Error Distribution in Maximum Likelihood Estimation of Radio Propagation Model Parameters

Andreas Achtzehn, Janne Riihijärvi, Petri Mähönen
Institute for Networked Systems, RWTH Aachen University
Kackertstrasse 9, D-52072 Aachen, Germany
Email: {aac, jar, pma}@inets.rwth-aachen.de

***Abstract*—A growing demand for bandwidth in wireless data networks has motivated policy makers to consider opening underused frequencies, e.g. in the TV broadcasting bands, to opportunistic secondary access. To support secondary network operations in these TV whitespaces (TVWS), centralized databases are envisioned that employ radio propagation models to derive the spectrum usage of the primary system. Precise estimation of the actual spectrum usage is highly important in this scenario, because maladjusted models can cause primary outage or hamper secondary capacity. The requirements for accurate propagation estimation have hence further increased.**

Using the example of a single TV transmitter, we study in this poster the prospects of maximum likelihood estimation for generic power law propagation models. Transmitter characteristics are considered known, but environment-specific parameters of the model are derived through fitting measurement results, obtained, e.g., through drive tests. We derive for the case of uncorrelated shadow fading the error distributions and their parameters for two parameters, namely pathloss coefficient and shadow variance. Our results are applicable to other large-scale scenarios where precise signal strength estimation is required. They can help to minimize the number of required measurements and to determine necessary protection policies.

I. Introduction

With the rapid growth of wireless communication services, demand for exclusively licensed spectrum bands has been constantly increasing. While the allocation of more spectrum to wireless services is limited by the availability of low-frequency spectrum, new plans for allowing access to otherwise licensed spectrum in an opportunistic manner are starting to gain momentum. The U.S. Federal Communications Commission (FCC) has, as forerunner among national regulators, opened bands used for TV broadcasting on a fair-access basis [1].

In order to control secondary access to these bands, regulators are envisioning a system of databases with primary spectrum usage information and allocation policies. The feasibility of such systems will be determined by their ability to correctly predict the signal strength of the primary TV transmitters in order to enforce protection requirements. Current prototypes tested in small-scale field trials use transmitter information and a fixed empirical propagation model for this calculation. We argue that for deployments in more complex radio environments, local peculiarities affecting the propagation such as building densities need to be taken into account as well. One method to improve prediction accuracy is to retrieve the parameter of a generic propagation model through maximum-likelihood fitting of measurements obtained, e.g., through extensive drive tests. In this poster, we study the feasibility of such approach in terms of the residual error of the employed mathematical framework. Building on a generic power law for propagation, we derive the error statistics for the estimated pathloss coefficient and the shadowing variance. Our contribution is not limited TVWS applications, but can generally help to minimize drive testing for a large-scale environment. It can also be used to adjust protection requirements of the primary system according to the observed environmental volatility.

The rest of this poster abstract is organized as follows: The system model with application to a single TV transmitter is presented in Section II. Our analysis of the error terms of pathloss coefficient and shadowing variance for the maximum likelihood estimation is given in Sections III and IV, respectively. Section V concludes this work.

II. System model

We consider a single TV transmitter that emits a broadcasting signal with known power P_{TX}. Its exposed position, e.g. on top of a hill, and carefully chosen measurement locations $\theta_{\mathrm{RX}} = \{\theta_i\}_{i=1}^{n}$ allow accurate estimation of the respective received signal powers $P_{\mathrm{dBm},i}$. Measurement points are located within a common radio environment but sufficiently separated to exhibit uncorrelated shadowing. We denote d_i as the Euclidean distance between the transmitter and receiver location θ_i. Our aim is to accurately estimate the received TV signal power at locations $\theta \notin \theta_{\mathrm{RX}}$ through a fitted propagation model.

Propagation models for such scenarios, e.g. the Okumura-Hata model [2], follow a common structure. The received power is assumed to decrease logarithmically with distance, whereby the extent of power decay depends on an environment-specific parameter α. Effects of antenna heights, frequency, and propagation zone are represented through a fixed, location-invariant offset L_{off} that may comprise multiple components. Values for these parameters are usually extrapolated from measurements in scenarios with *similar* propagation characteristics. The received signal power is derived as

$$P_{\mathrm{dBm},i} = P_{\mathrm{TX}} - L_{\mathrm{off}} - 10\alpha \log_{10}(d_i) - \mathcal{X}_i, \qquad (1)$$

978-1-4673-1905-8/12/$31.00 ©2012 IEEE

where the last term, $\mathcal{X}_i$, describes location-specific shadow fading caused, e.g., by buildings and trees. We model $\mathcal{X}_i$ as a i.i.d. random variable (RV) with $\mathcal{X}_i \sim \mathcal{N}(0, \sigma^2)$, i.e. in contrast to σ^2 which can be estimated, the specific value of $\mathcal{X}_i$ can only be determined through direct measurements. We assume that the offset parameter L_{off} is accurately derived through a small number of reference measurements, but that the parameter tuple (α, σ^2) needs to be acquired individually due to the strong dependency on the actual propagation environment.

A. Maximum likelihood estimation of parameters

Let

$$f(P, \theta|\alpha, \sigma^2) = \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}P} e^{\frac{(\ln P - \tilde{\mu}_i)^2}{2\tilde{\sigma}^2}} \tag{2}$$

be the conditional probability density function (PDF) that describes the probability that at location θ a power P (in linear scale) is measured given propagation parameters (α, σ^2). $f(\cdot)$ represents a log-normal distribution with parameters

$$\tilde{\sigma}^2 = c\sigma^2, c := 10/\ln(10), \tag{3}$$

and

$$\tilde{\mu}_i = c\left(P_{\text{TX}} - L_{\text{off}} - 10\alpha \log_{10}(d_i)\right). \tag{4}$$

As noted above, the random components of the measurements are independent, hence the joint probability density function for a set of observations can be expressed by multiplication of their respective PDFs. A consistent maximum-likelihood estimator for the propagation model parameters finds $(\hat{\alpha}, \hat{\sigma}^2)$ that maximizes this joint conditional probability. Using the fact that $f(\cdot) \geq 0 \ \forall \ P, d \in \mathbb{R}$ and the monotonicity of the logarithm, the fitting problem is alternatively described as

$$(\hat{\alpha}, \hat{\sigma}^2) = \underset{\alpha, \sigma^2}{\arg\max} \sum_{i=1}^{n} \ln f(P_{\text{mW},i}, d_i|\alpha, \sigma^2), \tag{5}$$

where $P_{\text{mW},i}$ is the received power at location θ_i in mW.

III. Error Distribution of Pathloss Coefficient α

The first derivative of (5) with respect to α is solved for zero and yields a simple estimator

$$\hat{\alpha} = \frac{\sum\limits_{i=1}^{n}(P_{\text{TX}} - L_{\text{off}} - P_{\text{dBm},i}) \ln d_i}{c\sum\limits_{i=1}^{n}(\ln d_i)^2}. \tag{6}$$

In the following, we let $\hat{\alpha} = \alpha^* + T$, with α^* being the true pathloss coefficient and T being the error term of the MLE estimator. We find that T is normally distributed with

$$T \sim \mathcal{N}\left(0, \frac{1}{c^2\sum\limits_{i=1}^{n}(\ln d_i)^2}(\sigma^*)^2\right), \tag{7}$$

where σ^* denotes the true standard deviation of the shadowing term. A proof of (7) is provided in Appendix A. The analysis shows that the estimator is unbiased, expressed through the zero-mean distribution of the error term, and for $n \to \infty$ converges to the true pathloss coefficient. The geometry of selected measurement points has a direct impact on the error term's standard deviation, with higher measurement distances yielding a smaller variance of the error term. Carefully note though that in a real measurement environment practical hardware limitations may lower the performance. For validation of our theoretical results we show the analytical PDF of the error term in comparison to PDFs derived through Monte-Carlo simulations in Figure 1. The average of the numerical results matches the analytical ones, and the confidence intervals show a tight fit, in particular for higher error offsets.

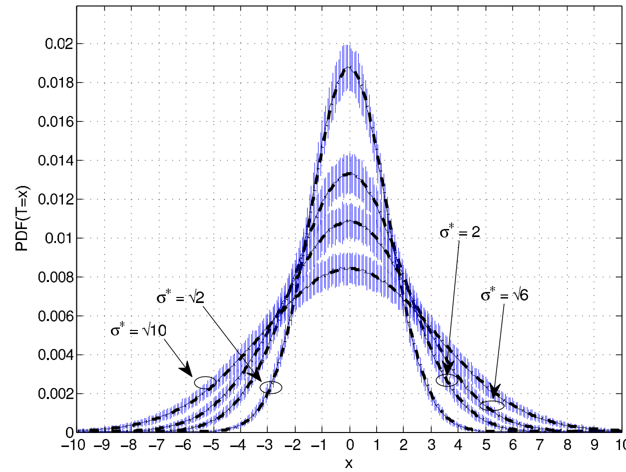

Fig. 1. Probability density function (PDF) of error term T of MLE estimator $\hat{\alpha}$ for different values of σ^*. Blue lines denote the 95% confidence intervals of the empirical PDF derived through Monte-Carlo simulation with 100 locations, repeated for 50,000 realizations of the shadowing terms. Locations have been selected so that $\sum_i^n(\ln d_i)^2 = 1/c^2$, i.e. $T \sim N\left(0, (\sigma^*)^2\right)$. Black lines denote the analytical results.

IV. Error Distribution of Shadowing Variance

Setting $\alpha^* = \hat{\alpha}$ and considering the error term of the pathloss estimator T as in (13), $\hat{\sigma}^2$ can be derived as

$$\hat{\sigma}^2 = \frac{(\sigma^*)^2}{n} \times \sum_{i=1}^{n} \underbrace{w_i \left(\frac{V_i}{\sqrt{w_i}\sigma^*}\right)^2}_{R_i}, \tag{8}$$

with RV $V_i = X_i - 10T\log_{10} d_i$, scaling term $w_i = A^{-1}(1 + 2\ln(d_i)^2) + A^{-2}\ln(d_i)^4 + 1$ and $A = \sum_i \ln(d_i)$. The sum consists of n correlated RVs R_i, whereby each RV is gamma-distributed with parameters

$$R_i \sim \Gamma\left(r = \frac{1}{2}, \theta_i = 2w_i\right), \tag{9}$$

see Appendix B, and $(R/n) \to 1$ for $n \to \infty$.

The multiplicative term $R = \sum_{i=1}^{n} R_i$ has a PDF that is a weighted gamma series with [3]

$$f_R(y) = \prod_{i=1}^{n} \sqrt{\frac{\lambda_1}{\lambda_i}} \sum_{k=0}^{\infty} \frac{\delta_k y^{n/2+k-1} e^{-y/\lambda_1}}{\lambda_1^{n/2+k}\Gamma(n/2+k)}, y \geq 0 \tag{10}$$

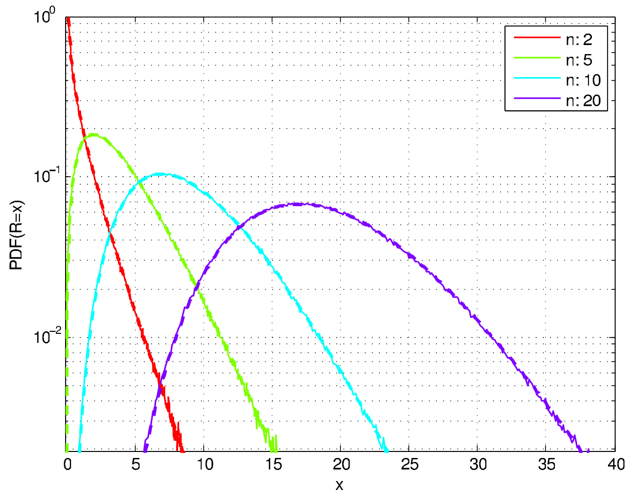

Fig. 2. Distribution of $R = \sum_i^n R_i$ for $(\sigma^*)^2 = 3$ for different number of measurement locations n. Sold lines depict results retrieved through Monte-Carlo simulations for different realizations of shadowing terms and locations. Dashed lines denote the analytical approximation if sum terms are considered uncorrelated.

where $\lambda_1 = \min\{\lambda_i\}_{i=1}^n$, with λ_i being the ith eigenvalue of the matrix $A = DC$ where D is the $N \times N$ diagonal matrix with elements $D_{i,i} = \theta_i$. C is an $N \times N$ positive definite matrix with the square roots of the correlation coefficients ρ_{R_i,R_j}, yielding

$$C_{i,j} = \frac{1}{\sqrt{2}} \sqrt{\frac{\mathbb{E}[V_i^2 V_j^2]}{(\sigma^*)^4 w_i w_j} - 1}. \quad (11)$$

The coefficients δ_k can be obtained recursively as

$$\delta_0 = 0$$
$$\delta_{k+1} = \frac{1}{2(k+1)} \sum_{i=1}^{k+1} \sum_{j=1}^{n} \left(1 - \frac{\lambda_1}{\lambda_j}\right)^i \delta_{k+1-i}. \quad (12)$$

The error term can be sufficiently approximated by a single gamma RV with $R \sim \Gamma(k = (n-1)/2, \theta = 2)$ if uncorrelatedness of the sum terms is assumed [4]. In Figure 2 we compare the results of Monte-Carlo simulation with this approximation and find a close fit. We furthermore observe that the expected value of the shadow variance estimator is $\mathbb{E}[\hat{\sigma}^2] = (\sigma^*)^2 \times (n-1)/n$, i.e. the estimator is biased.

V. Conclusion

In this poster we have analyzed the performance of maximum likelihood estimators for parameters of a generic power law propagation model. We have analytically derived the corresponding error distributions, and confirmed the results through extensive Monte-Carlo simulations. These results allow the evaluation of uncertainty in estimated propagation models for a given measurement configuration, and have number of applications including better estimation of protection zones for TV transmitters in whitespace applications.

Acknowledgment

We acknowledge a partial financial support from the European Union through the FP7 project INFSO-ICT-248303 QUASAR. We are also grateful for a partial funding from DFG received through UMIC research centre.

Appendix A
Proof of Equation 7

Substituting (1) into (6) yields

$$\hat{\alpha} = \alpha^* + \sum_{i=1}^{n} X_i \underbrace{\frac{\ln d_i}{c \sum_{i=1}^{n} (\ln d_i)^2}}_{Q_i}. \quad (13)$$

We observe that

$$Q_i \sim \mathcal{N}\left(\underbrace{0}_{\mu_{Q_i}}, \underbrace{\frac{(\ln d_i)^2}{c^2 (\sum_{i=1}^{n} (\ln d_i)^2)^2} (\sigma^*)^2}_{\sigma_{Q_i}^2} \right). \quad (14)$$

Then T is the sum of n independent, normally distributed random variables Q_i, i.e. $T \sim \mathcal{N}\left(\sum_{i=1}^n \mu_{Q_i}, \sum_{i=1}^n \sigma_{Q_i}^2\right)$.

Appendix B
Proof for Shadowing Approximation

First, we evaluate V_i for the error term T of (13). We derive

$$V_i = \left(1 + \frac{(\ln d_i)^2}{A}\right) \mathcal{X}_i + \sum_{j \neq i}^{n} \mathcal{X}_j \frac{\ln d_i \ln d_j}{A}, \quad (15)$$

i.e. V_i is the sum of n scaled normal-distributed RVs, each representing the shadowing variance at a particular location. Hence, the distribution is

$$V_i \sim \mathcal{N}\left(0, \underbrace{\left(\frac{1}{A}(1 + 2(\ln d_i)^2) + \frac{(\ln d_i)^4}{A^2} + 1\right)}_{w_i} (\sigma^*)^2 \right). \quad (16)$$

While $(V_i / \sqrt{w_i} \sigma^*)^2$ is distributed according to a χ_1^2 distribution, $R_i = w_i \times (V_i / \sqrt{w_i} \sigma^*)^2$ is Γ-distributed with $R_i \sim \Gamma(r = 1/2, \theta = 2 w_i)$.

References

[1] FCC, "In the matter of unlicensed operation in the TV broadcast bands: Second memorandum opinion and order," Tech. Rep. 10-174A1, 2010.

[2] M. Hata, "Empirical Formula for Propagation Loss in Land Mobile Radio," *Transactions on Vehicular Technology*, 1980.

[3] M.-S. Alouini, A. Abdi, and M. Kaveh, "Sum of gamma variates and performance of wireless communication systems over Nakagami-fading channels," *Vehicular Technology, IEEE Transactions on*, vol. 50, no. 6, pp. 1471 –1480, Nov 2001.

[4] P. Moschopoulos, "The distribution of the sum of independent gamma random variables," *Annals of the Institute of Statistical Mathematics*, vol. 37, pp. 541–544, 1985.