

Fuzzy Clustering Algorithm applied to the Radio Frequency Signals Prediction

Paulo Tibúrcio Pereira Glaucio Lopes Ramos

GAPEA - Antennas and Propagation Research Group, Federal University of São João del-Rei (UFSJ), Brazil,
{paulotiburciop,glopesr}@gmail.com

Abstract

In this work the Fuzzy Clustering technique was used to perform radio frequency signals prediction. This technique was used with georeferencing maps of topography and morphology for prediction of radio frequency power levels, at the region of Viçosa - MG. The performance of this method was evaluated through tests of propagation and mapping for a 879.660 MHz signal, used in cellular mobile telephony. This method of prediction showed excellent results in comparison with measurements of RF power levels with a success rate greater than the classical models of Okumura-Hata and Walfisch-Ikegami. Consequently, this method can be very useful to the telecommunications companies when making the RF cellular coverage prediction.

Keywords: RF prediction, fuzzy clustering, mobile telephony

1 Introduction

The electromagnetic wave propagation between transmitter and receiver antennas has its characteristics fundamentally defined by the transmission medium between them. The radio signal has to propagate with low distortion and the received power must be adequately above the noise level in order to be correctly decoded.

To analyse Radio Frequency (RF) signals we must consider the electromagnetic waves, but also the topography and morphology of the terrain and, in some cases, the meteorological, ionospheric and spacial waves conditions. If these RF waves propagate in free space conditions, with no occurrence of reflection, diffraction, refraction, absorption attenuation, we would have ideal conditions to obtain the signal prediction.

In fact, the parameters of the medium where the electromagnetic waves propagate are strongly dependent on the specific propagation area, as forests, deserts, lakes, mountains, buildings, cities, and they also frequently vary due to atmospheric conditions, as temperature, pressure, humidity and noise.

2 Classical RF Prediction

The power signal strength, for a specific frequency, can be determined using (1):

$$P_r = P_t - L_p, \quad (1)$$

where P_r is the received power signal [dBm], P_t is Effective Isotropic Radiated Power (EIRP) [dBm] and L_p computes the total losses between the transmitter and receiver [dB].

These power losses can vary significantly due to the environments and atmospheric conditions previously described. The most usual way to compute the diffraction power losses are the Fresnel-Kirchoff equations and classical propagation models are also usually implemented in RF software predictions, as: Okumura-Hata, Walfisch-Ikegami Lee, COST 231 Hata, and COST 231 Walfish-Ikegami (Parsons, 2000; Hata, 1980; Walfisch and Bertoni, 1988; Lee, 1980; Ikegami et al., 1984). The Okumura-Hata model is applied to urban and suburbs environments. The signal strength attenuation is obtained by (2):

$$L_p = 69.55 + 26.16 \log f_c - 13.82 \log h_t - a(h_r) + (44.9 - 6.55 \log h_t) \log d, \quad (2)$$

where L_p is the signal attenuation [dB], f_c is the frequency [150-1500 MHz], d is the distance between the base and mobile stations [1-20 km], h_r is the effective height of the base station [30-200 m] and $a(h_r)$ is the correction factor for mobile antenna height, as (3),

$$a(h_r) = (1.1 \log f_c 0.7) h_r (1.56 \log f_c 0.8), \quad (3)$$

where h_r is the effective height of the mobile station [1-10 m].

The Walfisch-Ikegami model is used to predict the received signal in urban and dense urban environments. The signal attenuation when in a line-of-sight (LOS) situation is estimated by (4):

$$L_p(LOS) = 42.6 + 20 \log f_c + 26 \log d. \quad (4)$$

When in a non-line-of-sight condition (NLOS), the attenuation is predicted by (5):

$$L_p(NLOS) = 32.4 + 20 \log f_c + 20 \log d + L_{diff} + L_{mult}, \quad (5)$$

where f_c is the frequency [MHz], d is the distance [km], L_{diff} accounts for the diffraction losses on rooftops and L_{mult} accounts for the multiple diffraction in multiple buildings (Ikegami et al., 1984; Walfisch and Bertoni, 1988).

The coefficients for the equations of these classical prediction models are based in experimental data and statistical analysis. Some problems arises due the fact that there are different propagation environments. The RF prediction in the traditional way, are carried out by adjusting the coefficients of the equations in order to adjust them to a particular region, aiming to decrease the prediction errors.

3 Fuzzy Clustering Prediction

The classic prediction models uses restricted and empirical approaches of the RF propagation problem in certain regions. As opposite, the Fuzzy Clustering (Besdek, 1981) prediction method does not use empirical equations for calculating the received RF power (Pereira, 2000).

Fuzzy technology has been used in various areas of knowledge, such as control, decision making, pattern recognition, prediction of time series and state estimation and also in RF prediction (Phaiboon, 2010; Pelusi et al., 2014, 2013; Pelusi, 2012). It circumvents the resolution of complex differential equations with multiple variables applying the fuzzy rules and the compositional inference.

To make RF predictions it was chosen the Viçosa/MG region, which has an irregular topography and diverse morphology, thus promoting the emergence of various phenomena that influence the propagation of RF signals. It was collected signal sample levels at a cellular frequency of 879.660 MHz. Some of these samples were used for training and others to test the success rate. In this work it was used as input variables for the Fuzzy Clustering processing (Besdek, 1981), some RF measured samples performed in the field and georeferenced maps of topography and morphology (Pereira, 2000).

The proposal of using the fuzzy method was to circumvent the difficult in the modeling process of the propagation physical problems, featuring an unknown RF propagation environment from a group of known measurements, providing continuous mapping of RF coverage throughout a region. The method of RF signal strength prediction using Fuzzy Clustering and mapping is illustrated in Fig. 1 (Pereira, 2000).



Figure 1. Fuzzy Clustering RF prediction technique (Pereira, 2000).

The first step of the method consists in measuring the RF signal. This was done through an automatic sampling process for the RF levels in some points of interest in the region under study. These measurements were performed

using an RF receiver coupled to a GPS within a moving vehicle, thereby linking each measured RF signal level with its longitude and latitude coordinates. The equipment used was the HP E7474A equipped with the RF receiver model E6452A controlled by the software Viper. This system was developed by Agilent and adjusted to perform RF power measurements by distance, every 10 meters of vehicle movements.

In the second step it was carried out the pre-processing, aiming the data standardization and formatting. Initially it was calculated the distances, in relation to the transmitting antenna, of the points at which the RF power measurements were carried out.

For each point it was associated its information of the altitude and the region type, extracted from georeferenced maps of topography and morphology, as shown in Figs. 2 and 3. For each morphological type it was associated a numeric value, as follows: Low vegetation (field) = 1, rural area (plantation) = 2, scarce buildings = 3, airport = 4, forest (low vegetation bush) = 5, suburban area = 6, water (fresh-water) = 7, urban area = 8 and forest (tree) = 9. The altimetry and morphology databases were in the terrestrial ellipsoid model DATUM SAD69, cylindrical projection LAT/LON and had the planimetric resolution of 30 meters.

The use of databases with low planimetric resolution may compromise the results of RF predictions. For the RF signals predictions with higher frequencies, where the wavelength is smaller, it will be required higher resolution databases in order to better characterize the propagation environment and thus preserve the RF quality prediction.

All these data form a numerical table, so that each measurement point has now information on: longitude, latitude, distance, altitude, region type and RF level. These are the input variables for the Fuzzy Clustering processing.

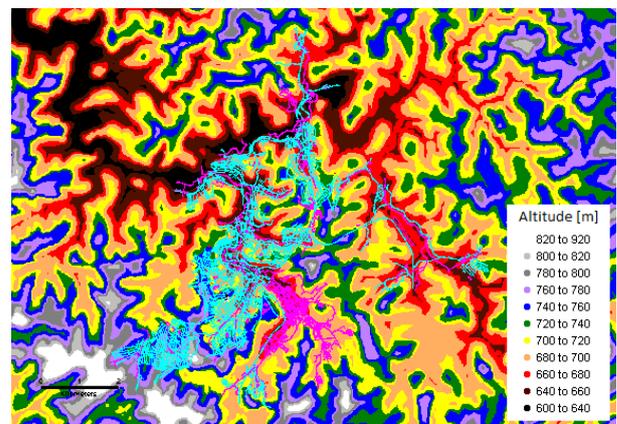


Figure 2. Georeferenced elevation contours map.

The variables previously obtained were used to train the Fuzzy Clustering process (Besdek, 1981), illustrated in Figs. 4 and 5, carried out in the two following distinct

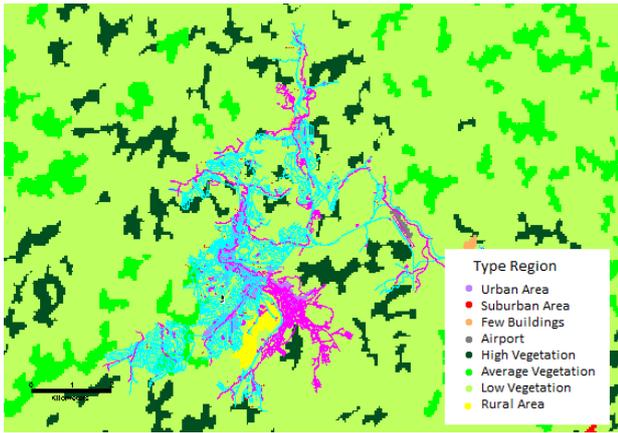


Figure 3. Georeferenced type region map.

steps.

In the first step, it was setted the number of fuzzy rules and their properties using the Grouping Estimation method that performs the following algorithm (Chiu, 1994; Yager and Filev, 1994):

1. A data vector with many neighboring data vectors have a high potential value;
2. The vector data with the greatest potential was the first assembly center;
3. An amount of potential from each data vector was subtracted as a function of distance from the first assembly center. The data vectors near from the first assembly center will have very limited potential, and therefore should not be selected as next assembly center;
4. It was selected the data vector with the largest remaining potential, as a second assembly center;
5. The number of clusters centroids was determined by (6):

$$P_k^* < \epsilon P_1^*, \tag{6}$$

where P_k^* is the potential value of the k -th cluster center, ϵ is a small decimal number between 0.15 and 0.50, and P_1^* is the potential value of the first assembly center.

In the second step, it was performed an optimization of the resulting rules. The Fuzzy Model Identification method was used to convert the parameters optimization of the resulting equations using a Linear Least Squares Estimation, as follows:

1. x_i^* are considered cluster centers, and will be the fuzzy rules that describe the system behavior;
2. x_i^* are decomposed in: y_i^* (input variables) and z_i^* (output variables). Being y the input vector, the value of relevance of the rule i is defined as (7):

$$\mu_i = e^{-\alpha \|y - y_i^*\|^2}, \tag{7}$$

where α is a constant. The output vector, z is calculated by (8):

$$z = \frac{\sum_{i=1}^c \mu_i z_i^*}{\sum_{i=1}^c \mu_i}, \tag{8}$$

z_i^* is considered a linear function of the input variables, as in the Takagi-Sugeno model (Sugeno, 1985):

$$z_i^* = G_i y + h_i, \tag{9}$$

where G_i is a constant matrix $(m - n) \cdot n$, and h_i is a constant column vector with $m - n$ elements. To convert the parameter optimization problem of (9) in a Linear Least Squares Estimation problem, it was defined:

$$\rho_i = \frac{\mu_i}{\sum_{i=1}^c \mu_i}. \tag{10}$$

Equation (8) can be rewritten as (11)

$$z = \sum_{i=1}^c \rho_i z_i^* = \sum_{i=1}^c \rho_i (G_i y + h_i), \tag{11}$$

or (12)

$$z^T = [\rho_1 y^T \quad \rho_1 \quad \dots \quad \rho_c y^T \quad \rho_c] \begin{bmatrix} G_1^T \\ h_1^T \\ \vdots \\ G_c^T \\ h_c^T \end{bmatrix}, \tag{12}$$

where z^T and y^T are line vectors. With a collections of n input data points y_1, y_2, \dots, y_n the resultant output collection is (13):

$$\begin{bmatrix} z_1^T \\ \vdots \\ z_n^T \end{bmatrix} = \begin{bmatrix} \rho_{1,1} y_1^T & \rho_{1,1} & \dots & \rho_{c,1} y_1^T & \rho_{c,1} \\ \rho_{1,n} y_n^T & \rho_{1,n} & \dots & \rho_{c,n} y_n^T & \rho_{c,n} \end{bmatrix} \begin{bmatrix} G_1^T \\ h_1^T \\ \vdots \\ G_c^T \\ h_c^T \end{bmatrix} \tag{13}$$

where $\rho_{i,j}$ denotes ρ_i valued in y_j .

The first matrix on the right side of (13) is constant, while the second contains all the parameters to be optimized. To minimize the squared error between the output model and the training data, it was applied to the problem the Linear Least Squares Estimation technique (Chiu, 1994).

In the Least Squares Estimation process, (13) was placed in the form (14):

$$AX = B, \tag{14}$$

where B is the output data matrix, A is a constant matrix and X is the matrix of the parameters to be estimated.

The pseudo-inverse solution that minimizes

$$\|AX - B\|^2 \tag{15}$$

is given by (16):

$$X = (A^T A)^{-1} A^T B. \tag{16}$$

The $(A^T A)^{-1}$ calculation is computationally time consuming when $A^T A$ is a large array. The size of $A^T A$ is given by $c(n+1) \cdot c(n+1)$. Numerical problems also arise when $A^T A$ is almost singular. To solve these problems it was used the Recursive Least Squares method and X is obtained via the iterative formula (17) (Chiu, 1994):

$$X_{i+1} = X_i + S_{i+1} a_{i+1} (b_{i+1}^T - a_{i+1}^T X_i) \tag{17}$$

and (18)

$$S_{i+1} = S_i - \frac{S_i a_{i+1} a_{i+1}^T S_i}{1 + a_{i+1}^T a_{i+1}}, \tag{18}$$

where X_i is the X value estimated at $i - th$ iteration, S_i is a covariance matrix $c(N+1) \cdot C(N+1)$, a_i^T is the $i - th$ row vector of A and b_i^T is the $i - th$ vector B . The Least Squares Estimation of X matches the X_n value.

After the training was performed the Fuzzy Clustering activation processing was activated with the Longitude, Latitude, Distance, Altitude and Region Type data. The training data are the blue crosses and the activation data are the green crosses, as illustrated in Fig. 6. The activation grid was obtained from topography and morphology georeferenced maps. The spacing between grid points is the planimetric resolution of the maps. The results of this step are the RF levels at each point of the grid, forming a continuous mapping (Pereira, 2000).

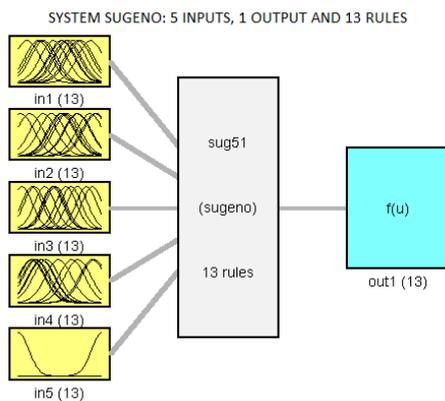


Figure 4. Processing Fuzzy Diagrams Blocks.

In the last step it was carried out the integration of the street map with the continuous mapping of RF levels. The RF levels were related with colors. These colors were superimposed on the streets in the region under study in order to facilitate the RF level preview of the important

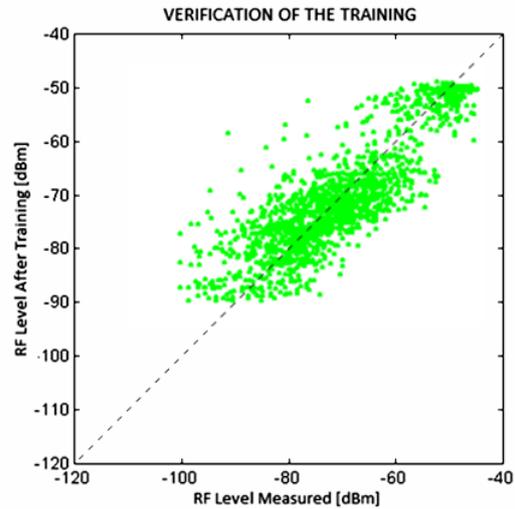


Figure 5. Verification Training.

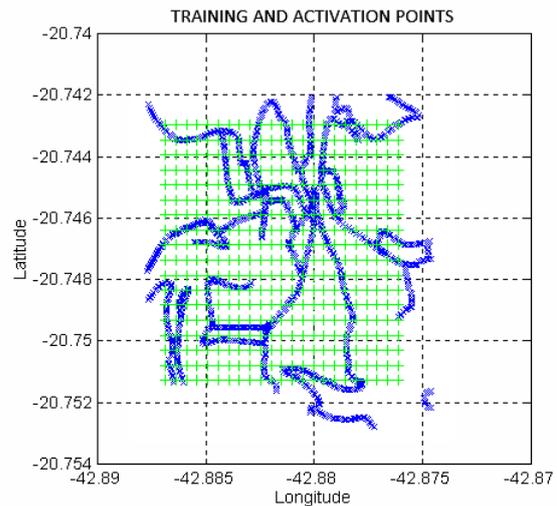


Figure 6. Training and Activation Points.

points. As a result, it was obtained the RF coverage map for the study area as in Fig. 7.

4 Analysis

The results of the RF Fuzzy Clustering Prediction method were compared with the results of the conventional prediction methods. To this end, it was used two measured data sets that were not used in the fuzzy prediction process. These two measured data sets were compared with the data obtained from the predictions. The mean absolute errors from the realized prediction methods can be seen in Table 1.

In Table 1 it is presented the comparison between the hit rate using the Fuzzy Clustering Prediction method and empirical techniques. The Fuzzy Clustering Prediction method was applied in three different scenarios: 3 (Longitude, Latitude and Distance), 4 (Longitude, Latitude, Distance and Altitude) and 5 (Longitude, Latitude, Distance,

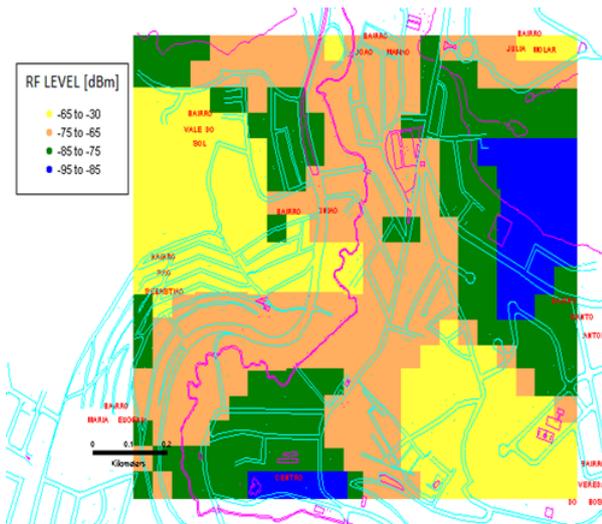


Figure 7. RF prediction with the Street Map.

Altitude and Region Type) input variables. It is evident the superiority of Fuzzy Clustering method in relation to the Okumura-Hata and Walfisch Ikegami classical methods.

Table 1. Arithmetic mean of the absolute errors of predictions

RF Prediction Method	Data Set 1 [dB]	Data Set 2 [dB]
Okumura-Hata	9.8596	18.0591
Walfisch-Ikegami	16.7718	10.8308
Fuzzy Clustering 3 input variables	5.9252	10.8028
Fuzzy Clustering 4 input variables	5.9537	9.5681
Fuzzy Clustering 5 input variables	5.9645	8.2682

5 Conclusions

The use of altimetry, morphology database and RF geo-referenced sample levels can characterize the propagation environment and in conjunction with the fuzzy logic processing provides small errors in RF predictions.

This method of prediction showed excellent results in comparison with the actual RF power levels, with a success rate greater than some classical models as Okumura-Hata and Walfisch-Ikegami.

Using the prediction method developed in this work it was generated a continuous RF coverage mapping integrated to the street layout of the study area, becoming an application of fuzzy logic in the telecommunications area.

The method described in this paper can be used by wireless telecommunications companies to perform RF predictions quickly and accurately. It also includes carries operating with different frequencies or other regions with different characteristics.

The Fuzzy Clustering RF prediction process, described in this work, together with high resolution databases may

assist in the deployment of future mobile telecommunications networks.

Acknowledgment

This work has been supported by the Brazilian agency CAPES/PROCAD 068419/2014-01.

References

- J. C. Besdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, 1981.
- S. Chiu. Fuzzy model identification based on cluster estimation. *Journal of Intelligent & Fuzzy Systems*, 2(3), 1994.
- M. Hata. Empirical formula for propagation loss in land mobile radio services. *IEEE Transactions on Vehicular Technology*, 29(3):317–325, August 1980.
- F. Ikegami, S. Yoshida, T. Takeuchi, and M. Umehira. Propagation factors controlling mean field strength on urban streets. *IEEE Trans. On Antennas and Propagation*, 32(8):822 – 829, August 1984.
- W. C. Y. Lee. Studies of base-station antenna height effects on mobile radio. *IEEE Trans. on Vehicular Technology*, 29(2): 252–260, May 1980.
- J. D. Parsons. *Mobile Radio Propagation Channel*. Wiley, 2000.
- D. Pelusi. PID and intelligent controllers for optimal timing performances of industrial actuators. *International Journal of Simulation: Systems, Science and Technology*, 13(2):65–71, 2012.
- D. Pelusi, L. Vazquez, D. Diaz, and R. Mascella. Fuzzy algorithm control effectiveness on drum boiler simulated dynamics. *36th International Conference on Telecommunications and Signal Processing*, pages 272–276, 2013.
- D. Pelusi, M. Tivegna, and P. Ippoliti. Intelligent algorithms for trading the euro-dollar in the foreign exchange market? *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, pages 243–252., 2014.
- P.T. Pereira. Predição de sinais de radiofrequência utilizando lógica fuzzy. Master's thesis, CEFET/MG, 2000.
- Supachai Phaiboon. RF Macro-cell Prediction Using Fuzzy Logic: Case study in Bangkok City - Thailand. In *Proceedings of the 4th International Conference on Communications and Information Technology*, CIT'10, pages 105–109, Stevens Point, Wisconsin, USA, 2010. World Scientific and Engineering Academy and Society (WSEAS). ISBN 978-960-474-207-3. URL <http://dl.acm.org/citation.cfm?id=1864098.1864117>.
- M. Sugeno. *Industrial applications of fuzzy control*. Elsevier Science Pub. Co., 1985.
- J. Walfisch and H. L. Bertoni. A theoretical model of UHF propagation in urban environments. *IEEE Transactions on Antennas and Propagation*, 36(12):1788–1796, December 1988.
- R. Yager and D. Filev. Generation of fuzzy rules by mountain clustering. *Journal of Intelligent & Fuzzy Systems*, 2(3):209–219, 1994.