

Stochastic modeling of Wireless Networks: A case study in time domain*

Zoltán Gál, Ágnes Nagy, Márton Szikszai, György Terdik

University of Debrecen
zgal@it.unideb.hu, agnesnagy@gmail.com, szikszai.marton@science.unideb.hu,
terdik.gyorgy@inf.unideb.hu

Abstract

In wireless communication networks, interference produced by simultaneous transmission of the undesired transmitters and useful signal sources is one of the central aspects in system designs, since it dominantly limits network performance. We sampled data from a centrally controlled IEEE 802.11agn wireless sensor network established at the main campus of the University of Debrecen consisting of 54 non-mobile access points communicating in 29 active channels. After a reasonable and practical transformation the captured channel noise level is analyzed both as random variables and as time series by statistical methods for several fixed channels and access points.

1. Introduction

The importance of wireless networks in our days is extreme. They offer a vast variety of uses like communication across difficult-to-wire places; effective connection for sharing printers, scanners; providing alternative for expensive physical networks; high internet speed communication. Of course, the big picture cannot be confined to these specific uses, many other options arise for the applications of wireless networks in different areas.

Wireless Fidelity (WiFi) networks usually consist densely distributed access points serving as sensor nodes as well, providing simultaneous controlling and monitoring functions of the neighboring environment. Such functions are useful in many areas like health care, quality of service (QoS) monitoring in any production process. To give a more specific example, applying such an intelligent system as the tool for measuring the noise level at wireless access points, considerable help is given to determine the optimal physical coordinates of new access points in a

*This research was supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project. The project has been supported by the European Union, co-financed by the European Social Fund.

university or enterprise campus environment. We mention that a comprehensive description about the theory, practice and applications of wireless sensor networks can be found in [2].

Despite the above mentioned positives, difficulties can and will arise in wireless networks, such as interference, absorption and reflection of signals; the so-called multipath fading; hidden node problems and many others. In this paper we are going to present a stochastic analysis of interference utilizing the sensor function of the wireless access points, making possible sampling of network traffic at the physical coordinates of the nodes.

Analysis of interference in wireless networks has already been considered by many authors in many ways. A nice mathematical framework is given in [8]. Here the authors consider a network in which the interferers are scattered in space according to a spatial Poisson process and are operating asynchronously in a wireless environment subject to path loss, shadowing, and multipath fading.

In [6, 7] the authors showed that for similar interference levels, the performance degradation between coexistent 802.11b and 802.11g wireless environments is slightly worse than exclusively 802.11b systems and slightly better than exclusively 802.11g systems. It is a common assumption that IEEE 802.11b/g transmissions can coexist without interfering if the minimum distance of the channels is at least 25MHz. While in practice such an assumption seems to be true for larger distances, the interference can cause serious problem for the design of wireless multichannel mesh networks, which is the main topic of the paper of Fuxjager et al. [3].

Statistical analysis of wireless network traffics was considered in the papers [4, 5]. Utilizing the sensor functions of access points, in [1] a frequency based analysis of interference was presented. Such a method helps planning of a physical extension of the existing wireless network.

The purpose of this paper is to give an insight into the stochastic properties of interference in a wireless sensor network through a case study. Measuring the interference as channel noise at the physical coordinates of the wireless access points we take two different ways in our analysis. Fixing pairs of channels and access points, on one hand we study our data as a sample for a random variable and on the other hand as time series. Most of our calculations and graphics are made using the MATLAB computational environment.

2. The measurement scenario

A nice and detailed description of the measurement scenario for the present study can be found in [1]. However, a less complete presentation of the environment and obtained data shall be more than enough for our goals. Hence in this section only a short survey will be given together with a practical and reasonable transformation of the sample.

2.1. The wireless environment

As it was previously mentioned in the abstract, we sampled data from a centrally controlled IEEE 802.11agn wireless sensor network established at the main campus of the University of Debrecen. This environment consists of 54 non-mobile access points communicating in 29 active channels of which 13 and 16 is in the 2.4 GHz and 5 GHz frequency band, respectively. We are considering this set of access points as a spatially distributed sensor network, where the sensors are the access points and the sink node is the controller. Each access point has as many sensors as the number of channels, one for each. These sensors are responsible to measure the noise level in all channels simultaneously.

The sampling was executed once in each 30 seconds. Naturally, the response time can differ from 30 seconds due to the stochastic characteristic of the environment. However, its relative variation is less than 2%. Thus, we consider the sampling process equidistant in time.

2.2. Transformation of the data

At the end of the scenario we had a data of the form

$$D = (D_j^i(t)) \quad (i = 1, \dots, 29; j = 1, \dots, 54; t = 1, \dots, 10703),$$

where $D_j^i(t)$ is the observation at time t for the i th channel and the j th station. Due to technical failure we removed the 35th access point from our sample. Since we would like to avoid complications with the above indexing of the data we simply write $j = 1, \dots, 53$ in the latter parts, where $j = 35$ denotes the 36th wireless station and, 36 denotes the 37th and so on.

Note that D is measured in decibels, which is a logarithmic value, further it was rounded to the closest integer. As one can imagine this significantly damages the variability of the data. For example, consider $D_{12}^2(t)$ and $D_{33}^6(t)$. If we plot the first 1000 observations, then on Figure 1 we clearly see so. Moreover, these samples take on 12 and 9 different values only, respectively.

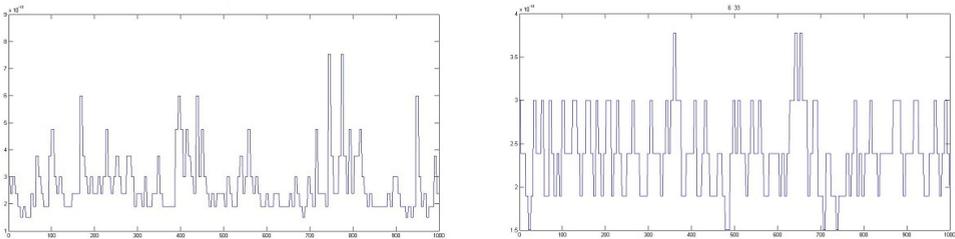


Figure 1: Channel noise in decibel: $D_{12}^2(t)$ and $D_{33}^6(t)$

In the followings we transform D into a more variable data of the form

$$M = (M_j^i(t)) = \left(\frac{1}{10} \sum_{k=10(t-1)+1}^{10t} \exp_{10}(D_j^i(k)) \right)$$

$$(i = 1, \dots, 29; j = 1, \dots, 53; t = 1, \dots, 1070)$$

Physically, our new sample is mean energy over five minutes. To compare with the original decibel data we consider $M_{12}^2(t)$ and $M_{33}^6(t)$, which are the corresponding cases of what have been shown before. Plotting the first 500 observations, on Figure 2 one can clearly see that a data with more variability is obtained and in addition we have that our samples take on 261 and 178 different values instead of 12 and 9, respectively.

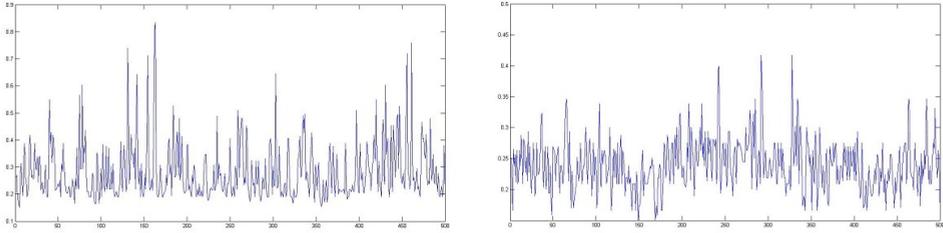


Figure 2: Channel noise in mean energy: $M_{12}^2(t)$ and $M_{33}^6(t)$

3. Statistical analysis of interference

In the followings we present two different approaches in the analysis of the mean energy data $M = (M_j^i(t))$. First, fix some channel and wireless access point, say the i th and j th ones, respectively. For every pair of parameters i and j we have that

$$M_j^i(t) \quad (t = 1, \dots, 1070)$$

can be considered either as a sample for a random variable or as a time series. We discuss both cases in different parts of the section, starting with the former.

3.1. Channel noise as random variable

We shall be interested in finding a well-fitting probability distribution for the sample. In the first step we use the kernel-smoothing density function estimate method to distinguish different classes of possible distributions. Foreshadowing results in this section we can say that such a class can consist of gamma-distributed samples. In this paper we confine our attention only to these cases.

A random variable X is said to be gamma-distributed with shape α and rate β if its corresponding probability density function is

$$f(x; \alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \quad (x \geq 0, \alpha, \beta > 0),$$

where $\Gamma(\cdot)$ stands for the gamma function. Our examples on Figure 3 show that the kernel-smoothing density function estimates (dash-dot line) are similar to the probability density functions of some gamma-distributed random variables. We already involved the probability density functions of the fitted gamma distributions (solid line) on the plot. The parameters of these distributions are the maximum likelihood estimates of α and β .

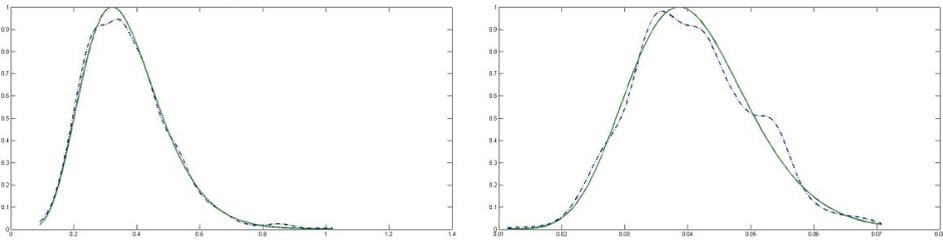


Figure 3: KSDensity function vs. fitted gamma: $M_{27}^{17}(t)$ and $M_2^{17}(t)$

Non-visual evidence for the good fitting of gamma distributions provided by statistical tests, for example the Kolmogorov-Smirnov test, which is a nonparametric test for the equality of continuous, one-dimensional probability distributions. Checking all the possible pairings of channels and stations it is acceptable that a significant number of them are gamma-distributed.

Questions can arise about the remaining cases. As we have mentioned before, our attention is confined to the gamma distributions. However, it can be said that the remaining samples show a wide variety of possible distributions and it seems to be far from sure that there is any class which covers such a large portion of the data.

3.2. Channel noise as time series

Considering the mean energy data for a given access point and channel, we are interested in time domain analysis. Before proceeding with our results we introduce some well-known definitions and notations concerning time series. While most of these are basics of the topic, there are many different ways of defining them and for the sake of clarity the proper versions are given.

Let $\{X_t\}$ be a time series. We say that $\{X_t\}$ is (weakly) stationary if

- $\mathbb{E}X_t$ is independent of t ;

- $\text{Cov}(X_t, X_{t+h})$ does not depend on t for each h .

A sequence $\{Z_t\}$ of random variables is referred to as a white-noise with variance σ^2 if

- each term has zero mean and variance σ^2 ;
- the terms are pairwise uncorrelated.

A common notation of such a process is $\{Z_t\} \sim WN(0, \sigma^2)$. The Gaussianity of the time series is not assumed here. In case $\{X_t\}$ is Gaussian then $\{Z_t\}$ is i.i.d.. It is clear that a white-noise is stationary in the above sense. We call the time series $\{X_t\}$ an autoregressive-moving-average process of order p and q if

- it is stationary;
- satisfies the linear difference equation

$$X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q},$$

where $\{Z_t\} \sim WN(0, \sigma^2)$.

We denote a process in the above sense by $\text{ARMA}(p, q)$. If $p = 0$ or $q = 0$, then we simply talk about a moving-average process of order q or an autoregressive process of order p , respectively. In these cases our notation reduces to $\text{MA}(q)$ and $\text{AR}(p)$ for the corresponding time series. An $\text{ARMA}(p, q)$ process is said to have a unit root if 1 or -1 is a root of the characteristic equation

$$x^p - \varphi_1 x^{p-1} - \dots - \varphi_{p-1} x - \varphi_p = 0.$$

If the process is stationary then all the roots of the above equation lie outside the unit circle.

First, we consider a time-domain analysis of $\{M_j^i(t)\}$, where i and j are fixed as before. In the followings we illustrate the fitting of some ARMA model through an example.

Let i and j be 11 and 1, respectively and simplify our notation by $\{M_{11}^1(t)\} = \{X_t\}$. To begin we observe that there is no trend or periodicity in the time series. Hence we start modeling the residuals. One can be given evidence of the stationarity of $\{X_t\}$ using the augmented Dickey-Fuller test. This method tests the existence of a unit root, which contradict the stationarity. Of course, there are other useful tests. At the confidence level of 95% the Dickey-Fuller test rejects the null hypothesis on the existence of a unit root. Thus, suppose that $\{X_t\}$ is stationary.

In the next step we estimate the orders p and q . It is common to use the sample autocorrelation and partial autocorrelation functions for this purpose. If the model is $\text{ARMA}(p, q)$ then the former vanishes at $q + 1$ and the latter does at $p + 1$. We find several possibilities like $(p, q) = (2, 0)$, $(0, 2)$, $(2, 1)$, $(2, 2)$, but consider just one case, the others go along the same way.

Let $(p, q) = (2, 1)$, that is, we are looking for a model of the form

$$X_t + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} = Z_t + \theta_1 Z_{t-1}.$$

Using the Innovations algorithm we have the estimations $\hat{\varphi}_1 = -0.3087$, $\hat{\varphi}_2 = 0.05429$, $\hat{\theta}_1 = 0.1506$, $\hat{\sigma}^2 = 0.005716$, where $\hat{\sigma}^2$ is the estimated variance of the white noise process $\{Z_t\}$.

Using the same methods for every possible pair (p, q) we have several different models. We apply the AICC value to select the final one. Interestingly, looking for well-fitting ARMA models for every pair of stations and channels, it seems that a significant number of them will have autoregressive and moving-average orders between 0-4, that is the orders are relatively low. Naturally, higher orders must be considered in some cases but those are mostly non-stationary time series and other tools shall be applied.

References

- [1] T. BALLA, Z. GÁL AND A. SZ. KARSAI, Sensor Based Analysis of the WiFi Interference, *Advances in Wireless Sensor Networks 2013, Conference Proceedings, Debrecen University Press*, 13-20.
- [2] W. DARGIE AND C. POELLABAUER, Fundamentals of wireless sensor networks: theory and practice, *John Wiley and Sons* (2010)
- [3] P. FUXJAGER, D. VALERIO, F. RICCIATO, The Myth of Non-Overlapping Channels: Interference Measurements in IEEE 802.11, *Fourth Annual Conference on Wireless on Demand Network Systems and Services - WONS '07* (2007) 1-8.
- [4] Z. GÁL AND GY. TERDIK, On the Statistical Analysis of Wireless Sensor vs Wired Data Network Traffics, *Carpathian Journal of Electronic and Computer Engineering Vol. 4, No. 1, 2011*, (2012) 41-47.
- [5] Z. GÁL AND GY. TERDIK, Statistical Analysis of Next Generation Network Traffics Based on Wavelets and Transformation ON/(ON+OFF), *Applied Computation Intelligence in Engineering and Information Technology, Springer-Verlag Berlin Heidelberg*, (2012) 107-122.
- [6] S. KAWADE, T. G. HODGKINSON, V. ABHAYAWARDHANA, Interference analysis of 802.11b and 802.11g wireless systems, *IEEE Vehicular Technology Conference - VTC* (2007) 787-791.
- [7] S. KAWADE, T. G. HODGKINSON, Analysis of Interference effects between co-existent 802.11b and 802.11g Wi-Fi systems, *IEEE Vehicular Technology Conference - VTC* (2008) 1881-1885.
- [8] WIN, MOE Z., PEDRO C. PINTO, AND LAWRENCE A. SHEPP, A Mathematical Theory of Network Interference and Its Applications, *Proceedings of the IEEE 97* (2009) 205-230.