Deconvolution-based indoor localization with WLAN signals and unknown access point locations

Shweta Shrestha, Jukka Talvitie, Elena Simona Lohan, *Senior Member, IEEE*

***Abstract*—In this paper, the problem of Received Signal Strength (RSS)-based WLAN positioning is newly formulated as a deconvolution problem and three deconvolution methods (namely Least Squares, Weighted Least Squares and Minimum Mean Square Error) are investigated with several RSS path loss models. The deconvolution approaches are compared with the fingerprinting approach in terms of performance and complexity. The main advantage of the deconvolution-based approaches versus the fingerprinting methods is the significant reduction in the size of the training database that need to be stored at the server side (and transferred to the mobile device) for the WLAN-based positioning. We will show that the deconvolution based estimation can decrease of the order of ten times the size of the training database, while still being able to achieve comparable root mean square errors in the distance estimation.**

***Index Terms*—Access Point (AP) estimation, Indoor localization, Least Squares (LS), Minimum Mean Square Error (MMSE), Weighted Least Squares (WLS), WLAN positioning**

I. Introduction and motivation

INDOOR localization is currently gaining more and more interest in both academic and industrial worlds, motivated by the fact that accurate three-dimensional (3D) indoor localization could open a myriad of new Location Based Services (LBS) and location-based business models [2], [3]. It is well-know that the Global Navigation Satellite Systems have their performance severely limited in indoor scenarios, due to the very weak received signal powers and to multipath propagation [12], [13] . Alternatively, personal and local area networks-based localization, such as Radio Frequency Identification tags (RF-ID), Ultra Wide-Band (UWB) or WLAN-based solutions are more suitable to indoor short-range communications. In particular, the WLAN-based localization using the signal Received Signal Strength (RSS) has the advantage of re-using existing structures with no additional cost associated with infrastructure deployments and being based on software-based solutions at the receiver side. All RSS-based localization solutions involve two stages:

1) A *training stage*: in here, information about the indoor environment is collected, typically in the shape of measurement points coordinates and received signal strengths. The measurement points are also called fingerprints, and their 3D coordinates $(x_i, y_i, z_i), i = 1, \ldots, N_{FP}$, need to be measured with the help of indoor maps in an off-line initial phase. N_{FP} is the total number of fingerprints measured per building. The measured RSS for each of those fingerprints depend on the heard Access Points (AP) in that particular location and are denoted here via $P_{i,ap}, i = 1, \ldots, N_{FP}, ap = 1, \ldots, N_{AP}$, with N_{AP} being the total number of AP per building. In the training phase, there are also two alternatives: i) we can either store the values $x_i, y_i, z_i, P_{i,ap}$ for each fingerprint and each heard AP (this is the typical *fingerprinting FP* approach) or ii) we can create a *probabilistic* model, for example *path-loss PL* based, and store only a reduced number of parameters (for example, few parameters per AP). In our paper we investigate the probabilistic PL approaches, as explained in Section III and we will compare them with FP approaches.

2) An *estimation stage*: this involves real-time processing, where the Mobile Station (MS) is estimating its position based on the data stored in the training phase and on its current received signal strength from various heard APs $P^{(MS)}_{ap_{heard}}, ap_{heard} \in \mathcal{A}_{heard}$ where $\mathcal{A}_{heard}$ is a sub-set $\mathcal{A} = [1 : N_{AP}]$ of all the AP in the building and it has N_{heard} elements. Typically, this is done via triangulation approaches, where data fusion algorithms are applied in order to combine the information coming from various APs. Differently from the classical triangulation methods, where the emitter position is known, in here the AP positions within a building are typically unknown and need to be estimated beforehand, in the training stage.

While FP approaches have been widely studied [2], [3], [4], [6], the PL approaches have only recently gained attention [1], [7]. Also, the usual PL model used in PL approaches is the traditional one-slope PL model [1], [8], [9]. Few authors investigated multi-slopes PL models, but in different contexts than indoor wireless localization [10], [11]. Moreover, the AP/transmitter locations are many times assumed known [5], [7], which is an assumption not generally valid for large-scale mobile indoor localization solutions based on WiFi. In this paper we introduce an innovative approach, based on deconvolution ideas, for probabilistic PL localization based on WiFi signals. Our approach is completely different from the approach [1], which was based on iterative Gauss Newton methods to solve the non-linearities in the PL model. In here, the non-linearities are tackled out by assuming various AP location and choosing the locations the minimize the deconvolution mean square errors. Moreover, we investigate the multi-slope path loss models in the context of WLAN indoor wireless localization and we show that a two-slope model may give better results than the one-slope models in

The authors are with the Department of Electronics and Communication Engineering, Tampere University of Technology, Finland, e-mails: {shweta.shrestha, jukka.talvitie, elena-simona.lohan}@tut.fi.

buildings with many open spaces between floors, such as shopping malls.

II. Problem formulation

A. Fingerprinting (FP) approaches

In the traditional fingerprinting approaches, used here as benchmarks, the given data are:

- The observations $x_i, y_i, z_i, P_{i,ap}, i = 1, \ldots, N_F, ap = 1, \ldots, N_{AP}$, available from the training database
- The received signal strengths measured in the unknown location at the mobile: $P^{(MS)}_{ap_{heard}}, ap_{heard} \in \mathcal{A}_{heard}$

The unknown MS position (x_{MS}, y_{MS}, z_{MS}) is estimated via pattern matching of $P^{(MS)}_{ap_{heard}}$ into the fingerprints database $P_{i,ap}$ and choosing the coordinates of the neareast neighbour point or an average over N_{neigh} nearest neighours:

$$[\hat{x}_{MS}, \hat{y}_{MS}, \hat{z}_{MS}] = \frac{1}{N_{neigh}}([\sum_{\hat{i}=1}^{N_{neigh}} x_{\hat{i}}, \sum_{\hat{i}=1}^{N_{neigh}} y_{\hat{i}}, \sum_{\hat{i}=1}^{N_{neigh}} y_{\hat{i}}]) \tag{1}$$

where index $\hat{i}$ of the nearest neigbours is found by minimizing (or maximizing) a certain cost function. The most typical cost functions ad their associated optimization criterion are:

1) minimizing the power differences between the observed RSS and the database RSS:

$$\hat{i} = argmin_i \frac{1}{N_{heard}} \sum_{ap_{heard} \in \mathcal{A}_{heard}} |P^{(MS)}_{ap_{heard}} - P_{i,ap_{heard}}|^2 \tag{2}$$

This is in fact the k-nearest neighbour method (NNM) that is currently considered the state-of-art in RSS approaches [3].

2) maximizing the number of commonly heard APs at the considered fingerprint and at the mobile side. This is also known as the test rank based method [4].

Based on our measurement data, we have observed that a combination of the two criteria gives the best result (that is, maximizing the number of commonly heard APs at the mobile and at the reference point, and, if there are several grid points where the same number of APs is heard, taking the one with minimum power difference between the observed RSS and the database RSS). Thus, we will use this comined approach in our measurement analysis from Section IV.

In the fingerprinting approaches, the positions of the APs are typically not known and not needed in the estimation process. The drawback of such approaches is the need of large databases in order to store the measured information $x_i, y_i, z_i, P_{i,ap}, i = 1, \ldots, N_F, ap = 1, \ldots, N_{AP}$. For example, for a multi-storey building where 1000 fingerprints were taken, and in each fingerprint we heard between 7 and 20 access points, we need to store in the fingerprinting database between 10000 and 23000 parameters. This is a prohibitive number of parameters if we aim at low-cost localization solutions, because these parameters would need to be transferred to the mobile when a positioning request is to be done.

B. Probabilistic path-loss PL approaches

In order to decrease the size of the database, instead of the original fingerprints, we could store only few parameters per AP. Let's assume that each AP can be characterized by a vector Θ_{AP} with M parameters (see (5)). The measured RSSs (in logarithmic scale) are non-linear functions $f(\cdot)$ of these parameters and of the Euclidian distance $d_{i,ap} = \sqrt{(x_i - x_{ap})^2 + (y_i - y_{ap})^2 + (z_i - z_{ap})^2}$ between the ap-th AP and the i-th measurement point:

$$P_{i,ap} = f(d_{i,ap}, \Theta_{AP}) + \eta_{i,ap}, \tag{3}$$

where $\eta_{i,ap}$ is a noise factor, typically assumed Gaussian distributed, of zero mean and standard deviation σ. The noise is typically due to shadowing, fading and measurements errors: $\eta_{i,ap} \propto \mathcal{N}(0, \sigma^2)$. In the absence of additional prior information, it may be assumed that the noise variance σ^2 is constant per building.

The **traditional path-loss model** is based on free space wave propagation [1] and involves two modeling parameters per AP: $\Theta_{AP} = [P_{T_{ap}} \; n_{ap}]$, where $P_{T_{ap}}$ is the ap-th AP transmit power and n_{ap} is the path-loss coefficient of the ap-th AP. Those two parameters are related to the RSS via:

$$P_{i,ap} = P_{T_{ap}} - 10 n_{ap} log_{10} d_{i,ap} + \eta_{i,ap}, \tag{4}$$

Alternatively, we can extend the above model to $M \geq 2$ parameters per AP and form a **multi-slope path-loss model** :

$$\Theta_{AP} = [P_{T_{ap}} \; n^{(1)}_{ap} \; n^{(2)}_{ap} \; \ldots n^{(M-1)}_{ap}], \tag{5}$$

by assuming that the path loss coefficients $n^{(m)}_{ap}$ varies with the distance between the transmitter and the receiver. In this case, we can model the RSS via:

$$P_{i,ap} = P_{T_{ap}} - \sum_{m=1}^{M-1} 10 w^{(m)}_{ap} n^{(m)}_{ap} log_{10} d_{i,ap} + \eta_{i,ap}, \tag{6}$$

where $w^{(m)}_{ap}$ is a distance-dependent flag which takes 0 or 1 values:

$$w^{(m)}_{ap} = \begin{cases} 1 & \text{if } \gamma_m \leq d_{i,ap} < \gamma_{m+1} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

and $\gamma_m, m = 1, \ldots, M-1$ are some distance thresholds between 0 and maximum hearable distance that define the changes in the path-loss coefficients. It can be straighforwardly seen that $\sum_{m=1}^{M-1} w^{(m)}_{ap} = 1, \; \forall ap$.

Two additional path loss models are derived from (4) and (6) by adding an additional floor loss parameter L_{pf} that could model the ceiling and walls in between floors:

$$P_{i,ap} = P_{T_{ap}} - 10 n_{ap} log_{10} d_{i,ap} - \xi_{i,ap} L_{pf} + \eta_{i,ap}, \tag{8}$$

and, respectively,

$$P_{i,ap} = P_{T_{ap}} - \sum_{m=1}^{M-1} 10 w^{(m)}_{ap} n^{(m)}_{ap} log_{10} d_{i,ap} - \xi_{i,ap} L_{pf} + \eta_{i,ap}, \tag{9}$$

where $\xi_{i,ap}$ is a factor showing the number of floors between the ap-th AP and the i-th grid point: 0 if the vertical distance between them is less than half floor height, 1 if the vertical distance between them is between half floor height and 1.5

times the floor height, and so on. $\xi_{i,ap}$ can be easily estimated based on the estimated AP coordinates, as it will be explained in Section III-A, and L_{pf} is a constant parameter, building specific that will take certain a priori value, then it can be enhanced through successive trials. In all these probabilistic PL approaches, we have to solve a two-step estimation problem:

1) In the training phase, being given the database $x_i, y_i, z_i, P_{i,ap}$, estimate the AP positions $[\hat{x}_{ap}, \hat{y}_{ap}, \hat{z}_{ap}]$ and the Θ_{AP} AP parameters (vector of length M).
2) In the mobile positioning phase, being given the database with $M+3$ parameters stored per AP: $[\hat{x}_{ap}, \hat{y}_{ap}, \hat{z}_{ap}, \Theta_{AP}]$, $ap = 1, \dots, N_{ap}$ and the measured RSS at the mobile $P^{(MS)}_{ap_{heard}}$, estimate the unknown MS location $[\hat{x}_{MS}, \hat{y}_{MS}, \hat{z}_{MS}]$.

In such approaches, the amount of stored data can be drastically reduced. For example, with 200 AP per building and a 2-parameter path-loss model, we only need to store 5 parameters per AP, namely $[\hat{x}_{ap}, \hat{y}_{ap}, \hat{z}_{ap}, P_{T_{ap}}, n_{ap}]$. Thus, in this example, we only need to store a total of 1000 parameters, achieving thus a reduction in the database of order of 10 or higher compared to the similar example given in the fingerprinting approach in Section II-A.

In the folowing section we formulate the above estimation problem as a deconvolution problem and we analyze three different deconvolution methods for the estimation of the unknown AP parameters and of the unknown MS position.

III. Deconvolution estimators

A. Access Point position estimation and AP parameters estimation (off-line training stage)

The PL models of (4) to (9) can be written in matricial form as

$$\mathbf{P_{ap}} = \mathbf{H_{ap}} \mathbf{\Theta_{ap}^T} + \mathbf{n} \tag{10}$$

where $\Theta_{ap} = [P_{T_{ap}}, n^{(1)}_{ap}, \dots, n^{(M-1)}_{ap}]$ are the unknown parameters per AP excepting the AP coordinates (namely the AP apparent transmit power and the path loss cofficients associated with different distances from AP), $\mathbf{P}_{ap} = [P_{1,ap}\ P_{2,ap}\ \dots\ P_{N_F,ap}]^T$ is the vector with power fingerprints in logarithmic scale coming from ap-th access point (corrected with a floor loss factor associated with each AP in case the path loss models from (8) or (9) are used), ^{T} is the transpose operator, $\mathbf{n}$ is a Gaussian distributed $N_F \times 1$ vector and

$$\mathbf{H}_{ap} = \begin{bmatrix} 1 & -10 w_1 log_{10} d_{1,ap} & \cdots - 10 w_{M-1} log_{10} d_{1,ap} \\ \dots & & \dots \\ 1 & -10 w_1 log_{10} d_{N_F,ap} & \cdots - 10 w_{M-1} log_{10} d_{N_F,ap} \end{bmatrix} \tag{11}$$

In (10) both $\mathbf{H}_{ap}$ and Θ_{ap} are unknown, thus we introduce the following solution in order to have a linear deconvolution problem:

1) For each measured fingerprint $[x_i, y_i, z_i]$, compute $\mathbf{H}_i$, where the unknown ap-th AP position in $\mathbf{H}_{ap}$ is replaced with $[x_i, y_i, z_i]$ coordinates.
2) Compute a tentative $\hat{\Theta}^T_{i,ap}$ that is dependent on the assumed AP location (i.e., the fingerprint position i) via one of the following deconvolution methods:
   - Least Squares (LS):
   $$\hat{\Theta}^T_{i,ap,LS} = (\mathbf{H}_i^T \mathbf{H}_i)^{-1} \mathbf{H}_i^T \mathbf{P_{ap}} \tag{12}$$
   - Weighted Least Squares (WLS)
   $$\hat{\Theta}^T_{i,ap,WLS} = (\mathbf{H}_i^T \mathbf{W}_{ap} \mathbf{H}_i)^{-1} \mathbf{H}_i^T \mathbf{W}_{ap} \mathbf{P_{ap}} \tag{13}$$
   where $\mathbf{W}_{ap} = \text{diag}(10^{-\mathbf{P_{ap}}/\mathbf{10}})$ is a diagonal weight matrix stating that we should put less weight on those fingerprints where the AP is heard at lower powers.
   - Minimum Mean Square Error (MMSE)
   $$\hat{\Theta}^T_{i,ap,MMSE} = (\mathbf{H}_i^T \mathbf{H}_i + \mathbf{I}_M/\sigma^2)^{-1} \mathbf{H}_i^T \mathbf{P_{ap}} \tag{14}$$

   where $\mathbf{I}_M$ is the identity matrix of size $M \times M$ and σ^2 is the assumed shadowing variance (in the absence of additional information, we set it to 5 dB).
3) Compute the expected observation vector $\hat{\mathbf{P}}_{i \to ap}$ if ap-th AP were situated in point i:
$$\hat{\mathbf{P}}_{i \to ap} = \mathbf{H}_i \hat{\Theta}^T_{i,ap,method} \tag{15}$$
where $method$ is one of the three deconvolution methods LS, WLS or MMSE.
4) Choose the **estimated AP position** $[\hat{x}_{ap}, \hat{y}_{ap}, \hat{z}_{ap}]$ as the fingerprint i that minimizes the mean square error between $\hat{\mathbf{P}}_{i \to ap}$ and $\mathbf{P_{ap}}$. Alternatively, we can take the average over N_{neigh} points that offer the lowest mean square error $||\hat{\mathbf{P}}_{i \to ap} - \mathbf{P}_{ap}||_2^2$, $||\cdot||_2$ being the Euclidian norm. We note that our approach of estimating the APs location will always estimate the APs as being within the building. In practice, some APs could be located outside the building (but close to it). However, we noticed that the apparent AP location (as done before) is enough to be used, because the apparent AP transmit power is scaled (implicitly) according to this apparent AP location, and thus the model will still capture the propagation effects of the wireless channel.
5) With the estimated AP position, re-compute via LS, WLS or MMSE the AP parameters, based on equations (12) to (14).

We remark that these steps are done off-line, at the server side, and thus their complexity (that depends on the size of matrices $\mathbf{H}_i$) is fully manageable.

B. MS position estimation (real-time estimation stage)

In this stage, we no longer have the positions of the original fingerprints stored in the database (the database size was reduced to only few parameters per AP). Thus, we first generate a synthetic grid per building. This grid can be either fixed for the whole building (an example with a two-floor building is shown in Fig. 1) or built around each access points. If a fixed grid per building is employed, in the fingerprinting database we also need to store some corner coordinate of the building, plus the maximum length, width and height of the building. However, these are only 6 additional parameters and do not increase much the size of the database, as it will be shown in the comparative figures from Table VI.

Thus, the estimation problem can be written as: being given the AP database $\{\hat{x}_{ap}, \hat{y}_{ap}, \hat{z}_{ap}, \Theta_{AP}\}$, $ap = 1, \dots, N_{ap}$, and

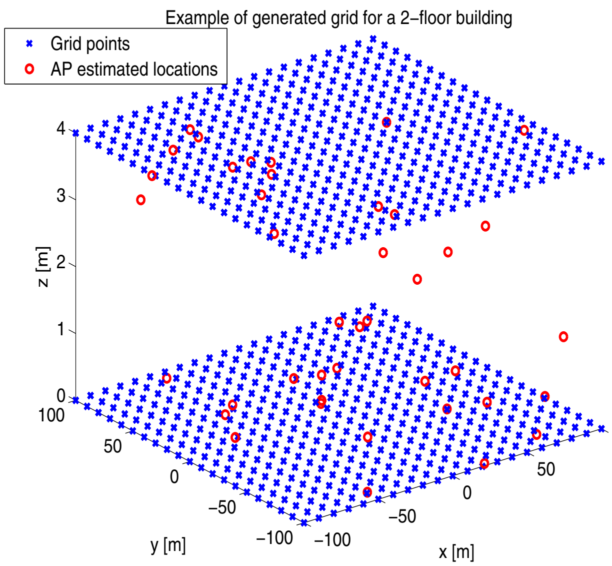

Fig. 1. Example of a generated grid for a 2-floor building .

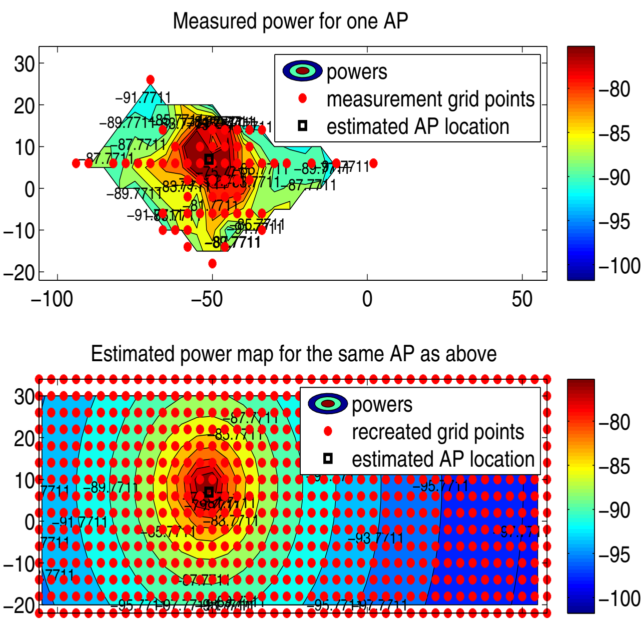

Fig. 2. Example of a measured and re-created power map coming from one AP in a university building; 5-m measurement grid .

being given the measured RSS at the mobile $P^{(MS)}_{ap_{heard}}$, the prolem is to estimate the unknown MS location $[\hat{x}_{MS}, \hat{y}_{MS}, \hat{z}_{MS}]$. This is done in the following steps

1) Create a synthetic grid $[x_i, y_i, z_i], i = 1, \ldots, N$ per building. This can be created for example based on maximum and mimimum coordinates stored for that building $[x_{min}, x_{max}, y_{min}, y_{max}, z_{min}, z_{max}]$, with certain grid steps in horizontal and vertical directions (and possibly allowing for a larger grid to capture also the MS possibly located outside the building, but close to it). Horizontal grid sizes $(\Delta g)_h$ can be for example between 1 and 10 meters, while vertical grid size can be equal or smaller to the building floor height (again stored from the measurement phase). For example, x_i values will span the set $[x_{min} : (\Delta g)_h : x_{max}]$ and so on. The grid step will determine the size of the synthetic grid N. Smaller grid steps would mean a larger size for the synthetic grid, and thus a slower data processing in the estimation phase.
2) For each tentative MS position estimate $[\hat{x}_i, \hat{y}_i, \hat{z}_i]$, compute the power difference (or cost function $J(\cdot)$) between the observed RSS at the mobile and the expected power based on the AP database:

$$J(\hat{x}_i, \hat{y}_i, \hat{z}_i) = \Bigg(\sum_{ap_{heard} \in \mathcal{A}_{heard}} |P^{(MS)}_{ap_{heard}} - P_{T_{ap}} - \sum_{m=1}^{M-1} 10 n^{(m)}_{ap} w^{(m)}_{ap} log_{10} \hat{d}_{i,ap}|^2 \Bigg) \frac{1}{N_{heard}} \quad (16)$$

where

$$\hat{d}_{i,ap} = \sqrt{(\hat{x}_i - \hat{x}_{ap})^2 + (\hat{y}_i - \hat{y}_{ap})^2 + (\hat{z}_i - \hat{z}_{ap})^2} \quad (17)$$

and N_{heard} is the number of heard AP by the MS or the cardinal of $\mathcal{A}_{heard}$. Note that the cost function can be improved if we consider the set $\mathcal{A}_{heard}$ as containing only those heard APs that are strong enough. A condition to test the AP's strength is for example whether the received signal strength corresponding to that particular AP is higher or equal to the median RSS values minus 3 dB. Another condition to test the AP's strength is to consider only those APs whose corresponding RSS at the MS is at least -85 dB. Both conditions have been implemented in our simulations and they showed slightly better results than when considering all heard APs.
3) Choose the estimated MS position as the value that minimizes the cost function $J(\cdot)$ or as an average over N_{neigh} neigbour points that offer the minimum values of $J(\cdot)$.

IV. Comparison of path-loss deconvolution approaches with fingerprinting

The deconvolution-based position estimators are compared here with fingerprinting approaches with measurement data coming from 4 different buildings in Tampere, Finland. The buildings and their main characteristics are described in Table I. After the fingerprinting phase, several tracks per building were measured during different days and those tracks were used for our mobile positioning analysis. The tracks were taken over various floors of each building, in such a way to cover all floors. All measurements were performed with a Windows tablet with WLAN receiver, where the user was selecting his/her own position via the touch screen, using the available building map on the tablet. The large Mall has a floor surface of about 22000 m^2, while the small Mall has a floor surface of about 18000 m^2 (small and large refering here to the building heights, rather than to the floor surface). The university buildings have a surface of about 9000 m^2 (univ. building 1) and 14000 m^2 (univ. building 2).

The results with the one-slope path loss models with and without floor loss corrections are shown in Tables II and, respectively, III for a 5m grid step. 4 nearest neighbours were used in the AP parameter estimation as well as in the fingerpriting case. The last column shows the results based on fingerprinting. The floor was determined as the nearest floor (Euclidian distance) to the estimated mobile position (the floor height is assumed known from the training phase and equal for all floors). The floor detection probability P_d is computed as the probability of estimating the exact floor, while the distance RMSE is taken over all estimated points (not only over those estimated at the correct floor).

We can see from Table II that for both university buildings, the performance of deconvolution approaches is similar with the one from FP. In the University Building 2, where deconvolution approaches outperform the FP approaches, the reason can be due to a lower number of fingerprints in that building, that were also taken with some higher measurement errors (due to user approximations) than in the case of University Building 1. Also, for the small Mall, the deconvolution and FP results are comparable, though the distance RMSE is higher in all cases compared to the university buildings (this is due to a lower number of fingerprints in that particular building, as well as to a much lower number of AP per building). In the 6-floor large mall of last row, while the distance RMSE is still comparable in the different approaches, the floor detection fails more often in the deconvolution cases than in the FP case. This is certainly due to the many open spaces in that particular mall, which make the distinction between floors much more challenging. Enhanced algorithms for floor detection based on PL approaches are currently under study. The results in Table III show that, while introducing a floor loss coefficient in the PL models may improve the floor detection probability in some cases (e.g., in University Building 1), in most cases such an additional parameter only deteriorates the results. Also, from Tables II and III, we can see that the performance of the three deconvolution approaches is very similar, with MMSE outperforming slightly the other two.

The results based on the PL models with multi-slope (i.e., two-slope) coefficients from (6) and (9) are shown in Tables IV and V, respectively. We notice that, for the university buildings, one-slope and two-slope models offer comparable results (e.g., comparing Table II with Table IV, and Table III Table V). The two-slope model works better than single-slope model for buildings with many open spaces, such as the two studied malls. Again, the models without the floor loss factor work better. The results point out towards the model of eq. (9) as being the best model to capture the indoor WLAN environment.

Scenario	Number of floors	Open spaces between floors	N_{AP}	Total nr. of points for all user tracks per building	Floor height [m]
Univ. Building 1 (628 fingerprints)	4	Yes, few	309	490 points in 9 tracks	3.7
Univ. Building 2 (437 fingerprints)	3	Yes, few	354	117 points in 6 tracks	3.5
Small Mall (318 fingerprints)	3	Yes, many	69	215 points in 3 tracks	5
Large Mall (1062 fingerprints)	6	Yes, many	326	205 points in 11 tracks	5

TABLE I
BUILDING DESCRIPTION FOR THE MEASUREMENT DATA.

Average values	LS RMSE/P_d	MMSE RMSE/P_d	WLS RMSE/P_d	FP RMSE/P_d
Univ. Building 1	9.21 m/ 77.46%	9.18 m/ 77.77%	9.31 m/ 77.26%	7.24 m/ 86.15%
Univ. Building 2	8.64 m/ 92.99%	8.68 m/ 92.99%	8.89 m/ 91.37%	12.48 m/ 77.48%
Small Mall	24.02 m/ 89.07%	23.53 m/ 88.72%	25.98 m/ 83.81%	22.25 m/ 95.96%
Large Mall	16.34 m/ 47.37%	15.47 m/ 50.21%	16.43 m/ 48.80%	13.70 m/ 83.66%

TABLE II
DISTANCE RMSE IN METERS AND FLOOR DETECTION PROBABILITY P_d, **traditional path loss model** OF (4).

Average values	LS RMSE/P_d	MMSE RMSE/P_d	WLS RMSE/P_d	FP RMSE/P_d
Univ. Building 1	9.77 m/ 83.99%	9.53 m/ 85.33%	9.92 m/ 82.66%	7.24 m/ 86.15%
Univ. Building 2	9.36 m/ 85.53%	9.29 m/ 85.53%	9.89 m/ 84.42%	12.48 m/ 77.48%
Small Mall	27.44 m/ 66.62%	26.41 m/ 64.86%	30.76 m/ 64.86%	22.25 m/ 95.96%
Large Mall	21.12 m/ 19.84%	19.42 m/ 24.83%	20.90 m/ 23.64%	13.70 m/ 83.66%

TABLE III
DISTANCE RMSE IN METERS AND FLOOR DETECTION PROBABILITY P_d, **one-slope path loss model** WITH NON-ZERO FLOOR LOSS OF (8).

Average values	LS RMSE/P_d	MMSE RMSE/P_d	WLS RMSE/P_d	FP RMSE/P_d
Univ. Building 1	9.19 m/ 74.40%	9.16 m/ 74.21%	9.27 m/ 73.17%	7.24 m/ 86.15%
Univ. Building 2	8.60 m/ 89.77%	8.60 m/ 89.69%	8.96 m/ 89.15%	12.48 m/ 77.48%
Small Mall	23.50 m/ 90.48%	23.60 m/ 88.72%	25.29 m/ 89.77%	22.25 m/ 95.96%
Large Mall	16.85 m/ 47.85%	15.51 m/ 50.35%	16.84 m/ 47.15%	13.70 m/ 83.66%

TABLE IV
DISTANCE RMSE IN METERS AND FLOOR DETECTION PROBABILITY P_d, **two-slope path loss model** WITH ZERO FLOOR LOSS, SEE EQ. (6).

Average values	LS RMSE/P_d	MMSE RMSE/P_d	WLS RMSE/P_d	FP RMSE/P_d
Univ. Building 1	9.99 m/ 81.82%	9.62 m/ 83.80%	10.13 m/ 78.23%	7.24 m/ 86.15%
Univ. Building 2	9.54 m/ 84.50%	9.35 m/ 84.50%	10.22 m/ 84.50%	12.48 m/ 77.48%
Small Mall	25.66 m/ 72.58%	25.20 m/ 81.70%	29.20 m/ 73.06%	22.25 m/ 95.96%
Large Mall	24.17 m/ 19.02%	19.83 m/ 23.58%	22.75 m/ 24.21%	13.70 m/ 83.66%

TABLE V
DISTANCE RMSE IN METERS AND FLOOR DETECTION PROBABILITY P_d, **two-slope path loss model** WITH NON-ZERO FLOOR LOSS, SEE EQ. (9).

In terms of training database sizes (i.e., the database size is equal to the number of stored parameters per building), the comparison between FP and deconvolution approaches is

Scenario	Database size in deconvolution approaches (one-slope models)	Database size in deconvolution approaches (two-slope models)	Database size FP approaches FP approaches	Database reduction factor (one-slope models)	Database reduction factor (two-slope models)
University Building 1	1551	1860	21921	13.73	11.40
University Building 2	1776	2130	22084	12.43	10.40
Small Mall	351	420	2827	8.05	6.73
Large Mall	1636	1962	21989	13.44	11.2

TABLE VI
DATABASE SIZES IN DECONVOLUTION APPROACHES VERSUS FP APPROACHES FOR A 5 M GRID SIZE.

shown in Table VI for a 5-m grid step that was used in our analysis. For smaller grid steps, the reduction factor due to deconvolution approaches is even higher. Clearly, we reach about 5-10 times reduction in the database size, and, as seen from Tables II (traditional PL model), the performance of deconvolution approaches is comparable with that one of FP approaches in most of the cases. The only exception occurs for the large mall case (6 floors), when the floor detection is significantly poorer in LS, MMSE and WLS cases than in FP approaches (despite the fact the the distance RMSE is still comparable, which means that we are typically one floor wrong in the estimation).

V. CONCLUSION

In this paper we proposed probabilistic/path-loss based approaches for WLAN localization when AP location is now known. Our probabilitic approaches are based on deconvolution algorithms, by formulating the estimation problem as a deconvolution problem and by tackling out the non-linearity of the problem in minimum mean square error approach. AP location was estimated based on the deconvolution approaches too. We have also investigated multi-slope path loss models and path loss models that take into account estimated floor attenuations and we showed that the traditional single-slope path loss approach is the most robust among all the tried variants when real-field measurement data is employed. In addition, we showed that MMSE approach gave the best performance among the deconvolution approaches, and we also showed that there is still place for optimizing the floor detection performance of such algorithms, especially in multi-storey buildings with large openings. Our approaches decrease significantly the database sizes (factors of 10 times) and, in the future, solutions to decrease the database sizes even more can be investigated (for example by taking into account the fact that some WLAN transmitters have multiple MAC addresses and same physical location).

ACKNOWLEDGMENT

This research was partly funded by Nokia Inc. and by the Academy of Finland, which are gratefully acknowledged. The authors are grateful to Dr. Tech. Lauri Wirola and Dr. Tech. Jari Syrjärinne for their support and advice. Elina Laitinen and Toni Fadjukoff are thanked for conducting parts of the measurements.

REFERENCES

[1] H. Nurminen, J. Talvitie, S. Ali-Loytty, P. Müller, E.S Lohan, R. Piche, and M. Renfors. Statistical Path Loss Parameter Estimation and Positioning Using RSS Measurements in Indoor Wireless Networks. In 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN2012), Nov 2012.

[2] H. Koyuncu and S. Hua Yang, "A Survey of Indoor Positioning and Object Locating Systems", IJCSNS International Journal of Computer Science and Network Security, VOL.10 No.5, May 2010

[3] V. Honkavirta, T. Perälä, Simo Ali-Löytty, and Robert Piche, "A comparative survey of WLAN location fingerprinting methods", in Proceedings of the 6th Workshop on Positioning, Navigation and Communication 2009 (WPNC'09), pages 243-251, March 2009.

[4] J. Machaj, R. Piché, and P. Brida, "Rank Based Fingerprinting Algorithm for Indoor Positioning", in Proc. of 2011 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Sep 2011, Portugal.

[5] Hyo-Sung Ahn and Wonpil Yu, "Environmental-Adaptive RSSI-Based Indoor Localization," IEEE Transactions on Automation Science and Engineering, Vol. 6, No. 4, Oct. 2009.

[6] N. Marques, F. Meneses, A. Moreira, "Combining similarity functions and majority rules for multi-building, multi-floor, WiFi positioning," Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on , vol., no., pp.1,9, 13-15 Nov. 2012, doi: 10.1109/IPIN.2012.6418937.

[7] N. Chang, R. Rashidzadeh, M. Ahmadi, "Robust indoor positioning using differential wi-fi access points," Consumer Electronics, IEEE Transactions on , vol.56, no.3, pp.1860,1867, Aug. 2010 doi: 10.1109/TCE.2010.5606338

[8] V. Erceg, L. J. Greenstein, S. Y. Tjandra, S. R. Parkoff, A. Gupta, B. Kulic, A. A. Julius, and R. Bianchi. 2006. An empirically based path loss model for wireless channels in suburban environments. IEEE J.Sel. A. Commun. 17, 7 (September 2006), 1205-1211. DOI=10.1109/49.778178 http://dx.doi.org/10.1109/49.778178

[9] Capulli, F.; Monti, C.; Vari, M.; Mazzenga, F., "Path Loss Models for IEEE 802.11a Wireless Local Area Networks," Wireless Communication Systems, 2006. ISWCS '06. 3rd International Symposium on , vol., no., pp.621,624, 6-8 Sept. 2006 doi: 10.1109/ISWCS.2006.4362375

[10] C. B. Andrade and R. P. F. Hoefel, IEEE 802.11 WLANs: A Comparison on Indoor Coverage Models, in 23rd Canadian Conference on Electrical and Computer Engineering, 2010.

[11] Solahuddin, Y.F.; Mardeni, R., "Indoor empirical path loss prediction model for 2.4 GHz 802.11n network," Control System, Computing and Engineering (ICCSCE), 2011 IEEE International Conference on , vol., no., pp.12,17, 25-27 Nov. 2011 doi: 10.1109/ICCSCE.2011.6190487

[12] H. Hurskainen, J. Raasakka, T. Ahonen, and J. Nurmi, "Multicore Software-Defined Radio Architecture for GNSS Receiver Signal Processing," in EURASIP Journal on Embedded Systems, Hindawi, vol. 2009, Article ID 543720, 10 pages, 2009. doi:10.1155/2009/543720.

[13] G. Seco-Granados, J. A. Lopez-Salcedo, D. Jimnez-Banos, G. López-Risueno, "Signal Processing Challenges in Indoor GNSS", IEEE Signal Processing Magazine, vol. 29, no. 2, pp. 108-131, Mar 2012