Location Estimation via Sparse Signal Reconstruction in Subsampled Overcomplete Dictionaries for Wireless 4G Networks

John Roth*, Murali Tummala†, John McEachen†, James Scrofani†, and Allison Hunt*
*Department of Electrical and Computer Engineering
United States Naval Academy
Annapolis, MD, USA
Email: jroth@usna.edu
†Department of Electrical and Computer Engineering
Naval Postgraduate School
Monterey, CA, USA

Abstract
We present a simple and effective means for position estimation designed to be deployed in urban and dense multipath environments characteristic of 4G wireless networks. To address the multipath channel of such environments a fingerprinting scheme is proposed. One of the drawbacks to this class of methods is the large initial cost associated with establishing a database matrix. This issue is addressed by using a multi-channel filtering method adapted from the H.264 video standard to recover the incomplete data. Position estimation is accomplished via a modified k-nearest neighbor approach to pattern matching. We show through simulation that not only are we able to achieve compelling fidelity in the reconstructed databases from highly incomplete data, but that we are able to do so at a relatively low computational cost. Finally, our results demonstrate that we are able to achieve accurate position estimates vis-à-vis severe undersampling and noisy channel conditions.

1. Introduction

Location based services (LBS) is an emerging market that has commanded the attention of a vast body of research. Indeed, it has been estimated that LBS will generate on the order of 15 billion US dollars annually [1]. Fueled by economic promise, the vast body of research has launched a multi-faceted approach on the problem of locationing outdoors in line of sight channels has been long studied for decades and excellent methods in accurate location finding are prolific [4], [6]. However, these well studied and established methods are all parameterized by the problem geometry. In other words, the physical layout of electromagnetic landmarks relative to the location of the target, termed a base station (BS) and mobile subscriber (MS) respectively from this point forward, dictate the accuracy of the estimate.

An alternative approach proposes to first collect a database of fingerprints throughout the tracking area. The goal is then to match current signal readings with the most likely signal exemplar and estimate the E911 mandate has provided significant impetus for work in this area [5].

Since the inevitable inception of LBS the problem has been studied widely. Initial navigational and military applications of location finding drove early interest in this area. The problem of locationing outdoors in line of sight channels has been long studied for decades and excellent methods in accurate location finding are prolific [4], [6]. However, these well studied and established methods are all parameterized by the problem geometry. In other words, the physical layout of electromagnetic landmarks relative to the location of the target, termed a base station (BS) and mobile subscriber (MS) respectively from this point forward, dictate the accuracy of the estimate.

Despite early success in these areas, subsequent conceptualization of the multipath channel and the movement toward urban centers around the world point to a different type of approach. Urban and indoor environments in particular offer propagation channels with the heaviest multipath influence. These environments, presenting non-line of sight (NLOS) conditions are rich in physical blockages causing a signal to reflect repeatedly while traveling some non-minimal distance enroute to its destination. This type of signal propagation induces significant error in the aforementioned techniques which have a heavy reliance on the geometric orientation of the infrastructure and an implicit assumption that signal propagation is line of sight [7].

An alternative approach proposes to first collect a database of fingerprints throughout the tracking area. The goal is then to match current signal readings with the most likely signal exemplar and estimate the
current position as the location where the exemplar was previously collected. This technique has been termed database correlation, scene analysis, or fingerprinting and has garnered wide acclaim in recent decades [1]–[3], [7]–[10]. We subsequently refer to this method as fingerprinting in congruence with the vast majority of the literature. The problem then is reduced to solving an underdetermined system of linear equations and can be concurrently viewed as a signal representation problem using a overcomplete, and necessarily redundant, dictionary. Knowledge that the signal is inherently sparse is the last piece of the puzzle that points to a unique solution from amongst infinite possibilities (i.e., the null space of the database matrix). Thus we wish to represent the observation with as few dictionary atoms as possible. The time varying nature of the channel makes this an interesting problem and with overwhelming probability guarantees that the solution to the aforementioned system of equations will have a support larger than unity. Even so, there is not necessarily a guarantee of a unique solution.

While effective in NLOS environments, the glaring disadvantage to this approach is the cost associated with building and maintaining a representative database. Many ways to circumvent this disadvantage have been suggested of which using propagation models or radio mapping tools and the emerging theory of compressive sensing have gained significant ground in the literature [11], [12]. The former solution attempts to utilize knowledge of electromagnetic propagation to build a database with only a subset of measurements or no measurements at all. The latter solution uses the continually varying nature of radio signals and their spatial correlation to recover a sparse representation of the signal via convex optimization [13], [14].

The contributions of this work are twofold. First, we propose a novel method by which this disadvantage may be mitigated. We adapt proven methods based on the ability to represent the database elements sparsely (i.e., spatial correlation) in multi-channel filtering to recover a database estimate. Second, we reformulate the pattern matching portion of the solution via an adaptation of the \(k\)-Nearest Neighbor (\(k\)-NN) algorithm and show that via our approach high fidelity in accuracy can be achieved. Throughout we remain mindful of the computational cost incurred at the hands of the techniques employed and show how we make significant headway in this territory while minimizing that lost in position accuracy.

The rest of this paper is organized as follows. We present the details of the method by which we formulate the two aspects of our contribution, database reconstruction and position estimation, in Section 2.
2.2. Position Estimation

The online portion of fingerprinting involves creating a position estimate by matching an observation to the most likely corresponding match in the database $X$. This is done by correlating the observation vector, $h_k \in \mathbb{R}^{N_t \times 1}$, taken at time $k$ with the database entries in $X$. Ultimately, we seek to solve

$$
\begin{align*}
\min & \| h_k - X \varphi_{fp}^k \|_2 \\
\text{s.t.} & \varphi_{fp}^k = e_i, \forall k
\end{align*}
$$

(2)

where $e_i \in \mathbb{R}^{N_t \times 1}$ is the trivial vector with all positions, except for the $i^{th}$ position are zero and $\| e_i \|_2 = 1$. Here we restrict the solution space to the case of extreme 1-sparse solutions only for simplicity. It follows then that there are $N_t$ possible answers to (2) and through exhaustive enumeration a global minimum can be found. This is essentially the $k$-NN grouping (with other obvious similarities to $k$-means) where we seek to minimize the euclidean distance between a test vector and the nearest cluster centroid or dictionary atom. The main difference between this approach and $k$-means is that we are offered no latitude in our dictionary atoms as they are fixed by the measurements made at each location, therefore they may not be an optimal means for representation of the observation (i.e., $X_j \neq X_i$ where $j \neq i$).

Our approach to this solution uses a method that has been observed to be robust [12]. We begin with a construct termed the likelihood vector, $L \in \mathbb{R}^{N_t \times 1}$. The index of each element of $L$ is defined as

$$
l_i = \frac{1}{\| y_k - X e_i \|}.
$$

(3)

Each entry, $L_i$, can be thought of as a measure of likelihood of the observation being take at the $i^{th}$ location as quantified by euclidean distance in $\mathbb{R}^{BS}$. The information in the vector entries lies in the contrast in values. The higher the value, the more likely the observation was made at that location. The solution to (2) is then obtained by

$$
\varphi_{fp}^k = e_i, \ i = \arg \max_i L_k 
$$

(4)

for which it is trivial to see that a global minimum is guaranteed.

It should be noted that the complexity of this solution method will grow with $O(n)$ yielding a low complexity solution to (2). An interesting alternative solution methodology to (2) is sparse signal recovery via convex optimization which has gained some traction in the literature [11], [14], [18]. With this method the inherent sparsity in the location signal is leveraged to recover the location via a recasting of the problem as a linear program solved by basis pursuit which runs in $O(n \log(n))$ time [19], [20]. We contrast our method with this one and note that we arrive at a solution via a method requiring significantly less computational prowess, satisfying our requirement for deployment on power limited mobile devices.

2.3. Sparse Signal Interpolation in Overcomplete Dictionary Representations

With the larger model framework in mind we turn to the focus of this paper, namely reconstruction of
the database from the incomplete database matrix, $X^\dagger$ which is randomly sub-sampled along the spatial dimension such that $n < N_t$ as shown in Figure 1. We propose a method of recovery that is based off the widely successful H.264 compression standard. This standard allows for prediction within frames based on surrounding blocks that have already been decoded. Here we have an advantage over H.264 in that the standard tends to decode linearly in raster format, meaning an interpolated pixel or block will only have those above and to its right from which to interpolate a value [21]. While, depending on the level of sub-sampling, our application will usually have more neighbors in various locations from which to interpolate a sample.

We begin with the subsampled database, $X^\dagger \in \mathbb{R}^{N_t \times n}$. Our goal is to provide an estimate, $\hat{X}$, of the fully sampled database, $X$. As each column index represents a physical location we can leverage the correlation in adjacent signals effected by the electromagnetic propagation phenomena. The problem can then be expressed as

$$\min \ f(\theta_i) \quad \text{s.t.} \quad AF^{-1}\theta_i = X^\dagger T, \ \forall i \tag{5}$$

where $\theta_i = F\{X_i\}$ (i.e., the Fourier transform of $i$th row of the reconstructed database), $A \in \mathbb{R}^{n \times N_t}$ is a sensing matrix, and $F^{-1} \in \mathbb{R}^{N_t \times N_t}$ is the inverse Fourier basis. Each row of $A$ is constructed from the trivial basis set, $e_k$, where $k$ is selected at random for each entry and also dictates the subsampling scheme. Because we expect correlation in neighboring samples, we can also expect $\theta_i$ to be sparse, thus any number of penalty functions which minimizes the support of $\theta_i$ is sufficient. Examples in the literature include the Manhattan and total variation norms [14], [22].

We propose solving (5) indirectly and recovering the missing data by creating a multi-channel weighted linear combination of the adjacent fingerprints

$$X^\dagger_j = \sum_{m \in A_j} \alpha_m X_m \tag{6}$$

where $A_j$ is the set of tiles adjacent to the $j$th tile and $\alpha$ is some weight. Our method utilizes an extremely simple scheme such that $\alpha_i = \alpha_j \forall i,j \leq N_t$ and $\sum_m \alpha_m = 1$. In other words, $\hat{X}_j$ is a linear average of the tiles surrounding it in $X^\dagger$. The problem then takes the shape of a multi-channel filtering approach where each BS signal serves as a different channel [23]. At first glance the weights selected may seem overly simplistic, but have been chosen to provide a solid foundation for the possible filtering approaches to database reconstruction. Additionally, besides the surprisingly good numerical results reported in Section 3 justification for this simple approach can be found in the H.264 video coding standard. This particular weighting scheme is a modification of proven method used by the H.264 intra prediction DC mode in which a block prediction is the mean of of the pixels directly above and to its left [21].

The set $A_j$ includes all locations physically adjacent to the $j$th tile such that

$$|A_j| \leq 8 \tag{7}$$

and $\cdot$ denotes the cardinality of the set. If data are missing for locations in $A_j$, then the inequality in (7) is strict.

3. Implementation

In this section results obtained by monte carlo methods are presented. First, we study the fidelity of the database $\hat{X}$ vis-à-vis the subsampling rate. A direct relationship between the fidelity of the database reconstruction is then made with the accuracy of the overall model by using the reconstructed databases as a radio map in online estimation.

3.1. Experimental Setup

The tracking area is a 144 square kilometer area divided uniformly into $N_t = 144$ tiles, each one square kilometer as shown in Figure 1. The tracking area consisted of $N_{BS} = 46$ base stations distributed similar to a previous network distribution in the vicinity of San Jose, CA. Each BS is treated as an omnidirectional transmitter where no sectoring is in use. The target is assumed to be traveling at a relatively high speed along a road network and travels over a comprehensive path in the tracking area covering a representative number of locations. The entire position estimation is conducted in $\mathbb{R}^{N_t}$. No effort is made to perform the estimation in a projected subspace such as $\mathbb{R}^\gamma$ where $\gamma < N_t$ as in [16], [17] so as not to dilute the impact of a reconstructed dictionary $\hat{X}$, although this would be a natural extension of this work.

3.2. Noise Model

In order to model a noisy channel from the BS to the MS we establish three subsets of tiles for each BS. The first subset is termed center-local, $T^\text{cl}_i$, and is described for the $i$th BS by $X_{i,j} = 1$ if the $j$th tile is center-local. These tiles are considered close enough
Figure 2. A comparison of database reconstruction is presented at a sub-sampling rate of approximately 0.28. Databases are presented as a grayscale image where the extrema 1 and 0 are represented as white and black respectively. The original database \( X \) (a) and the reconstructed database (b) is presented.

to the BS such that their connectivity is essentially deterministic in that an MS in this tile will always be able to connect to the given BS. The second subset of tiles is termed peripheral-local, \( T^\text{pl}_i \), and is described for the \( i \)th BS by \( \chi_{i,j} = y \) if the \( j \)th tile is peripheral-local. These tiles are considered far enough away from the BS such that their connectivity is probabilistic so as to model the shadowing effect. To this end this entry in \( X \) is drawn from the standard continuous probability simplex \( y \in (0, 1) \) where \( y \) is drawn from the pseudo-Gaussian distribution \( f_{Y}(y) \approx N(m, \sigma)I[0,1](y) \) \( \chi_{i,j} = 0 \).

Previous works have noted that actual multipath signal channels can produce non-trivial signal strength distributions. In particular, Kwon et al. documented in [24] the presence of outliers amidst a strong Gaussian component in realized errors based on ranging via received signal strength. Statistical filtering was then used to mitigate the detrimental effects of the outliers. Because of the strong Gaussian component noted in the empirical data, we believe the noise model presented by this work to provide a good baseline for the initial validation of the proposed scheme that assumes an effective statistical method of outlier abatement has been implemented. However, as with all methods initially validated in simulation, we acknowledge experimental migration to real world data as a natural extension of our work.

3.3. Database Reconstruction

The first portion of this study examines the appropriateness of the formulation of (6). The result of a representative reconstruction is shown in Figure 2. Here the sub-sampling is chosen to be conducted at a very low rate, specifically \( n = 40 \) in order to show the excellent reconstruction possible. The first striking observation is the level of fidelity that can be realized while collecting under 30% of the available information. While reconstructions vary somewhat due to the random nature in which the subsampling is conducted, this repeatable result shows that the problem is well formed, stable, and highly accurate even under disadvantageous conditions.

To better quantify the result we show the mean squared error (MSE) between the reconstructed database and the original in Figure 3. Here we use the Frobenius norm to this end

\[
\text{MSE} = \| X - \hat{X} \|_F .
\]

With the understanding that MSE is not a perfect measurement of a better database than another, its comparison of the error residual is an appropriate one in this study since ultimately the formulation of the likelihood vector in (3) is based on minimization of mean squared error. The results presented are intuitive. As the subsampling rate decreases the MSE increases.
establishing the inversely proportional relationship expected. The compelling result here is that the behavior does not demonstrate quadratic or exponential behavior but has more linear behavior.

3.4. Positioning Performance Based on Reconstructed Databases

With the general trend of MSE of the database in mind we turn to its effect on the position accuracy. The fingerprinting performance offered at the hands of the reconstructed database for various rates of subsampling relative to the fully constructed database is shown in Figure 4. Here we notice the dramatic level of positioning accuracy available with severely subsampled databases. Even at a subsampling rate of \( n = 40 \), we note similar performance to the fully sampled database. At \( n = 80 \) samples and above we note essentially the same performance realized by a fully sampled database with approximately the exact same circular error probable (CEP) up to 60% and also at and above roughly 95%.

In contrast to the results presented for larger \( n \) we note the poor performance seen in the case of extreme subsampling in which there is not sufficient information to correlate an acceptable interpolation. This steep drop in performance between \( n = 40 \) and \( n = 10 \) in this simulation suggests a lower bound on the amount of information that must be present in order to obtain an acceptable \( X \).

4. Conclusions

In this paper we have presented a simple and effective means for geolocation in wireless 4G networks. Due to the requirement for deployment in dense multi-path environments such as indoors and urban canyons, fingerprinting was chosen for the locationing framework. The scheme proposed reduces the otherwise unattractive complexity of the overall fingerprinting model in two ways. First, it significantly reduces the initial costs associated with database construction via subsampling fingerprints in the tracking area. Missing information is then recovered via a computationally efficient multi-channel filtering scheme. Second, the location problem, formulated as sparse signal recovery in overcomplete dictionaries, is solved with a modified \( k \)-NN approach. This stands in contradistinction to other more computationally costly formulations such as convex optimization which has been proposed in the literature. Throughout, we have maintained an ardent deference to positional accuracy. While some accuracy must be capitulated in the spirit of minimizing computational burden, we have shown that this sacrifice can be minimized in the face of tremendous gains.

Our approach was validated in a simulated environment modeled after a real network deployment. The excellent results expected were corroborated with those empirically obtained via monte carlo methods. First, the intuition of a monotonically increasing MSE against increasing severity of subsampling was validated. Concordantly, the relationship between subsampling and the expected fidelity in position estimation was established. We showed that excellent positional accuracy can be maintained despite a massively subsampled database and is principal among other results. Specifically, we noted similar performance above CEP 95% at subsampling rates of \( n = 80 \) and comparable performance at the subsampling rates of \( n = 40 \).

This study has demonstrated a simple efficient solution to the fingerprinting problem that can easily be deployed indoors or in urban environments with no hardware or system level modifications to existing infrastructure. The computational burden of the scheme is light, the level of offline effort is minimized, and the fidelity in position estimation is high. Further, it is derivative of highly acclaimed and proven techniques which provide a solid theoretical foundation upon which to stand. This work can be seen as a significant step forward in fingerprint positioning in dense multipath environments due to its characteristic simplicity, accuracy, and robustness.
References


