Available online at www.sciencedirect.com
ScienceDirect
Int. J. Electron. Commun. (AEÜ) 64 (2010) 521 – 530
International Journal of Electronics and Communications AEÜ
www.elsevier.de/aeue

Application of information criteria for the selection of the statistical small scale fading model of the radio mobile channel

Carlos Pereira[a], Guilhem Coq[b], Xiang Li[a], Yannis Pousset[a,*], Christian Olivier[a], Olivier Alata[a], Rodolphe Vauzelle[a], Marc Arnaudon[b], Pierre Combeau[a]

[a]*XLIM-Signal Image Communications (SIC) Department, UMR CNRS n° 6086, University of Poitiers, Bât. SP2MI Téléport 2, Bvd Marie et Pierre Curie B.P. 30179, 86962 Futuroscope Chasseneuil Cedex, France*
[b]*Laboratory of Mathematics and Applications (LMA), UMR CNRS n° 6086, University of Poitiers, Téléport 2 Bvd Marie et Pierre Curie B.P. 30179, 86962 Futuroscope Chasseneuil Cedex, France*

Received 10 July 2008; accepted 18 March 2009

Abstract

Modeling of radio channel behavior is very important for digital communication. It is the reason why several propagation models have been implemented to take into account the electromagnetic phenomena inherent to the radio wave channel. So, among these models is the family of statistical distributions, fast in computation time, to describe channel fast fading. In this study context, our problematic is to find among the different statistical laws the one which best coincides with radio channel behavior. Different research has been realized on this subject over about 20 years. Thus, the Kolmogorov–Smirnov (KS) test using cumulative distribution functions (CDF) is usually employed by the radio mobile community partly for its simplicity. Nevertheless, there exist other approaches like Kulback–Leibler (KL) divergence, which consider distances between probability density functions that may be respresented as histograms. The bins of those histograms are obtained empirically, i.e. the bins are same size and their number is arbitrary. The authors propose approaches allowing optimal histograms in terms of the number and size of bins. These approaches are based on information criteria and we show that the KL discriminative method is more stable with the optimal histograms than with the empirical one. Moreover, our approaches are more reliable than the KS test whatever the number of considered samples. This last point is directly connected to the campaign of measures or of deterministic simulations, the duration of which can be very significant.

Keywords: Radio mobile propagation channel; Simulation; Ray tracing software; Information criteria; Distribution law; Divergence

1. Introduction

In the radio mobile communication context, the prediction of the propagation channel is crucial. Indeed, the latter is associated with the multi-path phenomenon which produces fast fading variations of the narrow band received power during mobile communication. Therefore, this phenomenon presents the risk of inter-symbol interference, brutal rupture of the communication, etc. [1]. It is the reason why the modeling of radio channel fast fading has been investigated to predict the behavior of the latter during a short

* Corresponding author. Tel.: +33 5 49 49 69 88; fax: +33 5 49 49 65 70.
E-mail addresses: pereira@sic.sp2mi.univ-poitiers.fr (C. Pereira), Guilhem.Coq@math.univ-poitiers.fr (G. Coq), li@sic.sp2mi.univ-poitiers.fr (X. Li), pousset@sic.sp2mi.univ-poitiers.fr, pousset@sic.univ-poitiers.fr (Y. Pousset), olivier@sic.sp2mi.univ-poitiers.fr (C. Olivier), alata@sic.sp2mi.univ-poitiers.fr (O. Alata), vauzelle@sic.sp2mi.univ-poitiers.fr (R. Vauzelle), Marc.Arnaudon@math.univ-poitiers.fr (M. Arnaudon), combeau@sic.sp2mi.univ-poitiers.fr (P. Combeau).

1434-8411/$ - see front matter
doi:10.1016/j.aeue.2009.03.005

spatio-temporal lap [2]. Two family models exist. The first one consists in predicting these fast variations deterministically according to the studied environment characteristics [2,3]. Nevertheless, their drawback is the excessive computation time associated with precision, especially since the fast fading step is close to the half-wavelength and the studied environments are complex. So, the second family considers statistical distribution models [2]. These are very fast in computation time compared to the preceding ones. Moreover, they are based on data provided by measurements or from deterministic simulations. Several statistical models are used to predict the behavior of fast fading. Among them Nakagami, Weibull and Rayleigh distributions are generally employed. Most of all, the choice of one of these models is made via discriminative methods like Kolmogorov–Smirnov (KS) using cumulative distribution function (CDF), provided by the analyzed empirical data. But we can also consider divergence [4], as Kullback–Leibler (KL), between probability density functions (PDF). Classically, the PDF are represented by histograms at k bins. Nevertheless, these histograms are not optimized in terms of bins. That leads, in some cases, to difficulties in identifying the best statistical laws associated with fast fading behavior and especially since the number of samples decreases. So, we propose, in this article, new methods based on histograms and information criteria (IC), which allow discrimination in all cases of the best statistical law from among group. Much research has been carried out in this domain and for this application context [4–9] but, to our knowledge, nothing considering optimal histograms, in terms of variable size and minimal number of bins, has been approached in this case. This article is split into three sections. Firstly (Section 2), we introduce the measures between PDF and we recall the IC principle. Secondly, in Section 3, we present a comparison test allowing validation of our discriminative methods. Lastly, Section 4 considers realistic scenarios in which we evaluate the best statistical laws connected to fast fading behaviors in LOS (line of sight) and NLOS (non-line of sight) configurations.

2. Theoretical tools

2.1. Some measures between probability laws

Disposing of two random variables X and Y, those being theoretical or empirical, one would like to compute a distance between them. The most used tool to this end is the KS distance. Denoted by F_X and F_Y the CDF of those variables, theoretical or empirical according to X and Y, that distance is given by

$$KS(X,Y) = \sup_t\{|F_X(t) - F_Y(t)|\} \tag{1}$$

Even though that measure is easy to use, it lacks discriminative power as shall be seen in the following: a sample from a known Gaussian variable might be as close to this very variable as to another Rayleigh variable for instance. To solve this problem, and in the case where our variables have PDF f_X and f_Y, the family of Φ-divergence [10] also gives several tools to measure the distance between them. Among them, we will use the classical KL divergence, such as

$$KL(X,Y) = \frac{1}{2}\int (f_X - f_Y)\log\left(\frac{f_X}{f_Y}\right) dv \tag{2}$$

where v is, for instance, the Lebesgue measure. Note to the KL divergence is a distance. In the following, our data will be of an empirical nature (deterministic simulation). With such data, it is easy to compute KS tests via empirical CDF. However, if one wants to use KL divergence, the first problem is to estimate a PDF. We choose here to use an optimal histogram estimation.

2.2. Histograms estimating an unknown density function

Several methods exist to build a histogram of an unknown PDF f_X of a random variable X defined on an interval I of $\mathbb{R}$, from a sample or realisation $x = (x_1, \ldots, x_n)$. We discuss here the usual empirical method and optimal model selection methods.

2.2.1. Empirical method

One may always choose a number of bins, say k, to build a histogram with k regular bins. That k is often chosen at $\lfloor 2\sqrt{n} - 1 \rfloor$ to give the so-called empirical histogram. However, that choice of k does not provide an optimal representation of the PDF and this histogram might therefore not be adequate for computing KL distance as required.

2.2.2. Model selection methods

Model selection methods, and in particular information criteria, may be used here to avoid the empirical choice of k. They are mainly based on the minimization of a cost. If the Kullback–Leibler cost is used, we obtain the Akaike information criterion (AIC) [11]. But Schwarz [12] suggested the widely known Bayesian information criterion (BIC) based on the Bayesian justification. A different approach was introduced by Rissanen [13]. This criterion is referred to as minimum description length (MDL) and can be compared to BIC [14]. The use of model selection makes it possible to build a purely database histogram that retains only inherent information of the sample; which the empirical histogram does not since it always depends on the a priori choice of k.

2.2.3. The criterion used

Our use of IC to select a histogram has a strong theoretical background since it is based for the MDL criterion on the stochastic complexity ([13,15] and more lately in [16]) to which had been added an optimal coding information in Coq et al. [17]. We use the idea that any partition of I allows encoding of the data without losses. The best histogram is