

Power Delay Doppler Profile Fingerprinting for Mobile Localization in NLOS

Turgut Öktem and Dirk Slock

Mobile Communications Department, Eurecom

2229 route des Cretes, B.P. 193, 06560 Sophia Antipolis Cedex, FRANCE

Email: {oktem,slock}@eurecom.fr

Abstract—For existing localization algorithms, Non-Line-of-Sight (NLOS) propagation introduces some problems for the determination of the mobile position. This is because most of these algorithms depend on the information extracted from the Line-of-Sight (LOS) path such as Time-of-Arrival (ToA) or Time-Difference-of-Arrival (TDoA) received by either one or more Base Stations. On the contrary, algorithms based on Power Delay Profile Fingerprinting (PDP-F) take advantage of the uniqueness of the multipath channel between the Base Station (BS) and the Mobile Station (MS) over the geographical region of interest. The fingerprinting approach as its name implies performs a matching between a simulated database quantity and its corresponding measured quantity and indicates a match based on some cost function between the two. In this paper, we introduce an extension to this approach by including the effects of the Doppler shifts of the paths which we call the method as Power Delay Doppler Profile Fingerprinting (PDDP-F). With the inclusion of this extra information, we aim to increase the localization accuracy by resolving the paths not only in delay dimension but also in Doppler dimension.

keywords: mobile positioning, ray tracing, localization, power delay doppler profile fingerprinting.

I. INTRODUCTION

For the past few years, there has been a high interest in mobile positioning systems from both academic and industrial world [1], [2]. The primary motivation for the development of mobile positioning systems was due to the mandatory requirement of E-911 service by the U.S. Federal Communications Commission (FCC) [3]. Although the starting was because of security-emergency need, later it has found various applications in many fields. With the position information of the mobile, it is possible to make beamforming in the direction of the mobile to decrease the interference between the users in the cell, increase the range and throughput of the system and so on. Also intra and inter-system handoff can be handled more properly.

The localization algorithms are based on collecting position-dependent information from the MS, and then making a position estimate by processing this information. Depending on the technique used, either one or multiple BSs are required for identifiability of the position. Many of the well-known localization algorithms depend on the existence of the LOS path between the BS and MS. For example by using ToA information from three BSs obtained from the LOS paths, localization of the mobile is possible. Here the difficulty is to have three LOS paths at all the MS-BS links at the same time. Another technique which uses the combination of Angle of Arrival

(AoA) and ToA information makes the localization possible with one MS-BS link. However, obtaining AoA requires multiple antennas at the BS, also LOS path may not always be present. So to make the algorithm work even in NLOS conditions, Nájár et al. [4] proposes an alternative approach. During LOS condition, by estimating the TOA of the LOS and a NLOS path, the time offset (bias) between the two ToAs is calculated. When LOS condition is no longer present, the bias is then subtracted from the ToA of the NLOS path to estimate ToA of the LOS path. In general by adding a Kalman filtering stage, the accuracy of position estimate can be improved. The use of Kalman filter allows the tracking of the position trajectory, the velocity of the mobile and ToA bias caused by multipaths. As can be deduced from the examples, these techniques work robust in LOS conditions and try to minimize the effects of the NLOS conditions. However its effect cannot be eliminated completely, leading to some irreducible errors in the position estimates for the techniques relying on LOS paths. Therefore new techniques have been developed which try not to eliminate NLOS effects, but instead get use of it. The idea is to store the position dependent parameter of the environment in a database from the coverage area of the BS. Then a correlation algorithm is used to determine the position by comparing the measured parameter with the entries stored in the database. These kinds of techniques are generally called Fingerprinting techniques. Location Fingerprinting technique (LF) (introduced by U.S. Wireless Corp. of San Ramon, Calif.) relies on signal structure characteristics [1], [5]–[7]. By using multipath propagation pattern, the LF creates a signature unique to a given location. The position of the mobile is determined by matching measured signal characteristics from the BS-MS link to an entry of the database. The location corresponding to the highest match of the database entry is considered as the location of the mobile. For LF, it is enough to have only one BS-MS link (multiple BSs are not required) to determine the location of the mobile. Ahonen and Eskelinen suggest using the measured Power Delay Profiles (PDPs) in the database [8] for fingerprints, because amplitudes and delays of the multipath components create a unique position dependent signature. In [9], authors provide deterministic and Bayesian methods for PDP-F based localization. However, bandwidth limitations in the system might result in several rays arriving in the same sampling interval disturbing identifiability or the localization accuracy. For Ultra-WideBand (UWB) systems employing ToA algorithms, localization accuracy is quite good [10], [11] due to the very high bandwidth utilized in the system, but on the other hand its range is quite limited. Therefore it is mainly used in indoor localization applications. Outdoor localization which is also the scope of this paper, suffers from the considerably lower bandwidth used in the system. Due to lower bandwidth, number of resolvable paths of the multipath channel defined in [12] is less resulting in worse accuracy than UWB systems. We propose an alternative to increase the localization accuracy without changing the system. The main idea is to resolve the multipath components not just in one dimension but also in the Doppler dimension. So, the rays arriving in the same sampling duration but with different Doppler shifts can still

Eurecom's research is partially supported by its industrial members: BMW Group Research & Technology, Bouygues Telecom, Cisco, Hitachi, ORANGE, SFR, Sharp, STMicroelectronics, Swisscom, Thales. The work presented in this paper has also been partially supported by the European FP7 projects Where and Newcom++.

be discriminated if the difference between their Doppler frequencies are more than the discretization in the Doppler domain which we will explain later in the text. We call this technique as PDDP-F, and provide two versions of the algorithm, namely Frequency-Domain PDDP-F and the Time-Domain PDDP-F.

Notations: upper-case and lower-case boldface letters denote matrices and vectors, respectively. $(\cdot)^T$ and $(\cdot)^H$ represent the transpose and the transpose-conjugate operators. $E\{\cdot\}$ is the statistical expectation, and $\text{tr}\{\cdot\}$ is the trace operator defined for square matrices.

II. FREQUENCY-DOMAIN PDDP FINGERPRINTING FOR MOBILE LOCALIZATION

The database which we mention at the beginning is constructed offline either by ray tracing or ray launching simulation methods or by performing channel measurements over the geographical area of interest. The main objective while creating the database is to obtain location dependent channel information of the BS-MS link as we described before. The area is divided into several discrete sections, each section having a unique fingerprint. First we begin with the channel model. All the parameters for localization can be found or extracted from it. For the techniques relying on PDPs, PDP is just the magnitude squared version of the channel impulse response (CIR). But before using the measured PDPs, it must be averaged over some time duration. However, if the mobile moves rapidly and/or some paths are not resolvable (due to the limited bandwidth of the pulse-shape $p(t)$, paths contributions can overlap), the averaging gives a poor PDP estimation, and then a poor location accuracy. The time varying channel impulse response between the BS and MS can be written as:

$$h(t, \tau) = \sum_{i=1}^{N_p} A_i(t) e^{j2\pi f_i t} p(\tau - \tau_i(t)) \quad (1)$$

where N_p denotes the number of paths (rays), $p(t)$ is the convolution of the transmit and receive filters (pulse shape), f_i , $\tau_i(t)$, $A_i(t)$ denote the Doppler shift, delay and complex attenuation coefficient (amplitude and phase of the ray) of the i^{th} path respectively. Path delays and amplitudes vary slowly with the position, whereas the phase varies rapidly due to the high carrier frequency. Change of location on the order of a wavelength can result in a complete change of the phases. Thus one can assume that delays and fading amplitudes of individual paths are constant over a reasonable amount of (T) channel observations, but not the fading phases. The sampled estimated CIR with τ_s being the sampling period can be written as:

$$\hat{\mathbf{h}}(t, k\tau_s) = \sum_{l=1}^L A_l(t) \mathbf{p}(k\tau_s - \tau_l) + \mathbf{v}(t, k\tau_s) \quad (2)$$

where

$$A_l(t) = \sum_{k=1}^{K_l} A_{l,k} e^{j\varphi_{l,k}(t)} e^{j2\pi f_{l,k} t}, \quad (3)$$

$A_{l,k}$ and $\varphi_{l,k}(t)$ being the amplitude and phase of the ray respectively. In Equation (2), $\mathbf{v}(t, k\tau_s)$ is the additive white Gaussian noise vector, τ_l 's are in discrete sample units, and $\hat{\mathbf{h}}(t, k\tau_s) = [\hat{h}(t, t_0) \cdots \hat{h}(t, t_0 + (N-1)\tau_s)]^T$ is a vector of length N , N denoting the delay spread in samples. Equation (3) is just an expression of the superposition of rays arriving within the same sampling duration. It is easy to see that $K_1 + K_2 + \dots + K_L = N_p$. It is highly probable that in systems having low bandwidths, many paths will arrive in the same sampling durations which will decrease the resolvability of individual paths. So for such systems, PDP averaging

might result in a poor PDP estimation. This is our main motivation in the paper to include the Doppler effect, so that a 2D resolvability is possible. Equation (2) can also be written in matrix notation as:

$$\hat{\mathbf{h}}(t, k\tau_s) = \underbrace{[\mathbf{p}_{\tau_1} \cdots \mathbf{p}_{\tau_L}]_{\mathbf{P}_{\mathcal{T}}}}_{\mathbf{P}_{\mathcal{T}}} \underbrace{\begin{bmatrix} A_1(t) \\ \vdots \\ A_L(t) \end{bmatrix}}_{\mathbf{b}(t)} + \mathbf{v}(t, k\tau_s), \quad (4)$$

where \mathbf{p}_{τ_1} is the complex pulse delayed by τ_1 samples.

A. Obtaining PDDP from Ray Tracing Data

After expressing the time varying channel model, we can construct the PDDP from the ray tracing data. Here in this section, we will first introduce a formulation of PDDP for a general MIMO channel, then simplify the case for the SISO case which we are interested in the paper. Finally we will explain step by step how to apply the formulation to a given ray tracing data.

Consider a specular wireless MIMO channel model with multiple (N_t) transmit and (N_r) receive antennas. The time-varying channel impulse response is:

$$\mathbf{h}(\tau, t) = \sum_{i=1}^{N_p} A_i(t) e^{j2\pi f_i t} \mathbf{a}_R(\phi_i) \mathbf{a}_T^T(\theta_i) p(\tau - \tau_i) \quad (5)$$

where \mathbf{h} is rank 1 in 3 dimensions. The N_p pathwise contributions are characterized by these additional parameters:

- θ_i : angle of departure (AoD)
- ϕ_i : angle of arrival (AoA)
- $\mathbf{a}_R(\cdot)$, $\mathbf{a}_T(\cdot)$: (Rx/Tx side) antenna array response (if only a single antenna is present on one side or the other, then the corresponding $\mathbf{a}(\cdot) = 1$)

We shall assume here 2D propagation, an extension to 3D is immediate. Note: in case the Tx & Rx array responses are unknown, one should instead consider a parameterization of the following form:

$$\mathbf{h}(\tau, t) = \sum_{i=1}^{N_p} A_i(t) e^{j2\pi f_i t} \mathbf{a}_{R,i} \mathbf{a}_{T,i}^T p(\tau - \tau_i) \quad (6)$$

with $\mathbf{a}_{R,i}$, $\mathbf{a}_{T,i}$ unknown vectors. Note that also the pulse shape may need to be adjusted to measurements.

The channel impulse response in (5) results in fact from the propagation channel

$$c(\tau, t, \phi, \theta, v, \phi_v) = \sum_{i=1}^{N_p} A_i(t) e^{j2\pi f_i t} \delta(\phi - \phi_i) \delta(\theta - \theta_i) \delta(\tau - \tau_i) \quad (7)$$

where we shall assume the channel evolution over a short time period t so that the AoA ϕ , the AoD θ , the path delay τ_i and even the path complex amplitude A_i can be considered as constant.

Any Doppler shift in the propagation channel is actually assumed to be due to the mobility of the mobile terminal (any mobility in the environment would have to be captured by $A_i(t)$). Assume the terminal speed vector to have a magnitude μ and an orientation ϕ_μ (if $\phi_i = \phi_\mu$, then incoming wave and speed vector are aligned, but are evolving in opposite directions). Mobility of the terminal leads to a Doppler shift for path i as follows:

$$f_i = \cos(\phi_i - \phi_\mu) \mu / \lambda \quad (8)$$

where λ is the carrier wavelength. The channel impulse response in (5) is the convolution of the propagation channel with the system elements:

$$\mathbf{h}(\tau, t) = c(\tau, t, \phi, \theta, v, \phi_v) * p(\tau) * \mathbf{a}_R(\phi) * \mathbf{a}_T^T(\theta) \quad (9)$$

Consider now sampling the impulse response with a sampling period τ_s leading to N_τ samples and then vectorizing it:

$$\underbrace{\mathbf{h}(t)}_{N \times 1} = \begin{bmatrix} \mathbf{h}(\tau_s, t) \\ \mathbf{h}(2\tau_s, t) \\ \vdots \\ \mathbf{h}(N_\tau \tau_s, t) \end{bmatrix} = \sum_{i=1}^{N_p} A_i(t) e^{j2\pi f_i t} \mathbf{h}_i \quad (10)$$

where $\mathbf{h}(\tau_s, t)$ is the vectorized version of the $N_r \times N_t$ channel for the first delay element at time t and

$$\mathbf{h}_i = \underline{p}(\tau_i) \otimes \mathbf{a}_T(\theta_i) \otimes \mathbf{a}_R(\phi_i), \quad \underline{p}(\tau) = \begin{bmatrix} p(\tau_s - \tau) \\ p(2\tau_s - \tau) \\ \vdots \\ p(N_\tau \tau_s - \tau) \end{bmatrix} \quad (11)$$

where $N = N_t N_r N_\tau = \# \text{ TX antennas} \times \# \text{ RX antennas} \times \text{delay spread}$, and \otimes denotes the Kronecker product: for two matrices A and B , we get the block matrix $A \otimes B = [a_{ij}B]$. Two possible models can now be considered for the path amplitudes:

- Gaussian model: $A_i(t)$ Gaussian, characterized by a power (variance)
- deterministic model: $A_i(t)$ deterministic unknowns

We shall consider here the Gaussian case (other random models could be considered also, at least for the introduction of the profiles). We are now ready to introduce the Power Delay Doppler Space Profile (PDDSP). At the propagation level we get

$$\begin{aligned} \text{PDDSP}_c(\tau, f, \phi, \theta, v, \phi_v) \\ = \int E c(\tau, t_1 + t, \dots) c^*(\tau, t_1, \dots) e^{-j2\pi f t} dt \\ = \sum_{i=1}^{N_p} \sigma_i^2 \delta(\tau - \tau_i) \delta(f - f_i) \delta(\phi - \phi_i) \delta(\theta - \theta_i). \end{aligned} \quad (12)$$

where $\sigma_i^2 = E |A_i|^2$, and the expectation is at least over the (independent and uniformly distributed) random phases in the A_i , and possibly over the amplitudes also if they are not deterministic. At the channel response level, we get

$$\begin{aligned} \text{PDDSP}_h(\tau, f) \\ = \int E \mathbf{h}(\tau, t_1 + t) \mathbf{h}^H(\tau, t_1) e^{-j2\pi f t} dt \\ = \sum_{i=1}^{N_p} \sigma_i^2 |p(\tau - \tau_i)|^2 \delta(f - f_i) \mathbf{a}_i \mathbf{a}_i^H \\ = \sum_{i=1}^{N_p} \sigma_i^2 |p(\tau - \tau_i)|^2 \delta(f - f_i) R_T(\theta_i) \otimes R_R(\phi_i) \end{aligned} \quad (13)$$

where $\mathbf{a}_i = \mathbf{a}_T(\theta_i) \otimes \mathbf{a}_R(\phi_i)$ and we introduced the spatial covariances

$$R_T(\theta_i) = \mathbf{a}_T(\theta_i) \mathbf{a}_T^H(\theta_i), \quad R_R(\phi_i) = \mathbf{a}_R(\phi_i) \mathbf{a}_R^H(\phi_i). \quad (14)$$

In the case of a SISO channel, we get the Power Delay Doppler Profile (PDDP)

$$\text{PDDP}_h(\tau, f) = \sum_{i=1}^{N_p} \sigma_i^2 |p(\tau - \tau_i)|^2 \delta(f - f_i). \quad (15)$$

Now the formulation is complete and we explain how to apply the above formulation to the ray tracing data. The construction of the PDDP is explained step-by-step below:

- 1) First we create the 2D Delay-Doppler profile by only taking the rays into account (pulse shape and windowing effects not included yet). Delay, Doppler and power information of each ray is known. One thing to keep in mind is that delay and Doppler domains should be discretized properly according to the parameters in the measurement data.

The discretization in delay domain which we call $\Delta\tau$ is fixed, and equal to the sampling duration τ_s like in the measurement data defined before. For the Doppler domain discretization, channel estimations are carried out every t_s seconds for the measurement data, so the highest Doppler frequency +

Frequency Offset (FO) that can be observed is in the range $[-f_s/2, f_s/2]$ where $f_s = 1/t_s$. The FFT length we use in the measurement data is denoted by N_{FFT} , so the Doppler discretization in the measurement data is given by f_s/N_{FFT} . Therefore we use the same for the ray tracing data which we call $\Delta\Omega$.

The FO, that we mentioned above is due to the difference between the carrier frequencies of the local oscillators of the transmitter and receiver. Normally in a uniform scattering environment, the mean Doppler spread $\bar{\Omega}$ given by:

$$\bar{\Omega} = \frac{\int_{\Psi} \Omega D(\Omega) d\Omega}{\int_{\Psi} D(\Omega) d\Omega} \quad (16)$$

would be close to 0, where $D(\Omega)$ is the Doppler spectrum and Ψ denotes the range of possible Doppler shifts. So if there is FO, there will be a nonzero mean Doppler spread $\bar{\Omega}$. Unless the FO is too high, leading to aliasing, it is not a problem for our algorithm, also the delay offset is not a problem either.

- 2) Next step is summing up the power of the rays which are in the same grid. Here we directly sum up their individual powers, as we compute the expected average power by the following formula in the grid (we are averaging over the random phases assuming uniform distribution over $[0, 2\pi]$):

$$E_{\varphi_{1,1} \dots \varphi_{l,n}} |(A_{l,1} e^{j\varphi_{l,1}} + \dots + A_{l,n} e^{j\varphi_{l,n}})|^2 = \sum_{i=1}^n A_{l,i}^2 \quad (17)$$

where n is the number of rays in the grid, $A_{l,i}$ is the magnitude of the i^{th} ray. It is also easy to see that mean of the rays is equal to 0, where we will use this fact in the Bayesian PDDP-F part.

- 3) Last step is to include the effects of pulse shape in the delay domain, and the windowing in the Doppler domain. This is an easy process. It is enough to make linear convolution in the Delay domain of the PDDP matrix with the absolute squared pulse shape.

And the other operation as we mentioned before is for windowing effect in the measurement data. We have M channel estimates per point. And we choose a certain number of consecutive channel estimates (N_{Window}) among them for computing Fourier Transforms. This process is called windowing, and should be included in the ray tracing PDDP.

Also it is important to choose a suitable window and a window length N_{Window} . Although Rectangular window has the narrowest main lobe for a given N_{Window} , its side lobes are not negligible leading to spectral leakage [13]. Therefore we decided to choose the Hamming window whose side lobes decay much faster than the Rectangular window. To include the effect of windowing, it is enough to make cyclic convolution in the Doppler domain of the PDDP matrix with the absolute squared DFT of the window.

In Figure 1, we show a sample PDDP obtained from a ray tracing data. In fact, as we did not have very reliable ray tracing data, we have used synthetically generated ray tracing data instead to test our algorithms.

B. Obtaining PDDP from Measurement Data and The Fingerprinting Operation

Obtaining the PDDP from the measurement data is easier. N_{Window} consecutive channel estimates are chosen with the Hamming window explained before. Each channel estimate is a

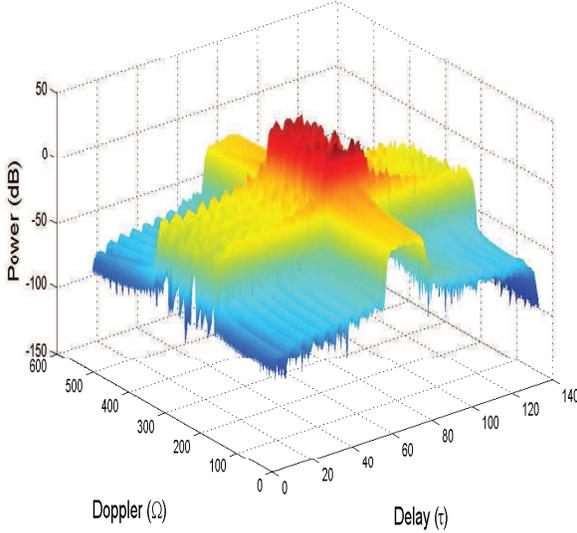


Fig. 1. A Sample PDDP obtained from Ray Tracing Data

vector of length N . For each delay element among N , absolute squared DFT is computed with respect to the time variable to see its variation in time (Doppler information). Consequently this gives the 2D Delay-Doppler profile of the measurement data. One thing to pay attention is that all hardware related effects must be removed from the channel estimates e.g. Automatic Gain Control (AGC) and others to make a true comparison.

After computing ray tracing and measurement data PDDPs, next step is to check the similarity between these matrix profiles. There are consecutive channel measurements, obtained in the BS, and the objective is to see which ray tracing data in the database will give the highest match to this measurement. The cost function is defined as the similarity between the matrices. Among the K element database, the one corresponding to the position of the k^{th} ray tracing data will be chosen:

$$\hat{k} = \arg \max_{k \in [1, K]} J(PDDP_M, PDDP_{RT_k}) \quad (18)$$

where J is the likelihood function, $PDDP_M$ is the PDDP obtained from the measurement data and $PDDP_{RT_k}$ is the PDDP from the k^{th} entry in the ray tracing database. For the likelihood function J , one reasonable candidate is to use the inner product criteria defined for matrices normalized by their norms as below:

$$\frac{\text{tr}(\mathbf{A}^T \mathbf{B})}{\sqrt{\text{tr}(\mathbf{A}^T \mathbf{A}) \text{tr}(\mathbf{B}^T \mathbf{B})}} = J(\mathbf{A}, \mathbf{B}). \quad (19)$$

But one thing to note is that, before using the above formula, the ray tracing and measurement PDDPs must be perfectly aligned in the delay and Doppler dimensions. Any delay offset or FO must be handled very precisely. Therefore it is reasonable to choose a computationally effective solution. 2D FFT operation to check the highest correlation between the two matrices normalized by their norms is a very fast operation given by:

$$\frac{\text{IFFT}(\text{FFT}(\mathbf{A}) \odot \text{conj}(\text{FFT}(\mathbf{B})))}{\sqrt{\text{tr}(\mathbf{A}^T \mathbf{A}) \text{tr}(\mathbf{B}^T \mathbf{B})}} \quad (20)$$

where \odot is the Hadamard (element-wise) multiplication and conj denotes conjugate. The maximum entry in the resulting matrix is the highest correlation between the two in the perfectly aligned case. Also the position of the maximum entry gives the delay offset and FO.

C. Simulation Results

Figure 2 is the simulation result for the comparison of deterministic PDP-F and Frequency-Domain PDDP-F over a range of SNR values. In our simulation environment, we have created 15 ray tracing points ($K = 15$), and then generated a measurement data from the first ray tracing point by adding noise. The objective is to see in how many cases the algorithms will match with the first point. The path loss exponent is taken as 2, and we generate more than 1000 rays in every iteration. There is no spatial relation between these 15 points, we just generate random channel parameters for each of them. For the system parameters, we have a sampling frequency of 9.1429 MHz and a wavelength of 0.4249 m. For the PDDP-F, the size of the 2D FFT is 1024 x 512 when we are computing the correlation. To compute the spectrum, we use an FFT length of 512. The delay spread N is taken as 135 samples. The channel estimates are obtained every 4 ms, window length is 200, speed magnitude is chosen randomly between 1-60 km/h and direction angle between 0-359 degrees. We see that the Frequency-Domain PDDP-F algorithm is almost independent of the SNR whereas the deterministic PDP-F improves with the increasing SNR over the range of interest. As we see in the plot, Frequency-Domain PDDP-F always outperforms the deterministic PDP-F algorithm in all the SNR values. This difference comes from the usage of the additional Doppler dimension which PDP-F cannot benefit from.

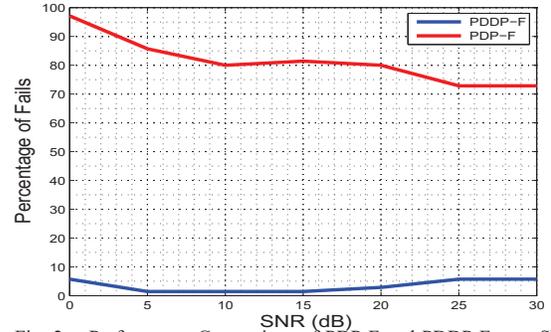


Fig. 2. Performance Comparison of PDP-F and PDDP-F over SNR

III. TIME-DOMAIN PDDP FINGERPRINTING FOR MOBILE LOCALIZATION

As in the Frequency-Domain version of the algorithm, here we first provide the formulation of the Time-Domain algorithm, then we explain in detail the application of the algorithm on the ray tracing data. Equation (3) shows that, sampled channel taps might be the superposition of several rays which arrive within the same sampling duration. As we explained before, mean of these channel taps is 0, due to averaging over random phases. Also the expected average power is just the summation of the individual powers of the incoming rays. The central limit theorem lets us model these taps as Gaussian random variables with mean 0. In this section, we propose the Time-Domain version of the PDDP-F algorithm which makes use of the second-order statistics of the channel. We assume that the complex fading vector $\mathbf{b}(t)$, and the additive noise $\mathbf{v}(t)$ are i.i.d. zero-mean Gaussian vector processes, i.e.,

$$\begin{aligned} \mathbf{b}(t) &\sim \mathcal{N}(\mathbf{0}, \mathbf{C}_b) \\ \mathbf{v}(t) &\sim \mathcal{N}(\mathbf{0}, \sigma_v^2 \mathbf{I}_N) \end{aligned} \quad (21)$$

where $\mathcal{N}(\mathbf{0}, \mathbf{C}_b)$ denotes the zero-mean complex Gaussian vector with covariance matrix \mathbf{C}_b (we will soon explain how to derive it), and σ_v^2 is the channel estimation error variance. With the statistical model of Equation (21), $\hat{\mathbf{h}}(t)$ is modeled as an i.i.d. complex Gaussian vector with $\hat{\mathbf{h}}(t) \sim \mathcal{N}(\mathbf{0}, \hat{\mathbf{C}}_{\hat{\mathbf{h}}})$, $\hat{\mathbf{C}}_{\hat{\mathbf{h}}} = \mathbf{P}_\tau \mathbf{C}_b \mathbf{P}_\tau^H + \sigma_v^2 \mathbf{I}_N$.

With the Bayesian modeling of $\hat{\mathbf{h}}(t)$, we can propose a Maximum Likelihood solution to the localization problem. Our aim is also to take into account the Doppler variation of the channel. Therefore we stack consecutive $\hat{\mathbf{h}}(t)$ channel estimates in a vector, instead of taking just one, and compute the covariance matrices based on this. Now consider the channel response at multiple consecutive time instants $t = t_s, 2t_s, \dots, nt_s$:

$$\underbrace{\hat{\mathbf{h}}}_{nN_\tau N_r N_t \times 1} = \begin{bmatrix} \hat{\mathbf{h}}(t_s) \\ \hat{\mathbf{h}}(2t_s) \\ \vdots \\ \hat{\mathbf{h}}(nt_s) \end{bmatrix} \quad (22)$$

Then we get

$$\hat{\mathbf{h}} = \sum_{i=1}^{N_p} A_i \underline{e}(f_i) \otimes \underline{\mathbf{h}}_i, \quad \underline{e}(f) = \begin{bmatrix} e^{j2\pi f t_s} \\ e^{j2\pi f 2t_s} \\ \vdots \\ e^{j2\pi f n t_s} \end{bmatrix} \quad (23)$$

We get for the covariance matrix of $\hat{\mathbf{h}}$

$$\mathbf{C}_{\hat{\mathbf{h}}} = \sum_{i=1}^{N_p} \sigma_i^2 R_f(f_i) \otimes R_\tau(\tau_i) \otimes R_T(\theta_i) \otimes R_R(\phi_i) \quad (24)$$

where

$$R_f(f) = \underline{e}(f)\underline{e}^H(f), \quad R_\tau(\tau) = \underline{p}(\tau)\underline{p}^H(\tau). \quad (25)$$

Note that R_f is Toeplitz. In the case of a SISO channel, we have $\mathbf{C}_{\hat{\mathbf{h}}} = \sum_{i=1}^{N_p} \sigma_i^2 R_f(f_i) \otimes R_\tau(\tau_i)$ and the PDDP is related to the diagonal part of this matrix, after taking DFT of the R_f part.

To be more specific, there are M channel estimates, and they are divided into $M - n + 1$ groups each group having n consecutive channel estimates. For example there can be 2 such groups for $M = 4$ and $n = 3$, i.e. $\hat{\mathbf{h}}(t_s, k\tau_s), \hat{\mathbf{h}}(2t_s, k\tau_s), \hat{\mathbf{h}}(3t_s, k\tau_s)$ for group 1 and $\hat{\mathbf{h}}(2t_s, k\tau_s), \hat{\mathbf{h}}(3t_s, k\tau_s), \hat{\mathbf{h}}(4t_s, k\tau_s)$ for group 2. We stack these vector groups into a longer vector as:

$$\hat{\mathbf{h}}(i) = \begin{bmatrix} \hat{\mathbf{h}}(i t_s, k\tau_s) \\ \vdots \\ \hat{\mathbf{h}}((i + n - 1) t_s, k\tau_s) \end{bmatrix}. \quad (26)$$

Now, the Gaussian Log-Likelihood can be constructed with $M - n + 1$ vectors as:

$$LL \propto - (M - n + 1) \ln \left(\det \mathbf{C}_{\hat{\mathbf{h}}} \right) - \sum_{i=1}^{M-n+1} \hat{\mathbf{h}}(i)^H \mathbf{C}_{\hat{\mathbf{h}}}^{-1} \hat{\mathbf{h}}(i) \quad (27)$$

where $\mathbf{C}_{\hat{\mathbf{h}}}$ is the covariance matrix of $\hat{\mathbf{h}}$. Instead of the usual Maximum Likelihood approaches to estimate the path parameters by maximizing the likelihood with respect to the parameters, the likelihood is evaluated by substituting the position dependent path parameters and hence it provides the likelihood of position. In other words, covariance matrices of the ray tracing database ($\mathbf{C}_{\hat{\mathbf{h}}}$) are created offline by the position dependent parameters (using delays, powers, Doppler shifts of the rays), then the likelihood is evaluated with the above formulation for the measurement data. The position giving the highest likelihood is the position estimate of the mobile. Equation (27) can be written equivalently as:

$$LL \propto - \ln \left(\det \mathbf{C}_{\hat{\mathbf{h}}} \right) - \text{tr} \left\{ \mathbf{C}_{\hat{\mathbf{h}}}^{-1} \hat{\mathbf{C}}_{\hat{\mathbf{h}}} \right\} \quad (28)$$

where $\hat{\mathbf{C}}_{\hat{\mathbf{h}}} = \frac{1}{M - n + 1} \sum_{i=1}^{M-n+1} \hat{\mathbf{h}}(i)\hat{\mathbf{h}}(i)^H$ is the observed sample covariance matrix which is asymptotically unbiased. Equation (28) clearly shows that the Log-Likelihood is just a Covariance Matching operation between the measurement covariance matrix and the pre-computed K covariance matrices of the database. One last thing to mention is the derivation of the covariance matrices of the ray tracing database ($\mathbf{C}_{\hat{\mathbf{h}}}$). We will begin with Equation (26). To simplify the notation, $\hat{\mathbf{h}}(it_s, k\tau_s)$ will be denoted as $\hat{\mathbf{h}}_i$, $\mathbf{v}(it_s, k\tau_s)$ as \mathbf{v}_i and we will just present the derivation for $n = 2$. Deriving for any n is straightforward afterwards.

$$E \left\{ \hat{\mathbf{h}}(i)\hat{\mathbf{h}}(i)^H \right\} = \begin{bmatrix} E \left\{ \hat{\mathbf{h}}_1 \hat{\mathbf{h}}_1^H \right\} & E \left\{ \hat{\mathbf{h}}_1 \hat{\mathbf{h}}_2^H \right\} \\ E \left\{ \hat{\mathbf{h}}_2 \hat{\mathbf{h}}_1^H \right\} & E \left\{ \hat{\mathbf{h}}_2 \hat{\mathbf{h}}_2^H \right\} \end{bmatrix} \quad (29)$$

where $E \left\{ \hat{\mathbf{h}}_1 \hat{\mathbf{h}}_1^H \right\} = E \left\{ \hat{\mathbf{h}}_2 \hat{\mathbf{h}}_2^H \right\} = \mathbf{C}_{\hat{\mathbf{h}}} = \mathbf{P}_\tau \mathbf{C}_b \mathbf{P}_\tau^H + \sigma_v^2 \mathbf{I}_N$, and \mathbf{C}_b is a diagonal matrix given as:

$$\mathbf{C}_b = \begin{bmatrix} \sum_{k=1}^{K_1} A_{1,k}^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sum_{k=1}^{K_L} A_{L,k}^2 \end{bmatrix}, \quad (30)$$

where we used Equation (3) to derive \mathbf{C}_b ($A_{l,k}$'s are the magnitudes of the rays). As can be seen from the formulation, there is no Doppler information yet because the calculated covariance is for the same time instant. The idea of stacking consecutive channel estimates brings the Doppler information which will be clear while deriving $E \left\{ \hat{\mathbf{h}}_1 \hat{\mathbf{h}}_2^H \right\}$.

We will not make the derivation for $E \left\{ \hat{\mathbf{h}}_2 \hat{\mathbf{h}}_1^H \right\}$ because it is just the transpose-conjugate of the other. $E \left\{ \hat{\mathbf{h}}_1 \hat{\mathbf{h}}_2^H \right\} = \left(E \left\{ \hat{\mathbf{h}}_2 \hat{\mathbf{h}}_1^H \right\} \right)^H = \mathbf{P}_\tau \mathbf{C}_d \mathbf{P}_\tau^H$, where \mathbf{C}_d is derived as:

$$\begin{bmatrix} \sum_{k=1}^{K_1} A_{1,k}^2 e^{-j2\pi f_{1,k} t_s} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sum_{k=1}^{K_L} A_{L,k}^2 e^{-j2\pi f_{L,k} t_s} \end{bmatrix}$$

by using Equation (3) again. Now, the Doppler contributions of each ray is visible in the covariance matrix. As can be seen, the overall covariance matrix is a function of delays, powers and Doppler shifts of rays. We aim to increase the localization accuracy by incorporating this additional Doppler information.

A. Simulation Results

Figure 3 is the simulation result for the comparison of Bayesian PDP-F and the Time-Domain PDDP-F over a range of SNR values. The simulation environment is the same as for the Frequency-Domain PDDP-F simulations except we do not use FFT here. We see 3 curves in the plot where $n = 1$ corresponds to the Bayesian PDP-F case. It is obvious that Time-Domain PDDP-F outperforms Bayesian PDP-F. Increasing n (number of consecutive channel estimates) also increases the success rate. If we also compare with the Frequency-Domain PDDP-F algorithm, we see that the Time-Domain PDDP-F is more robust and success rate is higher for $n \geq 3$. Also one drawback of the Frequency-Domain PDDP-F is that its non-parametric spectrum might suffer from limited resolution.

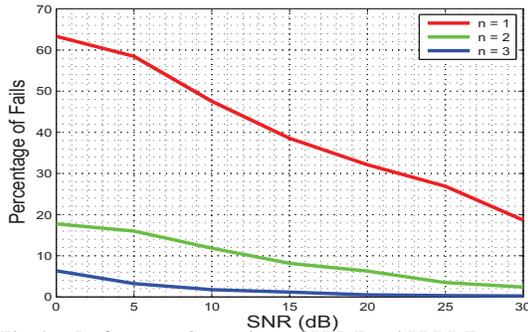


Fig. 3. Performance Comparison of PDP-F and PDDP-F over SNR

The Time-Domain PDDP-F approach can be seen as an elegant method for localization. Instead of trying to match only the PDPs (diagonal elements of the covariance matrices for the same time instant), the whole covariance matrices are compared, also by taking into account the Doppler information. However, there are some important things to note. First of all, the delay offset and FO problems are also present here. So care must be taken for them. Also another factor which is very important and not present for the Frequency-Domain case is the channel estimation error variance. It must be handled very carefully also. The advantage of the Time-Domain PDDP-F over its Frequency-Domain version lies in the weighting matrix (inverse of the covariance matrix) used. If we consider the noiseless case, the weighting matrix tries to balance the weak and strong rays. In other words, strong rays (rays with high power) are weighted by small coefficients whereas weak rays are weighted by higher coefficients. As a result, this gives the advantage of observing even very little details which would be mostly ignored in the Frequency-Domain case. However if we consider noise, this could also lead to noise amplification when the path contribution is below the noise level.

IV. CONCLUSION

In this paper, we have presented two new algorithms for localization. Classical localization methods take into account only powers and delays of the paths, but we also took the Doppler information into account. For the systems having small bandwidths, rays might not be resolvable in time. Therefore the Doppler dimension has the effect of increasing the accuracy, and also the identifiability. Having $N_p \geq 2$ paths is sufficient for the identifiability unless they arrive at the same Delay-Doppler grid. For a channel with delay spread N , the probability that two paths arrive in the same sampling duration is $1/N$, however the probability that both are located in the same Delay-Doppler grid is $1/(NN_{FFT})$. So we see that the resolution is increased by a factor of N_{FFT} by exploiting the Doppler information.

REFERENCES

- [1] H. Koshima, and J. Hoshen. "Personal Locator Services Emerge," *In IEEE Spectrum*, Vol.37, pp.41-48, Feb. 2000.
- [2] J.J. Caffery, and G.L. Stuber. "Overview of Radiolocation in CDMA Cellular Systems," *In IEEE Communications Magazine*, Vol.36, Issue 4, pp.38-45, Apr. 1998.
- [3] G. Sun, and W. Guo. "Robust Mobile Geo-Location Algorithm Based on LS-SVM," *In IEEE Trans. on Vehicular Technology*, Vol.54, Issues:3, pp. 1037-1041, May 2005.
- [4] M. Najjar, J.M. Huerta, J. Vidal, and J.A. Castro. "Mobile Location with Bias Tracking in Non-Line-Of-Sight," *In Proc. of IEEE ICASSP*, Vol.3, pp.956-959, May 2004.
- [5] O. Hilsenrath and M. Wax, "Radio Transmitter Location Finding for Wireless Communication Network Service and Management," *US Patent*, 6 026 304, Feb. 2000.
- [6] M. Wax, Y. Meng and O. Hilsenrath, "Subspace signature matching for location ambiguity resolution in wireless communication systems," *US Patent*, 6 064 339, May 2000.
- [7] M. Wax, O. Hilsenrath and A. Bar, "Radio Transmitter Location Finding in CDMA Wireless Communication Systems," *US Patent*, 6 249 680, June 2001.
- [8] S. Ahonen, and P. Eskelinen. "Mobile Terminal Location for UMTS," *In IEEE Aerospace and Electronic Systems Mag.*, Vol.18, Issue 2, pp.23-27, Feb. 2003.
- [9] M. Triki, and D. T. M. Slock. "Mobile Localization for NLOS Propagation," *In Proc. of IEEE PIMRC*, Sep. 2007.
- [10] H.V. Poor. "An Introduction to Signal Detection and Estimation, 2nd ed. New York," *Springer-Verlag*, 1994.
- [11] C.E. Cook, and M. Bernfeld. "Radar Signals: An Introduction to Theory and Applications. New York," *Academic*, 1970.
- [12] I. E. Telatar, and D. N. C. Tse. "Capacity and Mutual Information of Wideband Multipath Fading Channels," *In IEEE trans. on Information Theory*, Vol.46, Issue 4, pp.1384-1400, Jul. 2000.
- [13] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck. "Discrete-Time Signal Processing," *Prentice-Hall*, 1999.