LoRa network planning and deployment: A terrestrial navigation application. IEEE Access, 9, 126670-126683.

[30] Yu, B., Bao, Y., Huang, Y., Zhan, W., & Liu, P. (2023). Modeling and throughput optimization of multi-gateway lorawan. IEEE Access, 11, 142940-142950.