Explorematics
Journal of Innovative
Engineering and Technology

Volume: 06 No: 01 | April -2025
ISSN (Online) 2636 – 590
ISSN (Print) 2636 - 591X

IMPROVING THE PERFORMANCE OF 5G CELLULAR NETWORK USING OPTIMAL SMALL-CELL DEPLOYMENT TECHNIQUE

Onuorah Florence Chekwube[*1], Okorie Chikadibia Alphonsus[1], Abonyi D. O.[1]
1 Department of Electrical and Electronic Engineering, Enugu State University of Science and Technology
Author for correspondence: Onuorah F.C; **E-mail:** asumchek@gmail.com

Abstract - This paper presents the use of an optimal small cell deployment technique for improving the performance of a 5G cellular network. The research methodology used is the exploratory, experimental, and simulation methods. This encompasses several key steps aimed at improving the performance of a cellular system through optimal small cell deployment. Initially, an exploratory phase involves performing evaluation methodologies and cellular system optimization approaches. Based on the gathered data from MTN Nigeria, an optimal cell deployment strategy was developed using machine learning and a rule-based optimization algorithm, tailored specifically to the requirements of the cellular system under study, considering 3rd Generation Partnership Project Long-Term Evolution (3GPP LTE) design standards. The Machine Learning Algorithm utilized to develop the model is Linear Regression (LR). LR is a statistical modelling technique used to analyze the relationship between a dependent variable and one or more independent variables. The strategic framework for network planning and cell deployment presents the system integration of the trained LR model and the decision-based algorithm for cell deployment at Milliken Hills. The models and algorithms developed were implemented using the Regression App within MATLAB. From the results, it was observed that the predictor was able to correctly predict all data observations. What this means is that the LR predictor was able to correctly predict the network condition of a particular place for cell deployment. Overall, these results have demonstrated the effectiveness of LR in correctly predicting the network information of a particular area, to help network administrators take a position on the best-fit area for cell deployment. The result analysis shows that the network quality is excellent up to a distance of 1.7 km, fair from 1.8 km to 2.5 km, and poor beyond 2.6km. Based on the criteria of recommending small cell deployment only in areas with poor network quality, small cells should be deployed from a distance of 2.6km onward.

Keywords: Cell Deployment; 5G; Machine Learning; Logistic Regression; Rule-Based Optimization

1. Introduction

The main reason behind the invention of mobile phones was to facilitate systematic voice communication in the early seventies (70s). Currently, mobile devices and applications have grown so much that virtually all human activities depend on them. This growth of mobile applications and users has been fuelled by the rapid expansion of supporting technology, which has transformed from the First Generation (1G) to the present-day fifth Generation (5G) (Seiamak et al., 2015).

Over the years, network operators have proposed various administrative solutions to improve the quality of service and traffic management for the increased number of subscribers. One such solution identified by Chataut and Aki (2020) was to improve network densification and spectrum capacity to accommodate more carriers simultaneously and increase the number of antennas in a cell. However, Abonyi and Jonathan (2018) argue that the high cost of spectrum and a limited number of antennas per cell make this solution inefficient.

Recently, small cell (pico, femto, and micro) deployment has gained research attention due to its advantages over traditional larger cells (macro), such as lower cost, smaller space requirements for installation, and indoor installation feasibility. These small cells are installed to support macro cells and form a heterogeneous cellular network structure, with the potential to address problems arising from increased user traffic in conventional 4G networks (Shelly et al., 2012; Ghosh et al., 2012; Abonyi, 2016). However, achieving this goal requires an optimal cell deployment strategy.

According to Su et al. (2014), cell deployment is a process of planning and implementing distribution cells to provide network coverage and capacity for communication purposes. The strategy or this process of cell deployment can be classified into random deployment and deterministic deployment strategies (Su et al., 2014). Random deployment involves deploying small cell locations randomly within an area, while deterministic deployment positions cells based on predetermined factors (Abonyi and Jonathan, 2016). Both strategies have their advantages and disadvantages, but the use of the deterministic strategy provides a better solution to mobile network problems compared to its counterpart. Furthermore, Abonyi and Jonathan (2018) posited that distributed strategies for cell deployment can be categorized into uniformly distributed, cell edge, and user-aware strategies. However, Park et al. (2014), Hoadley and Maveddat (2012), and Bahceci (2014) have found that using the user-aware strategy for cell deployment, which considers key network and environmental information to determine the cell deployment location, outperforms the other strategies. However, a limitation of this approach is the inability to localize the best position for cell deployment. Therefore, there is a need for an optimal cell deployment strategy that can determine the best position for smart cell deployment.

Studies by Nika et al. (2014), Qutqut et al. (2014), Landstrom et al. (2011), Abonyi and Jonathan (2017), among others, have revealed that deploying cells in optimal positions can address issues in mobile heterogeneous networks, such as overload, interference, congestion, poor coverage, and signal strength. In addition, these studies provided various notable solutions for cell deployment, however, the complexity of the cell deployment problem makes it difficult for one solution in a given environment to be applicable for all environments.

Abonyi (2019) develop a novel clustered based two element antenna strategy which considered signal strength information, user clusters, and range information for cell planning and deployment and when validated experimentally considering 30 users, reported 100% accuracy in detecting position for cell deployment over 6 km; however despite the huge success recorded in the paper, the model may not apply to most African environments like Nigeria characterized with terrain such as mountains, valley and hills, as these actor were not mentioned in the work. Other studies, such as Oluwatoki et al. (2022), applied a path loss prediction approach for data collection and network planning in an urban environment, considering the terrain in Nigeria. Similarly, Enyi et al. (2021) applied a path loss model considering vegetation and terrain in Nigeria. In addition, Azubogu et al. (2010) predicted the pathloss information of localities in south-east Nigeria, considering Okumura-Hata, COST231-Hata, and free space models. While these studies (Azubogu et al., 2010; Oluwatoki et al., 2022; Enyi et al., 2021) captured the environmental characteristics of Nigeria while predicting the network information considering parameters like signal strength, distance, and network quality, the pathloss model developed and applied for the studies was not validated experimentally.

To this end, this research proposes to develop an optimal cell deployment model considering artificial intelligence techniques and then a smart decision-based strategy. The intention is to accurately predict the network information of Milliken Hills, Enugu, Nigeria, considering the complex environmental characteristics. The predicted information will serve as input for a decision-based system, which will assess key parameters such as signal strength, quality, coverage area, and interference to determine the suitability of the area for cell deployment and provide recommendations. Experimental validation and testing in a complex environment will be conducted to ensure the reliability and credibility of the results.

2. Research Methodology

The research methodology used is the exploratory, experimental, and simulation method. This encompasses several key steps aimed at improving the performance of a cellular system through optimal small cell deployment. Initially, an exploratory phase involves conducting a comprehensive literature review to gather relevant knowledge on small cell deployment techniques, performance evaluation methodologies, and cellular system optimization approaches. Following this, an experimental phase involves conducting a site survey test to collect data on various site characteristics that can influence the deployment of small cells, such as population density, terrain, building structures, and existing macro-cellular infrastructure.

Based on the gathered data, an optimal cell deployment strategy was developed using machine learning and a rule-based optimization algorithm, tailored specifically to the requirements of the cellular system under study, considering 3GPP LTE design standards. Extensive performance evaluations were conducted to assess the impact of the cell deployment technique on the cellular system, encompassing measurements of signal strength, coverage area, data transfer rates, call quality, and overall network capacity. Finally, the results obtained from the deployment were validated, and the effectiveness of the optimal small cell deployment technique was evaluated, providing insights into the regions where new cells are required for deployment.

3. Strategic Network Planning with Machine Learning

The method developed in this work for optimal cell deployment is a machine learning technique. This was adopted out of the numerous artificial intelligence techniques and traditional approaches due to the complexity of the requirements for cell planning and deployment. These complexities were due to the numerous environmental factors within the Milliken Hills, such as terrain, hills, vegetation, population density, traffic patterns, and interference levels, so as to effectively predict cell coverage performance. To develop the model for optimal cell deployment, several steps were utilized, which are data collection, selection of the machine learning algorithm, training, testing, evaluation, and generation of the prediction model for network planning.

3.1 Data collection

Data on network coverage and topological information for Milliken Hills, Coal Camp, Enugu, was collected from Airtel Nigeria, specifically tailored to the Airtel cell with serial number 401699. The dataset was generated from a fifteen (15) day, 40 km per minute interval drive test conducted on 3rd August, 2021, after the installation of the cell. The instrument used for the data collection is a BladeRFxA9 spectrum analyser powered by an external USB3 hub, and the dataset description was presented in Table 1.

Table 1: Dataset description

Data Name	Description	Data Type
Time (m:s:ms)	Time stamp indicating when the data was recorded	Time (mm:ss.ms)
Location	Geographical coordinates (latitude, longitude)	Latitude, Longitude
Coverage Area (km²)	The area covered by the cell site	Numeric (km²)
Signal Strength	Received Signal Strength Indicator (RSRP)	Numeric (dBm)
Signal Quality	Received Signal Quality Indicator (RSRQ)	Numeric (dB)
Cell Location	Location of the cell site	Text (address or code)
Transmit Power	Power level at which the cell site transmits	Numeric (dBm)
Interference Levels	Levels of interference experienced at the location	Numeric or Categorical

3.2 Machine learning algorithm

The machine learning algorithm utilized to develop the model is Linear Regression (LR). LR is a statistical modelling technique used to analyze the relationship between a dependent variable and one or more independent variables. It assumes a linear relationship between the independent variables and the dependent variable and aims to find the best-fit line that minimizes the difference between the predicted and actual values. The mathematical model of linear regression is represented as follows for multiple independent variables (Huang, 2020):

$$Y = \beta 0 + \beta 1X1 + \beta 2X2 + \ldots + \beta n * Xn + \varepsilon \quad (1)$$

Where:

Y is a dependent variable

X1, X2, …, Xn are independent variables

β0, β1, β2, ..., βn are regression coefficients

ε is error

The dependent variable, denoted as Y, represents the Signal strength, signal quality, and interference. The independent variables, denoted as X1, X2, ..., Xn, include cell location like terrain, geographical information, population density, and network planning data. The coefficients β0, β1, β2, ..., βn, also known as regression coefficients or weights, quantify the impact of each independent variable on the dependent variable. The error term, denoted as ε, accounts for the unexplained variability in the dependent variable. The objective of linear regression is to estimate the coefficients β0, β1, β2, ..., βn in a way that minimizes the disparity between the predicted values (Y) based on the independent variables (X1, X2, ..., Xn) and the actual values of the dependent variable. The estimation of the coefficients in linear regression typically involves employing the gradient descent algorithm, which seeks to minimize the sum of squared differences between the predicted values and the actual values. Once the coefficients are estimated, they can be utilized to make predictions for new data by inputting the values of the independent variables into the equation.

Configuration of the LR algorithm (Algorithm 1) (Huang, 2020)

1. Input:
   - X: Matrix of independent variables (cell locations, transmit powers, etc)
   - Y: Vector of the dependent variable (signal strength, interference levels, and signal quality.)
2. Initialize:
   - Initialize the coefficients β0, β1, β2, ..., βn to zeros.
   - Define the learning rate (α=0.001) and the maximum number of iterations (max iterations = 100).
3. Model Training:
   - For each iteration from 1 to max_iterations:
     - Compute the predicted values ($Y_{predicted}$) using the current coefficients:

$$Y_{predicted} = \beta 0 + \beta 1X1 + \beta 2X2 + \ldots + \beta n * Xn$$

- Compute the error (error) as the difference between the predicted and actual values: error = $Err = Y_{predicted} - Y_{actual}$
- Update the coefficients using the gradient descent algorithm: $\beta 0 = \beta 0 - \alpha * (1/m) * \Sigma(error)\, \beta 1 = \beta 1 - \alpha * (1/m) * \Sigma(error * X1)\, \beta 2 = \beta 2 - \alpha * (1/m) * \Sigma(error * X2) \ldots \beta n = \beta n - \alpha * (1/m) * \Sigma(error * Xn)$

- Repeat the steps (2-3) until convergence

4. Model Evaluation:
   - Compute the predicted values ($Y_{predicted}$) using the final coefficients.
   - Evaluate the performance of the model using appropriate metrics such as mean squared error (MSE), root mean squared error (RMSE), or R-squared (R2).
5. Model Deployment:
   - Deploy the trained LR model by using the obtained coefficients (β0, β1, β2, ..., βn) to predict the dependent variable for new data, which are RSRP, RSRQ, and interference.
6. End

3.3 Training of the LR algorithm

The training of the LR algorithm with network information data collected from Coal Camp, Enugu, begins by inputting the data consisting of a matrix X with independent variables like cell locations, transmit powers, interference levels, and a vector Y representing the dependent variable, such as signal strength. Before the training, the LR was configured by setting necessary parameters like coefficients, learning rate and iterations steps.

During the model training phase, the LR algorithm iterates through the training process from 1 to maximum iterations. In each iteration, the predicted values ($Y_{predict\ ed}$) are computed using the current coefficients and the input matrix X. The error is then calculated as the difference between the predicted values and the actual values (Y_{actual}). The coefficients are updated using the gradient descent algorithm, where each coefficient is adjusted by subtracting the learning rate multiplied by the average of the errors multiplied by the corresponding independent variable. These steps are repeated until convergence is achieved, allowing the LR model to learn the relationships between the independent and dependent variables. The figure 1 presents the flow chart of the training process.

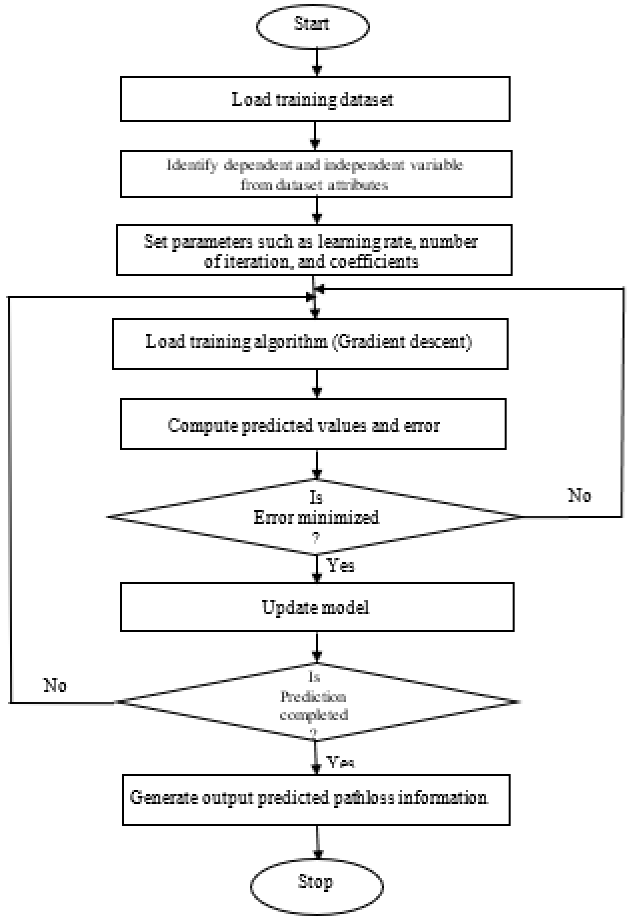

Figure 1: Flow chart of the trained LR for cell deployment

Once the model has been trained, the predicted values ($Y_{predicted}$) are computed using the final coefficients. Finally, the trained LR model was deployed by utilizing the obtained coefficients (β0, β1, β2, ..., βn) to predict the dependent variable for new data points or cell deployment. This enables the model to make predictions based on the learned relationships from the training process, providing valuable insights into network performance in Coal Camp, Enugu, and predicting the position with poor signal strength for cell deployment.

4. Smart Decision Framework for Cell Deployment

To develop the strategic framework, the ideas were to automate the decision for the network planning and cell deployment using the LR prediction model generated and 3GPP standards to develop a decision-based algorithm for the recommendation of positions for cell deployment. The DFCD was developed considering the predicted output from the LR model, which is the Reference Signal Receive Power(RSRP). Reference Signal Receive Quantity(RSRQ) and interference level, then key independent variables such as population density, network capacity, coverage area, and severalcells required to serve the locality.

The population density is an important factor in network planning as it affects the demand for network services. A higher population density generally requires higher network capacity to accommodate the increased traffic and is mathematically defined as;

$$Population\ Density = \frac{Total\ Population}{Area} \quad (2)$$

The Network capacity represents the maximum amount of data or traffic that a network can handle within a given timeframe and must be calculated considering interference from neighbouring cells.

$$Network\ Capacity = (Traffic\,Volume * Capacity\ per\ User) - Interference \quad (3)$$

The Coverage area, on the other hand, is influenced by factors such as transmit power, antenna characteristics, terrain conditions, other signal propagation characteristics, and interference; the model is defined as;

$$Coverage\ Area = \pi * (Cell\ Radius)^2 * Coverage\ Enhancement\ Factor \quad (4)$$

Where

$$\Pi = 3.142, and$$
$$coevrage\ enhancement\ factor = 1.2$$

Finally, the calculation for the number of cells to be deployed in the area is determined based on factors such as population density, traffic demand, network capacity requirements, and coverage area. In network dimensioning, capacity sharing is considered, where multiple cells share the available network resources, and overall is defined as;

$$Number\ of\ Cells\ Required = \left(\frac{Total\ Area}{Coverage\ Area}\right) * \mu$$

Where μ is the capacity sharing factor and is given as 0.7.

The algorithm is presented as (Algorithm 2)

1. Initialization:
2. Set parameters of dependent variables (Y): RSRP (< -90 dBm), RSQ (< -10 dB), coverage (1 km²), and interference level (> -100 dBm).
3. Initialize the LR (Logistic Regression) cell deployment model.
4. If Output of LR = true:
5. Identify all Y variables and read corresponding values.
6. Check X factors:
7. For each Check Population Density (CPD) over the target area:
8. Check Population Density (CPD) using Equation 2.
9. Check Network Capacity (NCC) of the micro cell using Equation 3.
10. Calculate the number of micro cells (NC) needed using Equation 5.
11. Assess the suitability of deploying cells based on the coverage area required to meet the capacity demands using cells.
12. Analyze the predicted quality of service metrics (such as RSRP, RSQ, and interference level) obtained from the LR model for the target area.
13. If RSRP < -90 dBm and RSQ < -10 dB and interference level > -100 dBm and coverage area 1km^2:
14. MsgBox (Flag the area as potentially suitable for cell deployment.)
15. Compute NC with Equation 5 to determine the recommended number of cells
16. Else if RSRP < -90 dBm and interference level > -100 dBm:
17. MsgBox (Flag the area as potentially suitable for cell deployment.)
18. Compute NC with Equation 5 to determine the recommended number of cells
19. Else if RSQ < -10 dB and interference level > -100 dBm:
20. MsgBox (Flag the area as potentially suitable for cell deployment.)
21. Compute NC with Equation 5 to determine the recommended number of cells.
22. Else
23. MsgBox (Flag the area as good network reception and hence not suitable for cell deployment.)
24. End the algorithm.

The algorithm for cell deployment aims to identify suitable areas for microcell deployment based on certain conditions and criteria. The initialization phase sets the parameters for dependent variables, such as RSRP being above -90 dBm, RSQ being below -10 dB, coverage area of 1 km^2, and interference level above 100 dBm. The algorithm then initializes an LR cell deployment model.

If the output of the LR model is true, the algorithm proceeds to identify all variables and read their corresponding values. It then checks various factors, denoted as X. For each Calculated Population Density (CPD) over the target area, the algorithm assesses the suitability of deploying micro cells. This assessment is based on the coverage area required to meet the capacity demands using micro cells and an analysis of predicted quality of service metrics obtained from the LR model. If the conditions are met (RSRP < -90dBm, RSQ < -10dB, interference level > -100dBm, and coverage area), the algorithm flags the area as potentially suitable for micro cell deployment. Similarly, if RSRP is below -90dBm and the interference level is above -100dBm, the area is flagged as potentially suitable. Alternatively, if RSQ is below -10dB and the interference level is above -100dBm, the algorithm computes the recommended number of micro cells needed using Equation 5. Figure 2 presents the flow chart of the strategic algorithm for cell deployment.

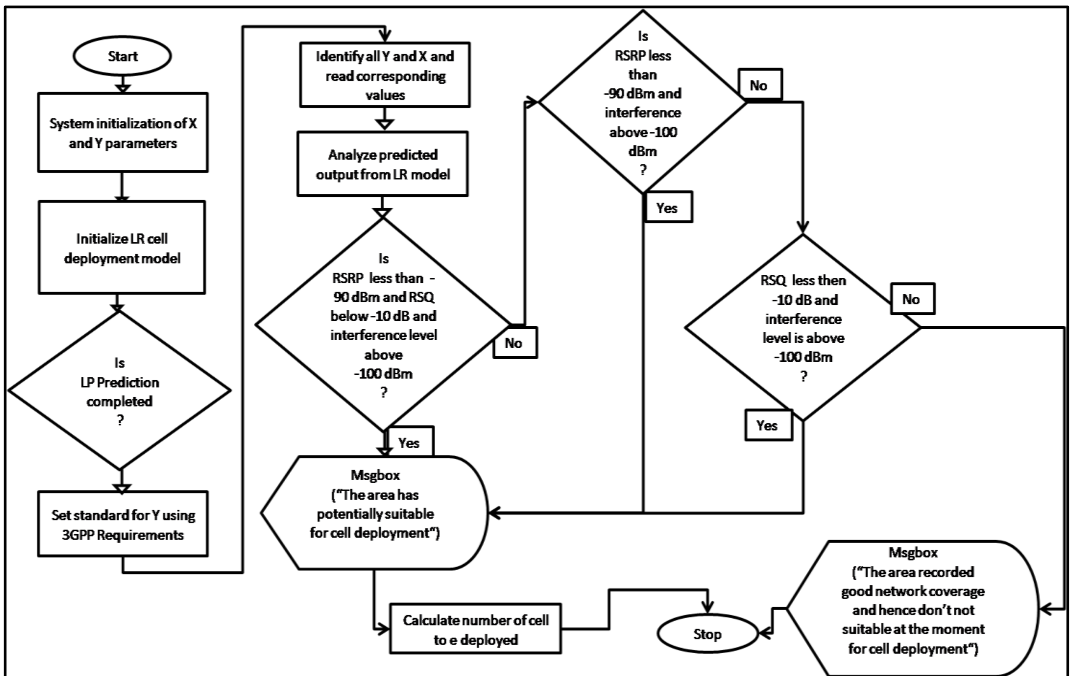

Figure 2: Flowchart of DFPD

Table 2: Parameters and standards for network information analysis

Parameter	Description	Ideal Value
Signal Strength (RSRP)	Power level of the received signal from an LTE cell	-70 dBm to -90dBm
Signal Quality (RSRQ)	Quality of the received signal from an LTE cell	-10 dB to -20dB
Coverage	Extent of network coverage area	Maximum coverage possible
Interference Level	Level of interference experienced in the network	-90dBm to -95dBm
Coverage size	The required microcell coverage	$1km^2$

4.1 Integrated Framework for Network Planning and Deployment

The strategic framework for network planning and cell deployment presents the system integration of the trained LR model and the decision-based algorithm for cell deployment at Milliken Hills. The flow chart showed how the test data, which contains geographical information of an area considered for the cell deployment, when loaded to the trained LR cell deployment model it was used to predict the quality-of-service information of the area, such as signal strength, signal quality, and level of interference. This output is analysed considering the 3GPP standard in Table 2 and the algorithm (2) to make a decision if the area is fit for cell deployment or not. The flow chart in Figure 3 presents the system workflow.

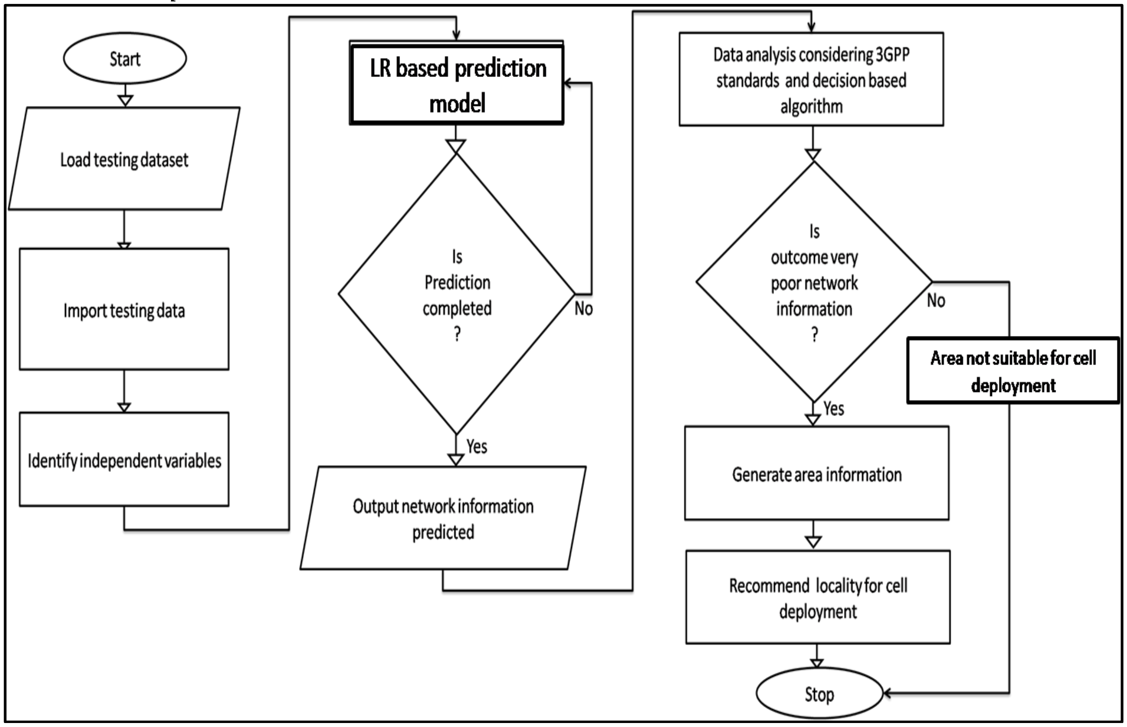

Figure 3: Flow chart of the ISPD

Overall, the data of the network information collected from the area for consideration are imported to the LR-based prediction model, which uses the dependent variables to predict the independent variables as model in equation 1. This was used to predict the network planning information requirements, providing data for RSRP, RSQ, interference, and transmitting power. These information servicesare used as input to the decision-based model developed considering network capacity, coverage area information, population density of the area, and the 3GPP standard for RSRP, RSQ, and interference for LTE to make a decision if the area is suitable for cell deployment or not.

4.2 Implementation of the ISPD

The models and algorithms developed were implemented using the Regression App within MATLAB. The process involved a series of steps aimed at ensuring accurate prediction and evaluation of cell deployment suitability.

Initially, data collection was undertaken, focusing on gathering test data that contained pertinent geographical information for the target area. This encompassed factors like terrain characteristics, population density, and other variables impacting cell coverage. The collected data underwent pre-processing techniques to cleanse, normalize, and transform it into a suitable format for subsequent analysis. Next, the pre-trained LR cell deployment model was loaded into the Regression App. This LR model had been previously trained on a labelled dataset, featuring geographical information as input features and quality of service metrics (signal strength, signal quality, interference level) as output. By integrating the LR model into the app, it became primed for generating predictions based on the test data.

Utilizing the loaded LR model, the test data was input into the app, enabling the prediction of quality-of-service information for the target area. Leveraging the provided geographical information, the LR model generated predicted values for signal strength, signal quality, and interference level. These predictions were then extracted from the app for further examination.

The output yielded by the LR model underwent analysis by comparing it against the applicable standards set forth by the 3rd Generation Partnership Project (3GPP). These standards establish acceptable ranges or thresholds for signal strength, signal quality, and interference level. Reference to the 3GPP standards, often presented in tabular form, facilitated the evaluation of the predicted quality of service values in terms of their adherence to the defined criteria as presented in Algorithm 2.

Based on the analysis of the predicted quality of service values and their alignment with the 3GPP standards, informed decisions were made concerning the suitability of the target area for cell deployment. This decision-making process involved assessing whether the predicted values fell within the acceptable ranges or thresholds specified by the 3GPP standards, guiding determinations of deployment suitability or the need for further considerations and modifications. Table 3 presents the criteria used for the data analysis and network planning to determine the area thatneeds network deployment.

Table 3: Criteria for network planning analysis (Isabona and Obahiagbon, 2013)

Criteria	RSRP	RSRQ	Interference
Excellent	≥ -80 dBm	≥ -10 dB	≤ -70 dBm
Very Good	-80 to -90 dBm	-10 to -15	-70 to -80 dBm
Good	-90 to -100 dBm	-15 to -20	-80 to -90 dBm
Fair	< -100 dBm	< -20 dB	≤ -100 dBm

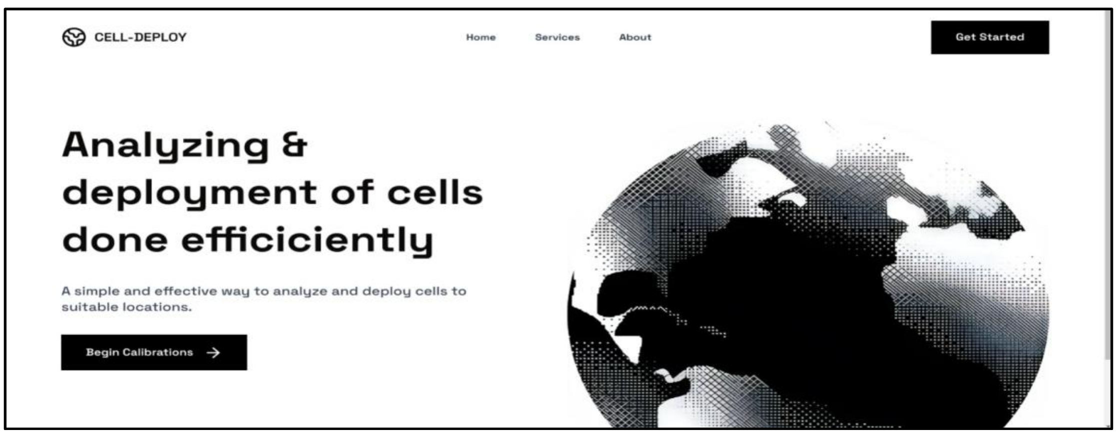

Figure 4: The application software for cell deployment

5. Results and Discussion

Having tested the LR model for the prediction of network information correctly, the model was integrated with the smart decision-based framework as a software application for cell deployment. Figure 4 shows the application software for cell deployment, while Figure 5 was used to test the software with data collected from different locations of the area considered for initial characterization.

Figure 4 presents the application software for cell deployment, which was developed using the decision-based framework model, which also considered the network information for interference, RSRP, and RSRQ in line with the LR algorithm toanalyze and then decide if the area is fit for cell deployment or not. The outcome of the analysis is stored in a database management system, which was also utilized in the future for further analysis and investigation of the network behaviour. Figure 5 shows the experimental test of the software with data collected from Milliken Hills, Coal Camp Enugu. To achieve this, the data utilized was the same data collected during characterization and was input to the software through the interface shown in figure 5.

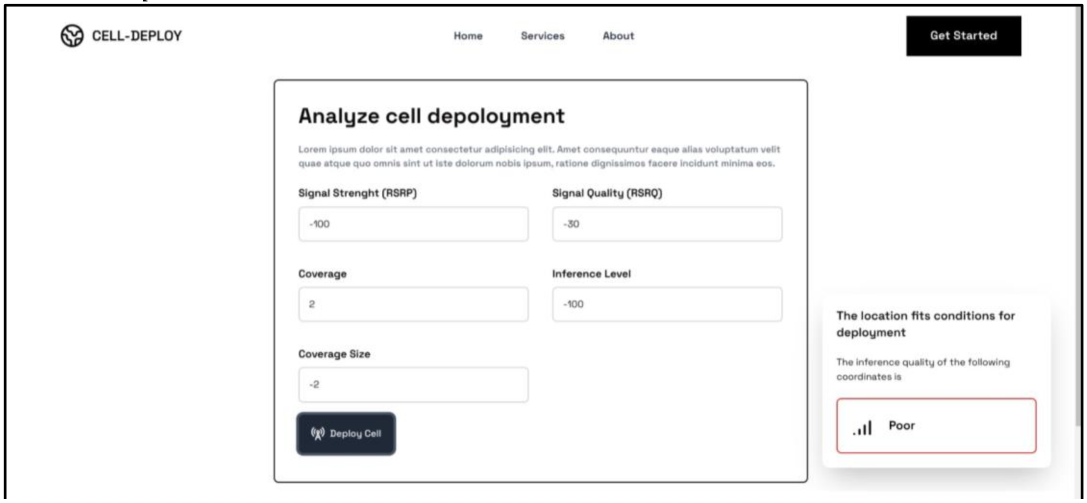

Figure 5: Testing of the software with a case study

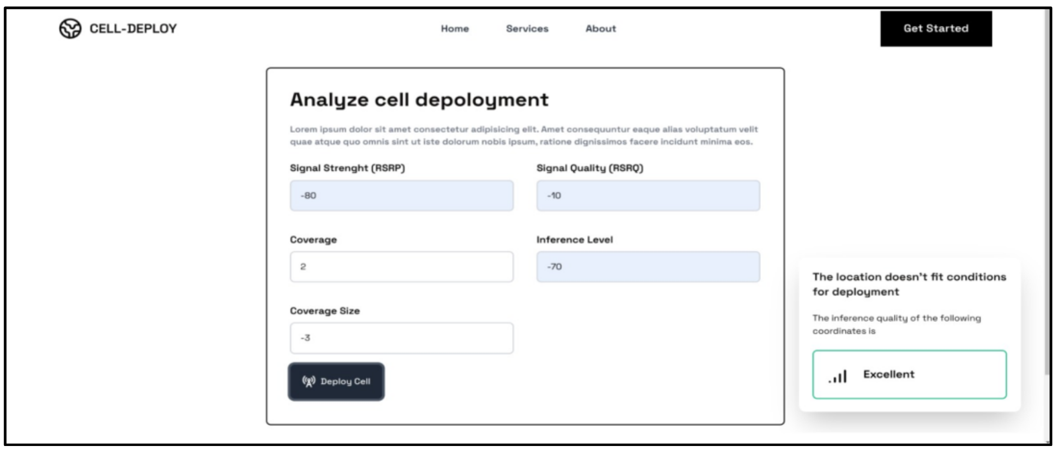

Figure 6: Testing of the cell deployment software

Figure 5 presents the result of the software testing using network parameters input to the system. The data loaded into the system was identified by the LR model as the dependent variable, and the predictor was used to predict the network condition in the area. In addition, the decision-based algorithm was also applied to ensure that the right position is selected for the cell deployment. From the results, it was observed that the model characterized the position of the data collected as an area with very poor quality of service, and hence declared it fit for cell deployment. Figure 6 shows the next results, where data from another location was collected and applied to test the software model, which utilized the decision-based model developed with LR to analyze the behavioural patterns in an environment and then decide if the location is fit for cell deployment or not.

Figure 6 showcases the result of the cell deployment software tested using network information from a particular area. The results after analysing the RSRP, RSRQ, and interference in the area showed that the quality of service in the location is excellent and hence not fit for cell deployment, as users in the area are already enjoying quality of service. In addition, after the analysis of the network information within the area and then decided by the software if a cell will be deployed or not, the data is also stored in a database as shown in Figure 7, for future referencing and also to facilitate easy decision making while installing future cells.

Figure 7: Database manager of cell deployment information

Figure 7 shows the result of the database manager, which is the software responsible for the management of network data collected of network information from a particular area.

6. Conclusion

Over the years, many notable studies have been submitted to improve network planning and cell deployment, however, the complexities and dynamic characteristics of geographical localities makeit impossible to generalize the solution to all areas. In other words, a solution developed that considers network coverage information like signal strength, signal quality, interference, and user density, for instance, may not be reliable for planning an environment with terrains such as hills, mountains, vegetation, among others. In addition, most of the existing systemsthat successfully captured these environmental terrains and predict the pathloss information for network planning were not experimentally validated and hence affecttheir trustworthiness for network planning. This paper presents the use of an optimal small cell deployment technique for improving the performance of a 5G cellular network. The research methodology used is the exploratory, experimental, and simulation method. This encompasses several key steps aimed at improving the performance of a cellular system through optimal small cell deployment. Initially, an exploratory phase involves performing evaluation methodologies and cellular system optimization approaches. Based on the gathered data, an optimal cell deployment strategy was developed using machine learning and a rule-based optimization algorithm, tailored specifically to the requirements of the cellular system under study, considering 3GPP LTE (Long-Term Evolution) design standards.

The machine learning algorithm utilized to develop the model is Linear Regression (LR). LR is a statistical modelling technique used to analyze the relationship between a dependent variable and one or more independent variables. The strategic framework for network planning and cell deployment presents the system integration of the trained LR model and the decision-based algorithm for cell deployment at Milliken Hills. The models and algorithms developed were implemented using the Regression App within MATLAB. The process involved a series of steps aimed at ensuring accurate prediction and evaluation of cell deployment suitability. Performance evaluation was conducted to assess the impact of the cell deployment technique on the cellular system, encompassing measurements of signal strength, coverage area, data transfer rates, call quality, and overall network capacity. From the results, it was observed that the predictor was able to correctly predict all data observations. What this means is that the LR predictor was able to correctly predict the network condition of a particular place for cell deployment. Overall, these results have demonstrated the effectiveness of LR in correctly predicting the network information of a particular area, to help network administrators take a position on the best-fit area for cell deployment. The result analysis shows that the network quality is excellent up to a distance of 1.7km, fair from 1.8km to 2.5km, and poor beyond 2.6km. Based on the criteria of recommending small cell deployment only in areas with poor network quality, small cells should be deployed from a distance of 2.6km onward.

References

Abonyi D., & Jonathan M. (2017) Cellular Network Load Balancing Strategy for Unexpected Hotspots in Developing Countries. International Conference of Women Engineers and Scientists (ICWES), New Dehli, India.

Abonyi D., & Jonathan M., (2018) A System for Optimizing Small-Cell Deployment in 2-Tier HetNets. International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), Barcelona, Spain.

Abonyi D., &Onoh G., (2022). development of a dynamic cell coordination framework for interference mitigation in 4G networks. IJSRED; Vol 5, Issue 2; pp.692-698.

Abonyi D., (2019) A Novel Strategy for Prompt Small Cell Deployment in Heterogeneous Networks. Advances in Science, Technology and Engineering Systems Journal; Special Issue on Advancement in Engineering and Computer Science; ASTES Journal; ISSN: 2415-6698; Vol. 4, No. 4, 265-270

Abonyi D., Jonathan M., (2016) Localisation system for network planning in 2-tier heterogeneous networks. Progress In Electromagnetics Research Symposium (PIERS), Shanghai, China.

Azubogu A., et al (2010) Empirical-Statistical Propagation Pathloss Model for Suburban environment of Nigeria at 800MHz band", The IUP Journal of Science and Technology, India.

Bahceci I., (2014). Femto-cell deployment strategies for randomly distributed hotspots in cellular networks. IEEE Wireless Communications and Networking Conference (WCNC). IEEE, pp. 2330–2335. [Online]. Available: http://ieeexplore.ieee.org/ document/6952713/

Chataut R., & Akl R., (2020) Massive MIMO Systems for 5G and Beyond Networks-Overview, Recent Trends, Challenges, and Future Research Direction. Sensors (Basel). 2020 May 12;20(10):2753. doi: 10.3390/s20102753. PMID: 32408531; PMCID: PMC7284607.

Enyi V., Hyginus U., Eze F., Ugwu C., & Chidubem C. (2021). Path Loss Model Predictions for Different GSM Networks in the University of Nigeria, Nsukka Campus Environment for Estimation of Propagation Loss. International Journal of Advanced Research in Computer and Communication Engineering; Vol. 10, Issue 8, ISSN (O) 2278-1021, ISSN (P) 2319-5940; DOI 10.17148/IJARCCE.2021.10816

Ghosh A., Mangalvedhe N., Ratasuk R., Mondal B., Cudak M, Visotsky E, Thomas A., Andrews G., Xia P., Jo S., Dhillon H., &Novlan T., (2012). Heterogeneous cellular networks: From theory to practice. Communications Magazine, IEEE, 50(6):54–64, doi: 10.1109/MCOM.2012.6211486.

Hoadley J., &Maveddat P., (2012) Enabling small cell deployment with Het- Net. IEEE Wireless Communications, vol. 19, no. 2, pp. 4–5, apr 2012

Huang M. (2020). Theory and Implementation of Linear Regression. 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL), Chongqing, China, 2020, pp. 210-217, doi: 10.1109/CVIDL51233.2020.00-99.

Landstrom S., Murai H., & Simonsson A. (2011). Deployment Aspects of LTE Pico Nodes. IEEE International Conference on Communications Workshops (ICC). IEEE, pp. 1–5. [Online].Available: http://ieeexplore.ieee.org/document/5963602/

Nika A., Ismail B., Zhao Y., Gaito S., Rossi P., & Zheng H. (2014). Understanding data hotspots in cellular networks. in the 10th International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness. IEEE, pp. 70–76. [Online]. Available: http://ieeexplore.ieee.org/document/6928662/

Oluwatoki O., Popoola J., & Busari S., (2022) Empirical Path Loss Modelling for Selected LTE Networks in FUTA Campus, Ondo State, Nigeria" Journal of Engineering Advancements Vol. 04(01) 2023, pp 1-7 https://doi.org/10.38032/jea.2023.01.001

Park Y., Heo J., Kim H., Wang S., Choi T., Yu T., & Hong D., (2014) Effective Small Cell Deployment with Interference and Traffic Consideration. in 2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall). IEEE, pp. 1–5. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper. htm?arnumber=6965951

Qutqut M., Abou-zeid H., Hassanein S., Rashwan M., & Al-Turjman F. (2014). Dynamic small cell placement strategies for LTE Heterogeneous Networks. IEEE Symposium on Computers and Communications (ISCC). IEEE, jun 2014, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/document/6912536/

Seiamak V., Rahim T., & Marcin F. (2015). Small Cells for 5G Mobile Networks. In Fundamentals of 5G Mobile Networks, pages 63–104. John Wiley & Sons, Ltd, Chichester, UK. ISBN 9781118867464. doi: 10.1002/ 9781118867464.ch3. URL http://doi.wiley.com/10.1002/9781118867464.ch3.

Shelly S., Christian O., &Sangman M. (2012). Heterogeneous Cellular Networks Superior to Homogeneous Ones? Lecture Notes in Computer Science, pages 751361–68, URL https://hal.inria.fr/hal-01551342.

Su G., Li L., Lin X., & Wang H., (2014) On the optimal small cell deployment for energy-e_cient heterogeneous cellular networks. Sixth International Conference on Ubiquitous and Future Networks (ICUFN). IEEE, pp. 172–175. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper. htm?arnumber=6876775