# Optimum Positioning of Base Station for Cellular Service Devices Using Discrete Knowledge Model

Cristiane R. Gomes ⓘ, Igor R. Gomes ⓘ, Ramz L. Fraiha Lopes ⓘ, Herminio S. Gomes ⓘ
Gervásio P. S. Cavalcante ⓘ
*Federal University of Pará- cris.ruiz.gomes@gmail.br,ruiz.igor@gmail.com, fraiha.ramz@gmail.com,*
*herminio.gomes@gmail.com, gervasiogmeio@gmail.com*

*Abstract*— **A good wireless network design depends on technical and financial viability and a number of other criteria that must be met. Following the emergence of new technologies and services, such as 5G transmission and the reuse of frequencies, new work is being carried out to ensure a better design for a particular area. This study examines a discrete radio propagation model which employs the K nearest neighbors classifier. The model takes into account the different characteristics of the environment. This article presents a case study for the optimum positioning of base stations in Federal University of Pará (Belém – Brazil), representing a typical Amazon environment. The mentioned scenario is heterogeneous, presenting edifications and considerable forest area. Measurement campaigns were conducted in three different frequencies for the design features of the model: 521 MHz (Brazilian digital TV system), 2100 MHz (Enhanced Data Rates for GSM Evolution), and 2600 MHz (Long Term Evolution). A study of the fading phenomenon in these frequencies was carried out to generalize the frequencies of application for the propagation loss model. When this model was ready, tests (computing simulations) were conducted in two scenarios to optimize the positioning of the radio base stations being studied.**

*Index Terms*— Antenna Positioning, KNN Classifier, Optimization, Radio Propagation Models, Refarming.

## I. INTRODUCTION

Currently, wireless network planning takes account of a number of factors such as: average/low radiation intensity, optimum positioning of new antenna transmitters and suitable models for signal intensity prediction. To meet all these requirements at the same time, the new antennas have less power which results in a lower coverage area; this requires new studies of signal propagation [1].

The big urban centers are exchanging large cell stations for small ones [2], this means that most of the studies of propagation are based in cities with a lot of high-rise buildings such as Tokyo, New York, Barcelona or São Paulo. These models are not suitable for signal prediction in cities where there is a combination of buildings and urban afforestation, like the environment examined in this study. In view of this, a particular study is needed not only to address the real circumstances of the Amazon towns and cities but also other cities characterized by equatorial forests or with a good deal of urban woodland.

On the premises of the Federal University of Pará (UFPA), and in several towns and cities in the Amazon, one often comes across shadow zones in various frequencies used in the service of telecommunications, especially in the cell phone services (links with, 3G, 4G, etc.) and digital television. Owing to shortcomings in telecommunication services, there has arisen the need and opportunity for the emergence of tools that are designed to improve this situation.

This article sets out an empirical model for outdoor mobile radio propagation applied to the optimum positioning of transmitter towers in an Amazon scenario. The main contributions made by this article are as follows: 1) Received signal strength prediction modelling in any area of a map through a generalized model and adapted to Amazon environments; 2) A methodology for finding the best positioning for transmitters with a view to maximizing the signal strength received by the network users.

## II. RELATED WORKS

Several modelling techniques have emerged with the aim of representing the phenomenon of electromagnetic wave propagation with its various nuances. It is worth citing some of the key studies - among these [3], [4], [5], [6] and [7] - which are related to this study and examine environments where there is a good deal of vegetation.

Ribeiro et al. [3] study the influence of vegetation in the 700 MHz band for Outdoor-to-Indoor paths. This work shows the loss of the signal (about 10 dB) and a bigger mean delay spread in environments whose path has more vegetation. The authors point out as the main causes or spreading and absorption of the signal.

Article [4] presents a three-layer deterministic model using Dyadic Green's Functions. In this, the authors emphasize differences in the received power for densely wooded urban scenarios and different climatic conditions in the UHF range.

The author of the study by [5] describes the main topics for planning the installation of new cellular technologies, in particular 5G. During the planning stage, the authors underline the importance of changing the large RBS (Radio Base Station) to a small cell format. They also highlight the need to optimize the positioning of the new RBS for better coverage, as well as propagation models designed for city maps, and the reuse of frequencies (refarming) among other factors.

In [6], the authors designed a propagation model for the Brazil digital TV band, by taking account of mixed pathways such as land-fresh water-land type in different seasons of the year, in this case, the Amazon summer and winter. The authors uses a machine learning model for the characterization of losses. The proposed model is a hybrid algorithm combining K nearest neighbors (KNN) classifier and knowledge-based agents where it uses attributes for each point of the scenario, according to the environment.

The study by [7] makes a comparison between the Okumura-Hata and COST231-Hata models, when planning the wireless network frequencies that are often used in the Long Term Evolution (LTE)

services. In its main results, it recommends the use of the COST231-Hata model when planning networks for transmitter towers with a height of at least 50 m.

In study [8], the authors modelled electromagnetic propagation on the frequency of 600 MHz by means of a hybrid ARIMA-ANN model. The received signal power was calculated as a function of the distance to the transmitter. In this work they analyzed part of the Brazilian DTV frequency range on a densely urbanized amazon city with equatorial climate (hot and wet). The results of this work suggests that the proposed modelling can be refined and, thus, applied widely in situations such as the one analyzed in the study.

In [9], the authors hold a discussion (review) about ray-tracing methods, their use in the world today and the opportunities for their use in the future by creating new algorithms. The authors recommend hybrid modelling by employing the ray-tracing method, together with empirical models and numerical methods in the planning of high frequency telecommunication services in "complicated propagation" conditions.

In [10], the authors calculate the value of the Signal to Interference Ratio (SIR) in LTE air-to-ground networks in low altitude flights. The results suggest that, for low altitudes, SIR is greater for macrocells than microcells. In higher altitudes, the case is the exact opposite.

In [11], a model for planning LTE networks is put forward in maritime regions by means of combinations of transmitters when they are not further than 100 km from each other.

The study carried out in [12] shows correlations between several propagation models found in the literature and applied to LTE networks in the city of Bogotá (Colombia). The results of this study show that geometric models represent the best propagation in the environment that is being analyzed.

The study carried out in [13] makes comparisons between the human and machine methods of learning based on the non-Markovian decision algorithm – that is, it depends on more than the immediately preceding state. The results show that the algorithm that follows these preconditions is able to faithfully represent the human learning curve.

In [14], the authors set out ways of improving precision and lowering the computational cost of this last category of algorithms by comparing the results of statistical methods with methods based on the theory of machine learning.

## III. MATERIALS AND METHODS

Initially, measurement campaigns were carried out in the 500 MHz, 2100 MHz and 2600 MHz bands with the aim of creating a database that can be implemented for the designed model. The three sets of data were used for a parabolic fitting to obtain the estimated frequency curves (700 MHz, 1800 MHz and 2400 MHz). Following this, these data were used as entries from the model which accepts heterogeneous routes. Finally, the optimum positioning was obtained for the transmission towers through an algorithm based on the KNN classifiers [15] to resolve a complete problem of combinatorial optimization [16].

The methodology adopted in this study is represented in Fig. 1 through a flowchart. Each of the stages of the flowchart will be described in detail in the following subsections.
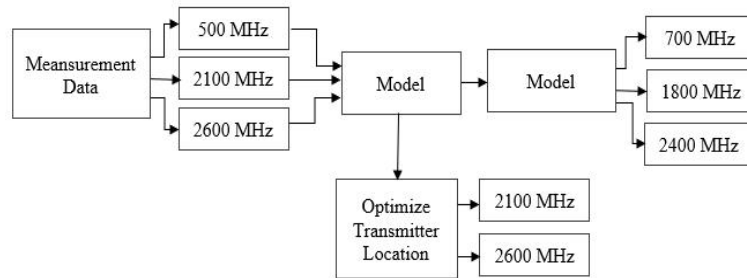


**Fig. 1.** Flowchart of the methodology

IV.  MEASUREMENT CAMPAIGN

*A. Measurement Campaign at 500 MHz:*

A dipole antenna Anritsu MPP651A, which picks up signals from frequencies between 470 MHz and 1700 MHz, gain 0 dB, was used to carry out the measurements. A calibrated portable spectrum analyzer (Anritsu Site Master S332E), a RGC 213 coaxial cable with a characteristic impedance of 50 Ohms and 3 m in length, was used to connect the analyzer to the antenna. A laptop and flash drives for the storage of the collected data were also used. Data were collected from 84 points in the city of Belém-PA at 521.14 MHz. Table I shows the main characteristics of the transmitter antenna. The spectrum analyzer provided 500 data for each measured point. These went through a data processing in which points with more than 2 standard deviations from the mean were excluded, thus removing the outliers.

TABLE I. TVD TRANSMITTING ANTENNA INFORMATION

| Location | Height (m) | Frequency (MHz) | Operation Power (W) | Effectively Radiated Power (ERP) (kW) |
|---|---|---|---|---|
| 01º27'43"S/48º29'28"O | 114 | 518-524 | 6000 | 52.15 |

*B. Measurement Campaign at 2100 MHz and 2600 MHz:*

Measurements at 2100 MHz and 2600 MHz were made on the premises of the UFPA for two reasons: 1) the transmission antennas of mobile phone systems radiate much less power than the towers that have lower frequencies, such as digital television; 2) the sites chosen for the installation of new towers must be within UFPA, or in a surrounding area situated less than 2 km from UFPA. Table II shows the main characteristics of the transmitter antenna used as a reference-point for these measurements.

TABLE II. CELLULAR TRANSMITTING ANTENNA INFORMATION

| Location | Height (m) | Frequency (MHz) | Operation Power (W) | Gain (dB) |
|---|---|---|---|---|
| 01º27'51,3"S 8º26'56.6"W | 40 | 2100 and 2600 | 40 | 15.4 |

Measurements were made at 24 points with the receiver placed in two positions: at a height of 1 m

and 2 m respectively. It continued to take measurements for 2 minutes at each point and for each of the frequencies, resulting in a total number of 60 samples per point. The average values of the samples of received signal strength were used for the estimates.

The measurements were made with the aid of the Cell Signal Monitor® software, which obtains the received signal strength at configurable intervals of time. This software also identifies the network and monitors the upload-download speed.

The measurement points were not spread out in radials since the area of UFPA makes this kind of deployment impossible because it would involve the existence of points outside the premises of the university. Two identical mobile devices connected to different networks (EDGE and LTE) were used to obtain the cellular signal.

Figure 2 shows a satellite image of the UFPA area and immediate vicinity, which are duly indicated. All the information regarding the transmission tower is available in [17].
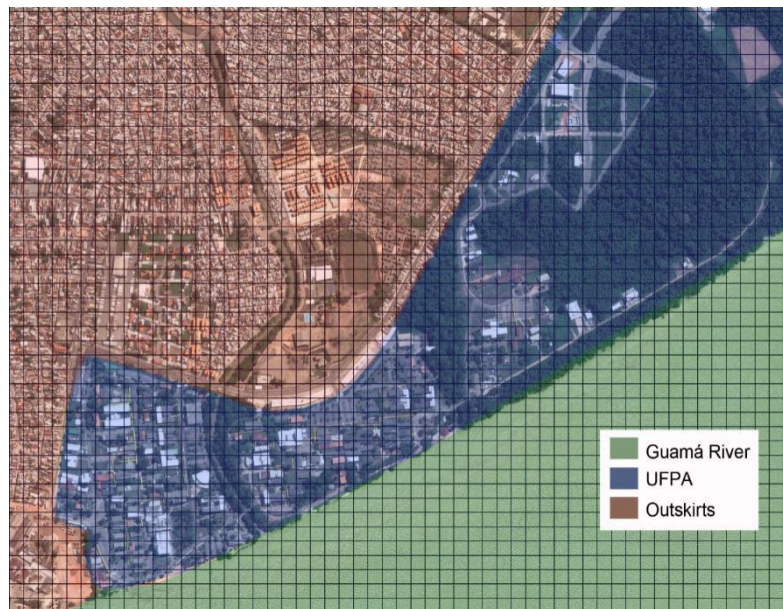


Fig. 2. UFPA and outskirts. Satellite image adapted from [18].

## V. ELECTROMAGNETIC WAVE PROPAGATION MODEL

The planned model takes account of the frequency of transmission. A study of the pattern of frequencies was needed with a view to determining a polynomial so that the model could be used for the frequencies that had not been measured.

Propagation trend curves were designed for the frequencies that were measured. The graph in Fig. 3 shows these kinds of curves. It should be noted that there is a greater degradation of the quality of the signal in higher frequencies.

On the basis of three trend curves, a curve can be found for a fourth frequency The trend curves are of the form expressed in (1):

$$y = a + b \, ln(x), \tag{1}$$

Where $y$ is the trendline of the received power for each frequency, $x$ the distance to the reference-point and $\boldsymbol{a_i}$, $\boldsymbol{b_i}$ , $i = 1,2,3$ are parameters to be determined by linear least square method. A fourth curve $y_4 = a_4 + b_4 \, ln(x)$ can be determined using parabolic fitting. Using the three parameters $a_i$, $i = 1,2,3$, it was calculated $a_4$. Then, using a second parabolic fitting with the parameters $b_i$, $i = 1,2,3$ it was calculated $b_4$. This fourth curve represents the trend line of a new frequency.
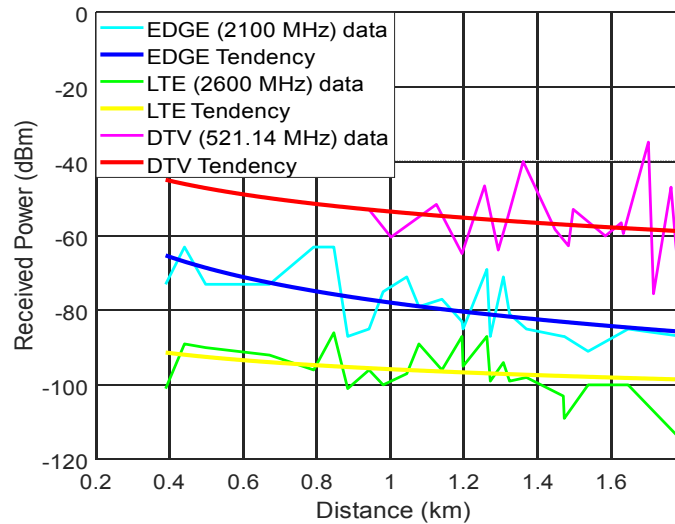


Fig. 3. Measured data and its trend curves.

Changes in technology in recent years have triggered a process of reusing the frequencies, which has begun to occur throughout the world since some services have begun to be deactivated. This process is known as refarming. In Brazil, it is followed because part of the population have mobile phones that cannot have access to 4G services. Refarming in Brazil is being carried out in frequencies of 700 MHz and 1800 MHz. Estimates for trend curves in these frequencies, and the frequency of 2400 MHz are made to test the model. If the estimated curves behave in the expected way, in accordance with electromagnetic theory, the trend model can be regarded as satisfactory.

Figure 4 shows the estimated curves for the frequencies of 700 MHz, 1800 MHz and 2400 MHz together with the trend curves of the measured data (500 MHz, 2100 MHz and 2600 MHz) so that an assessment can be made of these trend curves.
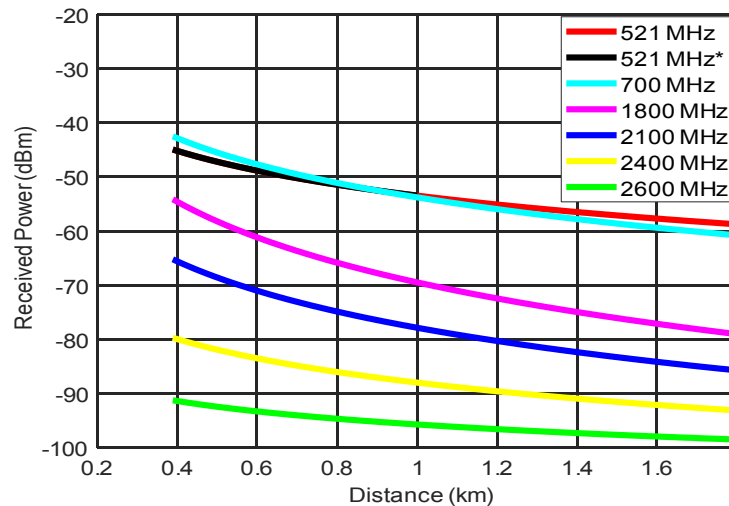
Fig. 4. Measured and estimated trend curves.

The curve for the 521 MHz trend was divided in two parts. The first one, in black, represents the estimated trend until 1 km from the transmitter. There were no measured data on this frequency at distances lower than 1 km from the transmitter. From this distance on, the red curve represents the trend obtained from the actual measured data in 521 MHz. When the curve turns red, it is possible to verify that the curve's decay is within the expected, since it is above the 700 MHz curve, thus confirming the quality of the measurements in this study. The degradation of the signal at a frequency of 1800 MHz also takes place in a predictable way since the curve lies within the curves of 2100 MHz and 521 MHz. The testing of 2400 MHz was conducted to find out if the curve could be adapted between the curves of 2100 MHz and 2600 MHz. Thus it should be noted that the estimated curves follow the expected pattern. Thus, it is possible to proceed to the stage of applying the equation of the trendline when calculating the loss of propagation.

## VI. PROPOSED TOOL

### A. Environment Discretization and Classification:

To design the model, the UFPA map was divided into squares with sides that measured 50 m. The discretization of the scenario was necessary so that the position of the measuring point(s) and transmitter(s) could be used in a matrix manner. Fig. 5 displays a satellite image with the mesh (discretization), transmission tower (red circle), measurement sites (blue squares) and candidate sites for the new towers (yellow triangles). The region of the map that was examined in this study is bordered by diagonal sections (straight lines in a light purplish red).
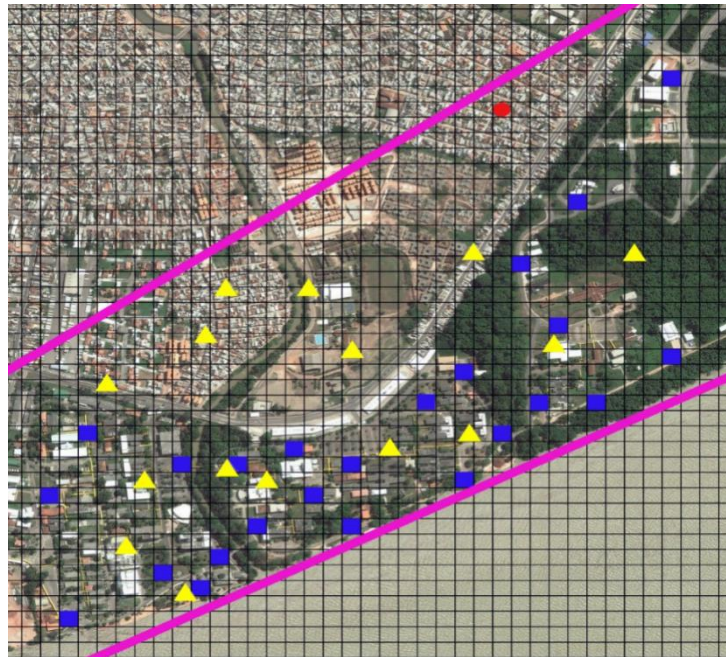
435



Fig. 5. Measured and candidate points in UFPA.

The size of the side of the square can be modified, although, if it is too large, the degree of precision might be impaired, and if it is too small, it might affect the feasibility of the model in terms of computing costs during the stage when the optimal positions are classified. The two purple lines (see Fig. 5) delimit the territory only to the studied area.

The area within the diagonals had a total of 652 squares. Each of the squares was given an attribute valued from 1 to 4, depending on the type of environment contained in it. The values provide the following information: 1) Bare – generally streets, areas of undergrowth or without any vegetation; 2) Afforestation – characterized by dense green areas; 3) Buildings – characterized by buildings and brickwork structures; 4) Afforestation + Buildings– a combination of environments 1) and 2). As each unit of checked mesh has a side of 50 m, there are squares which, in effect, have more than one characteristic. This is one of the aspects of the planned methodology that is susceptible to refinement.

The list of attributes can be modified because the knowledge model allows other categories to be included or those already included to be replaced. However, it should be noted that if many attributes are included, it reduces the reliability of the calculations.

Figure 6a shows an aerial picture of the analyzed region (and its outskirts) with the square grid used for the logical handling. Figure 6b shows the same discretized logical map, but with different colors for each attribute of each square of the analyzed region: black for "bare", green for "afforestation", ash gray for "buildings" and blue for "afforestation + buildings". The map in Fig. 6b was designed to allow a visual comparison to be made with the map of Fig. 6a.
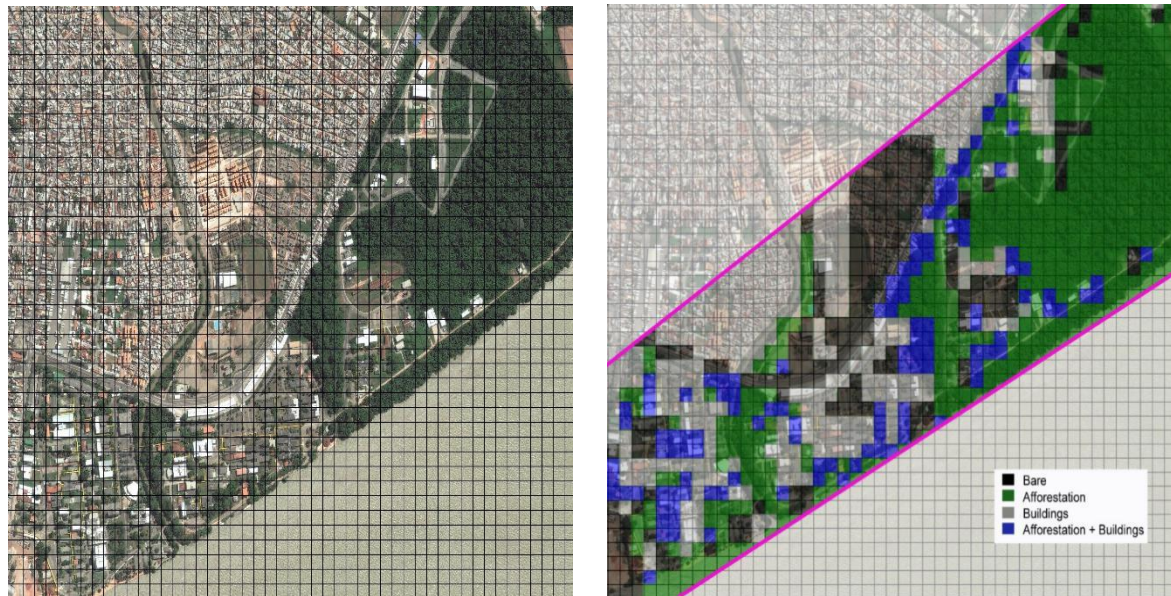
Fig. 6. (a) Aerial picture of studied region and outskirts; (b) Colored map of environmental attributes.

The inclusion of the values of the attributes was done manually – that is, the information about the 652 squares was inserted in the position matrix one by one. The delineation of the environment was carried out in a visual manner, starting from the satellite image and in loco. Thus it is possible that different people may analyze and classify the square in different ways. Despite this, in the terms established at this time, it is unlikely that possible variations affecting the classification of the squares, will significantly differ from the results shown here.

### B. Counting and Classification:

The calculation of received signal strength at each point depends on the total sum of losses in each square. Each measured point is at a different distance from the RBS, which means there is a need to determine a) how many and b) what kinds of squares were crossed.

The loss function or received strength depends on the number of points crossed and hence, the distance. Thus, this function can be expressed by the formula in (2):

$$Loss_{Ppoint} = F\left(\sum Loss\ by\ crossed\ squares\right) \tag{2}$$

In this study, the F function used was the identity function. The model can be improved by using another function which might depend in an explicit way, on the environment or even the distance. In other words, only the losses suffered in each square are taken into account.

The Euclidian distance was used to measure the distance from the transmission tower to the measured points. Since the scenario consists of a map with a discretization mesh, a straight line between two arbitrary points on the map can have a jagged or serrated (i.e. aliasing) line, in the shape of a stairway. What matters for the purposes of this study is to count how many and what types of squares were crossed by the abovementioned line, which connects the transmitter to the measured point.

In Fig. 7, there is a transmission tower and three receivers at different points in a way that shows

how the squares are counted, as well as how the "straight line" formed by the squares will appear. Each color along the straight line indicates the pathway followed by the signal to the different receivers. The appearance of the "stairway" should be noted in some of the straight lines.
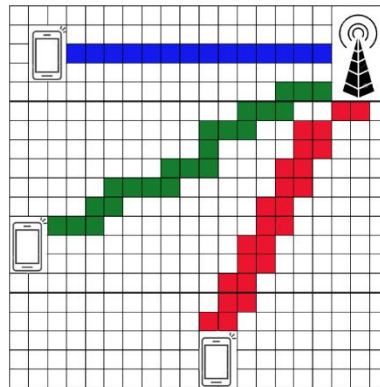


Fig. 7. Example of straight lines on a discrete map.

A linear system of equations was formed by means of the measurement data and the number of each type of square used. Then there is an equation of the system for each measured point. 24 measured points, implies there are 24 equations. In Fig. 7 each path forms an equation. The equations of the system are of the form shown in (3).

$$q_{i,1}\alpha_1 + q_{i,2}\alpha_2 + q_{i,3}\alpha_3 + q_{i,4}\alpha_4 = b_i , \qquad i = 1, \dots, 24 \qquad (3)$$

with: $\alpha_j$: losses in each type of square to be calculated (intrinsic loss); $q_{i,j}$: number of each type $j$ of the square crossed in Equation $i$; $b_i$: signal strength received at point $i$ (measured data).

The losses were calculated from the reference power close to the transmission tower, together with the trend (1). The system shown is incompatible of full column rank. For this kind of system, a Linear Least Squares (LLS) method can be applied to solve it. The formal solution is given by (4). Table III shows the results calculated for each α.

$$A^t \times (A\alpha = b)$$
$$\Downarrow$$
$$A^t A\alpha = A^t b$$
$$\Downarrow \qquad\qquad (4)$$
$$\alpha = (A^t A)^{-1} \times A^t b.$$

TABLE III. LLS SOLUTIONS.

| Square type | Intrinsic Loss EDGE (dB) | Intrinsic Loss LTE (dB) |
|---|---|---|
| $\alpha_1$: Bare | -0.49 | -0.80 |
| $\alpha_2$: Afforestation | -0.71 | -1.31 |
| $\alpha_3$: Buildings | -0.61 | -1.25 |
| $\alpha_4$: Afforestation + Buildings | -0.17 | -0.26 |

The results show the environment is predominantly of the "Afforestation" type. The loss of

"Afforestation" type was higher, followed by types of "Buildings" and "Bare". The value found for "Afforestation +Buildings" was the least of all because in the area under study, the buildings made up a lower fraction.

The measures were used to calculate the intrinsic losses of each type of square crossed, however in order to have a better accuracy of the values obtained, the KNN was used to classify the adjacent squares and corroborate with the visual classification made. With the KNN ready, it is possible to use the same model for different regions, as long as there is a specialist of knowledge and a set of data to feed the model since the KNN is a supervised learning technic.

The source code was designed in MATLAB, which sought flexibility when addressing the different scenarios for planning a wireless network and this allowed the following attributes to be aggregated: height of the transmitter, height of the receiver, operating frequency, transmission strength and number of candidate sites for the positioning of new tower transmitters.

All the solutions found make use of the collected measurements. Fig. 8a shows the levels of the power curves received for the 2100 MHz frequency, while Fig. 8b shows the levels of the power curves received for the frequency of 2600 MHz. These results already take account of the loss model with equations such as (4).



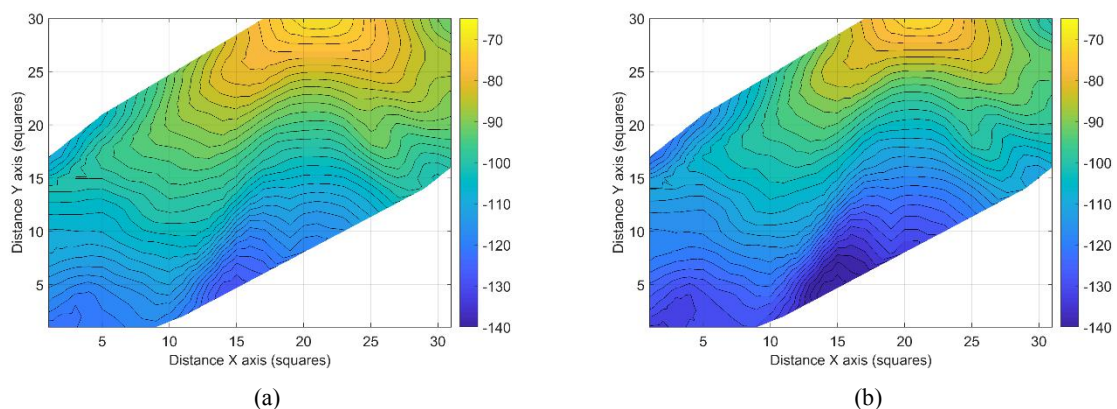(a)                                                      (b)

Fig. 8. EDGE and LTE contour lines

The weakening of received signal strength is evident from the distancing with regard to the tower transmitter. Regions near to the tower are represented in yellow. The units of the axes in the graphs refer to the number of squares in the mesh of the UFPA map. In the region near the (20,20) point, there is a large wooded area within UFPA, which is evidence of a larger loss in this zone, although there are regions at a further distance with a greater signal strength than the areas of the point (20,20). The model makes a prediction of the signal at each point and does not take account of the sensitivities of antennas from different devices. There are, thus, points in Figure 8b (2600 MHz) in which the reception of the signal attained values of the order of -140 dBm. In practice, most devices do not achieve connectivity at this level of received strength. One can find a pattern of behavior similar to the frequency of 2100 MHz, but with a more rapid fading, which is expected in higher frequencies.

This phenomenon is noted in the colour tones of the graph that are predominantly darker than the previous one.

## VII. RESULTS AND DISCUSSION

In this section, there is a discussion of the results obtained from optimizing the position of the tower transmitters within the territorial region of UFPA for maximizing the received signal strength in the map being considered. Two scenarios were simulated that included the frequencies of 2100 MHz and 2600 MHz. The scenarios examined are as follows: 1) two transmission towers – in this scenario, no account is taken of the influence of the already existing tower near to UFPA; 2) three transmission towers - one of them being in the immediate vicinity of UFPA, the same from which the measurement data were collected.

Fifteen candidate sites were chosen inside of UFPA, for the positioning of new transmission towers. These sites are the same in both of the scenarios employed. However, in the program that was carried out, both the number and position of the candidate sites could be altered. It was decided to restrict the number of candidate sites with a view to making it a more realistic project.

The value of the received signal strength /loss originating from each of the interior towers of UFPA was calculated for each mesh square (there are 652 altogether), on each of the possible combinations for the location of these towers. As there are 15 candidate sites for the positioning of the two transmission towers, the total number of possible scenarios is given by the value $C_{15,2} = 105$. The time needed for the optimization of each scenario is approximately 20 minutes. It is essentially a question of addressing a complete problem of combinatory optimization (NP-hard) in which all the solutions have to be calculated to determine the optimal solution. This problem is equivalent to find the k greatest numbers in a given list of numbers.

### A. Scenario 1 - Two Towers

The scenario of two towers can be interpreted as a brand new coverage problem. The criterion for the optimization was the received signal strength, in all the points, being maximized. This criterion seeks to maximize the coverage area, since in this configuration, no other transmission towers in the surrounding area are included.

Maps were designed for received signal strength at frequencies of 2100 MHz and 2600 MHz, which were the frequencies used in the measurement campaigns. Fig. 9a shows the result of optimization in the case of two towers with a frequency of 2100 MHz and Fig. 9b for the frequency of 2600 MHz.

The centralization of at least one of the towers was expected but it would not necessarily be the best configuration, because the model takes account of the losses for different environments. It is natural that the proximity of the set of squares would lead to greater losses because if the positioning was far from this set, the signal would be weak for this group.

The map obtained for 2600 MHz has the same positioning of the towers as the map for 2100 MHz, but as the fading of the signal is greater for higher frequencies, the power received ends up by being
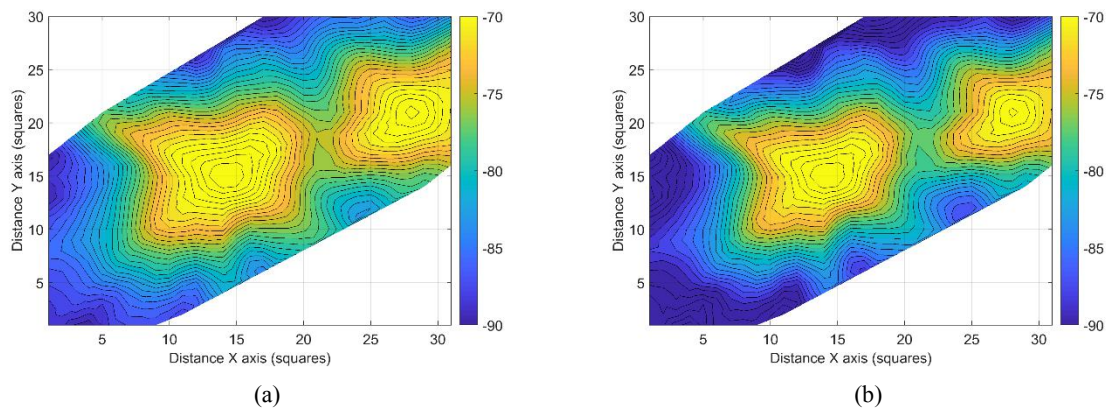
less than the case of 2100 MHz.



(a)                                                    (b)

Fig. 9. Two towers optimization.

### B. Scenario 2 - Three Towers:

The same criterion of maximum coverage was applied to the three towers. In this scenario, the tool optimizes the positioning but does not necessarily ensure the maximum coverage; this is mainly due to the existing fixed tower (the same transmission tower from which the measurement data were collected). To ensure that the maximum coverage area could, in fact be found, this would begin to be a multi-objective optimization, which is beyond the scope of this study. Even so, owing to the existence of "shadow zones" in UFPA, where it is not possible to achieve connectivity in the frequencies examined (and often even at lower frequencies), it was decided to display this scenario, in which the number of "shadow zones" would be reduced.

Figures 10a and 10b show the maps of coverage in the scenario of three towers for the frequencies of 2100 MHz and 2600 MHz, respectively.



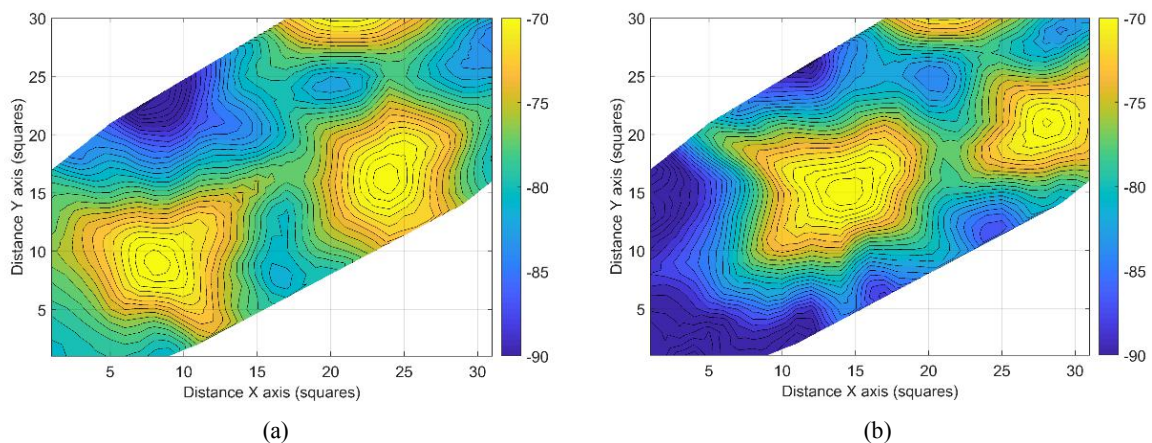(a)                                                    (b)

Fig. 10. Three towers optimization.

In this scenario, for 2100 MHz, the towers were not centralized. This is owing to the presence of the third tower, which serves to provide coverage for the north-eastern part of UFPA, and is involved in the displacement of the two others. As there is no kind of priority of service, the solution found for this scenario was to serve the needs of a public outside UFPA.

The small isolated areas of poor reception between the tower outside UFPA and the two others, are in the region where there is dense vegetation. It should be noted that there is also a region on the left of the map used for this study with a low level of received signal strength. When this is compared with the satellite image, it can be seen that its region includes buildings with lecture halls and administrative buildings. Hence, it can be concluded that although the well covered area is the best possible place for the candidate sites considered, in this scenario there is room for an improvement of the results. If there are future improvements with regard to the importance of each square mesh, it is expected that there can be a significant advance of the results shown here, particularly in this scenario which takes account of the three transmission towers.

## VIII. CONCLUSIONS

This study provided a planning tool for the telecommunications network services which makes use of the KNN classifier in a discretized form and a modelling for the intensity of signal strength in accordance with distance. Measurement campaigns were carried out in different frequencies to test the flexibility of the model for different frequencies. In addition, an attempt was made to optimize the positioning of new transmission towers for different scenarios with the aid of the designed tool.

The tool has several innovative features. The first of these is its hybrid character and the fact that it employs supervised machine learning in a theoretical/computational approach which makes it possible to handle different attributes of an environment.

Another innovative feature is the inclusion of a generalization for the use of frequencies based on a parabolic fitting of received signal strength in terms of the distance to the transmitter. Although the planned tool does not have a single defined mathematical formula, the algorithm that was created makes use of a lot of information originating from formulations that were well defined in their development.

The results of the model for propagation loss used in this planned tool, is proving to be suitable for the prediction of received signal strength in each point of the scenario, which distinguishes it from most propagation models which have a point by point characteristic.

Features of the environment were added to the model by the KNN classifier, where the attributes chosen represent a typical environment of an Amazon town or city. Since the incorporation of attributes depends on a knowledge specialist (of supervised machine learning), the model can be adapted to other types of environments.

The proposed tool can be used in planning signal coverage in various types of regions. The candidate sites can be chosen with greater flexibility if they are regarded as small cells, and provided that a) they have a post with a height of 15 m as their format, b) they incur low installation costs; this will thus increase the number of candidate sites.

Our intention in future studies is to include the following in the tool: 1) Important qualitative attributes of coverage, in which the question of the mesh squares will be given priority. As a result,

there will be a methodology of multi-objective optimization to ensure a greater coverage area, which will not only take account of the average value of received signal strength but also the importance of the squares on the terrain in question; 2) measurement campaigns in different frequencies, particularly the possible frequencies in 5G, can be carried out and, thus, data can be obtained to improve the model for received strength with regard to the generalization of the application; 3) Implementation of the Nelder-Mead method (in its binary variation) in the optimization stage and calculation of the results, with a view to refining the techniques for obtaining results; 4) Finally, the throughput of data can be associated with the power in the coverage maps.

## ACKNOWLEDGMENTS

## REFERENCES

[1] P. Muñoz, O. Sallent, J. Pérez-Romero, "Self-Dimensioning and Planning of Small Cell Capacity in Multitenant 5G Networks", *IEEE Trans. on Veh. Tech.*, vol. 67, no. 5, pp. 4552-4564, May. 2018.

[2] W. Ni. et al, "Graph theory and its applications to future network planning: software-defined online small cell management", *IEEE Wireless Commun.*, vol. 22, no. 1, pp. 52-60, Feb. 2015.

[3] L. G. Ribeiro, L. J. Matos, P. V. G. Castellanos, M. B. Moura, V. L. G. Mota, and W. D. T. Meza, "Influence of Vegetation on the Outdoor-to-Indoor Mobile Radio Propagation in 700 MHz Band", *JMOE*, vol. 18, no. 3, pp. 427-438, Apr. 2020. https://doi.org/10.1590/2179-10742019v18i31628

[4] C. R. Gomes, D. K. N. Silva, J. Araujo, H. S. Gomes, G. P. S. Cavalcante, "Radio-Wave Propagation Model for UHF Band in Different Climatic Conditions with Dyadic Green's Function", *J. Microw., Optoelectronics and Electromagn. Appl.*, vol. 14, pp. 60-72, 2015. doi: 10.1590/2179-10742015v14i1427

[5] A. Taufique, B. Jaber, A. Imran, Z. Dawy, E. Yacoub, "Planning Wireless Cellular Networks of Future: Outlook, Challenges and Opportunities," *IEEE Access*, vol. 5, pp. 4821-4845, 2017.

[6] I. R. Gomes, C. R. Gomes, H. S. Gomes, G. P. S. Cavalcante, "Empirical radio propagation model for DTV applied to non- homogeneous paths and different climates using machine learning techniques," *PLoS ONE* 13(3), 2019. e0194511. doi: 10.1371/journal.pone.0194511.

[7] N. S. Nkordeh et al. "LTE Network Planning using the Hata-Okumura and the COST-231 Hata Pathloss Models," presented at the World Congr. Engineering, London, 2014.

[8] R. L. Fraiha Lopes, S. G. C. Fraiha, H. S. Gomes, V. D. Lima, G. P. S. Cavalcante, "Application of Hybrid ARIMA and Artificial Neural Network Modelling for Electromagnetic Propagation: An Alternative to the Least Squares Method and ITU Recommendation P.1546-5 for Amazon Urbanized Cities," *Int. J. of Antennas and Propag.*, vol. 2020, Article ID 8494185, 12 pages, 2020. doi:10.1155/2020/8494185.

[9] Z. Yun, and M. F. Iskander, "Ray Tracing for Radio Propagation Modeling: Principles and Applications," *IEEE Access*, vol. 3, pp. 1089-1100, 2015.

[10] C. Xuesong, et al. "Interference Modeling for Low-Height Air-to-Ground Channels in Live LTE Networks", *IEEE Antennas and Wireless Propag. Letters*, vol. 18, no. 10, 2019.

[11] P. Minyoung, et al. "LTE Maritime Coverage Solution and Ocean Propagation Loss Model", Int. Conf. Performance Evaluation and Modeling in Wired and Wireless Networks (PEMWN), 2017.

[12] J. E. A. Peña, "Correlation Analysis of Propagation Models for the Design of a LTE - A Network". IEEE Int.

Engineering Summit, II Cumbre Internacional de las Ingenierias (IE-Summit), 2016.

[13] A. M. Clarke, J. Friedrich, E. M. Tartaglia, S. Marchesotti, W. Senn, M. H. Herzog, "Human and Machine Learning in Non-Markovian Decision Making," *PLoS ONE* 10(4): e0123105. doi: 10.1371/journal.pone.0123105.

[14] S. Makridakis, E. Spiliotis, V. Assimakopoulos, "Statistical and Machine Learning forecasting methods: Concerns and ways forward". *PLoS ONE* 13(3), 2018. e0194889. doi:10.1371/journal.pone.0194889.

[15] K. P. Murphy, *Machine learning: A probabilistic perspective*. Cambridge, MA, USA: MIT Press, 2012.

[16] B. Korte et al. *Combinatorial optimization*. 3rd ed. Heidelberg, Germany: Springer, 2012.

[17] Brazilian National Agency of Telecommunications (ANATEL)

www.anatel.gov.br, (access Dec. 1 2019).

[18] Arc Gis World Imagery.

www.arcgis.com/home/webmap/viewer.html?useExisting=1&layers=10df2279f9684e4a9f6a7f08febac2a9, (accessed Dec. 6, 2019).