Fingerprinting Based Localisation Revisited

a rigorous approach for comparing RSSI measurements coping with missed access points
and differing antenna attenuations

Christian Beder and Martin Klepal
Nimbus Centre for Embedded Systems Research
Cork Institute of Technology
Cork, Ireland
Email: {christian.beder,martin.klepal}@cit.ie

Abstract—Fingerprinting based localisation systems rely on taking a radio map of the environment and determine the position of the device by comparing this map to its current measurements. Therefore the performance of any such system heavily relies on the accuracy of this comparison. Typically two problems arise: first, access points can be missed in a scanning cycle both during the fingerprinting phase as well as during the localisation phase and second, if different devices are used for fingerprinting and for localisation their received signal strengths might not be the same due to differing antenna attenuation. A successful comparison function has to cope with both of these issues yielding repeatable high likelihoods for measurements taken in the same location while at the same time providing sufficiently high discrimination for measurements taken in different locations. In this paper we will propose such a likelihood observation function based on rigorous assumptions. Like most approaches we will compare RSSI values based on squared Euclidean distance of the log energies, which is essentially a Gaussian assumption on the distribution of the measurement error justifiable by arguments like maximum entropy or the law of large numbers acting on multiple additive error sources. However, this naive approach suffers from the varying dimensionality of the log-energy space caused by missing access points in the measurement. In order to overcome this in a rigorous manner we propose to model the access point pickup probabilities using Gibbs distributions enabling the introduction of rigorously motivated penalties for these dimension mismatches. As a further extension of the likelihood observation function we also propose to make it invariant to differences in antenna attenuation by estimating these explicitly from the log-energy observations and using the minimised squared observation residuals as invariant distance measure instead. We will discuss the properties of this improved likelihood observation function and compare its performance in a particle filter based WiFi localisation system.

Index Terms—WiFi localization, Fingerprinting, RSS based localisation

I. INTRODUCTION

Fingerprinting based localisation, i.e. determining the position based on comparing a current measurement of received signal strengths with a map based on previous measurements, is considered the methods of choice for WiFi based localisation techniques [1]. One key aspect common to all methods in this category is the need to determine the distance between a pair of signal strengths measurements, one taken from the fingerprint and the other being the current measurement. While most papers on fingerprinting based localisation focus on the optimisation of this measure, we will look at the measure itself only. Hence, the improvements we propose will be directly and easily applicable to most fingerprinting based localisation systems.

The early paper of Bahl et al. [2] already pointed out the importance of proper choice of distance metric and proposed the use of either the Euclidean or the Manhattan distance between both signal strength vectors, which has been replicated by a lot of subsequent approaches. Corte at al. [3] evaluated different distance metrics including Euclidean, Manhattan, Chi-Squared, Bray-Curtis, and Mahalanobis, the last of these achieving the best performance results. Also Biswas et al. [4] used Mahalanobis distance including variance estimates into the fingerprint map to improve its performance. Milioris et al. [5] went one step further by using a multivariate Gaussian estimating also the correlations between the beacons during fingerprinting.

Using the Mahalanobis distance (or equivalently a Gaussian likelihood observation function) became very common in a variety of approaches (e.g. [6], [7], [8]). Some improvements to this were proposed by Kaemarungsri et al. [9] who noted the non-centrality of the χ² distribution in case the fingerprint estimate is biased, which has also been noted by Seco et al. [10]. However all of these approaches essentially stick to the variance-normalised squared distance as comparison metric for received signal strength vectors. We will use this as well as a starting point for the improvements we propose in this paper.

Mirowski at al. [11] state that the signal strength distribution is not Gaussian at all, so they propose to model the entire distribution in the fingerprint and compare distributions using the KL-divergence as a metric instead of single measurements. Other authors also proposed to use non-Gaussian likelihood observation functions, for instance Elnahrawy et al. [12] used a t-distribution as a robustified version of a Gaussian, however their focus was on the optimisation algorithms not on the likelihoods used. We will not go down this route in this paper.

Another quite different approach to ours not requiring
explicit modeling of the distance metric are based on unsupervised machine learning techniques [13]. Brunato et al. [14] propose to use support vector machines, hence learning everything from examples rather than explicitly trying to implement domain knowledge about the distances in signal strength space. Other approaches in this direction include [15] using neural networks, or [16] proposing to use decision trees. We on the contrary will explicitly look at the distance metric and model it according to the particular problem domain of signal strengths.

We will take a fresh look at the problem of determining the likelihood of two signal strengths measurements being taken in the same location, which is the basis for essentially all fingerprinting based localisation algorithms. Like most approaches we will consider the measurement noise to be Gaussian distributed in log-energy space (which is despite all the criticism reasonable justifiable by maximum entropy arguments), but propose to extend it in two ways addressing two issues encountered by every localisation algorithm based on this type of likelihood observation function (or equivalently distance metric between signal strengths measurements).

The first issue we will address is the differing attenuation of signal strengths measurements occurring due to the use of different devices or even due to carrying the device differently. This effect has also been analysed by Vaupel et al. [17], who proposed a pre-calibration for different devices to increase the performance of their localisation system. Our proposed solution to this issue does not require any pre-calibration but will rely on the presence of enough beacons enabling the estimation of the relative offset of signal strengths at every point in time.

The second issue we will look at is much more important. It addresses the fact that not all beacons are visible everywhere in the environment, which causes dimensionality mismatches between the distance metrics that are used. Some attempts have been made to overcome this, for instance Kushki et al. [18] pointed out that selecting access points is crucial for the performance of localisation in order to avoid biased estimates. Also Gansemer et al. [19], [20] identified this issue distinguishing the different mismatch scenarios and thresholding the signal strengths values for inclusion into the distance metric in order to overcome it. Klepal et al. [21], then also reported in [22], tried to overcome the bias of missed access points by normalising the likelihood function with heuristic penalties for the number of missing beacons in either the fingerprint or the measurement. Our proposed solution to this problem is more rigorous in the sense that we model the probability of picking up a particular beacon explicitly in the likelihood observation function, which enables us to provide an explicit beacon pickup probability function. For this we chose to penalise these mismatches based on the missed energy and use Gibbs distributions (again justifiable by maximum entropy arguments) resulting in a distance metric between the fingerprint and the measurement which takes these mismatches into account. As we will see this will reduce the effect of the number of beacons on the results and allows comparing likelihoods with very different number of visible beacons.

We will finally compare the improvements we proposed with the more naive yet very common approach of a Gaussian likelihood observation function and show in what scenarios they differ and potentially improve the localisation results.

II. COMPAREING RSSI MEASUREMENTS

A. The naive approach

In this section we will start by describing the common approach to fingerprinting based localisation based on the Mahalanobis distance metric in log-energy space. We will use this to formalise the problem and introduce some notation that allows us to take a fresh look at the problem and address some of its issues.

We assume that in a given environment there are $K$ beacons (usually access points) installed in fixed locations. Each of these beacons emits a radio signal that we can measure in parts of the environment. We will collect these measurements of received signal strengths in a vector

$$s = \begin{pmatrix} s_1 & \cdots & s_K \end{pmatrix}^T$$

(1)

together with a boolean beacon-pickup indicator function

$$\tau : \{1, \ldots, K\} \rightarrow \{0, 1\}$$

(2)

that tells us whether a particular beacon has been picked up in this particular measurement or not. We will also assume an inverse covariance matrix $C^{-1}$ of these measurements to be known. In order to be consistent with the beacon-pickup indicator function the nullspace of this matrix has to reflect the infinite uncertainty about the signal strength values in case a beacon was not picked up, i.e. $\tau_i = 0 \Rightarrow C^{-1}_{ss} e_i = 0$ (with $e_i$ denoting a unit K-vector containing only zeros except in the $i$-th position). In order to further simplify the notation in the following we will also assume a single variance as well as no correlation between the measurements of the different beacons. In this case the inverse covariance matrix is simply given by

$$C^{-1}_{ss} = \sigma^{-2} \text{diag}(\tau)$$

(3)

Taking measurements in different locations and making assumptions on the propagation of the radio signals it is possible to create a location dependent map of received signal strengths, usually called a fingerprint, denoted here using the two location dependent functions

$$F[x] = \begin{pmatrix} F_1[x] & \cdots & F_K[x] \end{pmatrix}^T$$

(4)

for describing the received signal strengths in each location and

$$\phi[x] : \{1, \ldots, K\} \rightarrow \{0, 1\}$$

(5)

indicating the beacon-visibility in each location as before. This allows to easily describe the realistic case of environments where not every access point is visible everywhere. We will not go into detail on how to obtain this fingerprint in this paper but assume it to be given in the following.
The task of fingerprinting based localisation is now to determine the most likely position for a given signal strengths measurement based on comparing it to the previously obtained fingerprint map. We therefore need to maximise the probability of the position given the current measurement and the fingerprint map

$$\hat{x} = \text{argmax}_x p\{x|s, \tau, F, \phi\}$$

(6)

Applying Bayes’ theorem and noting that the location dependence is encoded in the fingerprint function this posterior location probability can be rephrased as follows

$$p\{x|s, \tau, F, \phi\} = \frac{p\{s, \tau|F[x], \phi[x]\} p\{x\}}{p\{s, \tau\}}$$

(7)

All components on the right hand side of this equation have to be considered. However, in this paper we will focus on the design of the likelihood probability density function $p\{s, \tau|F[x], \phi[x]\}$ only, which is a crucial sub-problem that has to be solved by every fingerprinting based localisation algorithm. Neither designing the actual maximisation nor discussing an appropriate prior evidence ratio $p\{x\}/p\{s, \tau\}$ (which is usually closely linked to the maximisation strategy used) is within the scope of this paper, although our final experiments will be carried out using a particle filter based approach employing a constant velocity motion model to address these two issues.

Looking closely into the likelihood function we observe that it can be split as follows

$$p\{s, \tau|F[x], \phi[x]\} = p\{s|F[x], \phi[x]\} p\{\tau|s, F[x], \phi[x]\}$$

(8)

We will look into both factors separately. The first part of this equation models the likelihood of received signal strengths measurements. It is common to assume a Gaussian distribution of the measured signal strengths around the fingerprint here, which is justifiable by arguments like maximum entropy or the law of large numbers acting on multiple additive error sources. Taking into account the possible singularity of the covariance matrix this measurement likelihood is then given by

$$p\{s|F[x], \phi[x]\} = \frac{\exp\left[-\frac{1}{2}(s - F)^T P^T C^{-1}_{ss} P(s - F)\right]}{\sqrt{(2\pi)^r \det [P^T C^{-1}_{ss} P]^{-1}}}$$

(9)

using the projection matrix $P = \text{diag}\{\phi[x]\}$ for considering only the subspace of visible beacons in the fingerprint, $r = \text{rank}[P^T C^{-1}_{ss} P]$ being the rank of the projected inverse covariance matrix, and $\det[P^T C^{-1}_{ss} P]$ denoting the product of its non-zero singular values. In essence this will use the Mahalanobis distance between the fingerprint and the measurement to evaluate the fit. Using the simplified covariance matrix from equation (3) it simplifies to

$$p\{s|F[x], \phi[x]\} = \frac{\exp\left[-\frac{1}{2} \sum_{i=1}^{K} \tau_i \phi_i[x] (s_i - F_i[x])^2 \right]}{\sqrt{2\pi \sigma^2 \sum_{i=1}^{K} \tau_i \phi_i[x]}}$$

(10)

This approach only permits zero-mean errors on the measurements. While this is a reasonable assumption if only a single device is used for fingerprinting and for subsequent localisation, differing antenna attenuation will potentially introduce a bias to the measurements, which will incur additional penalties for each measurement.

Most approaches do not model the second part of equation (8) at all, which is the beacon pickup probability. This means that these approaches implicitly assume a uniform distribution

$$p\{\tau|s, F[x], \phi[x]\} = \frac{1}{2\pi}$$

(11)

Doing so, however, will only consider the first part of the likelihood term and therefore incur a penalty for each beacon contained in both the fingerprint and the measurement. As we will discuss later this will favor areas covered by fewer access points, as the chance of incurring a penalty there is smaller. This effect is even greater if additional bias penalties per measurement as discussed above are present.

In the following we will introduce two improvements to this naive likelihood observation function addressing these issues by augmenting the classical approach accordingly in order to better cope with missed access points and differing antenna attenuation.

### B. Improving the likelihood observation function

1) Estimating the antenna attenuation: The first improvement on the likelihood observation function we propose is to take the differing antenna attenuation into account. This will allow to compensate for different devices being used for fingerprinting and for localisation. It also helps to compensate effects resulting from shielding the device differently, for instance by wearing it in the pocket. The basic idea is to estimate a common offset of the measured signal strengths compared to the signal strengths in the fingerprint for all received beacons assuming that both are not equal but offset by a common factor $\lambda$. The measured signal strengths and the fingerprints are then related as follows

$$\phi_i = \tau_i = 1 \Rightarrow s_i = F_i[x] + \lambda$$

(12)

This will align the log-energy measurements so that the measure becomes invariant against multiples of the received energy.

Using the simplified covariance matrix of equation (3) the best unbiased estimate of the offset factor (we will call it a factor because it acts on log-energies) is given by

$$\hat{\lambda} = \frac{\sum_{i=1}^{K} \tau_i \phi_i[x] (s_i - F_i[x])}{\sum_{i=1}^{K} \tau_i \phi_i[x]}$$

(13)

having a variance of

$$\sigma^2_{\hat{\lambda}} = \frac{\sigma^2}{\sum_{i=1}^{K} \tau_i \phi_i[x]}$$

(14)

Looking at the residual differences

$$\omega_i = s_i - F_i[x] - \hat{\lambda}$$

(15)
having the variance
\[ \sigma^2 = \frac{1}{\sum_{i=1}^{K} \tau_i \phi_i(x)} \]  \[ \sum_{i=1}^{K} \tau_i \phi_i(x) \sigma^2 \]  
(16)
we are now able to use this residual difference in the measurement likelihood function as follows
\[ p(s|F[x], \phi[x]) = \frac{\exp \left[ -\frac{1}{2} \sum_{i=1}^{K} \tau_i \phi_i(x) \omega^2 \right]}{\sqrt{2\pi \sigma^2} \sum_{i=1}^{K} \tau_i \phi_i(x)} \]  \[ \sum_{i=1}^{K} \tau_i \phi_i(x) \]  
(17)
This equation replaces the measurement likelihood function given in equation (10). The introduction of an additional degree of freedom into this function will of course broaden the likelihood and make it less discriminative for a single device. However, as we will show later we found that this approach allows to generalize better towards different devices during the localisation phase.

2) Dealing with missing beacons: The second and much more important improvement on the likelihood observation function we propose is to model the probability of missing beacons in the measurement (or equivalently areas not covered by certain access points in the fingerprint) explicitly. The idea is to penalize mismatches based on the expected energy being missed. This missed energy in the case of a mismatch is given by
\[ c_i[\tau_i] = \begin{cases} 0 & \text{if } \tau_i = \phi_i(x) \\ \alpha \tau_i \phi_i(x) & \text{if } \tau_i = 0 \land \phi_i(x) = 1 \\ \alpha \tau_i^2 & \text{if } \tau_i = 1 \land \phi_i(x) = 0 \end{cases} \]  
(18)
We will therefore use the corresponding maximum entropy Gibbs distribution for this energy as pickup probability
\[ p(\tau, s, F[x], \phi[x]) = \prod_{i=1}^{K} \frac{\exp \left[ -\beta c_i[\tau_i] \right]}{\exp \left[ -\beta c_i[0] \right] + \exp \left[ -\beta \tau_i c_i[1] \right]} \]  
(19)
instead of the uniform distribution given in equation (11), which is implicitly assumed by approaches taking only the differences of matched access points into account.

The energy unit \( \alpha \) and the temperature parameter \( \beta \) allow to control the offset and slope of the likelihood function. Figure 1 shows the likelihood observation function for a single beacon for all four pickup scenarios. As you can see there the penalty in case of a mismatch depends on the energy being missed as expected. Also note how the beacon pickup probability shapes the likelihood observation function in case of matches, making high energy matches more probable than low energy matches.

In the following we will show the properties of these two improvements to the likelihood observation function for a real example.

III. Evaluation

Having proposed two extensions to the commonly used Gaussian likelihood observation function for comparing signal strengths measurements to previously recorded fingerprints we will now demonstrate how these two improvements affect the shape of the likelihood function and thereby improve the localisation accuracy.

Our experiments were carried out in our office building depicted in figure 2. This figure also shows the area we fingerprinted with the number of beacons visible in each single location. The total number of access points is nine with at least three and at most eight of them being visible in every single location.

For the evaluation we used a different device and walked along a pre-defined path through the building from left to right as seen in figure 3. This figure also shows the location of

Fig. 1. The proposed likelihood observation function comparing a single beacon’s measurement to the fingerprint in case both measurements are present (top left), the beacon was missed in the measurement but is present in the fingerprint (top right), a beacon is not present in the fingerprint although it is measured (bottom left), and the beacon occurs neither in the measurement nor in the fingerprint (bottom right).

Fig. 2. The number if access points covered by our fingerprint. As you can see everything is covered by at least three beacons with up to eight beacons covering the central corridor. Obviously localisation performance should be close to optimal in this kind of environment.

Fig. 3. The path we walked in our experiments (green line) through our lab. The beacons we use are shown as red circles. The size of the building is approximately 60m from left to right. Also the residual errors for the estimated path (blue line) based on the likelihood observation function with both improvements are shown.
the access points as well as the residual errors for our best performing algorithm. The actual localisation was done using a particle filter (see figure 4).

Both the fingerprinting as well as the optimisation, i.e. the particle filter in this case, are not within the scope of this paper, as the proposed likelihood observation functions will be applicable to any other implementation of these as well. We will therefore now look into the different likelihood observation functions along this path as shown in figures 5 to 9.

It can be seen that the biggest problem of the naive purely Gaussian approach is the multi-modality of the likelihood observation function. This is caused by comparing likelihood values based on error metrics taken in different dimensions. Simply put each additional dimension, i.e. matched access point, can only incur a penalty; an effect that is amplified if the fingerprint value is not the actual mean of the underlying distribution. In this case areas covered by a smaller number of access points are favored, as not so much can go wrong there. The mismatch penalty terms from the Gibbs distribution counteract this effect in a rigorous way and thereby reduce the multi-modality of the likelihood observation function.

The second, however not so prominent observation is regarding the estimation of the device antenna attenuation in each step. While it of course broadens the likelihood observation function a bit the overall probability mass close to the actual position does increase in case of different devices. Because the distance to the fingerprint can only be reduced by this, it also reduces the effect of favoring areas of fewer access points as discussed in the previous paragraph.

To quantify the differences between these three likelihood observation functions we implemented a particle filter and measured the residual position error along the path. The cumulative histograms of these errors for all three approaches are shown in figure 10. While the attenuation estimation slightly improves the performance compared to the purely Gaussian likelihood observation function, both of these suffered from the multi-modality occurring at both ends of the building resulting from the likelihood observation function "hiding away" in areas covered by fewer access points, which is a problem commonly observed in fingerprinting based localisation systems. However the likelihood observation function explicitly addressing this issue by modeling the pickup probabilities as Gibbs distributions did not suffer from this so that the overall localisation performance was significantly improved. Another observable effect, however not so prominent, is the slightly increased accuracy as more information, i.e. the pickup of a beacon itself, is used in determining the value of the likelihood observation function. This is in contrast to the naive approach, where not picking up a beacon where it should have been picked up does not influence the value of the likelihood observation function other than potentially incurring less penalty.

IV. CONCLUSION

In this paper we took a fresh look at the likelihood observation function used in fingerprinting based localisation. We proposed two improvements on the commonly used Gaussian approach directly addressing two well-known issues. The first issue was the differing antenna attenuation between different devices, which we proposed to address by explicitly estimating a global offset thereby considering only relative differences between vectors of received signal strength measurements. The second and much more important issue was dealing with environments where not every beacon is visible everywhere,
Fig. 6. Second of the series of five consecutive snapshots as described in figure 5. The same effect regarding the multi-modality is still visible, however it is less prominent now so that all three approaches are now able to perform similarly.

Fig. 7. Third of the series of five consecutive snapshots as described in figure 5. Again all three approaches perform quite similar with the attenuation estimation spreading out the likelihood observation function a bit due to the introduction of an additional degree of freedom through the attenuation invariance. Interestingly modeling the pickup probabilities counteracts this effect a bit as more information is used compensating for this.

Fig. 8. Fourth of the series of five consecutive snapshots as described in figure 5. The multi-modality occurs again due to the symmetry of the distance metric used. In this open space area, however, the effect is much stronger now making the localisation algorithm based on the top two approaches perform very poorly.

Fig. 9. Fifth and final of the series of five consecutive snapshots as described in figure 5. Again the multi-modality is the problem being in the same situation as in figure 5, but now with the particles already concentrated on the wrong mode the particle filter cannot easily recover.
which we proposed to address by explicitly modeling the pickup probability of Beacons using Gibbs distributions. Both improvements of the likelihood observation function were shown to increase the performance of our particle filter based localisation system, however as a suitable distance metric between signal strength measurements is at the core of every fingerprinting based localisation system our findings are likely to improve all systems currently relying on a purely Gaussian approach.

ACKNOWLEDGMENT

This work has been supported by Enterprise Ireland though grant CF/2010/042 and CP/2011/0203.

REFERENCES


Fig. 10. Cumulative residual position error using the different likelihood observation functions. The pure Gaussian (red) as well as the Gaussian with attenuation estimation (green) did not allow a reliable position estimate at both ends of our building, however the attenuation estimation improved the results slightly when different devices are used for fingerprinting and for localisation. Modelling the pickup probability explicitly (blue) did not suffer from this deficiency in our experiments, hence the lowest residual error amongst the three approaches.