Improving RFID-Based Indoor Positioning Accuracy Using Gaussian Processes

Fernando Seco*, Christian Plagemann**, Antonio R. Jiménez* and Wolfram Burgard**

*Centro de Automática y Robótica (CSIC-UPM), Ctra. Campo Real km 0,200, 28500 Arganda del Rey, Madrid, Spain
E-mail: {fernando.seco,antonio.jimenez}@car.upm-csic.es
**University of Freiburg, Dept. for Computer Science, Georges-Koehler-Allee, Geb. 079, 79110 Freiburg, Germany
E-mail: {plagem,burgard}@informatik.uni-freiburg.de

Abstract—The received signal strength (RSS) of radio-frequency signals emitted from beacons placed at known locations in an environment, can be used by a local positioning system (LPS) to estimate the location of a person or a mobile object. In indoor environments, interference, multipath propagation of RF signals, and the presence of obstacles and people, lead to a complex spatial distribution of the RSS, which is inaccurately described by simple parametric models. In this work, we present a Bayesian method for an indoor RFID location system which uses an observation model based in Gaussian processes (GPs) nonparametric regression to represent the environment-specific RSS distributions for the individual RFID tags. The experimental results in an indoor environment demonstrate the effectiveness of GPs in order to increase positioning accuracy.

I. INTRODUCTION: RF-BASED INDOOR LPS

In many situations it is required to determine reliably the location of a person, an object or a robot, whether in the form of physical coordinates or as a symbolic label (e.g., "in the hall, near the exit door"). There is an intense research effort dedicated to the development of Local Positioning Systems (LPS) operative indoor, with the same degree of precision, reliability and availability already achieved by the Global Positioning System (GPS). A recent survey of the theoretical background on positioning estimation theory is presented in reference [1], while a description of current LPS technology is found in [2].

The majority of indoor localization systems utilize the propagation of radiofrequency (RF) signals emitted either by wireless communication nodes (Wifi, Bluetooth or Zigbee), or by dedicated beacons, notably ultrawideband radio (UWB) or radio-frequency identification (RFID) tags, whose location is previously known. In most cases, direct range-based multilateration positioning is not viable, since the hardware is not designed for accurate time-of-flight measurements; instead, location has to be inferred indirectly from the received signal strength (RSS). Unfortunately, in indoor environments, the RF signal strength depends not only on the range from emitter to receiver, but is also disturbed by phenomena like multipath propagation, interference, blocking caused by obstacles and people, etc. Many methods have been devised for determining accurate estimates of position in face of these difficulties [3]. The dependence of RSS with range can be modeled [4], so position can be estimated by minimization of a cost func-

tion of the difference between experimental and theoretical measurements; models can account for the presence of walls in the propagation path of the RF signal. Fingerprinting methods [5] create a database of recorded RSS signatures at every possible position in the displacement area; position is estimated later by comparison of current measurements to the entries of the database through statistical learning methods (k-nearest neighbors, decision trees, neural networks, etc). Although fingerprint techniques usually provide the highest location accuracy, there is a tradeoff in the time invested in the creation of the RSS database, as well as a high sensitivity to environment changes.

Bayesian localization methods [6] treat the experimental RSS data as a random variable dependent of the position, and produce optimal estimates of position by probabilistic techniques. Bayesian techniques are widely used in Robotics for localization and mapping problems [7], [8], [9]. In this research we will apply Bayesian positioning to an RFID-based LPS, using a regression technique to model the RSS-position dependence called Gaussian processes (GPs).

A. Previous work on Gaussian processes for localization

One of the pioneer works in RF-based indoor localization is the RADAR system [10], based in creating a fingerprint map of the strength of wifi signals, and the use of the nearest neighbor method in signal space for location of a person. The achieved accuracy is 8.3 m for 90% of estimations, using 3 access points covering 980 m^2. The Landmarc system [11], employs active RFID tags placed in a room, and combines the RSS measurements from a set of readers and those obtained from reference tags to estimate the user position. The attained accuracy was not very high due to the technological limitations at the time of the work (2004).

RFID technology has been extensively employed for simultaneous localization and mapping (SLAM) experiments in Robotics. In one representative example [9], a mobile robot determines the location of passive tags in an unknown building by Bayesian techniques and a laser ranger, and, in a second stage, finds its own location from the RFID readings themselves. A notable increase in positioning accuracy (error below 1 m) is achieved by fusion of RFID and laser ranger measurement data, without the use of odometry.

The first use of Gaussian processes for Bayesian loca-

tion reported in the literature is found in the work by Schwaighofer [12]. The case treated in this paper is localization in a DECT (digital portable phones) cellular network. The best accuracy reported is an average error of 7.5 m in an area of 280×180 m, with 15 base stations in sight, and 650 points used for calibration of the system. GPs show similar performance to simple nearest neighbor methods at a dense number of calibration points, but gets comparatively better as the number of calibration points decreases.

In the work by Ferris *et al* [13], Gaussian processes are applied to two different location problems: Wifi (indoor) and GSM (outdoor). In the first case, the displacement environment consisted in three floors of a building, comprising a total 54 rooms, and 75 different Wifi access points available. The calibration phase is intensive. The average error in localization tests is found to be 2.1 m, and a success rate in identifying the correct room of 81 %.

Kriging, a technique related to GPs, is used to generate NLOS (non-line of sight) correction maps for signal strength in mobile phone positioning [14].

B. Outline of paper

The next section will review briefly the theoretical basis for Bayesian location and GPs models. Section III describes our RFID-based experimental setup. In Sections IV and V, it will described how the measurement models and the positioning results, respectively, are obtained. Finally, we extract some conclusions and point to future work.

II. BAYESIAN LOCALIZATION METHODS FOR RFID POSITIONING

Suppose that a person equipped with an RFID-LPS device is located at an unknown position $\mathbf{x}$, and receives n RF signals in a given time interval. Decoding these signals, the system obtains the identities of the emitting tags, their corresponding positions in the environment $\{\mathbf{x}_{bi}\}, i = 1, \ldots, n$ from a database, and the received signal strength measurements $\{\text{RSS}_i\}$. The goal is producing an optimal estimate $\widehat{\mathbf{x}}$ of the current position of the user from that information.

A. Path-loss minimization solution

In general, the relationship between the measured signal strength RSS and position $\mathbf{x}$ can be written as a functional dependence and an error term affecting the measurements:

$$\text{RSS} = f(\mathbf{x}; \mathbf{x}_{bi}) + e. \tag{1}$$

If we know f and the probability density function (pdf) of the error, $p_e(e)$, then the maximum likelihood estimate (MLE) is given by [1]:

$$\widehat{\mathbf{x}} = \arg \min_{\mathbf{x}} p_e(\text{RSS} - f(\mathbf{x}; \mathbf{x}_{bi})). \tag{2}$$

In practice, an exact relationship between signal strength and position cannot be produced for RF propagation in normal indoor environments; however, an *approximate* relationship like the path-loss law [4] can be used:

$$\text{RSS}_{\text{PL}}(\mathbf{x}; \mathbf{x}_{bi}) = \text{RSS}_0 - 10\alpha \log_{10} \frac{\|\mathbf{x} - \mathbf{x}_{bi}\|}{r_0} + e, \tag{3}$$

where RSS_{PL} is the signal strength in a logarithmic scale (dB), $r = \|\mathbf{x} - \mathbf{x}_{bi}\|$ is the range between emitter and receiver, r_0 is a reference distance, and RSS_0 and α are constants which are determined from the calibration data. The term e is a random variable which accounts for the fluctuations of the signal strength caused by multipath propagation, reflections, interference, etc, and which is usually considered to be log-normally distributed with zero mean and variance σ_n^2. Note the rather strong assumption that the RSS depends only on the range between emitter and receiver. As the error $\mathbf{e}$ from a set of RSS measurements is distributed as $\mathbf{e} \sim \mathcal{N}(0, \mathbf{R})$, where $\mathbf{R}$ is the covariance matrix, the optimal estimation of position is obtained as [1]:

$$\widehat{\mathbf{x}}_{\text{LS}} = \arg \min_{\mathbf{x}} \|\text{RSS} - \text{RSS}_{\text{PL}}(\mathbf{x})\|^T \mathbf{R}^{-1} \|\text{RSS} - \text{RSS}_{\text{PL}}(\mathbf{x})\|. \tag{4}$$

Standard minimization methods, like Levenberg-Marquardt, using as initial position estimate the barycenter of all detected tags, weighted by their respective signal strengths, can be used to produce the LS position estimate.

B. Bayesian localization

In the Bayesian approach to position estimation it is assumed that both position $\mathbf{x}$ and measurements RSS are random variables related by a probabilistic law called the *observation model* $p(\text{RSS} \mid \mathbf{x})$. From a given experimental data set RSS, the MLE estimate of the unknown position $\mathbf{x}$ maximizes the conditional probability:

$$\widehat{\mathbf{x}}_{\text{MLE}} = \arg \max_{\mathbf{x}} p(\text{RSS} \mid \mathbf{x}), \tag{5}$$

assuming that there is no prior information of the position.

The computation of the conditional probability in Eq. 5 is usually made more tractable by supposing conditional independence of the measurements from different tags:

$$p(\text{RSS} \mid \mathbf{x}) = p(\text{RSS}_1, \ldots, \text{RSS}_n \mid \mathbf{x}) = \\ p(\text{RSS}_1 \mid \mathbf{x}) \times \cdots \times (\text{RSS}_n \mid \mathbf{x}). \tag{6}$$

In the case of RFID localization, different tag measurements are usually correlated. Fingerprint methods do not make the assumption of measurement independence, which is one of the reasons of their high precision.

The Bayesian inference method is readily extended from static to dynamic positioning with Bayesian filters [7]. The probability density function of the position is changed from a static pdf $p(\mathbf{x})$ to a sequence of pdfs $p(\mathbf{x}_t)$, sampled at consecutive time instants t. This pdf is updated from time $t - 1$ to time t, in the following way:

$$p(\mathbf{x}_t) = p(\text{RSS}_t \mid \mathbf{x}_t) \int p(\mathbf{x}_t \mid \mathbf{x}_{t-1}) \, p(\mathbf{x}_{t-1}) \, d\mathbf{x}_{t-1}, \tag{7}$$

where $p(\mathbf{x}_t)$ is actually a shorthand notation for $p(\mathbf{x}_t \mid \text{RSS}_1, \ldots, \text{RSS}_t)$, i.e., the conditional pdf of the position considering all the experimental measurements of signal strength since the beginning of the iterative process. The integral term in Eq. 7 fuses the prior pdf $p(\mathbf{x}_{t-1})$ with the so called motion model $p(\mathbf{x}_t \mid \mathbf{x}_{t-1})$, which incorporates

the available information about the current displacement. The multiplicative term $p(\text{RSS}_t \mid \mathbf{x}_t)$ is the Bayesian correction acting on the integral (prior) term, and describes how well the measured signal strength RSS_t fits the hypothesized position at time t, $\mathbf{x}_t$. Therefore, Eq. 7 is a *maximum a posteriori* (MAP) estimate of the position's pdf.

In this work we are mainly concerned with the improvement of positioning accuracy provided by Gaussian processes-based observation models; therefore we will only treat the static localization problem with the simple MLE of position of Eq. 5. In the future we will consider dynamic positioning with the full Bayesian formulation, with motion information obtained from an inertial motion unit (IMU) carried by the person to be localized (see companion paper [15] in this conference).

C. Range-only models for modeling the signal strength

Paramount for the performance of the Bayesian localization procedure is the accuracy of the probabilistic observation model $p(\text{RSS} \mid \mathbf{x})$. This model should fit the true distribution of signal strength measurements as tightly as possible, taking noise and the local characteristics of the environment into account. It is frequently assumed that, at a given point $\mathbf{x}$, the pdf of the received signal strength follows a Gaussian distribution:

$$p(\text{RSS} \mid \mathbf{x}) = \mathcal{N}(\mu_{\text{RSS}}(\mathbf{x}), \sigma_{\text{RSS}}^2(\mathbf{x})) . \tag{8}$$

The difficulty lies in producing an accurate estimate of the spatial distribution of the mean and variance of the signal strength, based on a training set of locations and measured RSS $\{\mathbf{x}_i, \text{RSS}(\mathbf{x}_i)\}$ obtained during a calibration stage.

One simple approach to model indoor RF signal propagation is to consider that the signal strength depends only with the range to the emitter, $p(\text{RSS} \mid r)$, where $r = \|\mathbf{x} - \mathbf{x}_b\|$. The path-loss model discussed above would be one particular case of a range-only model, although in general it is not needed to assume a particular form for the relationship between r and RSS. The main advantage of these models is their simplicity; also, the ability to incorporate new emitters without a previous calibration (all that is needed to know is their placement in the environment). However, the assumption of isotropy for the signal strength variation sets a limit to the attainable precision.

D. Gaussian processes for modeling the signal strength

Gaussian processes (GPs) are a non-parametric, probabilistic approach to function regression [16] that will be used in this work to model the dependence of the signal strength of RFID tags with position (Eq. 1), and obtain sensible estimates for $p(\text{RSS} \mid \mathbf{x})$.

Following the notation of [16], let $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n\}$ be the set of coordinates of the points where tag readings are recorded during calibration, and $Y = \{y_1, y_2, \ldots, y_n\} = \{\text{RSS}_1, \text{RSS}_2, \ldots, \text{RSS}_n\}$ the set of signal strengths RSS registered at these points for a given emitting tag.

GPs assume that the measurements y_p and y_q at two different locations $\mathbf{x}_p$ and $\mathbf{x}_q$ are random variables with a joint Gaussian distribution, and that their covariance is a function of the respective positions:

$$\text{cov}(y_p, y_q) = k(\mathbf{x}_p, \mathbf{x}_q). \tag{9}$$

The function k is called a kernel. One of the most used in the literature is the squared exponential kernel:

$$k(\mathbf{x}_p, \mathbf{x}_q) = \sigma_f^2 \exp(-\frac{1}{2l^2}\|\mathbf{x}_p - \mathbf{x}_q\|^2), \tag{10}$$

where σ_f^2 measures how strong is the correlation between the outputs, and l is a separation distance above which the outputs are no longer correlated.

GPs assume that a measurement y^* at a point $\mathbf{x}^*$ will be normally distributed:

$$p(y^* \mid \mathbf{x}^*, X, Y) = \mathcal{N}(\mu_{\mathbf{x}*}, \sigma_{\mathbf{x}*}^2),$$

and predict its mean and variance as:

$$\mu_{\mathbf{x}*} = \mathbf{k}_*^T(\mathbf{K} + \mathbf{R})^{-1}\mathbf{y}$$
$$\sigma_{\mathbf{x}*}^2 = k(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{k}_*^T(\mathbf{K} + \mathbf{R})^{-1}\mathbf{k}_* + \sigma_n^2, \tag{11}$$

where $\mathbf{k}_*$ is the vector of covariances between the point $\mathbf{x}^*$ and the calibration points, $\mathbf{K}$ is the covariance matrix of the calibration inputs X, and $\mathbf{R} = \sigma_n^2\mathbf{I}$ is the covariance of the measured noisy RSS. Because a constant (position independent) σ_n^2 is used to account for the Gaussian noise that adds up to the 'true' value of the signal strength, this formulation of GPs is called homoscedastic (i.e., constant noise). More realistic models must consider that the noise variance does indeed depend on the position: $\sigma_n^2(\mathbf{x})$, leading to the so-called heteroscedastic Gaussian process [17], [18].

Note that, from Eqs. 11, the predicted mean of the RSS can be viewed as a linear combination of the observations $\mathbf{y}$, or as a linear combination of n kernel functions, each one centered at a calibration point.

The final step before estimates can be produced with Eqs. 11, is obtaining suitable values of the so-called hyper-parameters $(\sigma_f^2, l, \sigma_n^2)$; this is usually done by maximizing the likelihood of the measured calibration data, following the procedure described in [16].

The main advantages of GPs-based regression of the RF signal strength variation in indoor environments, compared to fingerprinting are: (a) more flexible spatial sampling (measurements can be taken where possible, not at uniform points), (b) capacity to predict the RSS at locations outside of the calibration area, and (c) less accuracy degradation when the number of calibration points decreases.

III. RFID-LPS EXPERIMENTAL SETUP

Many different radiofrequency technologies (Wifi, Bluetooth, Zigbee, etc) can be used for the empirical demonstration of the proposed approach of this paper; indeed, from a theoretical point of view, which technology is chosen is irrelevant. We have built our indoor LPS system based in an active radiofrequency identification (RFID) system provided by RFcode (see Fig. 1). We use model M100 active tags, operating at 433 MHz, which are factory set to transmit

Fig. 1. RFID equipment used in this work. Model M100 active tag, and model M220 mobile reader with helical antennas, both by RFCode.
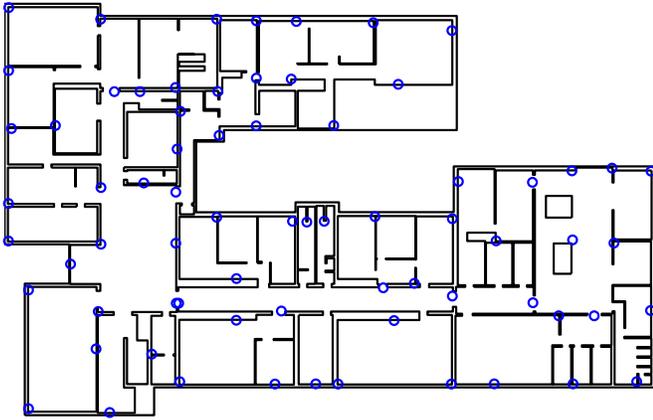

Fig. 2. Distribution of RFID tags in the displacement environment.

Fig. 3. Signal strength samples for one RFID tag (red point) at 40 calibration points in the building. RSS measurements are indicated by full circles, whose size corresponds to the obtained RSS value. X marks denote points where the tag was not detected.

their identification code at fixed intervals of one second, without being interrogated externally. The readers are model M220, portable, equipped with two 1/4 wave articulated helical antennas, and are light enough to be carried on a belt clip by a person, at an approximate height of 1 m. The readers decode the RF message transmitted by tags and, for every detection event, report the tag ID, the measured RSS at both antennas and a timestamp, to an HP ultraportable netbook through a Bluetooth link. Data collection, processing and location estimation is done in the Matlab environment.

A total of 71 tags are distributed in the building as shown in Fig. 2, attached to the walls at a height of 2 m, and covering a total area of 1600 m^2 (55 different rooms). The design of the localization system is thus privacy-oriented [2], in which the user can determine his location using RF signals from beacons placed in the infrastructure (much like the GPS system). Despite a relatively high tag density, the system is more scalable costwise than wifi technology, given the low cost of RFID tags. Besides, tags are easily placed in the environment, and, unlike wifi access points, they are untethered.

IV. EMPIRICAL RSS-POSITION MEASUREMENT MODELS

In this section we detail the procedure followed to build the experimental RSS-position measurement models which are used in the next section for positioning. We begin with a description of the calibration stage.

A. Calibration of the RFID-LPS system

For calibration, RFID measurements from all tags within range were collected at 40 different points in the building, at each of which a person stood for 60 seconds. Different random orientations were used at every point, to take into account the effect of shadowing of the RF signals by the human body, which caused a typical variation of 4 dB in the RSS. The dynamic range of the RFID signals registered by the reader was 50 dB (-100 to -50 dBm). The calibration process took about two hours to complete, and the sampling was not done at spatially uniform locations (as is usual in fingerprinting), but at random points in the environment. RSS results obtained during calibration from a particular tag, placed in the door that connects the lobby of the building with one of the corridors, are given in Fig. 3. On average, 30 unique tags per second are detected at the calibration points, which implies that a rather high density of tags has been used. Tags as far away as 30 m are consistently detected in some cases.

B. Path loss observation model

The collected data from all the calibration points is shown in Fig. 4 (a). While RSS clearly decreases with the range to the tag, it is obvious that individual measurements are not predictable. Fitting of the data to the path-loss model of Eq. 3 returns $\text{RSS}_0 = -51.9$ dB and $\alpha = 3.04$, which will be used for estimation of position with the least squares positioning method (the mean μ_{RSS} and the standard deviation σ_{RSS} are shown along with the data). Note that the linear dependence between RSS and the logarithm of the distance breaks down
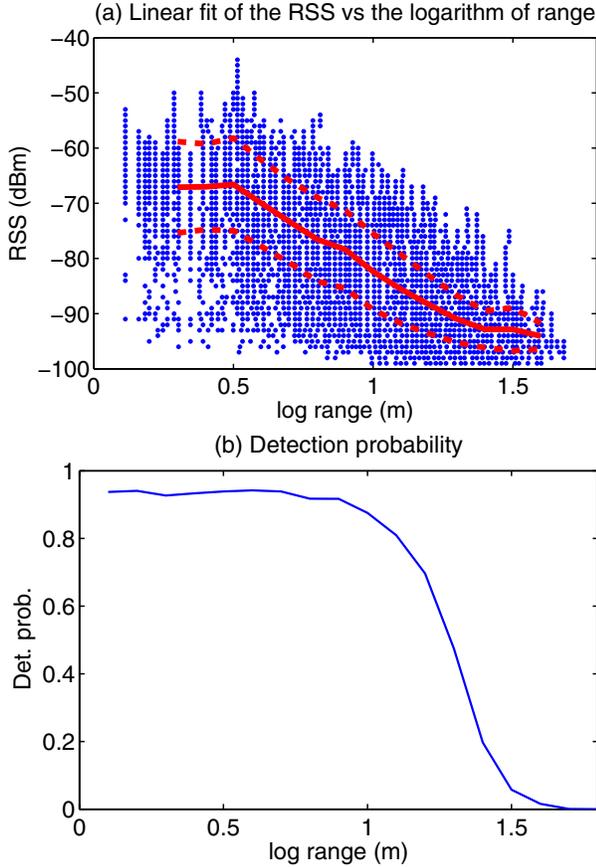
Fig. 4. Above: raw data of RSS vs. $\log r$ and linear fit (μ_{RSS}, and $\mu_{\text{RSS}} \pm \sigma_{\text{RSS}}$, in red) corresponding to the path-loss law model. Below: detection probability of a tag as a function of the logarithm of the range.

at very close ranges to the tag, and also for values of RSS near zero (the threshold level of the reader) due to truncation. Addressing these issues is out of the scope of this work and will be considered in future research.

Fig. 4 (b) shows the detection probability of a tag against the distance between the tag and the reader. As expected, the probability decreases with range, but on average, tags are still detected with 50 % probability at 20 m. Information from tag detection events can also be used to infer position estimates, along with RSS readings, as shown in [19].

C. Radial dependence only observation model

For computation of the range observation model $p(\text{RSS} \mid r)$ of Section II-C, the space surrounding a given tag is divided in discrete rings of 4 m width between zero and a maximum distance (24 m), generating a histogram for the RSS values within a given range. Once the histograms are produced, a normal distribution is fitted to the RSS data for each range interval, and used subsequently in the location estimation process. The advantage of using a Gaussian fit instead of a histogram of the RSS data is a smoothing effect which increases location robustness [20]. The statistical model for the RSS is shown for all seven range intervals in Fig. 5.
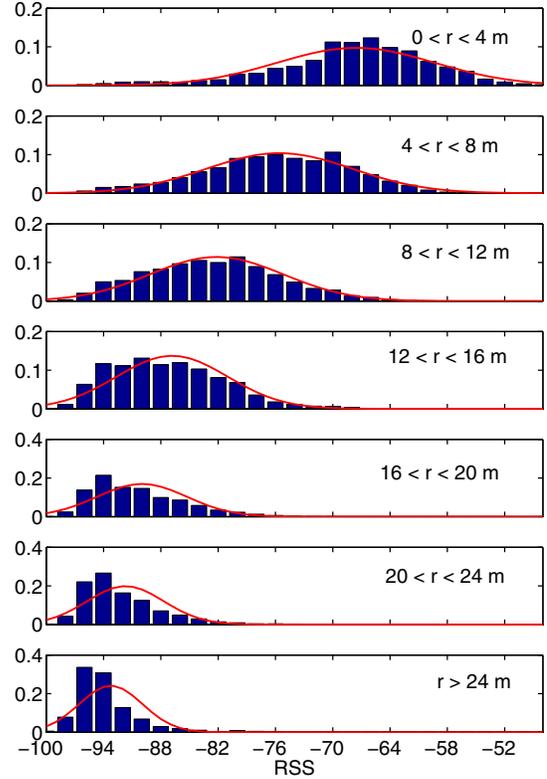

Fig. 5. The pdf of the received signal strength of tags at each range interval, $p(\text{RSS} \mid r)$, is plotted as a histogram (blue) and a continuous Gaussian distribution (red).

D. Gaussian processes-based observation model

For computation of the Gaussian processes models (one for each of the 71 tags), sets X and Y are created with empirical measurements of position and RSS taken at the calibration points. For every calibration point at which the tag is consistently detected (at least 10% of the emissions) the values $\mu_{\text{RSSI}} \pm \sigma_{\text{RSSI}}/\sqrt{2}$ are added to the set Y, and the calibration position $\mathbf{x}$ twice to the set X. To prevent numerical instabilities in the inversion of matrix $\mathbf{K} + \mathbf{R}$ in Eqs. 11, a small random value is added to each position sample $\mathbf{x}$.

For the reasons detailed in [13], the average dependence of RSS with the logarithm of range is subtracted from RSS samples, in order to achieve a zero-mean process. We have verified that doing so results in a GP estimate which fits better the experimental data. We have not found a significant difference in positioning results by using the Matérn kernels instead of the square exponential kernel in Eq. 10 for computation of the covariance matrix $\mathbf{K}$.

Optimal values of the hyperparameters are obtained by maximizing numerically the likelihood of the observed calibration data X, Y (see [16] for details). The results obtained for the different tags are similar: σ_n^2 is typically between 10 and 20, σ_f^2 between 5 and 20, and the length scale l varies between 5 and 10 m. For 5 of the 71 tags, the optimization process returned $\sigma_f^2 = 0$, indicating that no correlation between the RSS measurements at the calibration points could be found.
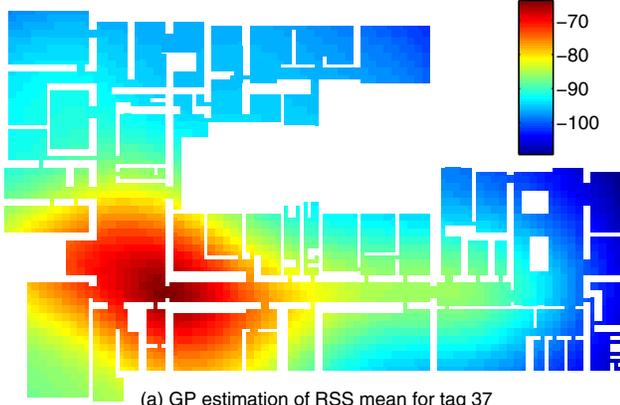
(a) GP estimation of RSS mean for tag 37
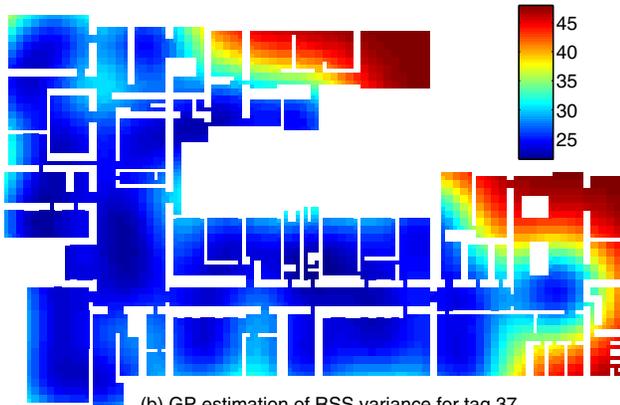
(b) GP estimation of RSS variance for tag 37

Fig. 6. Gaussian process model prediction for one particular RSS distribution on the complete building, based on the data from figure 3: (a) mean; (b) variance.

After obtention of the hyperparameters, the estimated values for $\mu_{GP}(\mathbf{x}*)$ and $\sigma^2_{GP}(\mathbf{x}*)$ are computed for all grid points $\mathbf{x}*$ and tags. From a computational point of view, this process is slow for a grid-based approach (due to the inversion of the covariance matrix in Eq. 11); however, the process needs to be done only once, during calibration.

Fig. 6 shows a particular GP sensor model learned from the empirical data for tag from Fig. 3, including the mean and variance predictions for the signal strength at all locations in the building. The GP estimate of RSS reflects the anisotropic nature of RF signal propagation, for example the existence of higher signal strength in the corridor that corresponds to a waveguide-like propagation. At places with a high density of calibration points, the estimated variance of the RSS is low, and in those rooms where calibration measurements were not available, the model can still make predictions for the signal strength; however, it cautiously assigns a high variance to those estimates. The influence of such estimates in the posterior computed with the Bayesian rule will be low, since

(a) GP predicted mean of the RSS vs. experimental data

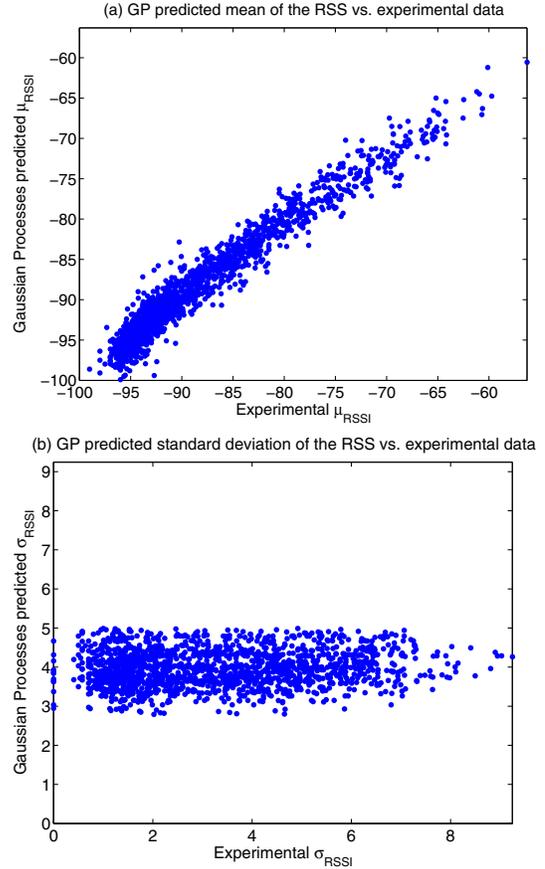(b) GP predicted standard deviation of the RSS vs. experimental data

Fig. 7. Correspondence of empirical and GP predicted values for RSS (a) mean and (b) standard deviation at the calibration points.

the corresponding Gaussian curve is relatively flat.

Fig. 7 contains a check of the accuracy of the GP model prediction of signal strength at the calibration points against the empirically measured values. While the mean μ_{RSS} fits reasonably well the empirical data, it is clear that the GP failed to model the variation of noise σ^2_{RSS} at different points of space, instead predicting a constant average noise. This is a feature of the homoscedastic Gaussian processes used in this work; in the future we will expand the current GP to a heteroscedastic model along the lines of references [17], [18], and check if this procedure provides more accurate positioning results.

V. POSITIONING RESULTS

This section provides empirical tests of the performance of the positioning methods described in Section II. The RSS data is taken from a set of 25 different locations, independent of the calibration set, but obtained with the same procedure.

For Bayesian methods, the distribution of position probability $p(\mathbf{x})$ was evaluated on a grid consisting of discrete square elements of size 0.8×0.8 m and 0.4×0.4 m covering the entire floorplan (two different sizes were used to obtain a connected grid in places where rooms or corridors were too narrow). Grid sampling was chosen because it is easy to implement,

Method	Mean error (m)	Median error (m)	90% error (m)
Least squares	2.28	2.17	3.74
Bayesian (range)	2.51	2.28	4.34
Bayesian (GP)	1.70	1.53	3.08

TABLE I
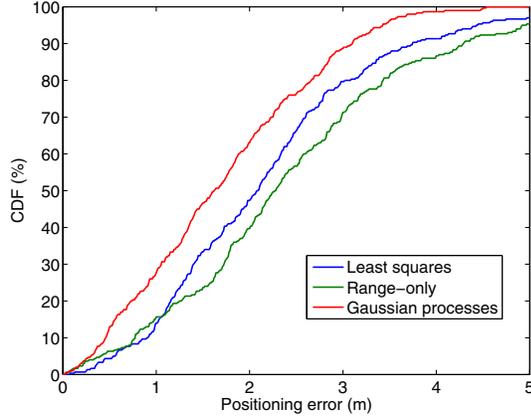
POSITIONING ERROR OBTAINED WITH DIFFERENT METHODS.

Fig. 9. Cumulative distribution function (CDF) of the positioning error of the three different methods.

and permits direct visualization of the estimated position's pdf, although it is not the most computationally efficient method [7]. For real time applications (tracking and guidance), more efficient schemes like particle filters are preferable [21].

Fig. 8 shows positioning results at 15 different locations. Each discrete point corresponds to a single position estimate combining measurements received in a 2 s interval by the reader. Also shown for each method is the confidence ellipse that contains 90% of the estimates at every point. The results are summarized in table I, and shown as a cumulative distribution function (CDF) of the error in Fig. 9.

As can be seen, minimization of the cost function of Eq. 4 and Bayesian location with the simple radial model achieve about the same precision, with a median error of $\simeq$ 2.2 m. This might indicate that both methods are limited by the assumption of isotropy in the RSS distribution. However, the median positioning error is reduced to 1.5 m by using Gaussian processes. The results are similar to those found in the literature (see for example [20] or [13]) although direct comparison is difficult because a coherent methodology does not exist between the different works (different RF technology, beacon spatial density, experimental conditions, etc). The important fact is that an improvement by 30% in the positioning accuracy is obtained by processing the same calibration data with GP regression rather than using one-dimensional RSS-range measurement models, and it is feasible that this accuracy can be further increased by using more complex GPs heteroscedastic methods.
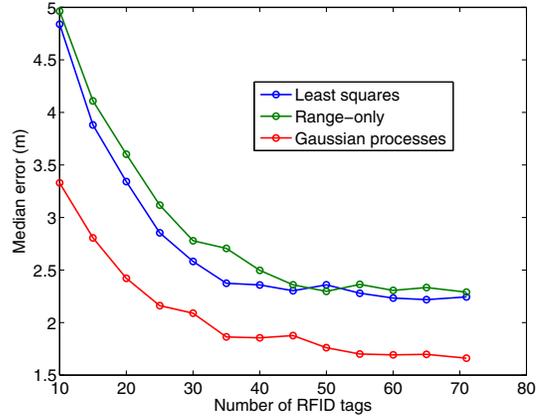

Fig. 10. Dependence of the median error of the three positioning techniques with the number of tags used.

Finally, a brief study of the dependence on the accuracy of the positioning methods with the number of tags used for the estimation is provided. For this purpose, the estimate of the position is computed with the readings of different subsets of tags, presenting the obtained median error in Fig. 10. As expected, the error decreases as $1/\sqrt{n}$, where n is the number of independent tags used for the estimation. The Gaussian processes approach shows consistently better accuracy that range-only Bayesian positioning and path-loss minimization for all tag densities.

VI. CONCLUSIONS AND FUTURE WORK

In this work we have discussed the application of Gaussian processes to compute measurement models that accurately describe the spatial dependence of the received signal strength (RSS) of RF signals propagating indoors. GPs are a non parametric regression technique which provides models readily usable in Bayesian positioning schemes, computed from a set of relatively sparse experimental data. The performance of the method is demonstrated empirically with an RFID positioning system operating in our research building, achieving a median positioning error of 1.5 m (with 71 tags covering 1600 m^2), which is 30% more accurate than least squares minimization with a path-loss model, or Bayesian positioning with a range-only model.

This preliminary work can be extended in several directions. The accuracy of the measurement model can probably be increased by using heteroscedastic (i.e., variable measurement noise) GPs rather than homoscedastic (constant noise) GPs. Use of different covariance kernels for the GP models deserves to be considered also. Finally, the ultimate goal of this research is to combine RFID tags with inertial motion units carried by a person to achieve accurate pose (position and orientation) estimates that permit location and guidance of people in indoor scenarios.
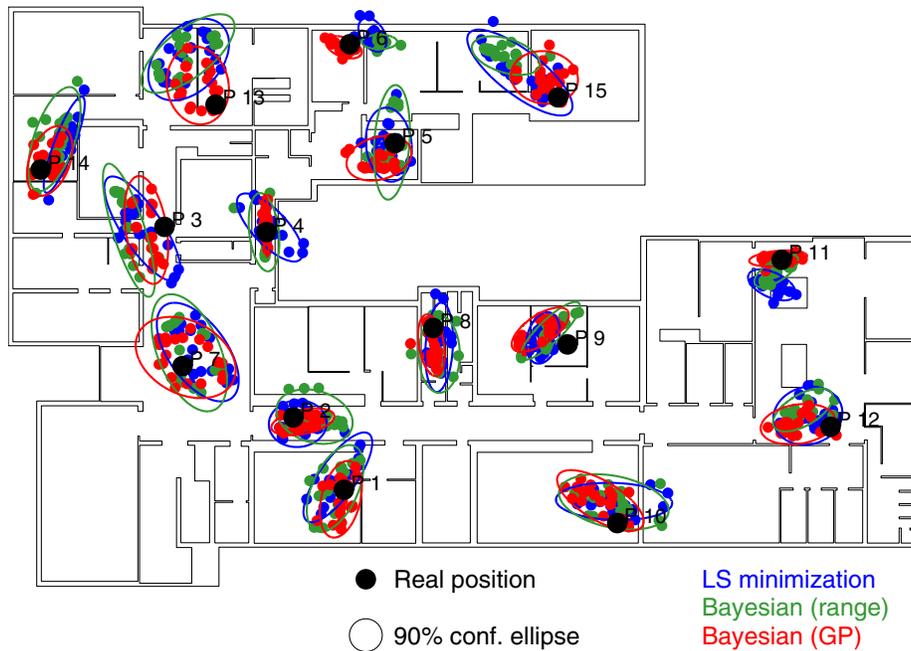
Fig. 8. Empirical location results at 15 different points in the building.

REFERENCES

[1] F. Gustafsson and F. Gunnarsson, "Mobile positioning using wireless networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 41–53, July 2005.

[2] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics Part C*, vol. 37, no. 6, pp. 1067–1080, Nov 2007.

[3] F. Seco, A. Jiménez, C. Prieto, J. Roa, and K. Koutsou, "A survey of mathematical methods for indoor localization," in *Intelligent Signal Processing, 2009. WISP 2009. IEEE International Symposium on*, aug. 2009, pp. 9–14.

[4] S. Y. Seidel and T. S. Rappaport, "914 MHz path loss prediction models for indoor wireless communications in multifloored buildings," *IEEE Trans. on Antennas and Propagation*, vol. 40, no. 2, pp. 207–217, 1992.

[5] K. Kaemarungsi and P. Krishnamurthy, "Modeling of indoor positioning systems based on location fingerprinting," in *23rd Conference of the IEEE Computer and Communications Societies*, vol. 2, March 2004, pp. 1012–1022.

[6] D. Fox, J. Hightower, L. Liao, D. Schulz, and G. Borriello, "Bayesian filtering for location estimation," *Pervasive computing*, vol. 2, no. 3, pp. 24–33, 2003.

[7] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. MIT Press, 2005.

[8] L. Liao, D. Fox, J. Hightower, H. Kautz, and D. Schulz, "Voronoi tracking: location estimation using sparse and noisy sensor data," in *Proceedings of the IROS*, 2003, pp. 723–728.

[9] D. Haehnel, W. Burgard, D. Fox, K. Fishkin, and M. Philipose, "Mapping and localization with RFID technology," in *Proceedings of the IEEE International Conference on Robotics and Automation*, New Orleans, USA, 2004.

[10] P. Bahl and V. Padmanabhan, "Radar: An in-building rf-based user location and tracking system," in *INFOCOM, IEEE Conference on Computer Communications*, Tel Aviv, Israel, 2000.

[11] L. Ni, L. Yunhao, L. Y. Cho, and A. Patil, "LANDMARC: indoor location sensing using active RFID," *Wireless networks*, vol. 8, no. 6, pp. 701–710, 2004.

[12] A. Schwaighofer, M. Grigoras, V. Tresp, and C. Hoffmann, "GPPS: A Gaussian Process positioning system for celullar networks," in *Advances in Neural Information Processing Systems*, 2003.

[13] B. Ferris, D. Haehnel, and D. Fox, "Gaussian processes for signal strength-based location estimation," in *Proc. of Robotics Science and Systems*, University of Pennsylvania, 2006.

[14] B. Li, C. Rizos, and H. K. Lee, "Utilizing kriging to generate a NLOS error correction map for network based mobile positioning," *Journal of Global Positioning Systems*, vol. 4, no. 1–2, pp. 27–35, 2005.

[15] A. Jiménez, F. Seco, J. Prieto, and J. Guevara, "Pedestrian indoor navigation by aiding a foot-mounted IMU with RFID signal strength measurements," in *2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Zurich, Switzerland, 2010.

[16] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[17] P. W. Goldberg, C. K. I. Williams, and C. M. Bishop, *Regression with input-dependent noise: a Gaussian process treatment*. MIT Press, 1998.

[18] K. Kersting, C. Plagemann, P. Pfaff, and W. Burgard, "Most likely heteroscedastic Gaussian process regresion," in *Proceedings of the 24th International Conference on Machine Learning*, Corvallis, OR, USA, 2007.

[19] D. Joho, C. Plagemann, and W. Burgard, "Modeling RFID signal strength and tag detection for localization and mapping," in *IEEE International Conference on Robotics and Automation, ICRA 2009*, 2009, pp. 3160–3165.

[20] A. Haeberlen, E. Flannery, A. M. Ladd, A. Rudys, D. S. Wallach, and L. E. Kavraki, "Practical robust localization over large-scale 802.11 wireless networks," in *Proceedings of the 10th ACM international conference on Mobile Computing and Networking*, Philadelphia, USA, 2004.

[21] A. Koutsou, F. Seco, A. Jimenez, J. Roa, J. Ealo, C. Prieto, and J. Guevara, "Preliminary localization results with an RFID based indoor guiding system," in *WISP 2007, IEEE International Symposium on Intelligent Signal Processing*, Oct. 2007, pp. 1–6.