# Generalization Capabilities Enhancement of a Learning System by Fuzzy Space Clustering

Zakaria Nouir, Berna Sayrac and Benoît Fourestié
France Telecom, Issy-les-Moulineaux, France
E-mail: zakaria.nouir@orange-ftgroup.com

Walid Tabbara and Françoise Brouaye
LSS Supélec, Gif sur Yvette, France
E-mail: tabbara@supelec.fr

***Abstract*— We have used measurements taken on real network to enhance the performance of our radio network planning tool. A distribution learning technique is adopted to realize this challenged task. To ensure better generalization capabilities of the learning algorithm, a preprocessing of data is required and involves the use of a clustering algorithm that divides the whole learning space into subspaces. In this paper we apply a new fuzzy clustering algorithm to a prediction tool of a third generation (3G) cellular radio network. Results show that the differences observed between simulations and measurements can be considerably diminished and the generalization capacity is enhanced thanks to the proposed clustering algorithm. This algorithm performs well than classical $k$-means algorithm. We can then predict with enhanced accuracy new configuration for which we don't have measurements, as long as they are not very different from learned configurations.**

***Index Terms*— Radio Network Prediction, Measurements, Distribution Learning, $k$-means, Fuzzy Clustering**

## I. INTRODUCTION

### A. General framework

The current mobile radio networks offer a wide class of services, from real time (RT) multimedia like video conference or television on mobile to non-real time (NRT) service of type best effort like web page download. Under a great competition, mobile operators are involved to undertake on the quality of service (QoS) that they offer. This QoS is becoming, for an operator, an major challenge to choose its infrastructure and the associated services. To attend these purposes, operators use *simulations* during the network conception phase as well as the optimization network phase.

The $3^{rd}$ generation (3G) radio access network relies upon novel, flexible and efficient communication methods; a consequence of which is that novel modeling and planning approaches became of prime importance to the network's rollout success. Thus, the need for accurate prediction has become a challenging task. The planning of any modern radiocommunication network is carried out by one of two methods. The first method, called *semi-empirical* simulation, is based on the *a priori* system information [1] like propagation model, traffic model and mobility model. This method uses computer algorithms that numerically calculate the desired network characteristics [2]. It uses mathematical and/or statistical models that represent real, physical phenomena like propagation, traffic, etc. These equations give a *good idea* on the described phenomena. The good idea notion is relative to the desired precision degree that we want: higher precision implies higher complexity model. Being inexact in nature, such models yield predictions that exhibit inevitable discrepancies with reality. Therefore this method is not robust. For example, using our 3G radio network planning tool to predict the UpLink and DownLink loads (ULL and DLL) for a given station, we remark that values given by the planning tool are slightly different from real values taken on real network as shown in figures 1 and 2. These figures show that the distribution given by the simulation does not correspond to the one taken from the real network.

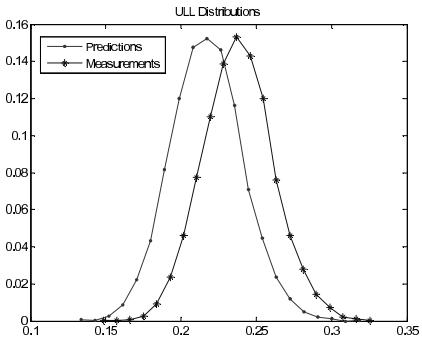


Figure 1. simulation and measurement for ULL distributions

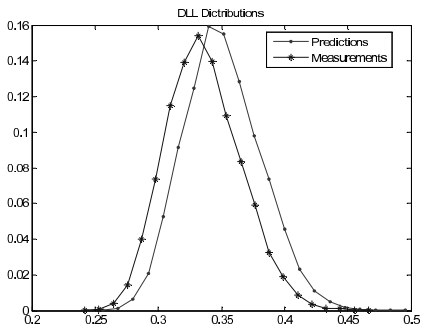


Figure 2. simulation and measurement for DLL distributions

This paper is based on "A New Fuzzy Bayesian Clustering Algorithm to Enhance Radio Planning Tool Predictions," by Z. Nouir, B. Sayrac, B. Fourestié, W. Tabbara and F. Brouaye; which appeared in the Proceedings of the third International Conference on Wireless and Mobile Communications (ICWMC), Guadeloupe, French Caribbean, March 2007. © 2007 IEEE.

The second method, called *empirical* simulation, uses *a posteriori* data (in the form of observations and/or measurements). It uses an automatic learning system like an Artificial Neural Network (ANN) that models the whole mobile network [3], [4]. The transfer function found by the automatic learning system is a simple mathematical operator without any physical reality, which may lead to stability problems.

Consequently, we have proposed in [5] an alternative solution based on a combination of the described two methods to benefit from their respective advantages. Table I compares the semi-empirical method, the empirical method and the proposed method.

### B. Authors' Previous Works

The proposed scheme [5] basically consists of finding the mapping between measurements and predictions via an ANN, thus enhancing the quality and precision of the predictions.

To make the scheme tractable and easy to use we have proposed to make the variables as independent as possible by using Independent Component Analysis (ICA). The method proposed in [5] was enhanced by the use of a clustering algorithm based on $k$-means [6]. The learning space is clustered into subspaces that guarantee good generalization capabilities. Generalization enables us to better predict new cases (new traffic, new network parameters, etc.) for which no measurement is available. Generalization with different traffic values was treated in [6] where the proposed scheme was shown to yield enhanced predictions that are much closer to measurement values than simulation results are. Generalization with different network parameters is handled in [7]. The chosen parameters are mechanical antenna tilts. Therefore the proposition in [7] enables us to correctly predict the mobile cellular network behavior in the case of an antenna tilt modification and saves us considerable amount of resources.

In [6], the clusters obtained by $k$-means presents a problem with objects located near the boundaries of clusters. These objects can belong to many groups until they are in the frontier of many clusters. This led us to propose the use of fuzzy classification for objects that are in the boundaries of two or more clusters [8].

The proposed fuzzy clustering algorithm is based on Bayes rule and assigns for each object in the learning space a membership probability vector. We called it fuzzy-Bayesian clustering algorithm (FBC). In our knowledge there is no work in the literature that proposes such an algorithm.

In [8] we have described the application of the proposed method to a third generation (3G) UMTS network to enhance the predictions of uplink and downlink base station loads. Comparisons between Radio Network Planning (RNP) tool simulation results and enhanced predictions are made using the analytical 2-D Kolmogorov-Smirnov (KS) distance [9] which is a measure of the degree of similarity between two multidimensional distributions. The results have shown that the proposed approach is able to correct the errors made by the RNP tool and the learning is successful.

### C. About this paper

In this paper we study the generalization capacity of the FBC in the case of tilt change. We are interested in two performance indicators: The UL Load (ULL) and the DL Load (DLL) whose expressions for an arbitrary base station $m$ can be written as [10]:

$$ULL^m = \sum_{i=1}^{K_m} \frac{1}{1+CIR_{ul}^i}(1+f) \quad (1)$$

$$DLL^m = \sum_{i=1}^{K_m} \left[ CIR_{dl}^i \left( (1+\alpha_i) + \sum_{\substack{b=1 \\ b\neq m}}^{B_m} \frac{L_{m_i}}{L_{b_i}} \right) \right] \quad (2)$$

where $K_m$ is the number of mobiles connected to base station $m$, $CIR_{ul(dl)}^i$ is the required UL (DL) CIR (carrier-to-interference ratio) for mobile $i$, $\alpha_i$ is the orthogonality factor, $L_{m_i}$ is the path loss from the serving base station $m$ to mobile $i$, $L_{b_i}$ is the path loss from the neighbouring base station $b$ to mobile $i$, $B_m$ is the number of neighbouring base stations of base station $m$, and $f$ is the proportionality factor between the power received from the mobiles connected to other base stations, $I_{oth}$, and the power received from the mobiles connected to base station $m$, $I_{own}$, i.e. $f = \frac{I_{oth}}{I_{own}}$. Note that equations 1 and 2 are valid when a single serve is considered and the handovers are not taken into account.

Results show that with FBC we can prevent overlearning and ameliorate generalization properties of the ANN.

The remainder of this paper is structured as follows: Section 2 gives a brief description of the fuzzy clustering algorithm as proposed in the literature and presents the proposed fuzzy-Bayesian clustering algorithm, its advantages and its difficulties. Section 3 describes the application of the fuzzy Bayesian clustering to a RNP tool and gives results. Finally, Section 4 gives the concluding remarks.

## II. Fuzzy and Fuzzy Bayesian Clustering

Clustering is the classification of similar objects into different groups (say $c$), or more precisely, the partitioning of a data set into subsets (clusters), so that the data in each subset (ideally) share some common trait, often proximity according to some defined distance measure [11]

### A. Fuzzy Clustering

In fuzzy clustering, each point has a degree of belonging to clusters, as in fuzzy logic [12], rather than belonging completely to just one cluster. Thus, points on the edge of a cluster may be in the cluster to a lesser degree than points in the center of cluster. For each point $\mathbf{x}_j$ we have a coefficient giving the degree of being in the $k^{th}$ cluster . Usually, the sum of those coefficients is

TABLE I.
COMPARISON OF EMPIRICAL AND SEMI EMPIRICAL METHODS

| | **semi-empirical method** | **empirical method** | **Proposed Method** |
|---|---|---|---|
| Principal | *Use a Monte-Carlo simulation based on a priori models* | *A learning system that fits the the system configuration to his characteristics* | *Combine empirical and semi empirical methods* |
| Strong aspects | ♠ ***Applicability***: valid for all configurations<br>♣ ***Coupling***: no data base required | ♦ ***Complexity***: Simple to do<br>★ ***Robustness***: no statistical models ⇒ no bias with regard to measurements | ♠ ***Applicability***: valid for all configurations<br>★ ***Robustness***: predictions are corrected using measures ⇒ no bias between measurements and simulations<br>♣ ***Coupling***: semi coupled method ⇒ no data base needed |
| Weak aspects | ♦ ***Complexity***: complex to do<br>★ ***Robustness***: simple models ⇒ non perfects models ⇒ bias in simulations | ♠ ***Applicability***: valid only for configuration seen on learning phase<br>♣ ***Coupling***: required data base which is hard to have | ♦ ***Complexity***: complex to process ⇒ inheritance from semi empirical method |

defined to be 1, so that $u_{k,j}(\mathbf{x}_j)$ denotes a probability of belonging to a certain cluster:

$$\forall \mathbf{x}_j \sum_{k=1}^{c} u_{k,j}(\mathbf{x}_j) = 1 \tag{3}$$

Fuzzy $c$-means (FCM) is one of the most popular fuzzy clustering algorithms [13] that was proposed by Bezdek [14]. With fuzzy $c$-means, the centroid ($\mathbf{m}_k$) of a cluster is the mean of all points, weighted by their degree of belonging to the cluster:

$$\mathbf{m}_k = \frac{\sum_j u_{k,j}(\mathbf{x}_j)^m \mathbf{x}_j}{\sum_j u_{k,j}(\mathbf{x}_j)^m} \tag{4}$$

The degree of belonging is related to the inverse of the distance to the cluster:

$$u_{k,j}(\mathbf{x}_j) = \frac{1}{D(\mathbf{m}_k, \mathbf{x}_j)} \tag{5}$$

Where $D(\mathbf{m}_k, \mathbf{x}_j)$ is the distance measure between $\mathbf{x}_j$ and $\mathbf{m}_k$.
Then the coefficients are normalized and fuzzyfied with a real parameter $m > 1$ so that their sum is 1. For $m$ equal to 2, this is equivalent to normalising the coefficient linearly to make their sum 1. When $m$ is close to 1, then cluster center closest to the point is given much more weight than the others, and the algorithm is similar to $k$-means.

The fuzzy $c$-means algorithm is very similar to the $k$-means algorithm:

1) Choose a number of clusters.
2) Assign randomly to each point coefficients for being in the clusters.
3) Repeat until the algorithm has converged (that is, the coefficients' change between two iterations is no more than $\epsilon$, the given sensitivity threshold):
   - Compute the centroid for each cluster, using formula (4).
   - For each point, compute its coefficients of being in the clusters, using formula (5).

The FCM algorithm minimizes intra-cluster variance as well, but has the same problems as $k$-means, the minimum is a local minimum, and the results depend on the initial choice of weights.

The FCM uses the Euclidean or $L_2$ norm distance function. Kersten suggested that city block distance (or norm) could improve the robustness of FCM to outliers [15]. Furthermore, Hathaway extended FCM to a more universal case by using Minkowski distance (or $L_p$ norm, $p > 1$ ) and seminorm ($0 < p < 1$ ) for the models that operate either directly on the data objects or indirectly on the dissimilarity measures [16].

### *B. Proposed Fuzzy Bayesian Clustering*

To cluster a set of data points $\mathbf{x}_j \in \Re^d$ into $c$ classes, we propose a new Bayesian framework, a free parameter scheme based on the Bayes formula. This new proposed fuzzy Bayesian algorithm works as fellows:

1) Perform a hard clustering on the data set using $c$ clusters. Hard clustering algorithms can be hierarchical clustering algorithms, squared error-based algorithms, pdf estimation via mixture densities, graph theory-based algorithms, combinatorial search techniques-based algorithms, etc.
2) Compute the cluster's probabilities $p(C_j)$, the prior probability
3) Perform the conditional of $\mathbf{x}_i$ given $C_j$, $p(\mathbf{x}_i/C_j)$
4) Compute the posterior probability, the probability of membership of $\mathbf{x}_i$ to the cluster $C_j$, i.e $p(Cj/\mathbf{x}_i)$

The fuzzy Bayesian framework is based on the formula:

$$p(C_j/\mathbf{x}_i) = \frac{p(\mathbf{x}_i/C_j)p(C_j)}{p(\mathbf{x}_i)} \tag{6}$$

Where: $p(C_j/\mathbf{x}_i)$ is the $\mathbf{x}_i$ membership probability to $C_j$.
$p(C_j)$ is the probability of the $j^{th}$ cluster.
$p(\mathbf{x}_i/C_j)$ is the probability of $\mathbf{x}_i$ if we consider only $C_j$

*1) Computation of cluster's probability:* After performing the *hard* clustering on the data set $\mathbf{x}_i$, we obtain $c$ ($j = 1, \ldots, c$) classes. Each cluster contains $n_j$ objects. Thus, the probability of each cluster can be computed as:

$$p(C_j) = \frac{n_j}{N} \tag{7}$$

TABLE II.
MEMBERSHIP MATRIX FOR 3 OBJECTS

| | $p(\mathbf{x}_i/C_1)$ | $p(\mathbf{x}_i/C_2)$ |
|---|---|---|
| $x_1$ | 0.788487 | 0.211513 |
| $x_2$ | 0.814868 | 0.185132 |
| $x_3$ | 0.956211 | 0.0437893 |

Where $N$ represents the total number of the objects:

$$N = \sum_{j=1}^{c} n_j \tag{8}$$

*2) Computation of the likelihood:* As we have $c$ clusters with $n_j$ elements each, the probability of $\mathbf{x}_i$ given $C_j$, $p(\mathbf{x}_i/C_j)$, is calculated after distribution computation of the $j^{th}$ cluster. This distribution is determined using the objects of the $j^{th}$ cluster, $\mathbf{x}_i^j$.
As $\mathbf{x}_i^j$ are multidimensional objects, we must calculate the multidimensional distribution of the cluster $C_j$. For example, if $d = 1$ ($x_i^j$ are one-dimensional objects) the calculation of the prior probability is performed by the computation of the histograms as shown in figure 3.

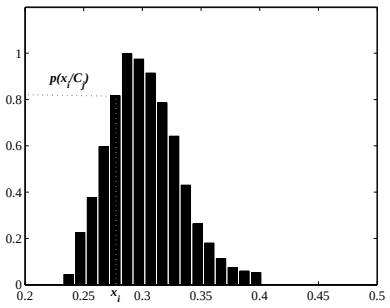


Figure 3. Likelihood Calculation

*3) Computation of $p(\mathbf{x}_i)$:*

$$p(\mathbf{x}_i) = \sum_{j=1}^{c} p(\mathbf{x}_i/C_j)p(C_j) \tag{9}$$

*4) Example:* In the context of the enhancement of 3G radio network planning tool predictions, we focus on the clustering of the 2-d learning space (ULL and DLL). The space to cluster is composed of 1000 objects and the number of cluster is equal to 2. Figure 4 shows the objects to be clustered and table II shows the membership matrix for the 3 first objects.

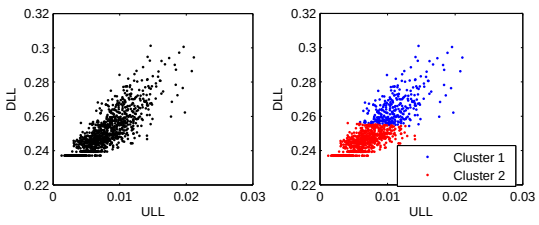


Figure 4. 1000 Objects and 2 Clusters

TABLE III.
COMPUTATION SPEED

| | FCM | FBC |
|---|---|---|
| Times (s) | 32 | 7 |

*5) Advantages and inconveniences:* Compared to others fuzzy algorithms, fuzzy Bayesian clustering is a very speed clustering algorithm. Table III gives a comparison of computation speed with the FCM algorithm for the example given above.

The computation speed of the fuzzy Bayesian algorithm is better than the classical $c$-means algorithm because the last one is based on convergence criteria while the fuzzy Bayesian algorithm works without convergence condition. It is a free parameter learning algorithm. The $c$-means algorithm converges when there is no change in the cluster prototype and this can take time for convergence.
However, fuzzy Bayesian algorithm can be applied only when we dispose of an abundant amount of objects. If the number of objects to be clustered, $N$, is not significant, the computation of the prior probabilities will not be accurate.

## III. METHODOLOGY AND RESULTS

### A. Architecture

Figures 5(a) and 5(b) depict the block diagrams of the learning and prediction phases of the proposed scheme. The core of the scheme is a Multi-Layer Perceptron (MLP). The training algorithm is the well-known Back Propagation algorithm, the generalization of the Least-Mean-Square algorithm that is based on gradient search [17].
In the learning phase, both measurement and simulation data are input to the system and the ANN is trained with these data. In the prediction phase, simulation data is input to the system and the trained ANN yields enhanced predictions.

The measurement data used in the learning phase comes from the field with specific traffic and network parameter values and the simulation data are produced by calibrating the RNP tool with the same traffic and network parameter values.
We are interested in two performance indicators: the ULL and the DLL. Let $\mathbf{c}_L^m$, $\hat{\mathbf{c}}_L^m$ and $\tilde{\mathbf{c}}_L^m$, where $L$ stands for either $UL$ or $DL$, represent the measurement, the simulation and the enhanced prediction vectors for a base station $m$:

$$\mathbf{c}_L^m = \left[c_{L,1}^m \dots c_{L,N}^m\right] \tag{10}$$

$$\hat{\mathbf{c}}_L^m = \left[\hat{c}_{L,1}^m \dots \hat{c}_{L,N}^m\right] \tag{11}$$

$$\tilde{\mathbf{c}}_L^m = \left[\tilde{c}_{L,1}^m \dots \tilde{c}_{L,N}^m\right] \tag{12}$$

where $c_{L,n}^m$ is the $n^{th}$ data sample coming from the measurement data, $\hat{c}_{L,n}^m$ is the data sample drawn from the $n^{th}$ snapshot of the Monte Carlo simulation, $\tilde{c}_{L,n}^m$ is

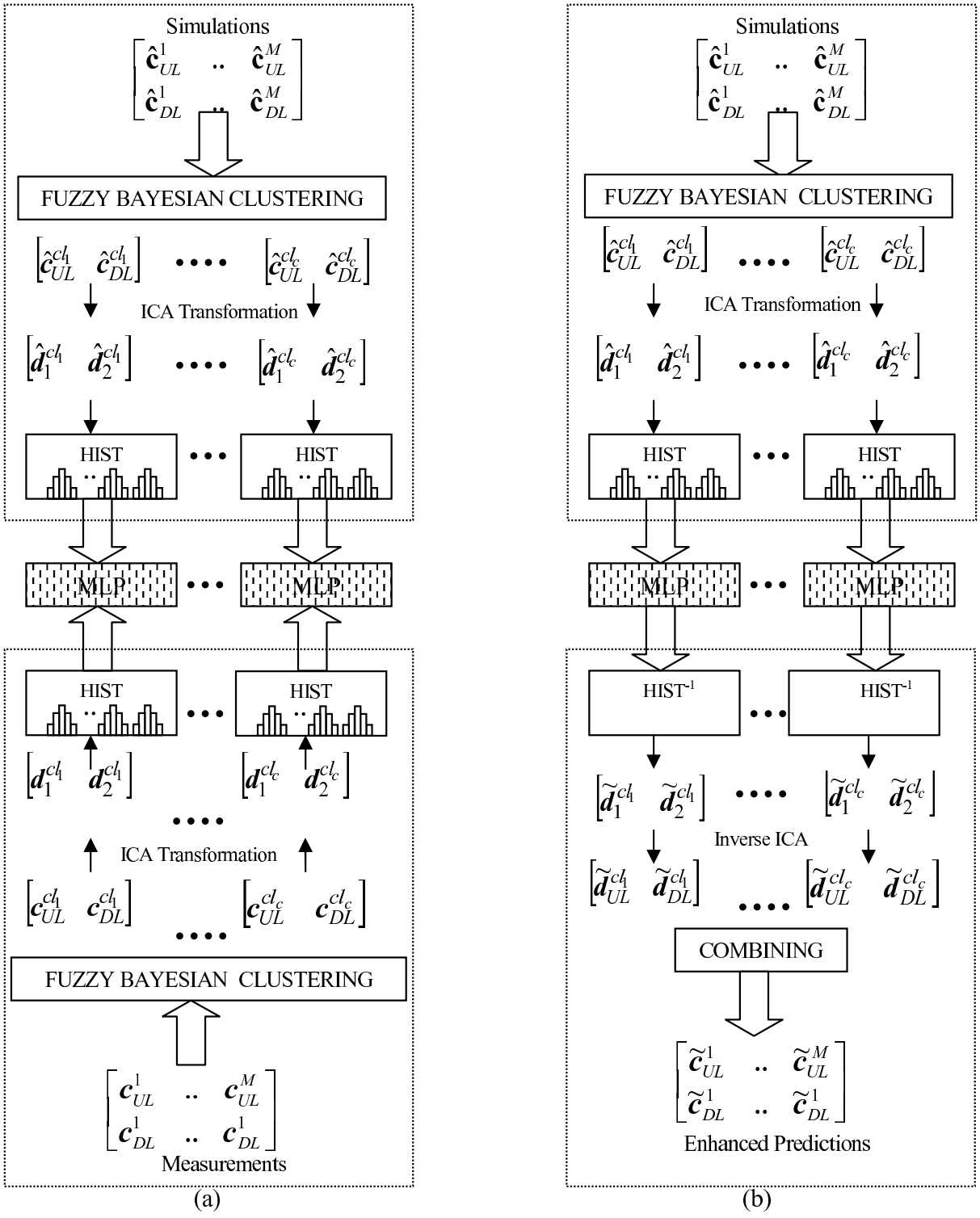


Figure 5. Global Architecture: Learning phase (a) and Prediction phase (b)

the $n^{th}$ data sample of the enhanced predictions, and $N$ is the number of data samples.

As we do not know the positions of the mobiles connected to the base station, we are not interested in the individual values of $c^m_{L,n}$ but on the statistical features of the vector $\mathbf{c}^m_L$ . So the ANN works on discrete statistical distributions.

However, a problem arises at this point. The performance indicators that are to be predicted are statistically dependent variables. Thus when we are interested in more than one indicator, we have to use the joint probability distributions (i.e. multi-dimensional histograms) in order to conserve the joint statistical features. Using multi-dimensional distributions makes the learning process excessively complex and intractable. To overcome this problem we propose to use ICA, a linear technique that extracts independent features from a data set. ICA aims at finding $\mathbf{W} \approx \mathbf{A}^{-1}$, the pseudo-inverse of $\mathbf{A}$, that maximizes the non-gaussianity of the components of the vector $\hat{\mathbf{s}} = \mathbf{W}\mathbf{x}$ .

ICA processing is performed for each cluster (figure 5) but for reasons of notation simplicity, we will focus on one cluster and drop the cluster indices (the superscripts $c_{lk}$ of the independent component vectors $\mathbf{D}$ in figure 5 in the following.

Note that we use the same matrix $\mathbf{W}$ for both measurements and simulations. $\mathbf{W}$ is then applied to measurement ( $\mathbf{c}^m_{UL}$ and $\mathbf{c}^m_{DL}$ ) and simulation ( $\hat{\mathbf{c}}^m_{UL}$ and $\hat{\mathbf{c}}^m_{DL}$ ) vectors coming from base station $m$ ($m = 1, \ldots, M$ and $M$ is the number of base stations of interest) to obtain the independent component vectors ($\mathbf{d}_1, \mathbf{d}_2$ for measurements and $\hat{\mathbf{d}}_1, \hat{\mathbf{d}}_2$ for simulations).

On the other hand, $ICA^{-1}$ is applied to the independent component vectors of the enhanced predictions during the prediction phase.

Using membership coefficients obtained in the learning phase the result of the prediction is:

$$\begin{aligned} \begin{bmatrix} \tilde{\mathbf{c}}^m_{UL} \\ \vdots \\ \tilde{\mathbf{c}}^m_{UL} \end{bmatrix} &= \mathbf{u}_1 \begin{bmatrix} \tilde{\mathbf{c}}^{cl_1}_{UL} \\ \vdots \\ \tilde{\mathbf{c}}^{cl_1}_{DL} \end{bmatrix} + \ldots + \mathbf{u}_c \begin{bmatrix} \tilde{\mathbf{c}}^{cl_c}_{UL} \\ \vdots \\ \tilde{\mathbf{c}}^{cl_c}_{DL} \end{bmatrix} \\ &= \sum_{i=1}^{c} \mathbf{u}_i \begin{bmatrix} \tilde{\mathbf{c}}^{cl_i}_{UL} \\ \vdots \\ \tilde{\mathbf{c}}^{cl_i}_{DL} \end{bmatrix} \end{aligned} \quad (13)$$

The real interest in using such a learning system lies in using the generalization capability of the ANN: the input simulation data of the prediction phase corresponds to a case that has never been encountered by the ANN during the learning phase (such as different traffic, different network parameters, etc.) and the ANN succeeds in correcting the simulations of the new case. However, the generalization is not always easy to achieve since it is an extrapolation operation that requires special attention. The proposition we make to facilitate the generalization process is to use *fuzzy Bayesian* clustering block to restrict the space on which the ANN carries out its learning and prediction. In a previous work, we showed that successful generalization to different traffic values was achieved thanks to the $k$-means clustering [6]. In [7] we took an additional measure to ameliorate the generalization: the training is performed for only one base station but its training data comprises of samples from several base stations. This aims at exposing the ANN to cases as rich as possible during learning in order to maximize the learning space and minimize the extrapolation failure risk.

### *B. Learning and Generalization Results*

*1) Leraning Results:* The proposed scheme is tested on a 3G network. For practical reasons, we have used synthetic measurement data that are produced by the RNP tool. To obtain the synthetic measurement data, we have modified the target UL CIR It is set to -20 dB for the simulation data and to -18 dB for the measurement data, accounting for a bias between field parameters and RNP parameters.

After clustering and ICA processing, each data stream is passed through a histogram transformation. Each histogram has 10 bins and requires at least 50 data samples for each bin.

In [8] we have shown that the learning phase is usefully done and results have shown that the enhanced prediction data are much closer to the measurement data than the simulation data are. Furthermore, we have shown in [8] that the statistical relation between simulations and measurements was successfully learned by the proposed scheme.

*2) Generalization Results:* In this study we focus on the generalization results for the case of tilt change. A new simulation data is generated by modifying the (mechanical) tilt of an antenna from $0^o$ to $10^o$. In practical cases, we would not dispose of measurement data for such a case and we would like to be able to obtain accurate predictions. This ability saves us from the cost of going out to the field to modify the tilt and to collect the measurements.

Thus, the new simulation data corresponding to a tilt value of $10^o$ is passed through the proposed scheme where all the coefficients and parameters of the trained ANN is preserved as well as those of the FBC algorithm and ICA blocks.

In order to evaluate the performance of the proposed scheme, the following error metrics are used:

$$\tilde{e} = D_{KS}\left( \begin{bmatrix} \mathbf{c}^m_{UL} \\ \mathbf{c}^m_{DL} \end{bmatrix}, \begin{bmatrix} \tilde{\mathbf{c}}^m_{UL} \\ \tilde{\mathbf{c}}^m_{DL} \end{bmatrix} \right) \quad (14)$$

$$\hat{e} = D_{KS}\left( \begin{bmatrix} \mathbf{c}^m_{UL} \\ \mathbf{c}^m_{DL} \end{bmatrix}, \begin{bmatrix} \hat{\mathbf{c}}^m_{UL} \\ \hat{\mathbf{c}}^m_{DL} \end{bmatrix} \right) \quad (15)$$

where $\tilde{e}$ is the error that represents the discrepancy between measurements and enhanced predictions, $\hat{e}$ is the error that represents the discrepancy between measurements and simulations and $D_{KS}$ is the 2-D Kolmogorov-Smirnov (KS) distance, a distance metric that measures the degree of similarity between two distributions [9].

Using KS distance, we obtain the results depicted in

table IV. We can see here that the enhanced prediction data for the Bayesian framework are much closer to measurements than are the enhanced prediction data for $k$-means algorithm. The KS distance has decreased from $\tilde{e} = 0.297$ to $\hat{e} = 0.239$ accounting for a gain of $7.7\%$.

TABLE IV.
GENERALIZATION RESULTS FOR FBC

| | $k$-**means** | ***FBC*** |
|---|---|---|
| $\hat{e}$ | 0.565 | |
| $\tilde{e}$ | 0.297 | 0.239 |

## IV. CONCLUSION

FBC algorithm is used to prevent over-learning and to enhance the generalization capabilities of a learning system that uses measurements to reduce the bias between simulations and reality. The results have shown that FBC performs well than the classical $k$-means algorithm and the scheme is able to learn and predict the stochastic transfer function between simulations and measurements. The prediction results are especially important since they imply that it is possible to reduce the discrepancy between simulations and measurements, and hence improve significantly the precision and quality of the simulations.

## ACKNOWLEDGMENT

The authors would like to thank Christophe Gay and Pierre Sailly for their helpful contributions in computing aspects and Othmane Guadi for his contribution in technical aspects.

## REFERENCES

[1] M. Centeno and M. Reyes, "So you have your model: what to do next? a tutorial on simulation output analysis," *Simulation Conference Proceedings*, vol. 1, pp. 23–29, Dec 1998.
[2] T. Gill, "Radio planning and optimisation - the challenge ahead," in *Proc. 4th International Conference on 3G Mobile Communication Technologies* , 2003, pp. 28–30.
[3] N. Andrea, C. Cecchetti, and A. Lipparwi, "Fast prediction of the performance of wireless links by simulation trained neural network," *Proc. IEEE MTT-S Digest 2000*, pp. 429–432, 2000.
[4] V. Litovski, . Mrèarica, and T. Iliæ, "Simulation of nonlinear magnetic circuits modelled using artificial neural network," *Journal Simulation Practice and Theory*, vol. 5, pp. 553–570, 1997.
[5] Z. Nouir, B. Sayrac, and B. Fourestie, "Independent component analysis for radio network prediction enhancement," *Proc. 20th European Conference on Modelling and Simulation*, pp. 51–55, 2006.
[6] ——, "Enhancement of network planning tool predictions through measurements," *Proc. 63rd IEEE Conference on Vehicular Technology*, vol. 46, pp. 155–160, 2006.
[7] Z. Nouir, B. Sayrac, B. Fourestie, W. Tabbara, and F. Brouaye, "Supervised prediction for radio network planning tool using measurements," *Proc. 17th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–5, 2006.
[8] Z. Nouir, B. Sayrac, and B. Fourestie, "A new fuzzy bayesian clustering algorithm to enhance radio planning tool predictions." in *Proc. Third International Conference on Wireless and Mobile Communications* , 2007, pp. 34–34.
[9] G. Fasano and A. Franceschini, "A multidimensional version of the kolmogorov-smirnov test," *Royal Astronomical Society*, vol. 255, pp. 155–170, 1987.
[10] L. Laiho, A. Wackerand, and T.Novosad, *Radio Network Planning and Optimisation for UMTS*. John Wiley & Sons, 2002.
[11] E. Backer and A. Jain, "A clustering performance measure based on fuzzy set decomposition," *IEEE Trans. Pattern. Anal. Mach. Intell*, vol. 3, pp. 66–75, 1981.
[12] L. Zadeh, "Fuzzy sets," *Inf. Control.*, vol. 8, pp. 338–353, 1965.
[13] F. Hoppner, F. Klawonn, and R. Kruse, *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis, and Image Recognition*. Wiley, 1999.
[14] J. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum, 1981.
[15] P. Kersten, "Implementation issues in the fuzzy $c$-medians clustering algorithm," in *Proc. 6th IEEE Int. Conf. Fuzzy Systems.* , vol. 2, 1997, pp. 957–962.
[16] R. Hathaway, J. Bezdek, and Y. Hu, "Generalized fuzzy $c$-means clustering strategies using $l$ norm distances." *IEEE Trans. Fuzzy Syst.*, vol. 8, pp. 576–582, 2000.
[17] S. Haykin, *Neural Networks: A comprehensive foundation*, 2nd ed. Prentice Hall, 1998.

**Zakaria Nouir** was born in Mahdia, Tunisia, in 1978. In 2001 he received his Engineering Diploma from the Polytechnic school of Tunisia (EPT). In 2002 he received the Master degree in Networks and Radio Systems from the Institut National des Télécommunications, Evry, France. He received the D.E.A degree in mobile networks from the Paris 6 university and the Ecole Nationale Supérieure des Télécommunications, Paris in 2003.
He is currently pursuing a PhD in the Research and Development Department in France Telecom. His research interests are in the area of wireless communication systems and precisely in planning, modeling and system performances. In addition to his active research activities, he is served as a TPC member of the Performance Modeling, QoS and Reliability Symposium in GLOBCOM 2007.

**Berna Sayrac** She received the B.S., M.S. and Ph.D. degrees from the Department of Electrical and Electronics Engineering of Middle East Technical University (METU), Ankara, Turkey, in 1990, 1992 and 1997, respectively. Between 1990 and 1999, she worked as a research assistant and as a lecturer at METU. Between April 1999 and August 2000, she was a postdoctoral fellow at France Telecom R&D (FTRD), Issy Les Moulineaux, France. She worked as an assistant professor at the Department of Electrical & Electronics Engineering of METU between August 2000 and May 2001, and as a research scientist at Philips Research France (PRF) between Mai 2001 and October 2002. Since October 2002, she is working as a research engineer at FTRD. Her research interests include mobile/wireless communications, radio access network planning and computational intelligence.

**Benoît Fourestié** was born in Paris, France, in 1975. In 1997, he received the diploma of the Institut National des Télécommunications, Evry, France, and the D.E.A. degree in digital telecommunication systems from the Ecole Nationale Supérieure des Télécommunications, Paris. He received the Ph.D. degree in physics from the University of Orsay, Paris XI, in 2000. Since then, he has been with the FTR&D, the R& D Centre of France Telecom. His research interests include mobile communications, antenna measurements, EMC, statistics, and signal processing.

**Walid Tabbara** was born in Beirut, Lebanon, in 1946. He received the Doctorat dès Sciences (Ph.D.) degree from the University of Paris VI, Paris, France, in 1976.

He is currently a Full Professor at the University of Paris VI, since 1981. His research is done at the Département de Recherche en Electromagnétisme (DRE), Supélec. From 1987 to 1996, he acted as a Deputy Director of the Laboratoire des Signaux et Systèmes (Supélec-CNRS). After working for a long period in the field of inverse scattering in electromagnetics and acoustics, his present fields of interest are Electromagnetic Compatibility (EM, remote sensing, indoor propagation and high-power microwaves (HPM). His present and passed researches has led to the publication of 48 papers in international journals and to 158 presentations at international and national conferences. He is on the editorial boards of the Journal of Electromagnetic Waves and Applications and the Journal of Computer Applications in Engineering Education.

Dr. Tabbara has organized and/or chaired sessions at national and international conferences. He was the French delegate to Commission B of URSI from 1984 to 1990. He is a member of URSI and the Electromagnetics Academy.

**Françoise Brouaye** was born in Paris, France. She received the Ph.D. degree in signal and system processing from the University of Paris XI-Orsay. Currently, she has worked with the 'Département de recherche en électromagnétisme' in the CEM domain. From 1972 to 1984, she was with the mathematical department of the University of Paris XI-Orsay, France. She was with the power system department of Supélec, Gif sur Yvette, France, and EDF, an electrical company in France, working on statistical modeling and signal-processing problems.