Statistical path loss parameter estimation and positioning using RSS measurements

Henri NURMINEN, Jukka TALVITIE, Simo ALI-LÖYTTY, Philipp MÜLLER, Elena-Simona LOHAN,
Robert PICHÉ and Markku RENFORS
Tampere University of Technology, Tampere, Finland
Emails: {henri.nurminen, jukka.talvitie, simo.ali-loytty, philipp.muller, elena-simona.lohan,
robert.piche, markku.renfors}@tut.fi

***Abstract*—An efficient Bayesian method for off-line estimation of the position and the path loss model parameters of a base station is presented. Two versions of three different on-line positioning methods are tested using real data collected from a cellular network. The tests confirm the superiority of the methods that use the estimated path loss parameter distributions compared to the conventional methods that only use point estimates for the path loss parameters. Taking the uncertainties into account is computationally demanding, but the Gauss–Newton optimization methods is shown to provide a good approximation with computational load that is reasonable for many real-time solutions.**

outdoor positioning; cellular network; received signal strength; path loss model; statistical estimation

I. Introduction

Hybrid navigation means navigation using measurements from different sources, such as Global Navigation Satellite Systems (e.g. GPS), Inertial Measurement Unit, and/or local wireless networks such as cellular networks, WLAN or Bluetooth. Range, pseudorange, deltarange, altitude, map constraint and heading are examples of measurements in hybrid navigation. This paper focuses on hybrid navigation using one or more wireless networks. The ranges from the network's base stations (BS) are computed using received signal strengths (RSS) as measurements.

A path loss (PL) model is a model for signal attenuation in space. In the literature, for example in [1] and [2], there are many different path loss modeling methods from deterministic and computationally heavy ray-tracing algorithms to empirical and semi-empirical channel models based on extensive measurement campaigns. Each model contains a set of tunable parameters which attempt to capture the nature of the investigated radio propagation environment.

This article uses a simple statistical path loss model. A method for dynamic estimation of the model parameters for each BS using learning data collected at known positions is presented. Estimation is based on the concepts of Bayesian statistics, which is a flexible and theoretically principled framework. The number of required path loss parameters is kept small in order to keep down the computational complexity and the amount of information required in the positioning phase. As a very important built-in property, the presented Bayesian method returns also a statistical description of the uncertainty for estimated parameter values.

Furthermore, this article shows the influence of PL parameter uncertainty on the positioning results. Three positioning methods are presented in this paper: grid method, Monte Carlo -based Metropolis–Hastings (MH) sampler and computationally lighter Gauss–Newton method (GN). For each of these, two versions were implemented: The first one uses point estimates for the path loss parameters and assumes them to be accurate. The second version assumes the parameters to follow specified probability distributions. Using collected real data sets, the latter is shown to be superior in consistency and similar or slightly better in accuracy. The latter also outperforms the coverage area solution which does not utilize RSS measurements.

The paper is organized as follows: First, in Section II the path loss model is introduced and the method for estimating the model parameters is presented. Then, in Section III a statistical measurement model for the positioning phase is presented, and the positioning algorithms are presented in Section IV. The results are presented in Section V. Finally, Section VI shows the conclusions.

Notations: Matrices are denoted with unitalicised uppercase letters. Vectors and scalars are not distinguished. $\mathrm{N}(m, \mathrm{P})$ refers to the (multivariate) normal distribution with mean m and covariance matrix P, and $\mathrm{N}_{\mathrm{P}}^{m}(x)$ refers to its probability density function (pdf) evaluated at x. Notation $u \leftarrow \mathrm{A}$ means that pseudo-random number u is generated from probability distribution A.

II. Path loss model

A. Path loss model input

The input for estimating a single base station (BS) path loss model is a set of RSS measurements of signals transmitted by the BS. The measurement set Ω includes N_m measurements given as

$$\Omega \in \{(x_i, P_i) \mid i = 1, 2, \ldots, N_m\}, \tag{1}$$

where $x_i \in \mathbb{R}^2$ includes the easting and northing of the i:th measurement point, and P_i is the received signal power of the i:th measurement point in dBm. We assume that the transmitter power and antenna gains are fixed during the measurements. This should be a valid assumption, since the received signal power values in cellular networks are measured through fixed

beacon channels whose transmission powers are not varying in time. According to this the received signal power is only dependent on the measurement coordinates x_i.

In this paper, we are using only 2-dimensional coordinates, since in many positioning applications this is enough to fulfill the use case expectations. Furthermore, in macro scale cellular environment the effect of altitude becomes considerable only near elevated BSs where the radio wave propagation distance does not approach to zero when approaching to the BS position in horizontal plane.

B. Path loss model definition

Friis's law determines the received signal power as a function of distance in a free space as

$$p_{rx}(d) = p_{tx} g_{tx} g_{rx} \left(\frac{\lambda}{4\pi d} \right)^2, \tag{2}$$

where p_{tx}, g_{tx}, g_{rx}, and λ are the transmitted signal power, transmitter antenna gain, receiver antenna gain, and signal wavelength, respectively. The distance between transmitter and receiver antenna is d. The square term is the actual channel dependent path loss term, while the other parameters are transmitter and receiver dependent. However, using the free-space model could be a practical approach only in line-of-sight scenarios, but not in real-life cellular networks where buildings and ground surface fluctuations act as obstacles to the radio signal path.

One of the most recognized outdoor path loss models is the classical log-distance model (or power law model) [3]. In the log-distance model which the received signal power is defined as

$$P_{rx}(d) = P_{rx}(d_0) - 10n \log_{10} \left(\frac{d}{d_0} \right) + w \tag{3}$$

where the power $P_{rx}(\cdot)$ is given in logarithmic scale, d_0 is a reference distance, n is a path loss exponent, and $w \sim \mathrm{N}(0, \sigma^2)$ is a normally distributed random variable which models the slow fading (shadowing) effects. Here the path loss exponent n and the slow fading standard deviation are dependent on the local propagation environment. Notice that since the term $P_{rx}(d_0)$ indicates the received signal power at the reference distance d_0, it automatically takes account of the transmission power along with the antenna gains and wavelength shown in (2). Moreover, apart from the slow fading, $P_{rx}(d)$ is only affected by the path loss exponent n whenever $d > d_0$. Now, by defining $d_0 = 1\,\mathrm{m}$, and denoting $P_{rx}(d_0) = A$, it is possible to write the final path loss model as

$$P_{rx}(d) = A - 10n \log_{10}(d) + w, \tag{4}$$

where the parameter A is referred to as apparent transmission power.

C. Estimation of BS position and path loss parameters

BS position and path loss parameters are estimated using the Iterative Reweighted Least Squares method (Gauss–Newton method). A similar kind of method is studied for positioning in simulated cases in [4]. The function to be minimized is

$$\phi(A, n, m) = \sum_{i=1}^{N_m} \left(P_i - h^i_{\text{est}}(A, n, m) \right)^2, \tag{5}$$

where

$$h^i_{\text{est}}(A, n, m) = A - 10n \log_{10}(||m - x_i||).$$

A is apparent transmission power, n path loss exponent and m BS position. The Jacobian matrix of the measurement model function h^i_{est} is

$$\mathrm{J} = \begin{bmatrix} 1 & -10\log_{10}(||m - x_1||) & -\frac{10}{\ln(10)} n \frac{(m-x_i)^{\mathrm{T}}}{||m-x_1||^2} \\ \vdots & \vdots & \vdots \\ 1 & -10\log_{10}(||m - x_{N_m}||) & -\frac{10}{\ln(10)} n \frac{(m-x_{N_m})^{\mathrm{T}}}{||m-x_{N_m}||^2} \end{bmatrix}. \tag{6}$$

The Bayesian Gauss–Newton algorithm is described in detail in Algorithm 3.

The measurement data of cellular networks tends to spatially correlated [5]. In order to reduce the effect of correlations, the measurements are mapped to a grid of pre-specified points before the estimation process. In this case, the distance between adjacent grid points is 50 meters, and the RSS value of a grid point is the mean of the RSSs observed in the proximity of the grid point.

To improve convergence properties of the Gauss–Newton algorithm, all the quantities are given an almost uninformative Gaussian prior, i.e. a Gaussian distribution with so large variance that the influence on the optimum is negligible. A suitable initial value for the BS position is the position of the strongest observed measurement, because otherwise the algorithm might locate the BS position to the area of the weakest RSSs. Initial values for A and n can be chosen more arbitrarily from the valid ranges, since the distribution is typically unimodal, if the number of data points is large.

Note that the algorithm also returns an approximation for the covariance matrix of each quantity. Consequently, we are potentially able to distinguish between trustworthy and untrustworthy path loss models. In the Bayesian sense, the algorithm tries to estimate the MAP value (*maximum a posteriori*), and the covariance matrix is the covariance of the linearized model.

In an ideal case without any slow fading variations the BS position would be found at the coordinate point where the received signal power reaches its maximum value. However, in practice there might be several clear peaks in the received signal power map or there might not be enough measurements to find even a single peak. Besides, the BS might not even be located inside the measured power map.

However, it should be emphasized that there is no need to know the exact true BS positions as long as the same estimated

BS positions are used also in the positioning phase. Thus, one could easily refer to a certain kind of pseudo BS positions. Furthermore, it can be shown that using correct BS positions may even result in worse modeling outcome. One reason is the used 2D model that does not work accurately in close proximity of the BS position where BS antenna height and antenna tilting have a considerable effect on the received signal power.

2D-projection effects are taken into account by increasing the covariance matrix of the BS position artificially with a diagonal constant matrix. This reflects also errors that stem from GPS errors in the learning data and measurement error correlations due to environmental effects. For this reason and for reducing the number of recorded parameters, the cross-covariances of BSs and path loss parameters are ignored. BS positions and path loss parameters are thus assumed to be uncorrelated. Fig. 1 shows power maps (interpolated between the measurement points) of two separate BSs, and the resulting BS position estimates along with the covariance ellipse.

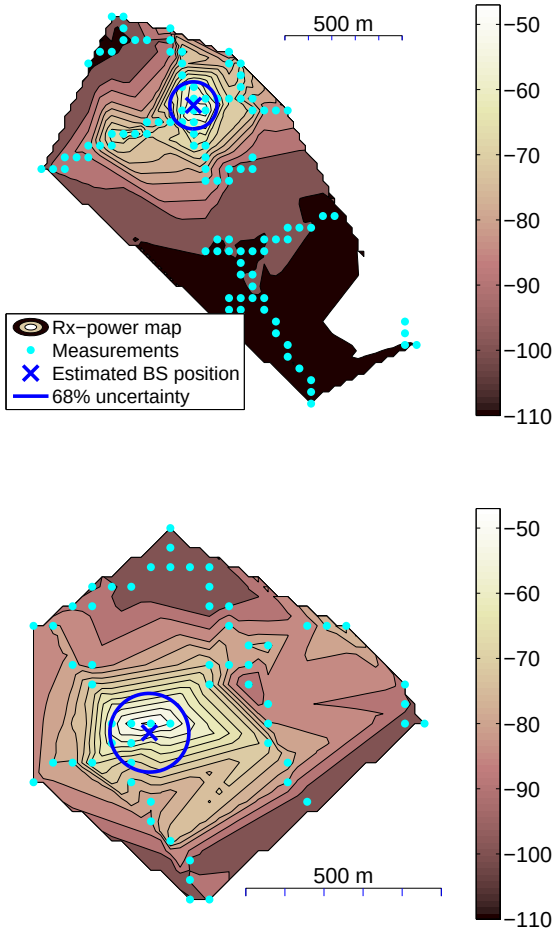

Figure 1: Power maps of two separate BS's and the BS position estimates along with covariance ellipses.

As pointed out before, path loss exponent n and the slow fading standard deviation are highly dependent on the radio propagation environment. For example, in a shadowed urban cellular radio network the typical values of n and σ are varying around $n = 1 - 8$ and $\sigma = 3 - 8\,\text{dBm}$, respectively [2, 6]. Examples of resulted path loss model curves can be found in Fig. 2, in which the path loss models are derived for the same BSs that were previously show in Fig. 1.

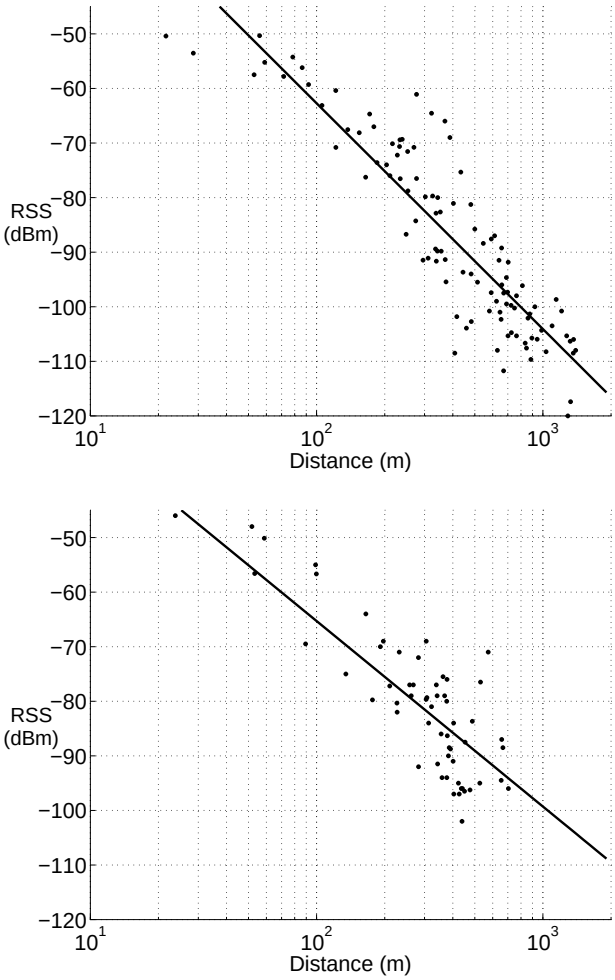

Figure 2: Path loss curves for two separate BSs.

III. Measurement Model

A. General form

Consider the nonlinear Gaussian measurement model

$$y = h(x, a) + v, \tag{7}$$

where $y \in \mathbb{R}^{N_y}$ is the vector of observations, $x \in \mathbb{R}^{N_x}$ represents the state of the system and $a \in \mathbb{R}^{N_a}$ represents nuisance parameters that have conditional prior distribution $p(a|x)$. The random noise term $v \sim \mathrm{N}(0, \mathrm{R})$ is assumed to be

zero mean and independent of state x and parameter vector a. The likelihood function of the model is

$$p(y|x) = \int p(y, a|x)\,\mathrm{d}a = \int p(y|a, x)p(a|x)\,\mathrm{d}a \quad (8)$$

Let the prior distribution of the state be $p(x)$. The prior should include all the information that is known before the measurement. In the case of time series, the prior can be, for example, the prediction estimate of a Bayesian filter. By Bayes' rule, the posterior pdf of the state is thus

$$p(x|y) = \frac{p(y|x)p(x)}{\int p(y|x)p(x)\,\mathrm{d}x} = \frac{\int p(y|a, x)p(a|x)p(x)\,\mathrm{d}a}{\int\int p(y|a, x)p(a|x)p(x)\,\mathrm{d}a\,\mathrm{d}x}. \quad (9)$$

B. Path loss model

The path loss model with uncertain parameters presented in Section II is a special case of model (7), where y is the vector of RSS measurements $P = \begin{bmatrix} P_1 & \dots & P_{N_P} \end{bmatrix}^{\mathrm{T}}$ and x is the user position. In addition, a contains unknown system parameters

$$a = \begin{bmatrix} A_1 & n_1 & m_1^{\mathrm{T}} & \cdots & A_{N_P} & n_{N_P} & m_{N_P}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}},$$

and the measurement model function is

$$h(x, a) = \begin{bmatrix} A_1 - 10n_1 \log_{10}(||m_1 - x||) \\ \vdots \\ A_{N_P} - 10n_{N_P} \log_{10}(||m_{N_P} - x||) \end{bmatrix},$$

where N_P is the number of observed BSs. Measurement noise covariance matrix is $\mathrm{R} = \sigma^2 \cdot \mathrm{I}_{N_P \times N_P}$. Note that because the path loss parameters of a BS are isotropic, $p(a|x) = p(a)$. For simplicity, this paper assumes that the estimated path loss parameters prior distributions are Gaussian and that BS position and PL parameters are independent a priori. Thus,

$$\begin{aligned} p(a) &= p(A_{1:N_P}, n_{1:N_P}, m_{1:N_P}) \\ &= \prod_{i=1}^{N_P} \mathrm{N}^{\hat{\mu}_{A_i,n_i}}_{\hat{\Sigma}_{A_i,n_i}} \left(\begin{bmatrix} A_i \\ n_i \end{bmatrix} \right) \cdot \mathrm{N}^{\hat{m}_i}_{\hat{\Sigma}_{m_i}}(m_i), \end{aligned} \quad (10)$$

where the parameters $\hat{\mu}_{A_i,n_i} = \begin{bmatrix} \hat{A}_i & \hat{n}_i \end{bmatrix}^{\mathrm{T}}$, $\hat{\Sigma}_{A_i,n_i}$, $\hat{m}_i$ and $\hat{\Sigma}_{m_i}$ are estimated from the learning data using the Gauss–Newton algorithm.

The priors are modeled to be normal, since the Gauss–Newton algorithm requires this in its basic form and the normal pdf of A and n is the conjugate prior of the likelihood. However, other prior distribution families such as Student's t-distribution could also be studied.

The next section considers the optimal state estimation problem for this model.

IV. Positioning algorithms

This section uses a Gaussian prior distribution $p(x) = \mathrm{N}^{\hat{x}}_{\hat{\Sigma}_x}(x)$ for the user's position. In case of multimodal likelihood function, the prior may function as a regularizer. Furthermore, it may reflect location information from other sources, and in case of time series filtering, the filter prediction is the prior mean $\hat{x}$ and covariance $\hat{\Sigma}_x$.

A. Grid method

By (9), the posterior pdf value at point x is

$$\begin{aligned} p(x|P_{1:N_P}) \propto \prod_{i=1}^{N_P} \iiint p(P_i|x, A_i, n_i, m_i)p(A_i, n_i)p(m_i) \\ \mathrm{d}A_i\,\mathrm{d}n_i\,\mathrm{d}m_i \cdot p(x), \end{aligned} \quad (11)$$

which can be approximated using standard Monte Carlo integration. The grid method is presented in Algorithm 1.

The most crucial implementation issues are the Monte Carlo sample size parameter N as well as grid size and density. Note that combining of the likelihood of each BS is done in logarithmic space to avoid numerical underflows.

Algorithm 1 Grid with Monte Carlo Integration

1) Set a grid $\{x_m \in \mathbb{R}^2 \mid m \in \{1, \dots, N_m\}\}$ that covers most of the prior probability mass.
2) For each heard base station $i = 1, \dots, N_P$, draw
$$\begin{bmatrix} A_i^{(k)} \\ n_i^{(k)} \end{bmatrix} \leftarrow \mathrm{N}\left(\begin{bmatrix} \hat{A}_i \\ \hat{n}_i \end{bmatrix}, \hat{\Sigma}_{A_i,n_i} \right)$$
$$m_i^{(k)} \leftarrow \mathrm{N}\left(\hat{m}_i, \hat{\Sigma}_{m_i} \right)$$
for $k = 1, \dots, N$.
3) At each grid point x_m compute for each AP $i = 1, \dots, N_P$ and for each each sample $k = 1, \dots, N$
$$I_{i,m}^{(k)} := \mathrm{N}^{P_i}_{\sigma^2}\left(A_i^{(k)} - 10n_i^{(k)} \log_{10}\left(\left|\left| m_i^{(k)} - x_m \right|\right| \right) \right),$$
and $I_{i,m} := \frac{1}{N}\sum_{k=1}^{N} I_{i,m}^{(k)}$. Then set
$$\ell_m := \ln\left(\mathrm{N}^{\hat{x}}_{\hat{\Sigma}_x}(x_m) \right) + \sum_{i=1}^{N_P} \ln(I_{i,m}), \quad L_m := \exp(\ell_m).$$
4) Normalize the grid to get a set of weights $w_m = \frac{L_m}{\sum_{m=1}^{n_m} L_m}$ and compute mean and covariance estimates
$$\hat{x}^+ := \sum_{m=1}^{N_m} w_m x_m$$
$$\hat{\Sigma}_x^+ := \sum_{m=1}^{N_m} w_m (x_m - \hat{x}^+)(x_m - \hat{x}^+)^{\mathrm{T}}.$$

B. Metropolis–Hastings method

The Metropolis–Hastings (MH) sampler generates Monte Carlo samples from an arbitrary posterior distribution of a multivariate random variable. It is an iterative algorithm that can be proved to converge towards the target distribution. The posterior mean and covariance can then be approximated by the sample mean and covariance of the sampled set. The algorithm is presented in Algorithm 2.

The MH sampler uses a so-called proposal distribution, from which it is straightforward to generate random numbers.

At each iteration of the algorithm, proposal values for the estimated variables are drawn from the proposal distribution. The proposal values are then accepted with the probability that is proportional to the ratio of the pdf values of the proposal value and the latest accepted value. [7, Ch. 5]

Note that

$$\begin{aligned} & p(x, m_{1:N_P}|P_{1:N_P}) \\ & \overset{(9)}{\propto} \iint p(P_{1:N_P}|x, A_{1:N_P}, n_{1:N_P}, m_{1:N_P}) \\ & \quad \cdot p(A_{1:N_P}, n_{1:N_P})\, \mathrm{d}A_{1:N_P}\, \mathrm{d}n_{1:N_P} \cdot p(x) \cdot p(m_{1:N_P}) \\ & \propto p(x) \prod_{i=1}^{N_P} p(m_i) \det(\check{\Sigma}_{A_i,n_i})^{\frac{1}{2}} \exp\left(\frac{1}{2}\check{\mu}_{A_i,n_i}^{\mathrm{T}} \check{\Sigma}_{A_i,n_i}^{-1} \check{\mu}_{A_i,n_i}\right) \end{aligned} \tag{12}$$

where

$$\begin{aligned} \check{\Sigma}_{A_i,n_i} &:= \left(\frac{1}{\sigma^2}{\mathrm{B}_i}^{\mathrm{T}}\mathrm{B}_i + \hat{\Sigma}_{A_i,n_i}^{-1}\right)^{-1} \\ \check{\mu}_{A_i,n_i} &:= \check{\Sigma}_{A_i,n_i}\left(\frac{1}{\sigma^2}{\mathrm{B}_i}^{\mathrm{T}} P_i + \hat{\Sigma}_{A_i,n_i}^{-1}\begin{bmatrix}\hat{A}_i \\ \hat{n}_i\end{bmatrix}\right), \end{aligned}$$

where $\mathrm{B}_i = \begin{bmatrix}1 & -10\log_{10}(||m_i - x||)\end{bmatrix}$. The simple form of this formula enables analytical integration over PL parameters A and n.

In the implementation phase, great care must be taken when setting the proposal distributions to make the algorithm converge in a computationally feasible number of iterations. For convenience, the proposal distributions are chosen to be multivariate normal with the latest accepted value as the mean and the covariance matrices P_x and P_{m_i} tuned from prior covariances of x and $m_{1:N_P}$.

C. Gauss–Newton method

With suitable measurement models, iterative state estimation methods can be as accurate as any closed form solution but simpler and easier to implement [8]. The Gauss–Newton method, also known as the Iterative Reweighted Least Squares method, is tested for positioning with the presented path loss model. The detailed description is in Algorithm 3.

The iteration does not converge globally, but including good enough prior information and initial values prevents the method from diverging. To improve convergence properties further, the step length in the state-space is adaptive so that the objective function value decreases at every iteration. As an exception to this, the while loop is exited after a number of iterations to ensure stability. The global convergence results of the Gauss–Newton method with adaptive step length are discussed in [9, 10].

For formulating the Jacobian matrix that is needed in the Gauss–Newton algorithm, the analytical partial derivatives of the measurement function h are formed:

$$\begin{aligned} & \frac{\partial h_i}{\partial x} = \frac{10}{\ln(10)} n_i \frac{(m_i - x)^{\mathrm{T}}}{||m_i - x||^2}, \quad \frac{\partial h_i}{\partial A_i} = 1, \\ & \frac{\partial h_i}{\partial n_i} = -10\log_{10}(||m_i - x||), \frac{\partial h_i}{\partial m_i} = -\frac{10}{\ln(10)} n_i \frac{(m_i - x)^{\mathrm{T}}}{||m_i - x||^2}. \end{aligned}$$

Algorithm 2 Metropolis–Hastings algorithm

1) Set $x^{(0)} := \hat{x}$, $A_i^{(0)} := \hat{A}_i$, $n_i^{(0)} := \hat{n}_i$ and $m_i^{(0)} := \hat{m}_i^{(0)}$ for $i = 1, \ldots, N_P$. Set $p^{(0)}$ using the formulae in step 3. Set $k = 1$.
2) Generate $x'^{(k)} \leftarrow \mathrm{N}(x^{(k-1)}, \mathrm{P}_x)$, and for each AP $i = 1, \ldots, N_P$, generate $m_i'^{(k)} \leftarrow \mathrm{N}(m_i^{(k)}, \mathrm{P}_{m_i})$.
3) For each $i = 1, \ldots, N_P$, compute $\mathrm{B}_i^{(k)} = \begin{bmatrix}1 & -10\log_{10}\left(\left|\left|m_i'^{(k)} - x'^{(k)}\right|\right|\right)\end{bmatrix}$ and

$$\begin{aligned} \check{\Sigma}_{A_i,n_i}^{(k)} &:= \left(\frac{1}{\sigma^2}{\mathrm{B}_i^{(k)}}^{\mathrm{T}}\mathrm{B}_i^{(k)} + \hat{\Sigma}_{A_i,n_i}^{-1}\right)^{-1} \\ \check{\mu}_{A_i,n_i}^{(k)} &:= \check{\Sigma}_{A_i,n_i}^{(k)}\left(\frac{1}{\sigma^2}{\mathrm{B}_i^{(k)}}^{\mathrm{T}} P_i + \hat{\Sigma}_{A_i,n_i}^{-1}\begin{bmatrix}\hat{A}_i \\ \hat{n}_i\end{bmatrix}\right) \\ p'^{(k)} &:= \ln\left(\mathrm{N}_{\hat{\Sigma}_x}^{\hat{x}}(x'^{(k)})\right) + \sum_{i=1}^{N_P}\left[\frac{1}{2}\ln\left(\det(\check{\Sigma}_{A_i,n_i}^{(k)})\right)\right. \\ &\quad \left. + \frac{1}{2}\,{\check{\mu}_{A_i,n_i}^{(k)}}^{\mathrm{T}}\,{\check{\Sigma}_{A_i,n_i}^{(k)}}^{-1}\,\check{\mu}_{A_i,n_i}^{(k)} + \ln\left(\mathrm{N}_{\hat{\Sigma}_{m_i}}^{\hat{m}_i}(m_i'^{(k)})\right)\right] \end{aligned}$$

4) Set $r := \exp(p'^{(k)} - p^{(k-1)})$. Generate $u \leftarrow \mathrm{Uni}(0, 1)$. Compute
   if $r > u$ **then**
      for $i = 1 : N_P$ **do**
         $m_i^{(k)} := m_i'^{(k)}$
      end for
      $x^{(k)} := x'^{(k)}$, $p^{(k)} := p'^{(k)}$
   else
      for $i = 1 : N_P$ **do**
         $m_i^{(k)} := m_i^{(k-1)}$
      end for
      $x^{(k)} := x^{(k-1)}$, $p^{(k)} := p^{(k-1)}$
   end if
5) Set $k := k + 1$. If $k < N$, go to step 2. Otherwise, set

$$\begin{aligned} \hat{x}^+ &:= \frac{1}{N}\sum_{k=1}^{N} x^{(k)} \\ \hat{\Sigma}_x^+ &:= \frac{1}{N}\sum_{k=1}^{N}(x^{(k)} - \hat{x}^+)(x^{(k)} - \hat{x}^+)^{\mathrm{T}}. \end{aligned}$$

The remaining partial derivatives are zeros. Note that the Jacobian matrix J of the measurement model is always full rank so the least-squares estimation can be performed. The measurement covariance matrix R is the diagonal matrix of the measurement variances. In Algorithm 3 the complete state is denoted with $z = \begin{bmatrix}x \\ a\end{bmatrix}^{\mathrm{T}}$. As in the PL parameter estimation phase, the output of the algorithm contains estimates for the MAP and the covariance matrix of the posterior of the linearized model.

Algorithm 3 Gauss–Newton algorithm

1) Choose the stopping tolerance δ. Let

$$\hat{\Sigma}_z := \text{blkdiag}(\hat{\Sigma}_x, \hat{\Sigma}_{A_1,n_1}, \hat{\Sigma}_{m_1}, \ldots, \hat{\Sigma}_{A_{N_P},n_{N_P}}, \hat{\Sigma}_{m_{N_P}})$$

and

$$\hat{z} := \begin{bmatrix} \hat{x}^{\mathrm{T}} & \hat{A}_1 & \hat{n}_1 & \hat{m}_1^{\mathrm{T}} & \ldots & \hat{A}_{N_P} & \hat{n}_{N_P} & \hat{m}_{N_P}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$$

be the prior covariance and mean. Let the initial guess be $z_0 := \hat{z}$. Additionally, measurement variance σ^2 is required. Set $k := 0$. Denote the objective function with

$$\theta(z) := (z - \hat{z})^{\mathrm{T}} {\hat{\Sigma}_z}^{-1} (z - \hat{z}) + \sum_{i=1}^{N_P} \frac{h_i(z) - P_i}{\sigma^2}.$$

2) Compute the Jacobian

$$\mathrm{J}_k := \begin{bmatrix} \frac{\partial h_1}{\partial x} & \frac{\partial h_1}{\partial A_1} & \frac{\partial h_1}{\partial n_1} & \frac{\partial h_1}{\partial m_1} & \mathbf{0}^{\mathrm{T}}_{4(N_P-1)} \\ \vdots & & \ddots & & \vdots \\ \frac{\partial h_{N_P}}{\partial x} & \mathbf{0}^{\mathrm{T}}_{4(N_P-1)} & \frac{\partial h_{N_P}}{\partial A_{N_P}} & \frac{\partial h_{N_P}}{\partial n_{N_P}} & \frac{\partial h_{N_P}}{\partial m_{N_P}} \end{bmatrix}.$$

3) Set

$$\Delta z_k := -\left(\hat{\Sigma}_z^{-1} + \frac{1}{\sigma^2}\mathrm{J}_k^{\mathrm{T}}\mathrm{J}_k\right)^{-1} \cdot \left(\hat{\Sigma}_z^{-1}(z_k - \hat{z}) + \frac{1}{\sigma^2}\mathrm{J}^{\mathrm{T}}(h(z_k) - P)\right).$$

4) Adapt step length:

$\alpha := 1$
while $||\theta(z_k + \alpha\Delta z_k)|| \geq ||\theta(z_k)||$ and $\alpha > \alpha_0$ **do**
$\alpha := \frac{\alpha}{2}$
end while

where α_0 is a configuration parameter, e.g. 0.05. Set $z_{k+1} := z_k + \alpha\Delta z_k$.

5) If stopping condition $||\Delta z_k|| < \delta$ is not satisfied and $k \leq k_{\max}$, increment k and repeat from Step 2. Otherwise compute $\mathrm{P} := \left(\hat{\Sigma}_z^{-1} + \frac{1}{\sigma^2}\mathrm{J}_k^{\mathrm{T}}\mathrm{J}_k\right)^{-1}$ and set the state estimate

$$\hat{x}^+ := z_{k+1,1:2}, \quad \hat{\Sigma}_x^+ := \mathrm{P}_{1:2,1:2}$$

V. Positioning Tests

A. Tests with artificial data

Fig. 3 illustrates the influence of the uncertain parameters on the likelihood. In the present case $\begin{bmatrix} A \\ n \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} 2 \\ 3.2 \end{bmatrix}, \begin{bmatrix} 100 & 3.5 \\ 3.5 & 0.15 \end{bmatrix}\right)$ and $m \sim \mathrm{N}\left(\mathbf{0}, 2 \cdot 10^4 \cdot \mathrm{I}\right)$. These values are based on our experimental knowledge of the Finnish cellular network. The likelihoods are calculated using the Monte Carlo based grid algorithm. The upper row illustrates the likelihoods of the model that takes the parameter uncertainties into account, and the lower row shows the likelihoods assumed that the path loss parameter values are correct. The RSS values corresponding to the likelihoods are -50 (on the left), -75 and -90 dBm. Path loss model standard deviation is $\sigma = 6$ dBm. It can be seen that with strong signals, the RSS likelihood is unimodal or almost unimodal when the parameter uncertainties are taken into account.

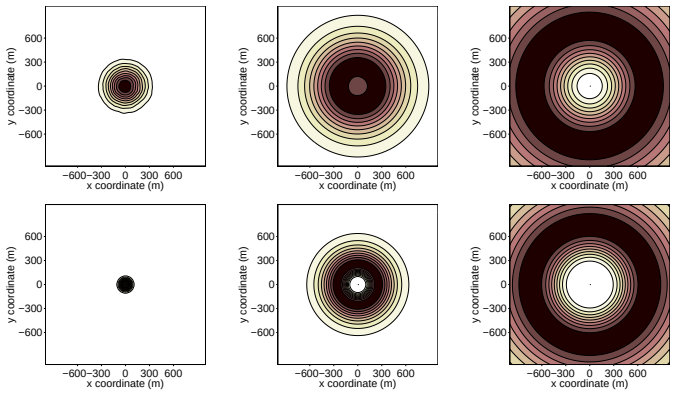

Figure 3: The likelihoods of measurements -50, -75 and -90 dBm. In the upper row, parameter uncertainties have been taken into account.

In the case of radially symmetric BS position distribution, the posterior density depends only on the distance from the mean of the BS position estimate. Fig. 4 illustrates the likelihoods of the user's position as a function of this distance. They have been computed using standard Monte Carlo integration and normalized so that the maximal likelihood value is one. Curve "N" represents algorithms that assume PL parameters to be normally distributed *a priori*, and curve "acc" the algorithms that assume that the parameters are known accurately. Fig. 4 shows that the tails of the "N" curve are considerably heavier.

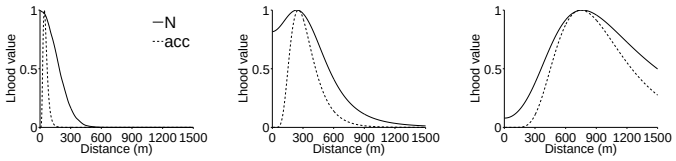

Figure 4: Likelihoods of measurements -50, -75 and -90 dBm as a function of the distance from the mean of the BS position estimate. Curve "N" represents algorithms that assume PL parameters to be normally distributed *a priori*, and curve "acc" the algorithm that assumes that the parameters are known accurately.

In Fig. 5 the likelihood of two RSS measurements of -80 dBm is presented. The PL parameters are similar to the ones in Fig. 3, and the distance between BS's is 400 meters. If the parameter uncertainties are not taken into account (on the right), the support of the likelihood consists of two separate parts, whereas in the left figure there is significant amount of likelihood mass also in the BS positions' surroundings.

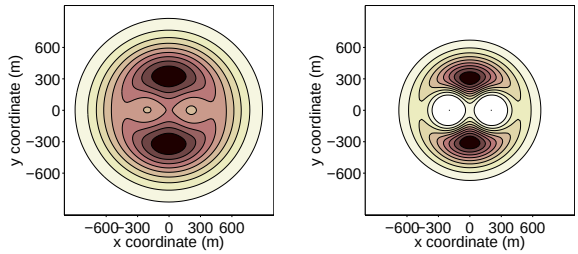

Figure 5: The combined likelihood of two BS's with signal strengths -80 dBm. On the left, parameter uncertainties have been taken into account.

B. Tests with real data

A measurement campaign was accomplished to evaluate the performance of different algorithms in a real use case. First, a large set of fingerprints was collected from the 3G cellular network in Tampere urban area, Finland for learning the radiomap. The measured RSS values are based on the measured Received Signal Code Power (RSCP) indicator reported by the user equipment. The coverage area of each BS was estimated by fitting a normal distribution to the data [11, 12]. Furthermore, path loss model parameters were estimated using the method that was presented in Section II of this paper.

There are three separate test tracks. The first track (Hervanta) was collected by a pedestrian in a densely populated urban/suburban area. In the second track (Lukonmäki) the measurer rode a bicycle with a low velocity in a suburban area. The third track (Linnainmaa) is a higher velocity suburban bicycle case. In all the cases the true user positions were tracked down using conventional GPS positioning. By plotting the GPS solutions on the map, it was confirmed that the GPS error on the area is small compared to the cellular positioning accuracy.

In all these tracks, the prior distribution is computed in each estimation point using the estimated coverage areas of the BS's that are observed at the point concerned. The prior is the product distribution of Gaussian coverage areas, so it is also a Gaussian distribution.

C. Results and discussion

The results of the true data tests are in Table I. Abbreviation "N" stands for the algorithms that assume the path loss parameters to be normally distributed *a priori* whereas "acc" indicates that the parameter values are assumed known. "CA" refers to the product of coverage areas. Columns "Mean", "Med" and "95% err." are mean error, median error and empirical 95% percentile of errors in meters. "Time" is the average running time of our MATLAB implementation in seconds. Note that the codes are not highly optimized so the running time values have to be considered only roughly indicative. The times are also highly dependent on the chosen configuration parameters.

Presented "N" algorithms seem to outperform "acc" algorithms in the positioning accuracy. In the Lukonmäki case

Table I: Results for the real data tests.

Solver	Mean (m)	Med (m)	95% err. (m)	Cons. (%)	Time (s)
Hervanta					
grid, N	247	204	432	96	49
grid, acc	250	225	444	85	22
MH, N	258	215	473	81	38
MH, acc	273	233	517	38	23
GN, N	243	186	438	94	0.6
GN, acc	232	186	409	81	0.4
CA	258	212	458	99	0.24
Lukonmäki					
grid, N	219	144	520	93	82
grid, acc	226	158	530	82	38
MH, N	238	161	572	83	67
MH, acc	286	172	837	40	36
GN, N	223	162	502	95	0.8
GN, acc	267	220	599	81	0.5
CA	251	203	532	96	0.26
Linnainmaa					
grid, N	195	159	412	96	83
grid, acc	197	177	420	91	36
MH, N	213	179	456	90	45
MH, acc	260	222	600	39	25
GN, N	218	176	544	89	0.6
GN, acc	229	170	523	82	0.3
CA	271	214	741	97	0.16

it is questionable whether the RSS measurements should be used at all, if the uncertainties are not in the model. The only case where the "acc" slightly outperforms "N" is Hervanta with Gauss–Newton positioning. This is likely to stem from the linearization approximation of the method, since the phenomenon is not visible with the other methods.

Column "Cons." displays the 95% consistency that was determined using Gaussian consistency test [13, pp. 235] with risk level 5%. The solver is deemed to be consistent at a certain time step, if the true position is within the 95%-ellipse of the posterior distribution, assuming normality of the posterior. The closer this number is to 95%, the more realistic the covariance matrix estimation is. From the figures it can be seen that taking the parameter uncertainties into account improves the consistency remarkably and independent of the estimation method. Thus, the "N" method seems to be beneficial especially if location information from other sources is combined with RSS measurements or when positioning is done with Bayesian time-series filters.

Among the three estimation methods, the grid and MH sampler approach the exact Bayesian model posterior distribution. The grid gives the precise posterior values in the grid points assuming that the Monte Carlo integration's accuracy is adequate. The MH sampler converges theoretically to the true posterior as the sample size parameter N approaches infinity. In practice, however, the rate of convergence in MH algorithms is highly dependent on the form and parameters of the proposal distributions. With the chosen configuration the method usually fails to compete with the grid especially in consistency, still providing a good estimate with slightly lighter computation.

The Gauss–Newton method lacks global convergence properties and the covariance matrix estimate is based on iterative

linearization procedure and has thus a less clear Bayesian interpretation. Indeed, the real data tests showed that the algorithm's convergence is more dependent on the quality of the prior distribution. However, the presented results are comparable with those of the other methods, and the GN is clearly the computationally lightest one of these algorithms and is applicable in many real-time solutions.

VI. Conclusions

It was shown that estimating the path loss parameter uncertainties and taking them into account in the positioning phase has significant influence on the positioning performance. Both accuracy and consistency are improved compared to the conventional methods where the path loss parameters are assumed to be known accurately. If the uncertainties are not used, it seems to be questionable, whether the path loss modeling brings any benefit compared with the coverage area positioning. Furthermore, it was shown that Gauss–Newton optimization algorithm provides satisfactory accuracy and consistency, being also computationally feasible for many applications. The comparison methods were Monte Carlo - based Grid method and Metropolis–Hastings sampler.

Acknowledgment

This research was partly funded by Nokia Inc. and the Academy of Finland. The authors are grateful to Dr. Tech. Lauri Wirola and Dr. Tech. Jari Syrjärinne for their support and advice.

References

[1] A. F. Molisch, *Wireless Communications*, 2nd ed. Wiley - IEEE, January 2011.

[2] T. S. Rappaport, *Wireless communications: Principles and practice, 2nd edition*. Prentice Hall PTR, 2008, ch. Mobile Radio Propagation: Large-Scale Path Loss.

[3] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *Vehicular Technology, IEEE Transactions on*, vol. 29, no. 3, pp. 317–325, August 1980.

[4] X. Li, "RSS-based location estimation with unknown pathloss model," *IEEE Transactions on Wireless Communications*, vol. 5, no. 12, pp. 3626–3633, December 2006.

[5] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electronics Letters*, vol. 27, no. 23, pp. 2145 –2146, nov. 1991.

[6] S. Shrestha, E. Laitinen, J. Talvitie, and E. S. Lohan, "RSSI channel effects in cellular and WLAN positioning," in *Proceedings of the 9th Workshop on Positioning, Navigation and Communication 2012 (WPNC'12)*, March 2012.

[7] J. S. Liu, *Monte Carlo Strategies in Scientific Computing*. Springer, 2001.

[8] N. Sirola, "Closed-form algorithms in mobile positioning: Myths and misconceptions," in *Proceedings of the 7th Workshop on Positioning, Navigation and Communication 2010 (WPNC'10)*, March 2010, pp. 38–44.

[9] Å. Björck, *Numerical methods for least squares problems*. Society for Industrial and Applied Mathematics, 1996.

[10] A. Ruhe, "Accelerated Gauss-Newton algorithms for nonlinear least squares problems," *BIT Numerical Mathematics*, vol. 19, no. 3, pp. 356–367, 1979.

[11] L. Koski, R. Piché, V. Kaseva, S. Ali-Löytty, and M. Hännikäinen, "Positioning with coverage area estimates generated from location fingerprints," in *Proceedings of the 7th Workshop on Positioning, Navigation and Communication 2010 (WPNC'10)*, March 2010, pp. 99–106.

[12] L. Koski, T. Perälä, and R. Piché, "Indoor positioning using WLAN coverage area estimates," in *2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, September 2010.

[13] Y. Bar-Shalom, R. X. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation, Theory Algorithms and Software*. John Wiley & Sons, 2001.