5G Network Coverage Hole Detection & Resolution: A Review

TOBECHUKWU C. OBIEFUNA[1], MATHEW EHIKHAMENLE[2]
[1]*Center for Information and Telecommunication Engineering, University of Port Harcourt*
[2]*Department of Electrical/Electronic Engineering, University of Port Harcourt*

Abstract*—The evolution of network technologies such as the 5G network is inextricably linked to the rise in demands for mobile devices. The deployment of 5G systems seeks to provide high throughput and ultra-low communications latencies, to improve users' quality of experience (QoE). To meet these demands, Conventional sub-6 GHz cellular systems are incapable of delivering the rapid data speeds and low latency needed by millimeter wave (mmWave) networks. However, mmWave signals are more vulnerable to blocking than lower frequency bands, leading to a higher number of coverage holes (CH) in a radio environment. Traditionally, cellular coverage hole detection is performed through drive tests, which consist of geographically measuring different network coverage metrics with a motor vehicle equipped with mobile radio measurement facilities. The collected network measurements need to be processed by radio experts for network coverage optimization, e.g., by tuning network parameters such as transmission power, antenna orientations and tilts, etc. The use of drive tests implies large Operational Expenditure (OPEX) and delays in detecting the problems, and they cannot offer a complete and reliable picture of the network situation. When users have poor wireless performance, the Key Performance Indicators (KPIs) for those clients are reported to a central manager, who converts them into a visual client-side perspective map. In this paper, other more efficient methods of detecting and resolving coverage holes such as Topographical, Probabilistic, UMAP and ML are explored.

Indexed Terms*—5G, Coverage hole, 5G KPIs, UMAP and ML.

I. INTRODUCTION

A coverage hole is an area within a sensor network where no sensor node is present. This area is unsupervised because no sensor node is present. As the dynamics of radio frequency (RF) change, wireless infrastructure elements frequently struggle to adapt to optimal coverage of the spaces, leaving coverage gaps or areas with poor RF performance. With the significant role of network intuitiveness, recognizing coverage gaps and areas with RF issues requires a

II. 5G NETWORKS

A. The Components of 5G Technology

While the evolution to 4G/LTE was driven by increased mobile data speed, the 5G system faces more stringent and diverse requirements. The deployment of 5G systems aims to provide high throughput and ultra-low communication latencies in order to improve the user experience (QoE). To meet these demands, 5G focuses on three evolution axes to support new application fields such as autonomous cars/driving, industrial automation/smart manufacturing, virtual reality, e-health, and so on. These are the axes: Improved mobile broadband: Allows for the development of new bandwidth-hungry applications with extremely high data rate demands across a uniform coverage area. Ultra-high-definition video streaming and virtual/augmented reality (VR/AR) are twoexamples.Massive machine-type communications (Mmtc): Scalable connectivity demand for expanding the number of wireless devices with efficient transmission of small amounts of data over extended coverage areas is a key feature of 5G communication services. This type of traffic will be generated by applications such as body-area networks, smart homes, IoT, and drone delivery.

Mmtc must be capable of supporting massive new uses as well as other uses that may emerge in the future. Connected healthcare, remote surgery, mission-critical applications, autonomous driving, unmanned aerial vehicles (UAV), vehicle-to-vehicle (V2V) communications, high-speed train connectivity, and smart industry applications all benefit from ultra-reliable low-latency communications (URLLC).

III. NETWORK COVERAGE HOLE

A coverage hole is defined as an area where the received signal strength or/and quality level of the serving and neighboring cells is less than what is required to maintain basic services. Coverage gaps in mobile networks are a common issue for mobile operators. It is a visible aspect that has a significant impact on the user experience. It is nearly impossible to completely avoid the existence of coverage gaps in cellular networks during the planning phase, so coverage optimization processes are usually required during the operational phase. [2].

- Coverage Hole Formation

CH, as previously defined, is the area of radio coverage that is below a threshold level required for robust radio performance. The primary cause of a CH is a lack of significantly powered (or no) MPCs reaching that point (due to environment geometry), resulting in RSS below the required threshold. Outside of the CH, however, there are MPCs with sufficient power, resulting in RSS suitable for a robust radio connection. The CH region LCH can be represented mathematically as:

$$\mathcal{L}_{\mathrm{CH}} = \{\ell_m \mid RSS(\ell_m) < RSS_0, \forall m \in [M]\} \quad (1)$$

and it can be a null set if the CHs do not exist. Thus,there will be a drastic change in the propagation characteristics across the CH boundary.

IV. COVERAGE HOLE DETECTION

The process of identifying areas where the signal strength or/and quality level is less than the required value for the given specified service standards using various techniques and tools is known as network coverage-hole detection. It is a specific case of signal degradation in a specific area caused by various factors such as cell-overload, a malfunctioning base station, congestion, or a planning problem

A. Methods of Detecting Coverage Hole

Collecting client radio and contextual data is one example step [1]. Using information fed back to the infrastructure from clients, this invention employs a wayfinding-like concept to discover poorly performing wireless areas and holes in a coverage area (rather than an infra-only perspective). As clients move around an area, their signal is tracked and their location is recorded. They report on performance issues that arise on a regular basis (e.g., a high number of retries, poor SNR, declining application quality, etc.).

Another approach, according to [1] is to analyze coverage.

Coverage Hole Detection via UMAP

The UMAP (User Modeling Adaptation and Personalization) algorithm is a new non-linear dimensionality reduction algorithm**.**

It attempts to preserve a similar structure in the low dimensional output space while learning the manifold structure in the high dimensional input data. It accomplishes this by optimizing the layout of the low-dimensional data space so that the structure's cross-entropy between the high and low dimensional space is minimized. [3] describes the hyper-parameters that are critical to understanding how the UMAP works. It is denoted as:

$$\mathcal{X} = \{X_1, X_2, \ldots, X_M\} \quad (2)$$

as the high dimensional input to UMAP and

$$\mathcal{Y} = \{Y_1, Y_2, \ldots, Y_M\} \subseteq \mathbb{R}^{dim} \quad (3)$$

as the low-dimensional output of dimension dim.

The algorithm's critical hyperparameters are: 1) (dis)similarity measure metric d; 2) neighborhood size n; and 3) minimum distance d_{min}. The parameter d quantifies the (dis)similarity of any two samples, and in this work, we use the Euclidean distance as the metric. The neighborhood size and takes into account the number of neighbors to take into account for a local metric approximation of d. The selection of neighbors represents a trade-off between capturing local versus global structure. Smaller values ensure that local-structure is accurately captured at the expense of global-structure, whereas larger values of n have the opposite effect. Furthermore, d_{min} is an aesthetic parameter that controls the minimum distance between the points in the output space [3. This parameter prevents potential overlapping of points at the output space, which can occur, for example, with the t-SNE technique. Low d_{min} values result in densely packed output space, whereas higher values result in loosely packed output space.

Probabilistic Methods

Coverage gaps in the network can be detected using statistical attributes when nodes are distributed uniformly and densely. In [4] an algorithm for determining the boundary node structure of a region is presented, but high-density nodes are required. Assume that the unit disk graph model in [4] determines node connectivity, and that a linear-time algorithm is proposed to identify the boundary of the holes. However, the algorithm cannot tell the difference between two holes that are close together.

B. The Various Thresholds and Values of Coverage Hole

Threshold estimation: In order to determine whether an indicator of a cell is degraded, the cell's normal performance should be characterized in order to determine the reference conditions. Then, for each metric, a threshold is established, above which the indicator is considered degraded. There are various methods for automatically designing these thresholds based on a historical dataset created from metrics and indicators provided by the OSS or UME of both the LTE network and its Urat (e.g., 3G) over time. Those historical datasets, in particular, are made up of the specific values of each indicator for each cell.

- Topological Methods

Topological methods identify the boundaries of a hole without knowing the exact location of nodes by using topological attributes such as connectivity information. The combinatorial Laplacians, according to [5], are the best tools for computing distributed homologous groups. Although distributed hole detection is possible, the holes cannot be accurately located. Gathering wireless data from client devices is described in the methods discussed above in order to provide a realistic picture of an RF coverage model. Thanks to the wireless infrastructure's (controller's) RF coverage model, users can be alerted to coverage gaps in their area and receive historical feedback on the quality of coverage in different areas.

C. Remedies for Coverage Holes (CHs) or CH Repair

To repair a detected coverage hole, Neighbor Intervention by Farthest Point (NIFP)employs cascaded movement [6]. When a node fails, the one-hop neighbors of the failed node calculate their intersection points, which are assumed to be the boundaries of the coverage hole. Only one node is chosen from among the neighbors to repair the hole. The node is chosen based on three factors: required moving distance, overlapping area with neighbors, and residual energy. The selected node moves to a new target location, and the moved node's original location becomes a coverage hole. In a cascaded movement of nodes, the algorithm recurs, and a new node is chosen to repair the newer hole. This process is repeated until no more holes are formed, or until the hole is small enough to be considered negligible. HEAL, according to [7], operates in two stages: hole detection and hole healing.

V. COMPUTER AIDED TECHNIQUES FOR ASSESSING COVERAGE HOLES

- Data Mining and Machine Learning

Data mining is the science and technology of exploring data to discover previously unknown patterns, and it is a component of the overall process of gaining knowledge from databases. Data mining is the process of extracting knowledge from large databases. Data mining tasks are classified into two types: descriptive and predictive.

- Supervised Learning

When a system is trained, it receives a set of inputs and outputs, i.e., a data set with labels, and creates a link between them. This learning algorithm predicts outputs by providing dependency links and relationships between inputs. Table 1 compares various machine learning techniques with various parameters.

- Unsupervised Learning

The unsupervised learning technique is used to classify data into similar patterns, reduce data size, form clusters, and detect anomalies. It is associated with given inputs and thus has no unlabeled output. This method addresses issues with connectivity, routing, data aggregation, and anomaly detection. Dimensionality reduction methods include singular value decomposition, independent component analysis, principal component analysis, and clustering methods such as fuzzy c-means, k-means, and hierarchical clustering.

- Semi-Supervised Learning

Semi-supervised learning can be applied to both supervised (labelled) and unsupervised (unlabeled) data sets. In real-world semi-supervised learning applications, classification is performed partially on labelled data and regression on unlabeled data. The important factor is predicting whether data in training and future datasets is labelled or unlabeled. This learning technique is used in video surveillance, speech recognition, web content classification, natural language processing, protein sequence classification, and spam filtering applications, as well as to solve fault detection and localization in wireless networks.

- Data Mining and Machine Learning

Data mining is the science and technology of exploring data to discover previously unknown patterns, and it is a component of the overall process of gaining knowledge from databases. Data mining is the process of extracting knowledge from large databases. Data mining tasks are classified into two types: descriptive and predictive.

- Supervised Learning

When a system is trained, it receives a set of inputs and outputs, i.e., a data set with labels, and creates a link between them. This learning algorithm predicts outputs by providing dependency links and relationships between inputs. Table 1 compares various machine learning techniques with various parameters.

- Unsupervised Learning

The unsupervised learning technique is used to classify data into similar patterns, reduce data size, form clusters, and detect anomalies. It is associated with given inputs and thus has no unlabeled output. This method addresses issues with connectivity, routing, data aggregation, and anomaly detection. Dimensionality reduction methods include singular value decomposition, independent component analysis, principal component analysis, and clustering methods such as fuzzy c-means, k-means, and hierarchical clustering.

- Semi-Supervised Learning

Semi-supervised learning can be applied to both supervised (labelled) and unsupervised (unlabeled) data sets. In real-world semi-supervised learning applications, classification is performed partially on labelled data and regression on unlabeled data. The important factor is predicting whether data in training and future datasets is labelled or unlabeled. This learning technique is used in video surveillance, speech recognition, web content classification, natural language processing, protein sequence classification, and spam filtering applications, as well as to solve fault detection and localization in wireless networks

VI. APPLICATION OF MACHINE LEARNING ALGORITHMS IN WIRELESS NETWORK ANALYSIS

This section discusses machine learning techniques for dealing with challenges in wireless networks, as well as their benefits, as well as existing approaches depicted in their respective tabular forms.Localization is the process of manually determining the geographical, physical location of a wireless node or by a global positioning system by sending beacon or anchor nodes. This can be determined by node proximity, distance and angle, range, or location. Continuous configuration and programming are required for a dynamically changing network, where machine learning techniques must be used to improve location accuracy. This has several advantages, such as the ability to easily find anchor and unknown nodes in a network by using machine learning algorithms to create clusters and train them separately.

VII. NETWORK CONNECTIVITY AND COVERAGE

Connectivity refers to any node that sends information to a receiver via relays or directly and does not include isolated nodes. 'Coverage' refers to monitoring as well as all effectively deployed area nodes. When compared to deterministic deployment, random node placement is feasible. There are two types of coverage: full coverage and partial coverage. Sweep, barrier, target, and focused are additional partial coverage classifications. Machine learning techniques for connectivity and coverage are shown in Table 2.

- Quality of Service (QoS)

The level of service provided by a network is referred to as its quality of service. This could be related to a specific application, such as active nodes, node measurements, and deployment, or network-specific aspects, such as bandwidth or rate of energy utilization. Unbalanced traffic, dynamic networks, data redundancy, resource constraints, scalability, energy balancing, and traffic type variations all have an impact.

- Artificial Neural Networks (ANN)

ANN is a connected input output network with weights assigned to each connection. It has a single input layer, one or more intermediate layers, and a single output layer. The neural network learns by adjusting the weight of the connections. Iteratively updating the weight improves network performance.

Table 1:

Specification Type	Decision tree	Reinforcement Learning	ANN	Deep learning	SVM	Bayesian	K-NN
Parameter Handling	Very good	Very good	Poor	Good	Poor	Best	Very good
Speed of learning	Very good	Good	Poor	Poor	Poor	Best	Best
Accuracy	Good	Good	Very good	Very good	Best	Poor	Good
Speed of classification	Best	Best	Best	Best	Best	Best	Poor
Missing values handling	Very good	Good	Best	Good	Good	Best	Good
Redundant variables handling	Good	Good	Good	Good	Very good	Poor	Good
Noise handling	Good	Very good	Good	Very good	Good	Very good	Poor
Independent variables handling	Good	Good	Very good	Very good	Very good	Poor	Poor
Irrelevant variables handling	Very good	Very good	Poor	Good	Best	Good	Good
Dealing over fitting	Good	Good	Poor	Poor	Good	Very good	Very good

Table 2:

Machine learning	Complexity	Connectivity or coverage	Network	Mobility of nodes	Contribution	References

Regression	Low	Connectivity	Centralized	Static	Reliability and quality of network improved	Sun et al., 2017
SVM	Moderate	Connectivity	Distributed	Static	Efficiency is improved	Kim et al., 2015
Random forest	Moderate	Coverage	Distributed	Static	Accuracy is improved	Elghazel et al., 2015
Bayesian	Moderate	Coverage	Distributed	Static	Time complexity is reduced	Yang et al., 2016
k-means & fuzzy c-means	Low	Connectivity	Distributed	Static	Workload is reduced	Qin et al., 2017
Reinforcement learning	Low	Coverage	Distributed	Static	Network lifetime is improved	Chen et al., 2016

. ANNs are classified into two types based on their connections: feed-forward networks and recurrent networks. In a feed forward neural network, connections between units do not form a cycle, whereas in a recurrent neural network, connections form a cycle [8]. The learning rule, architecture, and transfer function all have an impact on neural network behavior. The weighted sum of input activates neurons in a neural network. The activation signal is routed through a transfer function to produce a single neuron output. This transfer function causes the network to be nonlinear. The interconnection weights are optimized during training until the network achieves the desired level of accuracy. It has many advantages, such as parallelism, being less affected by noise, and having a high learning ability. [8]

An example of supervised learning is an artificial neural network. The knowledge is acquired by an artificial neural network in the form of a connected network unit. This knowledge is difficult for humans to extract. This factor prompted the extraction of a classification rule in data mining. The classification procedure begins with a dataset. The data set is split into two parts: training and testing samples. The training sample is used to train the network, while the test sample is used to assess the classifier's accuracy.
5G Key Performance Indicators (International Telecommunication Union, ITU, 2021)

- Peak data rates
- Peak spectral efficiency
- Area traffic capacity
- Latency
- Connection density
- Energy efficiency
- Reliability
- Mobility
- Bandwidth

- Coverage Hole (CH) Repair

Machine Learning has been widely used to improve network performance, particularly in recent years. The authors of [9] proposed a Deep Neural Network-based method to mitigate link failure caused by failed handovers and congested cells, among other things. The article's goal in [10]is to detect network intrusion by developing a technique based on RF and SVM.While the work [11] uses an approach based on Deep Learning. Moreover, focusing on QoS and Quality of Experience (QoE) prediction by means of machine learning, we also found a few research. The article [12] uses KNN, Decision Tree (DT), RF and Artificial Neural Network (ANN) to predict the QoE of Software Defined Networks, comparing their performances. The work [13] proposes to foresee the users' QoE for an LTE video streaming using ANNs. In its turn, the paper [14] conceives a system based on DTs to predict the QoE of end users of popular smartphone applications. When considering specifically the mobility management area, the techniques based on Machine Learning have recently been the object of some contributions. In [15], a handover mechanism for unmanned aerial vehicles is developed. The paper [16] proposes a scheme based on SVM to predict the mobile equipment location in an UDN within 5 seconds. None of the previously mentioned works has proposed machine learning handover management strategies focused on LTE networks in an HCS network. However, the paper [17] proposes a Self-Organizing Networks (SON) to make a handover scheme consisting of preselecting the Enb according to user speed and demanded QoS. Nevertheless, their solution does not delegate the handover decision to a machine learning technique. On the other hand, the article [18] proposes an ANN framework to make such decisions in an LTE network with a coverage hole, scenario presented by [18].In order to deal with the coverage hole scenario [18], a modified version of algorithm introduced in [18], and its performance was compared with the A2A4RSRP algorithms. Additionally, we also propose differently structured handover frameworks based on ANNs, KNNs, SVMs, and RFs. These machine learning frameworks vary in processing demands and scalability, allowing us to evaluate the cost-effectiveness trade-off of the proposed schemes. Furthermore, another scenario, with more severe propagation conditions than the first one (due to shadowing), is used in the analysis to better evidence the effects of the proposed schemes in the complexities of the current urban environments.

REFERENCES

[1] S. Ramanathan, G. Salgueiro, M. Natarajan, R. Barton, and ..., "Crowd Sourcing-Based True Coverage Hole and Radio Frequency Issue Detection," *Tech. Discl. Commons*, pp. 0–6, 2018, [Online]. Available: https://www.tdcommons.org/dpubs_series/1255/ .

[2] A. Galindo-Serrano, B. Sayrac, S. Ben Jemaa, J. Riihijärvi, and P. Mähönen, "Automated coverage hole detection for cellular networks using radio environment maps," in *2013 11th International Symposium and Workshops on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, 2013, pp. 35–40.

[3] C. K. Anjinappa and İ. Güvenç, "Coverage Hole Detection for mmWave Networks: An Unsupervised Learning Approach," *IEEE Commun. Lett.*, vol. 25, no. 11, pp. 3580–3584, 2021, doi: 10.1109/LCOMM.2021.3106251.

[4] S. P. Fekete, A. Kröller, D. Pfisterer, S. Fischer, and C. Buschmann, "Neighborhood-Based Topology Recognition in Sensor Networks BT - Algorithmic Aspects of Wireless Sensor Networks," 2004, pp. 123–136.

[5] F. Yan, P. Martins, and L. Decreusefond, "Connectivity-Based Distributed Coverage Hole Detection in Wireless Sensor Networks," in *2011 IEEE Global Telecommunications Conference - GLOBECOM 2011*, 2011, pp. 1–6, doi: 10.1109/GLOCOM.2011.6134128.

[6] B. Khalifa, Z. Al Aghbari, A. M. Khedr, and J. H. Abawajy, "Coverage Hole Repair in WSNs Using Cascaded Neighbor Intervention," *IEEE Sens. J.*, vol. 17, no. 21, pp. 7209–7216, 2017, doi: 10.1109/JSEN.2017.2755122.

[7] M. R. Senouci, A. Mellouk, and K. Assnoune, "Localized Movement-Assisted SensorDeployment Algorithm for HoleDetection and Healing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 5, pp. 1267–1277, 2014, doi: 10.1109/TPDS.2013.137.

[8] I. A. Basheer and M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application," *J. Microbiol. Methods*, vol. 43, no. 1, pp. 3–31, 2000, doi: https://doi.org/10.1016/S0167-7012(00)00201-3.

[9] S. Khunteta and A. K. R. Chavva, "Deep Learning Based Link Failure Mitigation," in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2017, pp. 806–811, doi: 10.1109/ICMLA.2017.00-58.

[10] Y. Chang, W. Li, and Z. Yang, "Network Intrusion Detection Based on Random Forest and Support Vector Machine," in *2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, 2017, vol. 1, pp.

635–638, doi: 10.1109/CSE-EUC.2017.118.

[11] C. Yin, Y. Zhu, J. Fei, and X. He, “A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks,” *IEEE Access*, vol. 5, pp. 21954–21961, 2017, doi: 10.1109/ACCESS.2017.2762418.

[12] T. Abar, A. Ben Letaifa, and S. El Asmi, “Machine learning based QoE prediction in SDN networks,” in *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2017, pp. 1395–1400, doi: 10.1109/IWCMC.2017.7986488.

[13] T. Begluk, J. B. Husić, and S. Baraković, “Machine learning-based QoE prediction for video streaming over LTE network,” in *2018 17th International Symposium INFOTEH-JAHORINA (INFOTEH)*, 2018, pp. 1–5, doi: 10.1109/INFOTEH.2018.8345519.

[14] P. Casas *et al.*, “Predicting QoE in cellular networks using machine learning and in-smartphone measurements,” in *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*, 2017, pp. 1–6, doi: 10.1109/QoMEX.2017.7965687.

[15] [15] H. Yang, B. Hu, and L. Wang, “A deep learning based handover mechanism for UAV networks,” in *2017 20th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, 2017, pp. 380–384, doi: 10.1109/WPMC.2017.8301842.

[16] J. Yang, C. Dai, and Z. Ding, “A scheme of terminal mobility prediction of Ultra Dense Network based on SVM,” in *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)*, 2017, pp. 837–842, doi: 10.1109/ICBDA.2017.8078755.

[17] H.-W. Ferng and Y.-Y. Huang, “Handover scheme with enode-B pre-selection and parameter self-optimization for LTE-A heterogeneous networks,” in *2016 International Conference on Machine Learning and Cybernetics (ICMLC)*, 2016, vol. 2, pp. 594–599, doi: 10.1109/ICMLC.2016.7872954.

[18] Z. Ali, N. Baldo, J. Mangues-Bafalluy, and L. Giupponi, *Simulating LTE Mobility Management in presence of coverage holes with ns-3*, vol. 2. 2015.