Improving Gaussian Processes based mapping of wireless signals using path loss models

Renato Miyagusuku, Atsushi Yamashita and Hajime Asama[1]

***Abstract*— Indoor robot localization systems using wireless signal measurements have gained popularity in recent years, as wireless Local Area Networks can be found practically everywhere. In this field, a popular approach is the use of fingerprinting techniques, such as Gaussian Processes. In our approach, we improve Gaussian Processes based mapping using path loss models as priors. Path loss models encode information regarding the signal propagation phenomena into the mapping. Our approach first fits training data to a simple path loss model, and then trains a zero-mean Gaussian Process with the mismatches between the models and the data. Signal strength mean predictions are done using both the path loss model and the Gaussian Process output, while variances are calculated by bounding the Gaussian Process variance using the path loss models. Notably, the main improvement generated by our approach is not an enhanced mean value prediction, but rather a better model variance prediction. This translates into better likelihood estimations, leading to higher localization accuracy.**

Experiments using data acquired in an indoor environment and our approach as the perceptual likelihood of a dual Monte Carlo localization algorithm are used to demonstrate this improvement. Furthermore, this idea can be extrapolated to other fingerprinting techniques and to applications other than wireless-based localization.

I. INTRODUCTION

Robot localization or position estimation is the problem of determining a robot's pose relative to a given map of the environment. Robot's knowledge of its pose is essential for most non-trivial tasks; hence, its importance.

The use of wireless signals for robot localization in indoor, GPS-denied locations has gained popularity in recent years; being, probably, the main cause, the almost ubiquitous presence of wireless local area networks (WLANs) in most buildings. Although wireless signals-based localization systems do not achieve as high accuracy as those based on sensors such as laser rangefinders or RGB-D cameras; they possess certain characteristics that make their usage appealing. These are: signals' uniqueness (due to each access point's MAC address); relative low computation as compared with computer vision approaches; and readily available hardware (as most robots already possess wireless capabilities and WLANs infrastructures).

In wireless signals-based localization, the challenges arise due to the difficulty of modeling signal strength propagation, and the noisy nature of the signals themselves. Wireless signal strength propagation through space can be accurately described by Maxwell's equations; however, these are rarely used in practice because of their complexity. Simpler models can be obtained by decomposing wireless signals path loss, shadowing and multipath components. While shadowing and multipath are location specific and hard to model without a detailed description of the environment, path loss is relatively easy to model. Our main contribution in this paper is the modeling of signal strength propagation using such path loss models as informative priors over Gaussian Processes (GPs) and its use for bounding the model's variance (Fig. 1 shows the overview of our proposed approach).

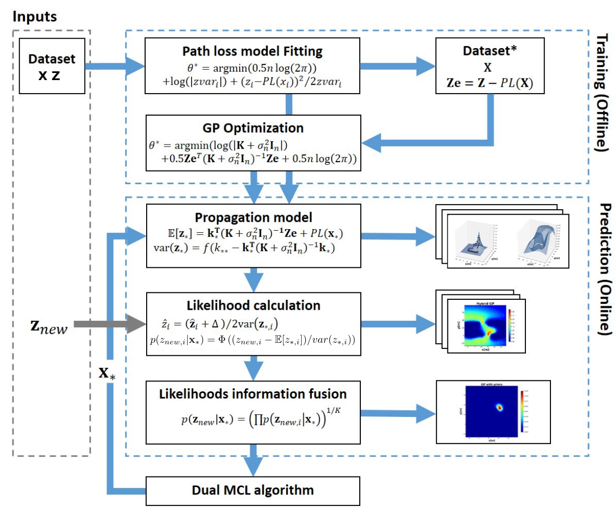

Fig. 1. Overview of our proposed approach.

In our experiments, we focus our analysis on how this refined variance improves localization accuracy. This is done by comparing the localization accuracy of a dual Monte Carlo Localization algorithm that uses our approach as perceptual likelihood, to one that uses Gaussian Processes with no priors. For the testing we use data we collected in an indoor environment. Additionally, we assess how sparser training datasets affect our localization accuracy, to verify the applicability of our approach when fewer data points in the training dataset are available.

The rest of this paper is organized as follows: the next section discusses related works and briefly introduces basic concepts of GPs; section 3 describes our proposed approach and can be considered as the core of this work. Section 4 describes our experiments and results; and section 5 gives our conclusions.

*This work was funded by Tough Robotics Challenge, ImPACT Program of the Council for Science, Technology and Innovation (Cabinet Office, Government of Japan).

[1]R. Miyagusuku, A. Yamashita and H.Asama are with the department of Precision Engineering, Graduate School of Engineering, The University of Tokyo, Japan. {miyagusuku, yamashita, asama} at robot.t.u-tokyo.ac.jp

II. PRELIMINARIES

A. Related work

Wireless signals-based localization techniques can be roughly classified into proximity, triangulation and fingerprinting - a more detailed survey can be found at [1].

Proximity techniques rely on the connectivity of the robot to its neighboring nodes. These techniques are often simple to implement, but require dense WLANs to have adequate accuracy, which is not an assumption we consider. Triangulation techniques use geometry to calculate the robot's location - several techniques are presented at [2]. These techniques are computationally fast, but often outperformed by fingerprinting ones.

Fingerprinting techniques use a training dataset of samples taken at known locations in the environment, to predict robot's location. This can be done by matching new measurements to the most samples in the training dataset, becoming a classification problem. Current work with classification techniques include the usage of grids [3], [4], kNN [5], [6], and random forests [7]. Fingerprinting can also be casted as a regression problem, where a location-measurement mapping is learned using the training dataset. Given new measurements, the likelihood of any candidate location is computed and can be used in a MCL approach. Current work with regression techniques include linear interpolation in graphs [8], smoothing [9], [10], and GPs [11], [12]. Among these techniques, GPs have the advantage of not only being able to model the complex behaviors of signal strengths, but also directly calculating the prediction variances; both necessary to construct perceptual likelihoods that can be employed by MCL-based algorithms, such as the dual MCL we use for testing. Work has also been done to extend WiFi GPs to handle the self-localization and mapping (SLAM) problem [13] and to handle heteroscedastic noise in the measurements [14]. In this work we extend the approach to include path loss models as priors, as well as refining the measurement model used for computing the likelihood of new measurements.

B. Gaussian Processes

A complete treatment of GPs can be found at [15], we will give a brief explanation and show the main equations used in our specific problem (wireless signal strength-based localization), for completeness of ideas.

GPs are a generalization of normal distributions to functions, describing functions of finite-dimensional random variables. Given some training points, a GP generalizes these points into a continuous function where each point is considered to have normal distribution, hence a mean and a variance. The essence of the method resides is assuming a correlation between values at different points, this correlation is characterized by a covariance function or a kernel.

Formally, given some training data $(\mathbf{X}, \mathbf{Z})$ where $\mathbf{X} \in \mathbb{R}^{n\times d}$ is the matrix of n input samples $\mathbf{x}_i, \in \mathbb{R}^d$; and $\mathbf{Z} \in \mathbb{R}^{n\times m}$ the matrix of corresponding output samples $\mathbf{z}_i \in \mathbb{R}^m$. Two assumptions are made. First, each data pair $(\mathbf{x}_i, \mathbf{z}_i)$ is assumed to be drawn from a noisy process:

$$\mathbf{z}_i = f(\mathbf{x}_i) + \epsilon, \tag{1}$$

where ϵ is the noise generated from a Gaussian distribution with known variance σ_n^2. Second, any two output values, $\mathbf{z}_p$ and $\mathbf{z}_q$, are assumed to be correlated by a covariance function based on their input values $\mathbf{x}_p$ and $\mathbf{x}_q$:

$$cov(\mathbf{z}_p, \mathbf{z}_q) = k(\mathbf{x}_p, \mathbf{x}_q) + \sigma_n^2 \delta_{pq}, \tag{2}$$

where $k(\mathbf{x}_p, \mathbf{x}_q)$ is a kernel, σ_n^2 the variance of ϵ and δ_{pq} is one only if $p = q$ and zero otherwise.

Given these assumptions, for any finite number of data points, the GP can be considered to have a multivariate Gaussian distribution:

$$\mathbf{z} \sim \mathcal{N}(m(\mathbf{x}),\, k(\mathbf{x}_p, \mathbf{x}_q) + \sigma_n^2 \delta_{pq}), \tag{3}$$

and therefore be fully defined by a mean function $m(\mathbf{x})$ and a kernel function $k(\mathbf{x}_p, \mathbf{x}_q)$.

Predictions $\mathbf{z}_*$ for an unknown data point $\mathbf{x}_*$, can be done by conditioning $\mathbf{z}_*$ to $\mathbf{x}_*, \mathbf{X}$ and $\mathbf{Z}$, obtaining:

$$p(\mathbf{z}_* | \mathbf{x}_*, \mathbf{X}, \mathbf{Z}) \sim \mathcal{N}(\mathbb{E}[\mathbf{z}_*], \mathrm{var}(\mathbf{z}_*)), \tag{4}$$

where,

$$\mathbb{E}[\mathbf{z}_*] = \mathbf{k}_*^T (\mathbf{K} + \sigma_n^2 \mathbf{I}_n)^{-1} \mathbf{z}, \tag{5}$$

$$\mathrm{var}(\mathbf{z}_*) = k_{**} - \mathbf{k}_*^T (\mathbf{K} + \sigma_n^2 \mathbf{I}_n)^{-1} \mathbf{k}_*, \tag{6}$$

with $\mathbf{K} = cov(\mathbf{X}, \mathbf{X})$ being the $n \times n$ covariance matrix between all training points $\mathbf{X}$, $\mathbf{k}_* = cov(\mathbf{X}, \mathbf{x}_*)$ the covariance vector that relates training points $\mathbf{X}$ to $\mathbf{x}_*$, and $k_{**} = cov(\mathbf{x}_*, \mathbf{x}_*)$ the variance of the test point. For making predictions, only these last two equations need to be computed.

III. GP USING PATH LOSS MODELS AS INFORMATIVE PRIORS

A. Problem formulation

In this work we addresses indoor robot localization. Regarding the indoor environment, the existence of several WLANs is assumed. However, there is no assumption regarding the WLANs' access points spatial distribution - there is no need for the access points to be homogeneously distributed, and there is no prior knowledge of their locations. Regarding the robot that wants to be localized. It is assumed that it has wireless capabilities, specifically a 802.11-compliant wireless network interface controller (WNIC). Robot's motion is considered to be planar, and localization is performed using only odometry information and measurements from the robot's WNIC. Odometry information is assumed to be readily available from the robot's sensors such as encoders. Wireless signals strengths are assumed to be sensed using its WNIC's built-in Received Signal Strength (RSS) indicator. It is also assumed that a training dataset composed of data pairs of RSS measurements means and the locations where they were taken is available. Formally, given an arbitrary number m of access points. The training dataset is considered as $(\mathbf{X},\ \mathbf{Z})$. With $\mathbf{X} \in \mathbb{R}^{n\times 2}$ the matrix of n training input samples $\mathbf{x}_i$ that correspond to the $x - y$

Cartesian coordinates where the samples were taken; and $\mathbf{Z} \in \mathbb{R}^{n \times m}$, the matrix created from the sampled mean of RSS measurements taken from each of the m access points at each of the n different locations. RSS measurements are normalized so 1 corresponds to 0 dB and 0 to -95 dB (the lower limit for most WNICs). Whenever no RSS measurements from an access point j are sensed at a certain location i, the value of $z_{i,j}$ is set to zero.

B. Models and training

Wireless signals are electromagnetic waves, and as such when propagating through environments they can be reflected, scattered and diffracted by walls, furniture and moving objects. It is common to consider that this propagation phenomenon is constituted by three components: path loss, shadowing and multipath. Path loss is caused by the dissipation of the power radiated by the transmitter, and it is a function of the distance between the transmitter and the receiver; shadowing effects are the result of power absorption by obstacles, and will persist as long as the obstacle remains, in case of walls or other fixed obstacles, they can be assumed to be constant in time; multipath effects are caused by signals reaching the receiver by several paths, and like the shadowing effects, they are location dependent. Our approach considers that these components can be learned using a parametric function as the path loss component, and a GP as the shadowing and multipath components.

1) RSS sensor model: We make a difference between the "sensor model" which models RSS as captured by WNICs, and the "propagation model" which models the signal strength propagation phenomena previously introduced.

When knowing the propagation model's predicted mean, the computation of the sensor model's predictive mean is straight forward. Considering WNICs have a lower sensing limit, we simply consider the sensor model's means to be the propagation model's bounded bellow by zero.

Variance computation, on the other hand, is different. For non zero RSS measurements, we consider the sensor model predictions to be corrupted by the same noise as that of the propagation model (this assumption holds when sensing noise is neglected, which is viable as the propagation model noise is comparatively much higher). However, for zero value RSS measurements, the noise variance is considered different. This difference arises from the sensor's inability to sense signal strengths under its lower sensing limit. The main intuition is the following: if the propagation model predicts an RSS value with mean -1 and a variance 0.04, the sensor model should output a zero value with high confidence (low variance) as even 3 standard deviations above the propagation model's mean predicted value, the signal strength is still lower than zero (-1 + 0.6). The equations used for mean and variance predictions using the sensor model are introduced in section III-C.

2) Path loss model: A common approach for modeling RSS is to use a simplified path loss model and consider shadowing and multipath as system noise. These simplified path loss models only aim to capture the essence of signal propagation without resorting to complex ray-tracing models or other geometric considerations. Such models often take the form of a linear function using log distance.

In our approach we use one of such models as a prior for the GP. Specifically, for our implementation we characterize the RSS at location $\mathbf{x}$, for access point j, as:

$$PL_j(\mathbf{x}) = Pt_j - k_j \log_{10}(d_j) + \epsilon, \quad (7)$$

with Pt_j being the RSS signal 1 m from access point j, d_j the Euclidean distance between $\mathbf{x}$ and the access point j- i.e. $||\mathbf{x} - (apx_j, apy_j)||$, k_j a positive constant and ϵ being Gaussian noise with variance σ^2_{PL}.

To find the value of model parameters $\theta_{PL,j} = (Pt_j, k_j, apx_j, apy_j)$, we fit the parameters to acquired training data. For this fitting, we consider zero values to have an additional variance penalty σ^2_{zero} as they encode less information regarding the path loss parameters - any negative value would be encoded as zero due to WNICs inability to sense them. Given this consideration we define the negative log likelihood of the training data given the path loss model parameters as:

$$nll_{\bar{PL}j} = 0.5n\log(2\pi) + \sum_{i=1}^{n} \log(\sigma_{i,j}) + \sum_{i=1}^{n} \frac{(z_{i,j} - \bar{PL}_j(\mathbf{x}_i, \theta_{PL,j}))^2}{2\sigma^2_{i,j}}. \quad (8)$$

for n training data points, where $\sigma^2_{i,j} = \sigma^2_{PL}$ for non-zero values and the penalized version $\sigma^2_{PL} + \sigma^2_{zero}$ otherwise. This negative log likelihood is then optimized using conjugate gradient descend. It was found experimentally that best results were obtained when initializing the access points origin parameters (apx_j, apy_j) to be placed in the vicinity of the location of the highest $\mathbf{z}_j$ value - conjugate gradient descend is sensitive to initialization.

3) GP for modeling mismatches: For characterizing the shadowing and multipath components, our approach uses GPs fed with the difference between the training dataset and the path loss prediction ($\mathbf{Ze} = \mathbf{Z} - PL(\mathbf{X})$). For there are two options: either learn one GP per access point or learn a single GP for all access points. The main advantage of using a GP per access point is that the leaned model fits each particular access point data much better, however it risks overfitting. The main advantage of learning a single GP is that predictions are faster to compute. In our particular case, as both components are highly location dependent, and locations are common to all access points, the second option seemed a more logical choice. Nevertheless, it was tested if using separate GPs would yield improvements in localization accuracy; but no noticeable difference in performance was obtained. This ratified our choice of using a single GP.

To fully define these GPs, it is also necessary to select the mean function and the kernel to be used. For our case a zero mean function is selected, which makes shadowing and multipath tend to a zero when no data contradicts this hypothesis. The selected kernel was the squared exponential one, as it has been successfully used in previous related works [12], [13], [14]; and its training was done by minimizing its negative log likelihood via conjugate gradient descend.

C. Predictions using the sensor model

For the propagation model, the predicted mean value, denoted as $\mathbb{E}[z_{*,j}]$, is calculated as the sum of the path loss model $PL_j(\mathbf{x}_*)$ and the predicted mean of the GP ($\mathbb{E}[ze_{*,j}]$), while the predicted variance, denoted as $\text{var}(z_{*,j})$, is considered equal to the variance outputted by the GP $\text{var}(ze_{*,j})$ - as our approach does not estimate noise levels for the path loss models, letting the GP handle the variance estimation. Therefore we have:

$$\mathbb{E}[z_{*,j}] = PL_j(\mathbf{x}_*) + \mathbb{E}[ze_{*,j}], \quad (9)$$

$$\text{var}(z_{*,j}) = \text{var}(ze_*), \quad (10)$$

with $PL_j(\mathbf{x}_*)$ calculated from eq. (7), and $\mathbb{E}[ze_{*,j}]$ and $\text{var}(ze_*)$ from eqs. (5,6).

For the sensor model, the propagation model's mean is bounded bellow by zero, yielding:

$$\mathbb{E}[z_{*,j}] = max(PL_j(\mathbf{x}_*) + \mathbb{E}[ze_{*,j}], 0). \quad (11)$$

Same as with the predicted mean, for positive $z_{*,j}$ values, the variance remains unchanged; however, as mentioned in section III-B.1 for negative values the variance needs to be adjusted. We do this by adding 3 standard deviations to the predicted mean, bound the result bellow by zero, subtract the bounded path loss prediction, divide the residue by three, and consider this value as the new standard deviation. 3 standard deviations means that with a 99.7% of confidence, we expect the values to be correctly predicted.

$$\text{var}(z_{*,j}) = ((max(PL_j(\mathbf{x}_*) + 3 * \sqrt{\text{var}(ze_*)}, 0) - \bar{P}L_j(\mathbf{x}_*))/3)^2, \quad (12)$$

Given these rectified statistics, the likelihood for each individual access point can be computed. Figure 2 shows sensor model predictions for one access point.

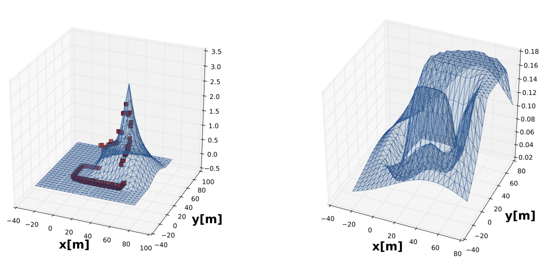

Fig. 2. Propagation model used for likelihood calculation of new RSS measurements. Figure shows (left) the predicted mean and (right) the predicted variance for the access point.

D. Likelihood estimation

In order to use our model as the perceptual likelihood of a MCL-based approach, it is necessary to calculate the likelihood of any new set of measurements $\mathbf{z}_{new}$ to have been obtained from any candidate locations $\mathbf{x}_*$ - i.e., $p(\mathbf{z}_{new}|\mathbf{x}_*)$. This is done using the sensor model's probability distribution and its predicted values. For each access point j, this likelihood is calculated as:

$$p(z_{newj}|\mathbf{x}_*) = \Phi\left(\frac{z_{new,j} - \mathbb{E}[z_{*j}]}{var(z_{*j})}\right), \quad (13)$$

If each access point is considered independent given the location $\mathbf{x}_*$, the integrated likelihood $p(\mathbf{z}_{new}|\mathbf{x}_*)$ should be obtained by multiplying the individual likelihoods for all access points. However, it has been observed in practice that this leads to overconfident estimates, yielding suboptimal results [12]. A simple remedy is to replace it by a "weaker" version $p(\mathbf{z}_{new}|\mathbf{x}_*)^\alpha$ with $\alpha < 1$ - similar to what is suggested for laser range finders beam models in [16].

The value of α selected is the inverse of the number of access points used. Making integration of individual likelihoods not the product but the geometric average of the individual ones:

$$p(\mathbf{z}_{new}|\mathbf{x}_*) = \left(\prod_{j=1}^{m} p(z_{newj}|\mathbf{x}_*)\right)^{1/m} . \quad (14)$$

IV. Results

In this section we evaluate the performance of our approach compared to GPs without priors. For this, we collected data using a commercial laptop placed on top of a Pioneer 3 DX mobile robot. This laptop used *tcpdump* version 4.5.1 as wireless packet analyzer, to collect RSS samples in monitor mode. Only beacon frames were considered to guarantee signals came from access points.

A first run with the robot was performed in order to collect the training dataset. In this run the robot was fully stopped every time a data sample was taken. Data samples consisted of 1000 RSS measurements and odometry information. The average distance between consecutive samples was 1.69 m. At the end of the run, odometry was manually rectified using the area's blueprint as reference. A second run was performed at a later time, in order to acquire the testing dataset that is used for the evaluation of our approach. During this second run, the robot was constantly operated at speeds close to its maximum (1.2 m/s), acquiring samples on the run. Each data sample consisted of 250 RSS measurements and odometry information - sampling frequency was 0.5 Hz. Odometry was not rectified when used as input for the dual MCL, and the ground truth of the robot's location used for computing the localization error was manually calculated, which induces errors in the results presented; however, as these errors are lower than those induced by the localization algorithm they were overlooked. Figure 3 shows the blueprint of the building used for testing.

In the following sections, our approach will be compared to a model based on Gaussian Processes with no priors - similar to the one developed at [12]. For all tests, the models were trained using the described training dataset and evaluated with the testing dataset.

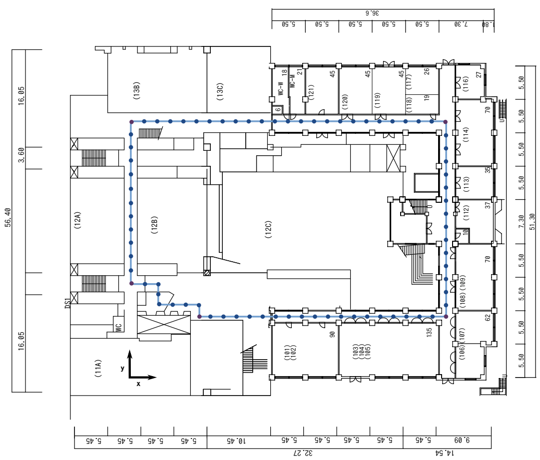

Fig. 3. Blue prints of the environment used for testing. The blue line and blue dots show the route taken by the robot when mapping and localizing, and purple dots indicate the *anchor* points used for rectifying odometry.

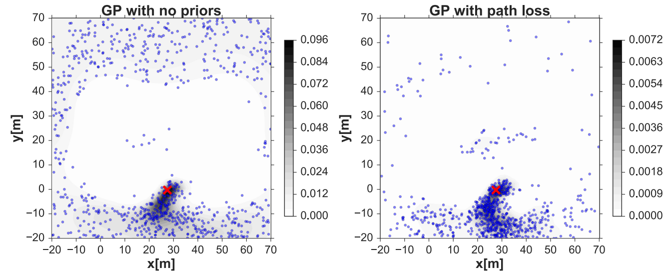

Fig. 4. Posterior probability distributions for (left) a GP with no priors and (right) our approach. Red $\mathbf{X}$ represents true location, shaded areas represent posterior probability values (lighter ones represent low probability, while darker ones, high), blue dots represent samples taken from the distribution.

A. Mean prediction

First, we assessed if the use of path loss priors significantly enhanced RSS mean prediction. For this, we calculated the average error between the testing dataset RSS values and both approaches' mean predictions. For the GP with no priors, the average error was 2.30 ± 4.1 dB, for our approach it was 2.22 ± 4.22 dB, and the average difference between both approaches was 0.46 ± 1.0 dB. Hence, it can be concluded that there was no statistical difference between the mean predictions. This is mainly due to the great flexibility of GPs approach, which, from data alone, is capable of generating adequate mappings without the need of priors.

B. Posterior distributions

However, posterior distributions generated from the models do significantly vary. This was mostly due to the variance bounding using path loss information done in eq. (12). In this section we compare these posteriors in a qualitative manner, and in the next one numerically assess the improvements this difference causes in localization accuracy.

Figure 4 shows the posterior distributions of the robot's location (red X) as computed by both approaches for a single testing point, with light shades representing low probabilities and darker ones higher. From the figure it can be easily observed that our approach assigns lower probabilities for points farther from the robot's true location than those assigned by the other method. This is a consistent result obtained for all testing points. As previously mentioned, this behavior emerges from adequate the handling of variance performed as it has already been stated in the previous section that mean predictions were not significantly different, and posterior distributions under the GP assumption, only depend on predicted means and variances.

Figure 4 also shows samples taken from each distribution (blue dots). As it can be observed for our approach samples are much more concentrated around robots true position. This is extremely important, as the main role of these distribution is to provide adequate samples for our localization algorithm.

C. Dual MCL using GP with and without priors

From the qualitative analysis it was stated that our approach computed better posterior distributions. In order to quantify how this improvement contributes to the generation of more accurate localization algorithms, this section evaluates the localization accuracy of a 800 particles dual MCL algorithm [17] when using each of the approaches as perceptual likelihoods. Readers are encouraged to see the accompanying video for a better visualization as well as our source code and datasets[1].

The localization errors for this test have been computed as the root of the mean squared error between the predicted localization and the ground truth. Due to the randomized nature of the dual MCL algorithm, all tests have been computed 25 different times using different random seeds. Figure 5 shows the mean of all tests in darker color, and each individual result in a lighter one. For the GP without priors the average localization error, once it converged around s=50 was 1.98 m, while for our approach it was of 1.31 m.

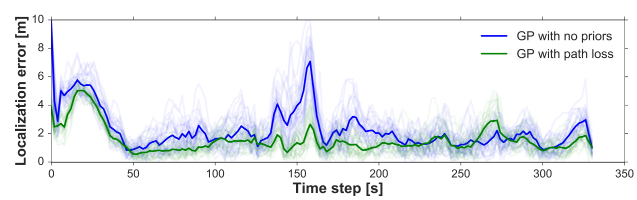

Fig. 5. Localization accuracy of our approach and a GP with zero mean when used as the perceptual likelihood of a dual MCL algorithm.

Figure 6 shows the cumulative localization error probability for the same test, where it can be observed that the GP without priors has an error lower than 2.26 m 80% of the time, while our approach has an error lower than 1.63 m.

D. Localization with sparser datasets

Finally, we tested the performance of our approach with sparser datasets. Sparser dataset provide several advantages, being the main one faster computation - as the algorithm has complexity $O(n^3)$ with respect to the training data

[1] http://www.robot.t.u-tokyo.ac.jp/~miyagusuku/iros2016

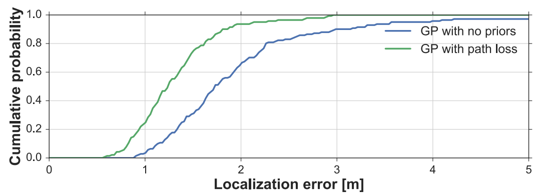

Fig. 6. Cumulative probability of localization accuracy once the dual MCL has converged, comparing our approach and a GP without priors.

points, therefore, sparser training datasets greatly speed up computation. For this test the original training dataset was modified by eliminating training samples, so the average distances between adjacent training points goes from 1.66 in the original dataset, to 3.28 in the sparse dataset 1 and 4.85 m in the sparse dataset 2. Using these sparse datasets our approach is trained and the same full testing dataset used before is employed again. Figure 7 shows the cumulative localization errors for the original dataset, the two sparse datasets as well as for the GP without priors (for reference), with 80% of localization errors under 1.63, 1.68, 1.68 and 2.26 m respectively. As it can be observed, the localization accuracy almost does not drop when using sparser datasets.

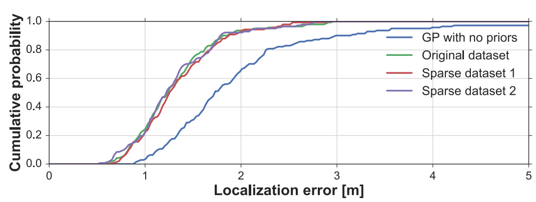

Fig. 7. Cumulative probability of localization accuracy for training datasets with 1.658, 3.276 and 4.849 m.

V. Conclusions

The main purpose of this study was the development of a novel approach for modeling wireless signal strength propagation through space. The main goal was for such novel approach to improve current state of the art wireless signals strengths-based localization. In summary, our approach first learns simple path loss models from data, and then feeds the mismatches to a GP. Predictive means are calculating as the addition of both models. Moreover, predictive variances are computed based on the GP and bounded considering path loss' encoded information about the signal strength propagation phenomena.

Through experiments we have demonstrated that the use of path loss priors notably improves localization accuracy. We mainly attribute this to the bounding of predictive variances rather than an increase in accuracy of the mean predicted signal strength. This bounding of predicted variances generates better likelihood distributions than previous GP-based approaches. Likelihood distributions generated by our approach output lower probabilities for points farther from the robot's true location. This improvement is directly translated into better localization accuracy when used in conjunction with a dual MCL algorithm. Our experiments showed an increase of accuracy from an average error of 1.98 m with errors lower than 2.26 m for 80% of the time, to an average error of 1.31 m with errors lower than 1.63 m for 80% of the time. Although it has only been tested with GPs in the specific case of wireless-based localization, the idea of using priors not only to enhance mean predictions but also variances can be potentially used with any fingerprinting technique and sensor whose model is known.

We also tested the performance of our approach with sparse training datasets. In this test it was shown that the localization accuracy when using training points separated, on average, by 1.6, 3.2 and 4.8 m, was almost the same. However, it remains for future work to evaluate to what extend path loss models could enhance prediction in much sparser datasets.

References

[1] 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 (Applications and Reviews)*, vol. 37, no. 6, pp. 1067–1080, Nov 2007.

[2] D. Koks, "Numerical calculations for passive geolocation scenarios," DTIC Document, Tech. Rep., 2007.

[3] J. S. Gutmann, E. Eade, P. Fong, and M. E. Munich, "Vector field slam - localization by learning the spatial variation of continuous signals," *IEEE Transactions on Robotics*, vol. 28, no. 3, pp. 650–667, June 2012.

[4] P. Kontkanen, P. Myllymäki, T. Roos, H. Tirri, K. Valtonen, and H. Wettig, "Topics in probabilistic location estimation in wireless networks." in *PIMRC*. Citeseer, 2004, pp. 1052–1056.

[5] P. Bahl and V. N. Padmanabhan, "Radar: an in-building rf-based user location and tracking system," in *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 2, 2000, pp. 775–784 vol.2.

[6] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil, "LANDMARC: indoor location sensing using active RFID," *Wireless networks*, vol. 10, no. 6, pp. 701–710, 2004.

[7] B. Benjamin, G. Erinc, and S. Carpin, "Real-time wifi localization of heterogeneous robot teams using an online random forest," *Autonomous Robots*, vol. 39, no. 2, pp. 155–167, 2015.

[8] J. Biswas and M. Veloso, "Wifi localization and navigation for autonomous indoor mobile robots," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, May 2010, pp. 4379–4384.

[9] A. LaMarca, J. Hightower, I. Smith, and S. Consolvo, "Self-mapping in 802.11 location systems," in *UbiComp 2005: Ubiquitous Computing*. Springer, 2005, pp. 87–104.

[10] J. Letchner, D. Fox, and A. LaMarca, "Large-scale localization from wireless signal strength," in *Proceedings of the national conference on artificial intelligence*, vol. 20, no. 1. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2005, p. 15.

[11] A. Schwaighofer, M. Grigoras, V. Tresp, and C. Hoffmann, "GPPS: A Gaussian Process Positioning System for Cellular Networks." in *NIPS*, 2003.

[12] B. Ferris, D. Haehnel, and D. Fox, "Gaussian processes for signal strength-based location estimation," in *In proc. of robotics science and systems*. Citeseer, 2006.

[13] B. Ferris, D. Fox, and N. D. Lawrence, "WiFi-SLAM Using Gaussian Process Latent Variable Models." in *IJCAI*, vol. 7, 2007, pp. 2480–2485.

[14] R. Miyagusuku, A. Yamashita, and H. Asama, "Gaussian processes with input-dependent noise variance for wireless signal strength-based localization," in *Proc. of the 13th IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR2015)*, 2015.

[15] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. the MIT Press, 2006.

[16] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. MIT press, 2005.

[17] S. Thrun, D. Fox, W. Burgard, *et al.*, "Monte carlo localization with mixture proposal distribution," in *AAAI/IAAI*, 2000, pp. 859–865.