

Impact of Position Errors on Path Loss Model Estimation for Device-to-Device Channels

Pedro M. Santos, *Student Member, IEEE*, Traian E. Abrudan, *Member, IEEE*,
Ana Aguiar, *Member, IEEE*, and João Barros, *Senior Member, IEEE*

Abstract—Many wireless applications require a propagation model that describes the attenuation of the transmitted signal as a function of the distance between devices. Such channel models are derived commonly from signal strength measurements, and assume that the true distances between wireless terminals are known. In practice, however, the true distances may be unavailable or difficult to obtain, for instance in mobile scenarios or in the absence of line-of-sight. These conditions typically occur in forested environments, urban areas, etc. This paper addresses the problem of path loss model parameter estimation in presence of erroneous distance measurements, such as the ones derived from the GPS positions. We provide a model for the uncertainties, and study the impact of distance errors on the estimation of a log-distance channel model.

Our main conclusion is that the path loss model can be estimated with a reasonable accuracy from unreliable distances, provided that the measurements are taken at distances beyond a few standard deviations of the GPS positioning error. In case the maximum communication range does not allow such large distances, we provide a method to correct the erroneous channel model. Real-world measurements are used in order to validate our approach.

Index Terms—Propagation, channel characterization and modeling, GPS, ad-hoc networks, wireless LANs.

I. INTRODUCTION

SMART phones may be used to set up communication where infrastructure support is unavailable or not dependable, such as in remote areas or catastrophe scenarios [1]. Likewise, vehicles may form ad-hoc networks (VANET) to provide increased road safety and improved traffic management [2]. In wireless networks such as these, it is often helpful to estimate the relationship between device-to-device distance and received power. This is typically expressed in the form of a propagation model, which can be used to design adequate wireless protocols.

Various types of propagation models exist (see Section 2) with different degrees of accuracy and complexity, as well as specific details about the environment. For the applications mentioned, generic models are more useful, as they are applicable to a wide range of scenarios. For this reason, we

focus on the estimation of the log-distance path loss model in this paper.

Many mobile devices used today are equipped with some form of positioning information, typically GPS, which can be used directly to estimate the path loss model. However, GPS positioning is prone to errors, that have a negative impact in distance estimation and in turn diminish the accuracy of the estimated propagation model.

We address this problem in the context of a practical use-case, that of ad-hoc communication in a forested environment among smart phones using WiFi. This use-case is motivated by a feasibility study for a smart phone-based information system for firefighters [3]. After analyzing the impact of GPS positioning errors on the estimation of the range between devices and consequently on the estimated propagation model, we derive guidelines for the design of future device-to-device path loss measurement campaigns, and propose a practical method to correct those errors based on Monte Carlo simulations. The conclusions we present are applicable to any outdoor scenario, provided that a good estimate of the standard variation of the GPS positioning error is available.

Our main contributions are as follows:

- 1) a model of the impact of GPS positioning errors on the ranging and path loss model estimation;
- 2) a method to improve path loss estimation using only GPS distances;
- 3) guidelines for designing measurement campaigns for path loss model estimation that reduce the impact of GPS errors.

In the next section, we present an overview of related work. In Section III, we describe the methodology for channel data collection and modeling, and present results using real-world measurements. In Section IV, we discuss the impact of GPS errors on the path loss model estimation. In Section V, we provide guidelines for measurement data collection and a path loss model parameter retrieval methodology. Finally, we conclude in Section VI.

II. RELATED WORK

The log-distance path loss model [4] is one of the most widely used models to describe the behavior of radio wave propagation. Many of the works in this field address outdoor communication between a mobile user and a base station, which is usually tall or positioned in a high location. This is the case in [5], which presents path loss, scattering and

Manuscript received June 17, 2013; revised October 28, 2013 and January 21, 2014; accepted March 2, 2014. The associate editor coordinating the review of this paper and approving it for publication was C. Yang.

P. Santos, A. Aguiar and J. Barros are with Instituto de Telecomunicações, Departamento de Engenharia Electrotécnica e de Computadores, Faculdade de Engenharia da Universidade do Porto, Portugal (e-mail: {pedro.salgueiro, anaa, jbarros}@fe.up.pt).

T. Abrudan is with Department of Computer Science, University of Oxford, UK. (email: traian.abrudan@cs.ox.ac.uk).

Digital Object Identifier 10.1109/TWC.2014.040214.131082

multipath delay statistics for digital cellular telephony, measured in urban context with distances in the range of 1.5 to 6.5 kilometers. The channel models in [6] are estimated for the 5.3 GHz range in urban mobile communications, with distances up to a few hundred meters. The map-based method is used to determine the distances precisely in those papers. Another relevant work is [7], which tackles user-to-user communication in a multi-floored building scenario. Its approach to modeling is to start from the standard path loss model formula and then expand it with additional factors in order to account for the effects of walls and office partitions.

A comprehensive survey of empirical path loss models for forested environments may be found in [8]. Typically such models are additive with respect to the free space path loss model [4]. The modified exponential decay (MED) model [9] is the basis for most empirical models concerning propagation in forests. MED uses the formula $A = \alpha f^\beta d^\gamma$, where the extra attenuation A is given in dB, and f identifies the signal frequency and d the tree depth. Parameters α , β and γ may be estimated from measured data. Given the wide variety of factors affecting propagation in forests (such as tree species, disposition of the trees, foliation), it is very difficult to find an universal set of values. The contribution of the various empirical models found in literature is to propose values that aim to be the most general possible, but these are strongly conditioned by the scenario in which data was obtained. Table I lists some of those models and corresponding parameter values. There is also a family of models based on the modified gradient model, such as the Maximum Attenuation (MA) [10] and Nonzero Gradient (NZG) [10] models. These models require additional parameters that are specific to the measurement geometry and/or methodology. Both modified exponential decay and modified gradient models describe essentially propagation through canopies or at canopy-level. In [11], the authors analyze propagation at trunk level. Based on an extensive data set, the log-distance path loss model is shown to be the most accurate of existing models. An improved version of this model is proposed by incorporating scenario-specific parameters, namely tree density and trunk diameter. In summary, existing models for propagation in forests are tightly dependent on the particular conditions in which measurements take place. The application scope of our approach is not limited to forests, so we opted not to use such specialized models and use the more general log-distance path loss model.

Concerning measurement campaigns in forests, a large number of works, such as [15] and [16], focus on the propagation of GSM signals, in the 1900 MHz band. Propagation studies on the 2.4 GHz band are usually related to wireless sensor networks, which use lower transmission power and thus have limited range compared to typical WiFi or VANET communications [17]. None of these works clarifies which distance measuring method was used, nor account for distance measurement errors or for their impact on the accuracy of the estimated model. The only factor deemed relevant is the range of distances. The use of GPS positioning for distance calculation finds substantial application in many VANET propagation studies, but the reality of GPS errors and their impact in model estimation is also not mentioned [18]–[20].

TABLE I: Parameters for various empirical models based on the MED model. f is set in GHz for the Weissberger model, and MHz for all others; d is in meters for all models.

Model	Note	α	β	γ
Weissberger [9]	$0 < d < 14\text{m}$	1.33	0.284	0.588
	$14 < d < 400\text{m}$	0.45	0.284	1
ITU [12]	$d < 400\text{m}$	0.2	0.3	0.6
COST 235 [13]	Not foliated	26.6	-0.2	0.5
	Foliated	15.6	-0.009	0.26
FITU [14]	Not foliated	0.37	0.18	0.59
	Foliated	0.39	0.39	0.25

Regarding positioning technology for outdoor scenarios, the Global Positioning System (GPS) [21] has become one of the most widely used technologies. It allows a user equipped with a GPS receiver to know its location with some degree of accuracy by receiving the signal broadcasted by a constellation of satellites. Some empirical studies address the performance and accuracy of consumer-grade GPS receivers in forest settings. In [22], six commercial receivers in the range of \$150 to \$320 (values of 2005) were tested in static conditions. Measurements were taken in open sky, under young forest canopy, and under closed forest canopy. For the later scenario, the reported values of average positioning error vary between 2.7 and 11.4 meters, depending on the device. The work in [23] provides an updated version of these results, but only for two devices and under heavy canopy conditions. Reported values for the average positioning error were 4.0 and 6.9 meters. Studies about GPS position accuracy in smart phones are scarce. One example is the work of [24], which analyses the location accuracy of the iPhone obtained from different information sources (GPS, cell ID, WiFi) in different settings (rural and urban, and indoor).

To the best of our knowledge, the problem of estimating a path loss model from distances that have errors has not yet been addressed in the existing literature.

III. MEASUREMENT COLLECTION AND PARAMETER ESTIMATION

We model the received signal power using a log-distance path loss model, whose formula in the logarithmic domain is given by

$$\rho_{[\text{dBm}]}(d) = \rho_0 - 10 \alpha \log\left(\frac{d}{d_0}\right) + X_\rho, \quad (1)$$

where ρ is the received signal strength (in dBm units) at an arbitrary distance d from the transmitter, and ρ_0 is the received signal strength at reference distance d_0 in the far field (typically 1 meter). The logarithmic signal strength measurements ρ are affected by normal fading, $X_\rho \sim \mathcal{N}(0, \sigma_\rho)$. In the following discussion, the term RSSI (Received Signal Strength Indicator) refers to the discrete readings of the received signal power ρ that are typically delivered by the drivers of wireless interface cards. The values of parameters α , ρ_0 and σ_ρ for each specific scenario are estimated using distance-RSSI data pairs obtained in a measurement campaign, and then calculating the line that best fits the data.

A laser-beam range meter could be used to determine the distances, but that would require line-of-sight between devices.



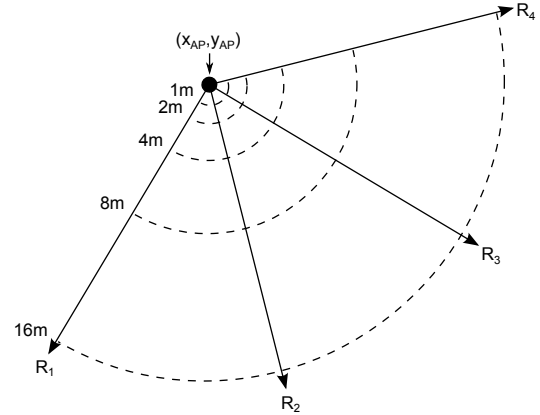
Fig. 1: Experimental setting.

This becomes impractical for very large numbers of measurements or in mobile scenarios. A more practical solution is to record the coordinates obtained from the GPS receiver incorporated in the devices at different test distances, along with RSSI. However, since GPS positions contain errors, this solution requires a clearer understanding of how GPS-based distances affect the estimation of path loss model parameters, compared to the case when the actual distances are used. In Section III-A, we explain the methodology we used to collect measurement sets of RSSI and distance data. Results on path loss model estimation in the presence of GPS errors can be found in Section III-B.

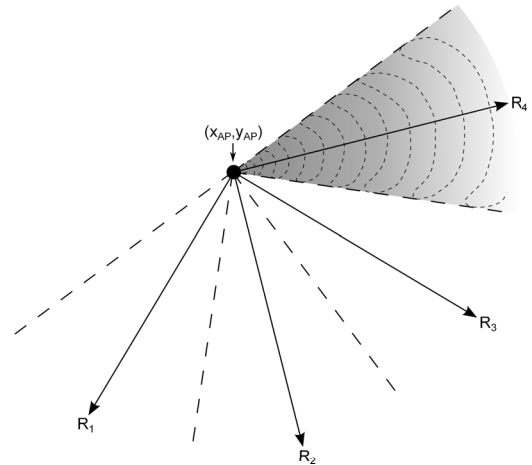
A. Data Collection Methodology

We divided our measurement campaign into two phases. The first phase consisted of collecting pairs of RSSI and GPS measurements at known distances, to study the impact of GPS errors. In the second phase, we collected only RSSI and GPS samples. As a motivating example, we took measurements in a forested area where the majority of trees were stone pines (*Pinus pinea*). The height of the devices was considerably lower than the begin of the canopies (~ 1.5 m vs. ~ 6 m). The ground between transmitter and receiver was mostly covered with grass and small weeds. Most readings were taken in line-of-sight, or with a small number of tree trunks between receiver and transmitter. A small subset of the readings was taken with vegetation in-between, specifically bushes and smaller trees, most of them slightly taller than a human. Their density ranged from a single plant to a compact set of these. Figure 1 shows the various settings. The measurement equipment consisted of three standard, off-the-shelf, consumer electronics smart phones: one Samsung Galaxy Nexus S to act as access point (AP), which we call Device A, and two Samsung Galaxy Nexus to act as mobile receivers, referred to as devices B_1 and B_2 . All models ran Android OS, and ran an application that stored GPS and WiFi RSSI values obtained from the Android APIs.

In the first phase of the measurements (henceforth called *Phase 1*), we placed device A at a fixed location, on top of a tripod of 1.5 meters, and set it in AP mode. This device periodically sent beacons announcing its presence to devices that aspired to join its network. It also recorded GPS measurements at a rate of 1 Hz. We then sequentially placed devices B_1 and B_2 together at several pre-defined distances from device A. For convenience, we call these distances “true distances”. Their values were 0, 1, 2, 4, 8, 16, 32 and 64



(a) *Phase 1* of measurements – static case. The true distances were measured using a laser-based range meter. Solid lines show the radials along which readings were taken. Dashed lines mark the distances at which data was taken.



(b) *Phase 2* of the measurements – zig-zag walk. No true distances are available. Solid lines are the radials along which readings were taken. Dashed lines limit the data collection area for each radial. Gradient area marks data collection area for radial R_4 .

Fig. 2: Spatial arrangement of measurements.

meters, and we verified them *in loco* using a laser range meter. We chose these specific distances because they are equally spaced in logarithmic scale. At each true distance, we held devices B_1 and B_2 at a height of 1.5 meters, and each device recorded roughly 3 minutes of GPS and RSSI data, at an average rate of one GPS measurement every second and one RSSI measurement every two seconds. We repeated the same procedure for four radials roughly 45 degrees apart, as shown

in Figure 2a. Device A recorded GPS measurements during the entire duration of each radial.

In the second phase of the measurements (henceforth called *Phase 2*), we again placed device A at a fixed location, set it in AP mode and activated it to record GPS measurements. We held devices B_1 and B_2 at a height of 1.5 meters next to the AP and initiated the recording application, collecting pairs of RSSI and GPS coordinates with the same frequencies as in *Phase 1*. We then carried them away from device A, following a zig-zag path that oscillated around each radial within a 45° angle approximately. We repeated the same procedure for the other three radials. Device A recorded constantly GPS coordinates during each radial measurement. Figure 2b shows the corresponding spatial arrangement.

B. Path Loss Model Parameter Estimation

From the data collected in *Phase 1* of the measurements, we obtain a set of GPS coordinates and RSSI values for each true distance. We compute the distances by pairing, via time-stamp, the coordinates recorded by device A and each device B ($B \in \{B_1, B_2\}$), and applying the Great Circle distance formula. This formula allows accurate computation of distances between points in a sphere whose positions are defined by decimal degrees. For convenience, we call distances calculated in this manner “GPS distances”. During the analysis of the data, we found large clusters of consecutive, repeated GPS coordinates. We conclude that this range of commercial-grade devices may tend to fix on a set of coordinates if they do not detect significant movement for some time, in order to save energy. We substituted these clusters by a single measurement, to which we associated the median of the RSSI values of that cluster.

It is now possible to compute the parameters of the path loss model using two different data sets from *Phase 1*. In one case, we pair the RSSI measurements with the actual distances (obtained with a laser range meter). We call this the “true model”. In the other case, we pair the RSSI values with the GPS distances, and we call this the “GPS model”. Equations (2) and (3) describe both models as follows:

$$\rho(d) = \rho_0 - 10 \alpha \log(d) + X_\rho, \quad (2)$$

$$\rho(d) = \tilde{\rho}_0 - 10 \tilde{\alpha} \log(d_{\text{GPS}}) + Y_\rho. \quad (3)$$

We perform regression over the two data sets using the Least Squares (LS) estimator which, for the true model, is also the maximum likelihood estimator. The parameters obtained from the measured data are shown in Table II.

In Figure 3, our estimated model is shown against models proposed in literature. While for small distances the literature models are consistent with our data, they stray at larger distances. As mentioned in Section II, most models reported in literature focus on propagation through canopies. For large distances, the extra-attenuation factor of the canopy models, which grows exponentially with distance, takes the overhand in the total attenuation. Given that our measurements were taken mostly in line-of-sight, with only a very few occasional trunks or canopies between devices, our data does not show the effect of the additional attenuation, being best fitted by the

TABLE II: Model parameters derived from true and GPS distances using the Least Squares Estimator.

Phase	Model	ρ_0	α	σ_ρ
1st	True	-42.63	2.22	5.64
	GPS	-24.82	3.00	9.56
2nd	GPS	-42.86	2.13	7.34

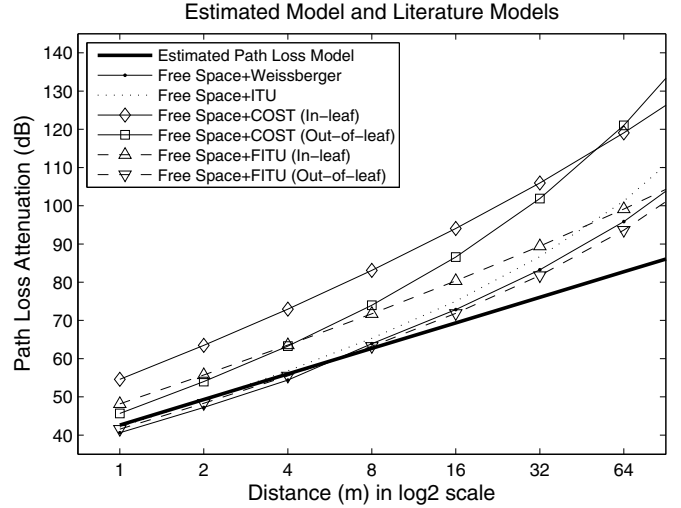


Fig. 3: Measured RSSI data and “true” path loss model compared to literature models. See Table I for references.

log-distance path loss model. This explanation is supported also by the results in [11].

The GPS model exhibits overestimated parameter values α , ρ_0 and σ_ρ compared to the true model. Figure 4a shows both models derived from *Phase 1* measurement data. It also depicts the RSSI measurements at the distances provided by the GPS measurements, to give further insight on how their positions on the RSSI-distance plane condition the regression method. In Section IV, we provide a model for the erroneous distances, and analyse in more detail the impact of GPS errors on distance estimation.

As for the *Phase 2* measurement data (see Figure 4b), for which the exact distance was not recorded, we observe fairly different parameters when compared with the GPS model in the first phase (see Figure 4a). We expected them to be more similar, because both are computed using error prone GPS distances. However, unaccounted factors, such as user mobility, may explain this behavior. In the second phase of the measurements, more states of the channel fading are being captured due to the obstruction by trees, the user’s body, and different device orientations.

IV. DISTANCE ESTIMATION IN PRESENCE OF GPS ERRORS

After discussing the previous motivating example, we now address the problem of estimating the distance between two GPS-equipped devices from error prone coordinates. Towards this end, we characterize the nature of the GPS errors, and provide a model for these errors. This model will help explain the difference between the true path loss model and the GPS path loss model.

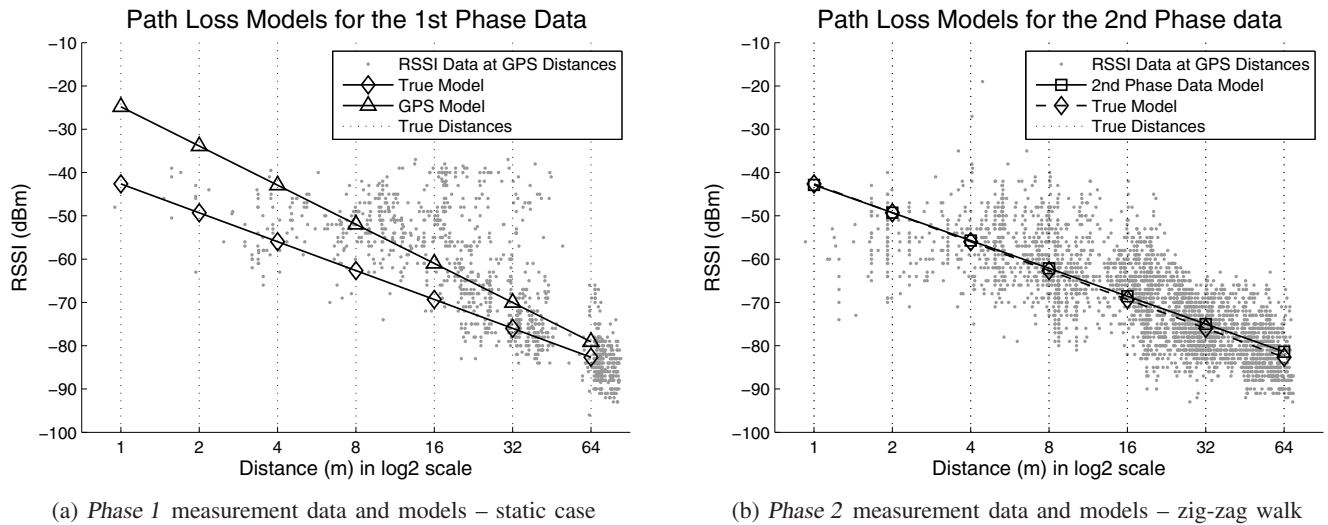


Fig. 4: Path loss models using the Least-Squares estimator. The true distances are marked by vertical dashed lines.

The accuracy of GPS is affected mainly by two factors: the geometry of the satellites visible to the user, and the quality of the distance estimate between satellite and user (pseudo-range). Concerning the first factor, the geometrical arrangement of the satellites in the sky impacts the space of possible user locations. A numerical metric named Geometric Dilution of Precision (GDOP) is often used to quantify the quality of a particular arrangement. On its turn, pseudo-range estimates are affected by errors that have a variety of sources. These can be the quality of the user's GPS receiver, attenuation of the satellite signal by the ionosphere and troposphere, noise and multipath effects, and accuracy of the ephemeris data contained in the satellite signal. Due to the diversified nature of the error sources, positioning errors can be categorized into *systematic* and *non-systematic* errors. Systematic errors affect all receivers within a certain area in similar manner, and hence are modelled as an identical position bias for all receivers. They are caused by atmosphere, quality of GDOP and ephemeris errors. Non-systematic errors are random in their nature, affecting each receiver and each measurement in a unique way. They are caused mainly by pseudo-range errors, multipath propagation, receiver noise, clock jitters and numerical errors.

We now derive a model for the error that affects distances computed from GPS coordinates. In this discussion, we assume local Euclidean coordinates, given that the scale of the distances we are using is small enough for the curvature of the Earth to be neglected. Our GPS position error model (see Figure 5) accounts for the distinction between the two types of errors mentioned earlier. Systematic positioning errors are modeled as a bias vector with respect to the actual position that is equal for all devices. Non-systematic positioning errors are modeled independently for each Euclidean coordinate as zero-mean circularly symmetric Gaussian random variables.

Given the exact Euclidean coordinates of the two terminals A and B, (x_A, y_A) and (x_B, y_B) respectively, the Euclidean coordinates corresponding to the measured GPS positions may

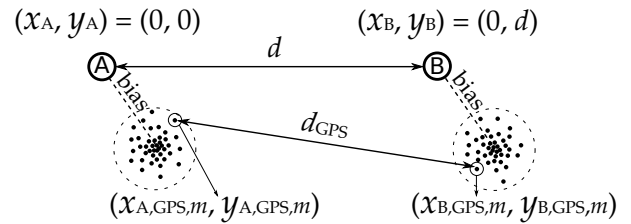


Fig. 5: Characterization of the measured GPS distances.

be written as

$$\begin{aligned} (x_{A,GPS}, y_{A,GPS}) &= (x_A, y_A) + (b_{x,A}, b_{y,A}) + (\epsilon_{x,A}, \epsilon_{y,A}), \\ (x_{B,GPS}, y_{B,GPS}) &= (x_B, y_B) + (b_{x,B}, b_{y,B}) + (\epsilon_{x,B}, \epsilon_{y,B}), \end{aligned}$$

where the errors $\epsilon_{x,A}, \epsilon_{y,A}, \epsilon_{x,B}, \epsilon_{y,B} \sim \mathcal{N}(0, \sigma_{GPS})$ are assumed to be mutually independent, and σ_{GPS} is the standard deviation of the GPS positioning errors in each coordinate (x, y) . The aggregated systematic errors along each axis are equal for both terminals, i.e., $b_{x,A} = b_{x,B}$ and $b_{y,A} = b_{y,B}$, effecting a translation of the terminal positions with no impact on the distance, as shown in Figure 5. For simplicity, we choose the local two-dimensional Euclidean system of coordinates with the origin centered at the exact location of device A, and with the abscissa-axis pointing in the direction of the device B, i.e., $(x_A, y_A) = (0, 0)$ and $(x_B, y_B) = (d, 0)$. Therefore, the expression of the GPS-based Euclidean distance between A and B reduces to

$$d_{GPS} = \sqrt{(\epsilon_{y_A} + \epsilon_{y_B})^2 + (\epsilon_{x_A} + \epsilon_{x_B} + d)^2}. \quad (4)$$

Consequently, d_{GPS} follows a Rice distribution with the location parameter being the actual distance d , and the scale parameter $\sqrt{2}\sigma_{GPS}$, i.e., $d_{GPS} \sim \text{Rice}(d, \sqrt{2}\sigma_{GPS})$. The probability density function (p.d.f.) of the GPS distances given the actual distance is

$$p(d_{GPS}|d) = \frac{d}{2\sigma_{GPS}^2} \exp\left(-\frac{d_{GPS}^2 + d^2}{4\sigma_{GPS}^2}\right) I_0\left(\frac{d \cdot d_{GPS}}{2\sigma_{GPS}^2}\right), \quad (5)$$

where $I_0(\cdot)$ is the zero-order modified Bessel function of first kind.

Next, we provide a simple method to estimate the variance of the GPS error in each coordinate. The advantage of the proposed method is that the true coordinates of A and B are not required, the true distance being sufficient. Some receivers provide reliability information on the estimated position that can be used, for example, in a weighted LS estimation of the path loss model. Here, we assume that such information is unavailable, and we estimate an overall “average” reliability of the GPS position estimates. The second raw moment of the Rice distribution, μ'_2 , can be analytically related to the variance of the GPS coordinates as

$$\mu'_2 \triangleq E[d_{\text{GPS}}^2] = 4\sigma_{\text{GPS}}^2 + d^2. \quad (6)$$

Empirically, the second raw moment of the GPS-based distances μ'_2 can be obtained from the measured data by simply averaging the squared GPS distances. However, there are multiple true distances d_i . We compute the empirical second raw moment for each of the true distance, $\mu'_2(d_i)$, and use LS to estimate the overall variance of the GPS coordinates

$$\widehat{\sigma_{\text{GPS}}^2} = \frac{1}{4N} \sum_{i=1}^N \mu'_2(d_i) - d_i^2. \quad (7)$$

Finally, we compare the histogram of the measured GPS distances obtained in *Phase 1* with the Rice p.d.f. predicted by our model for each true distance. Figure 6 shows the normalized histograms of the GPS-based distances (dotted lines), and the Rice p.d.f. corresponding to that true distance $d \in \{1, 2, 4, 8, 16, 32, 64\}$ (solid lines). The true distance at which we took the respective GPS measurements is marked by a thick vertical solid line. Our model proves able to predict the major trends of the data. The overall standard deviation for GPS coordinates estimated using Equation (7) is $\widehat{\sigma_{\text{GPS}}} = 10.09$ meters. It may be noticed that for very small distances, i.e., $d \ll 3\sqrt{2}\sigma_{\text{GPS}} \approx 42$ meters¹, the GPS distances are overestimated by far, i.e., the mode of the Rice p.d.f. corresponds to a value much larger than the true distance. For larger distances, i.e., $d > 3\sqrt{2}\sigma_{\text{GPS}}$, the Rice p.d.f. is very close to a normal p.d.f. with the mode slightly larger than the true distance. In conclusion, we observe that GPS errors hamper significantly the distance estimation between two devices if the actual distance is smaller than $3\sqrt{2}\sigma_{\text{GPS}}$.

V. COPING WITH DISTANCE ERRORS IN PATH LOSS MODEL ESTIMATION

The error model for GPS distances introduced in the previous section can be used to improve path loss model parameter estimation. We begin this section by detailing the process by which GPS errors condition the estimation of the model parameters, and use this analysis to learn how can measurement campaigns be designed in order to mitigate the impact of such errors. Then, we provide a method based on Monte Carlo simulations to retrieve the true model parameters from measurements affected by GPS errors.

¹The factor $\sqrt{2}$ appears due to the fact that the scale parameter of the standard Rice p.d.f. would be σ_{GPS} , whereas in our case, it is $\sqrt{2}\sigma_{\text{GPS}}$. The factor of three corresponds to the ratio d/σ_{GPS} for which a standard Rice p.d.f. can be approximated by a Gaussian p.d.f..

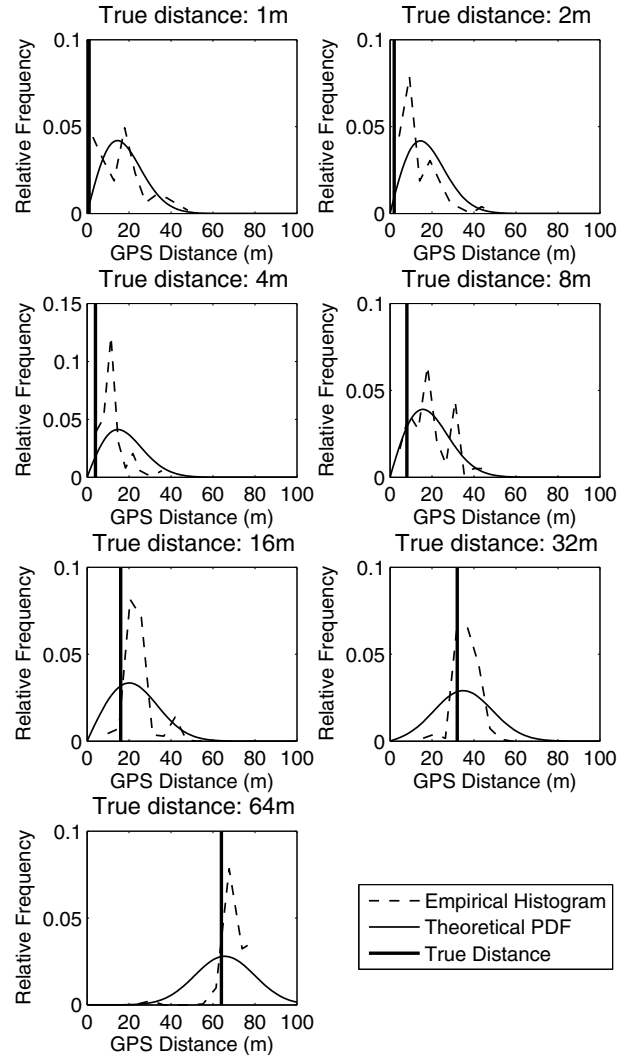


Fig. 6: Normalized histograms of the GPS distances measured at each true distance.

A. Guidelines for Selecting the Measurement Distances

In this section, we provide a selection policy concerning the distances at which measurements should be taken to mitigate the impact of GPS errors. We start by explaining the mechanism by which GPS errors impact parameter estimation using a linear regression method. In the previous section, we have seen that GPS distances tend to over-estimate by far small true distances. Consequently, for such distances, the RSSI values are paired with GPS distances larger than those at which we actually took them. This causes the distance-RSSI data pairs to be shifted to right side of the distance-RSSI plane. Due to this phenomenon, the estimated line departs from the true model, resulting in the GPS model. For large distances, the GPS distances are closer to the true distances. The corresponding distance-RSSI pairs are not shifted significantly, and therefore the perturbation they introduce in the path loss model estimation is little. Figure 7 illustrates the perturbed data, the mean and standard deviations for the RSSI-GPS distances pairs taken at each true distance, and the true and GPS models.

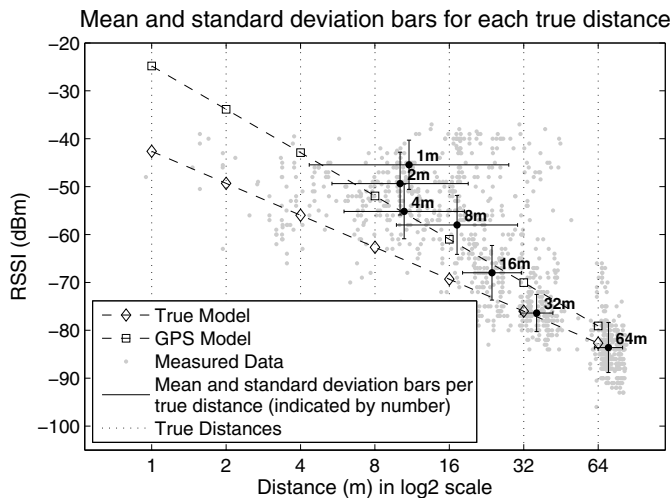


Fig. 7: Error metrics for the data measured at each true distance.

TABLE III: Model parameters and Root Mean Square Error (RMSE) if using the measurements taken at subsets of the true distances.

Model	d [m]	ρ_0	α	σ_ρ	RMSE [dBm]
True	1...64	-42.63	2.22	5.64	-
Range	1...64	-24.82	3.00	9.30	11.70
	2...64	-25.20	3.06	9.32	11.03
	4...64	-26.27	3.05	9.31	10.19
	8...64	-24.63	3.17	9.37	11.00
	16...64	-30.85	2.85	9.27	7.18
	32...64	-38.44	2.44	9.36	2.55

Based on this knowledge, a simple approach to improve the model estimation is to take more measurements at large distances, since those are less affected by GPS errors. Table III shows the regression using only measurements taken at different subsets of the true distances d , where $d \in \{1, 2, 4, 8, 16, 32, 64\}$. We keep the maximum true distance constant, while varying the minimum true distance to be considered. We observe that, as we restrict the data set to the measurements taken at larger true distances, the estimated model tends to the true model. From our data sets, the largest distances available are 32 and 64 meters. Although the condition $d > 3\sqrt{2}\sigma_{\text{GPS}} \approx 42$ meters (see Section IV) is satisfied just approximately for the lower distance, the corresponding estimated GPS model is the closest to the true model.

This conclusion allows us to propose guidelines for the selection of the distances at which measurements should be taken. If the communication range allows, the measurements should be taken at true distances larger than $3\sqrt{2}\sigma_{\text{GPS}}$. Otherwise, the model derived from the GPS distances needs to be corrected. A correction method is provided in the next section.

B. Retrieving the True Model from Imprecise Distances

In many practical scenarios, true distance measurements are not available and GPS distances need to be used instead. This is the typical case when the GPS positioning that comes

embedded in wireless devices and RSSI measurements from the wireless driver are used to derive a channel model. Note that direct estimation of the true distances from the GPS distances is not possible due to the insufficient number of i.i.d. samples that are taken at a fixed distance, especially in mobile scenarios. Therefore, we propose a simulation-based method² to improve the path loss model estimation. The essence of this method lies in the fact that the RSSI samples used for estimating both the true and the GPS model are the same. This equality is shown in the following equations.

$$\rho(d) = \rho_0 - 10 \alpha \log(d) \quad (8)$$

$$\rho(d) = \tilde{\rho}_0 - 10 \tilde{\alpha} \log(d_{\text{GPS}}) \quad (9)$$

Both equations hold in Least Squares sense, and therefore fading variance needs not to be included. As we assume that only GPS distances were recorded during the measurements, we can only compute the parameters of the GPS model $\tilde{\rho}_0$ and $\tilde{\alpha}$. Our method consists of defining a set of known reference distances, which are then perturbed according to the GPS error model of Section IV. The resulting simulated GPS distances are mapped into RSSI samples using the GPS model obtained from the measured data. Using the equality between (8) and (9), we pair the simulated RSSI values with the reference distances. By applying a regression method, we are able to retrieve corrected path loss parameters α and ρ_0 that are closer to the true model than if erroneous distance measurements had been used.

A crucial aspect of this procedure is the selection of the set of known reference distances. As seen in the previous subsection, the quality of the model parameters output by the regression method is very much affected by the distribution of the measured data with respect to the corresponding true distances. Therefore, we use Monte Carlo simulation to generate a set of simulated true distances \hat{d}_i that resemble the actual true distances d_i as closely as possible. This ensures that the estimated model is correct. Our method consists of selecting the simulated true distances in such a way that the corresponding erroneous distances \hat{d}_{GPS_i} resemble the compound p.d.f. $p(d_{\text{GPS}_i} | d_i)$ associated with the measured GPS distances d_{GPS_i} . We use the direct and inverse cumulative distribution function (c.d.f.) methods [25, Sec. 3.3] for this purpose. The direct c.d.f. method maps samples taken from a random variable X with some proposal distribution $p_X(x)$ into a uniformly distributed random variable $U \sim \mathcal{U}(0;1)$ using the c.d.f. of X as transformation function (by the uniform transformation theorem [25, Sec. 3.3.1]). Then, the inverse c.d.f. method maps U into a random variable Y that follows a target distribution $p_Y(y)$ using the inverse c.d.f. of Y as transformation function (by the inverse transformation theorem [25, Sec. 3.3.1]). In other words, $p_X(x)$ is used as a proposal distribution in order to sample from $p_Y(y)$. In our case, we use this procedure to sample from a proposal compound distribution of simulated erroneous distances \hat{d}_{GPS_i} in a way that the distribution of the transformed samples matches the compound p.d.f. $p(d_{\text{GPS}_i} | d_i)$ corresponding to the measured GPS distances d_{GPS_i} . In the

²Our attempt to derive a closed-form, or iterative estimator for the path loss model parameters (e.g. maximum likelihood) led to intractable calculations. Many of the Rice p.d.f. parameters have complicated expression (e.g. moments are expressed in terms of Laguerre polynomials), and the expression of the density itself contains a Bessel function.

TABLE IV: The sampling procedure that generates samples from the p.d.f. of the measured GPS distances, and the corresponding simulated true distances pairs.

1.	Generate initial reference distances: $\hat{\delta}_i \sim \mathcal{U}(0, \max_i d_{\text{GPS}_i}), i = 1, \dots, N$
2.	Simulate the GPS errors: $\hat{\delta}_{\text{GPS}_i} \sim \text{Rice}(\hat{\delta}_i, \sqrt{2}\sigma_{\text{GPS}})$
3.	Transform the resulting distances: $\hat{\delta}_{\text{GPS}_i} \rightarrow v_i = P_D(\hat{\delta}_{\text{GPS}_i}) \sim \mathcal{U}(0, 1)$
4.	Transform $v_i \rightarrow \hat{d}_{\text{GPS}_i} = P_V^{-1}(v_i) \sim p(d_{\text{GPS}_i} d_i)$
5.	Transform the uniform true distances to the simulated true distances: $\hat{d}_i = P_V^{-1}(P_D(\hat{\delta}_i))$
6.	Return: simulated true distances \hat{d}_i and the simulated GPS distances \hat{d}_{GPS_i}

end, our procedure outputs a set of simulated true distances \hat{d}_i paired with a set of simulated GPS distances \hat{d}_{GPS_i} whose distribution matches the one of the measured GPS distances. However, given that the mapping of the true distances d_i into GPS distances d_{GPS_i} is not bijective, there is no guarantee that the distribution of the true distances will be reproduced exactly.

We now describe the procedure in more detail. We start by generating an initial set of reference distances uniformly distributed between zero and the maximum GPS distance recorded, i.e., $\hat{\delta}_i \sim \mathcal{U}(0, \max_i d_{\text{GPS}_i}), i = 1, \dots, N$. We then perturb each reference distance $\hat{\delta}_i$ according to a Rice p.d.f. with the location parameter $\hat{\delta}_i$ and the scale parameter $\sqrt{2}\sigma_{\text{GPS}}$, i.e., $\hat{\delta}_{\text{GPS}_i} \sim \text{Rice}(\hat{\delta}_i, \sqrt{2}\sigma_{\text{GPS}})$. The standard deviation σ_{GPS} is the one computed from the empirical data in Section IV, using Equation (7). We then map the samples $\hat{\delta}_{\text{GPS}_i}$ of the proposal distribution to a standard uniform random variable $v_i \sim \mathcal{U}(0, 1)$, using the c.d.f. of the samples themselves (by the uniform transformation theorem). We denote the corresponding direct c.d.f. transformation $\hat{\delta}_{\text{GPS}_i} \rightarrow v_i$, by $P_D(\cdot)$. Then, we map the samples v_i to the target distribution of the measured GPS distances by using the inverse c.d.f. method, with the corresponding transformation $v_i \rightarrow \hat{d}_{\text{GPS}_i}$, denoted by $P_V^{-1}(\cdot)$, where $P_V(\cdot)$ is the empirical c.d.f. of the measured GPS distances. We obtain the simulated true distances by applying the same direct and inverse transformations to the original uniformly distributed reference distances $\hat{\delta}_i$. The sampling procedure is summarized in Table IV.

This approach is able to approximate the true model parameters with good accuracy. Figure 8 shows the histograms of the path loss parameters α and ρ_0 estimated by 10000 independent runs of the Monte Carlo simulations. Based on the shape of the histogram, we assumed a normal distribution for the estimation error, and the confidence intervals were computed accordingly. The standard deviation for the estimated path loss exponent $\hat{\alpha}$ is around 0.1, whereas for the estimated reference RSSI $\hat{\rho}_0$ is around 2dB. The true model parameters lie between one and two standard deviations with respect to the mean of the distribution of the estimated parameters. Table V presents the true model parameters and the mean values for the parameters estimated by the Monte Carlo simulations. Although the

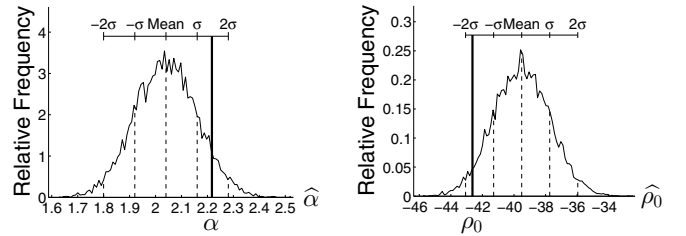


Fig. 8: Histograms of occurrences for the $\hat{\alpha}$ and $\hat{\rho}_0$ values in 10000 runs of the Monte Carlo simulation. Thick vertical line corresponds to true value.

TABLE V: The true model parameter values, the mean of the estimated parameters obtained from the Monte Carlo simulation, and the parameters derived from plain GPS distances.

Model	ρ_0	α
True	-42.63	2.22
Mean of MC runs	-39.51	2.04
GPS	-24.82	3.00

procedure does not guarantee retrieval of the exact true model parameters, it provides considerable improvement over the GPS model parameters.

VI. CONCLUSION

This work analysed the impact of positioning and distance errors on the estimation of a path loss model, using consumer-electronics (smart phones) communicating in the 2.4 GHz ISM band. We conclude that distances obtained from GPS measurements lead to the overestimation of communication range, and show that this is caused by the distance errors that result from GPS inaccuracy. We further modeled the impact of position errors on the distance estimation and provide guidelines for the selection of the distances at which measurements should be taken for path loss model parameter estimation. As a rule of thumb, if the communication range is sufficiently large, the measurements should be taken at distances larger than $3\sqrt{2}\sigma_{\text{GPS}}$. Otherwise, the path loss model derived from error prone GPS distances has to be corrected. A simulation-based method to correct the model estimated from erroneous distances was provided. A forested scenario was used as an example, but the presented analysis is applicable to any outdoor scenario. The only requirement is a good estimate of the standard deviation of the GPS positioning error.

ACKNOWLEDGMENT

This research work was funded by the Fundação para Ciência e Tecnologia through the individual grants SFRH/BD/67178/2009 and SFRH/BPD/81104/2011. Further support was provided by the CMU—Portugal project Vital Responder CMU-P/CPS/0046/2008, project Vital Responder 2 PTDC/EEI-ELC/2760/2012, and IT strategic funds PEst-OE/EEI/LA0008/2013.

REFERENCES

- [1] D. R. Bild, Y. Liu, R. P. Dick, Z. M. Mao, and D. S. Wallach, "Using predictable mobility patterns to support scalable and secure MANETs of handheld devices," in *Proc. 2011 ACM International Workshop on Mobility in the Evolving Internet Architecture*, pp. 13–18.

- [2] P. Papadimitratos, A. La Fortelle, K. Evenssen, R. Brignolo, and S. Cosenza, "Vehicular communication systems: enabling technologies, applications, and future outlook on intelligent transportation," *IEEE Commun. Mag.*, vol. 47, no. 11, pp. 84–95, Nov. 2009.
- [3] D. C. Teles, M. F. M. Colunas, J. M. Fernandes, I. C. Oliveira, and J. P. S. Cunha, "iVital: a real time monitoring mobile system for first responder teams," in *Proc. 2011 MONAMI*.
- [4] J. D. Parsons, *The Mobile Radio Propagation Channel*, 2nd ed. Baffins Lane, Chichester, England: Wiley, Nov. 2000.
- [5] S. Seidel, T. Rappaport, S. Jain, M. Lord, and R. Singh, "Path loss, scattering and multipath delay statistics in four european cities for digital cellular and microcellular radiotelephone," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 721–730, 1991.
- [6] X. Zhao, J. Kivinen, P. Vainikainen, and K. Skog, "Propagation characteristics for wideband outdoor mobile communications at 5.3 GHz," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 507–514, 2002.
- [7] S. Seidel and T. Rappaport, "900 MHz path loss measurements and prediction techniques for in-building communication system design," in *Proc. 1991 IEEE Vehicular Technology Conference*, pp. 613–618.
- [8] Y. S. Meng and Y. H. Lee, "Investigations of foliage effect on modern wireless communication systems: a review," *Progress In Electromagnetics Research*, vol. 105, pp. 313–332, 2010.
- [9] M. Weissberger, "An initial critical summary of models for predicting the attenuation of radio waves by trees," *Final Report Electromagnetic Compatibility Analysis Center*, vol. 1, 1982.
- [10] A. Seville and K. H. Craig, "Semi-empirical model for millimetre-wave vegetation attenuation rates," *Electron. Lett.*, vol. 31, no. 17, pp. 1507–1508, 1995.
- [11] J. A. R. Azevedo and F. Santos, "An empirical propagation model for forest environments at tree trunk level," *IEEE Trans. Antennas Propag.*, vol. 59, no. 6, pp. 2357–2367, 2011.
- [12] CCIR, "Influences of terrain irregularities and vegetation on troposphere propagation," Geneva, 1986.
- [13] M. Hall, "COST 235 activities on radiowave propagation effects on next-generation fixed-service terrestrial telecommunication systems," in *Proc. 1993 International Conference on Antennas and Propagation*, vol. 2, pp. 655–659.
- [14] M. Al-Nuaimi and A. Hammoudeh, "Measurements and predictions of attenuation and scatter of microwave signals by trees," *IEE Proceedings - Microwaves, Antennas and Propagation*, vol. 141, no. 2, pp. 70–76, 1994.
- [15] G. Joshi, C. Dietrich, C. Anderson, W. Newhall, W. Davis, J. Isaacs, and G. Barnett, "Near-ground channel measurements over line-of-sight and forested paths," *IEE Proceedings - Microwaves, Antennas and Propagation*, vol. 152, no. 6, pp. 589–596, Dec. 2005.
- [16] C. Oestges, B. Villaceros, and D. Vanhoenacker-Janvier, "Radio channel characterization for moderate antenna heights in forest areas," *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4031–4035, Oct. 2009.
- [17] J. Fernandez, I. Cuiñandas, and M. Sánchez, "Radioelectric propagation in a deciduous tree forest at wireless networks frequency bands," in *Proc. 2011 European Conference on Antennas and Propagation*, pp. 3274–3278.
- [18] T. Mangel, O. Klemp, and H. Hartenstein, "5.9 GHz inter-vehicle communication at intersections: a validated non-line-of-sight path-loss and fading model," *EURASIP J. Wireless Commun. Netw.*, vol. 2011, no. 1, p. 182, Nov. 2011.
- [19] M. Boban, T. T. V. Vinhoza, M. Ferreira, J. Barros, and O. Tonguz, "Impact of vehicles as obstacles in vehicular ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 15–28, Jan. 2011.
- [20] C. Sommer, S. Joerer, and F. Dressler, "On the applicability of two-ray path loss models for vehicular network simulation," in *Proc. 2012 IEEE Vehicular Networking Conference*, pp. 64–69.
- [21] National Coordination Office for Space-Based Positioning, Navigation, and Timing. Official U.S. Government information about the Global Positioning System (GPS) and related topics. Url: <http://www.gps.gov/>
- [22] M. G. Wing, A. Eklund, and L. D. Kellogg, "Consumer-grade global positioning system (gps) accuracy and reliability," *J. Forestry*, vol. 103, no. 4, pp. 169–173, 2005.
- [23] M. G. Wing and A. Eklund, "Performance comparison of a low-cost mapping grade global positioning systems (GPS) receiver and consumer grade GPS receiver under dense forest canopy," *J. Forestry*, vol. 105, no. 1, pp. 9–14, 2007.
- [24] P. A. Zandbergen, "Accuracy of iphone locations: a comparison of assisted GPS, WiFi and cellular positioning," *Trans. in GIS*, vol. 13, pp. 5–25, 2009.
- [25] J. V. Candy, *Bayesian Signal Processing: Classical, Modern and Particle Filtering Methods*. New York, NY, USA: Wiley-Interscience, 2009.



Pedro M. Santos (S'13) received his M.Sc. degree in Electrical and Computer Engineering from the University of Porto, Portugal, in 2009, and since then is pursuing the Ph.D. degree at the same institution. He develops his work at Instituto de Telecomunicações, and the main topics of his thesis are channel modelling and wireless ad hoc communications. He was awarded a Doctoral Scholarship from the Portuguese Foundation for Science and Technology in 2009. His general interests are wireless communications and computer architectures.



Traian E. Abrudan (S'02-M'09) received the D.Sc. degree (with honors) from Aalto University, Finland (formerly known as Helsinki University of Technology) in 2008, and the M.Sc. degree from the Technical University of Cluj-Napoca, Romania in 2000.

During 2001-2010, he was a member of SMARAD (Finnish Centre of Excellence in SMART RADios and Wireless Research) which has been selected as Center of Excellence in research by The Academy of Finland. He was awarded with Nokia

Scholarship in 2004 and 2006.

During September 2010-2013, he was a postdoctoral researcher at Faculty of Engineering, University of Porto (FEUP), Portugal, and member of Instituto de Telecomunicações (IT) Porto.

Since with October 2013, he has been a postdoctoral researcher at the Department of Computer Science, University of Oxford. Dr. Abrudan's research focuses on practical indoor localization and navigation systems, wireless transceiver algorithms, and sensing. His fundamental research topics include applied parameter estimation, numerical optimization and sensor array signal processing.



Ana Aguiar (S'94-M'98-S'02-M'09) graduated in Electrical and Computer Engineering from the Faculty of Engineering University of Porto (FEUP), Portugal, in 1998, and received her Ph.D. in Telecommunication Networks from the Technical University of Berlin, Germany, in 2008. She is an assistant professor at FEUP since 2009, with research interests in wireless networking and mobile sensing systems, specifically vehicular networks, crowd sensing, and machine-to-machine communications. She also contributes to several inter-disciplinary projects in the

fields of intelligent transportation systems and well-being (stress). She began her career as an RF engineer working for cellular operators, and she worked at Fraunhofer Portugal AICOS on service-oriented architectures and wireless technologies applied to ambient assisted living. She has published in and is reviewer for several IEEE and ACM conferences and journals.



João Barros (S'98-M'04-SM'11) is an Associate Professor of Electrical and Computer Engineering at the University of Porto and Founding Director of the Institute for Telecommunications (IT) in Porto, Portugal. He also teaches at the Porto Business School and co-founded two recent startups, Streambolico and Veniam, commercializing wireless video and vehicular communication technologies, respectively. He received his undergraduate education in Electrical and Computer Engineering from the Universidade do Porto (UP), Portugal and Universitaet

Karlsruhe, Germany, and the Ph.D. degree in Electrical Engineering and Information Technology from the Technische Universitaet Muenchen (TUM), Germany.