EUCLIDEAN MATRIX COMPLETION PROBLEMS IN TRACKING AND GEO-LOCALIZATION

Raghuram Rangarajan, Raviv Raich, and Alfred O. Hero III

Department of EECS, University of Michigan, Ann Arbor, MI 48109-2122, USA
{rangaraj, ravivr, hero}@eecs.umich.edu

ABSTRACT

We consider the problem of emitter tracking using received signal strengths (RSS) measured at a number of in-range access points (AP) when some of the AP locations are unknown. This can be formulated as a Euclidean distance matrix completion problem (EDMCP) to which an iterative distributed weighted multidimensional scaling (dwMDS) algorithm can be applied to simultaneously track emitters and localize APs. The algorithm is illustrated using real-time data collected by the University of California San Diego (UCSD) wireless topology discovery (WTD) project.

Index Terms— distributed multidimensional scaling, sparsity constrained tracking, wireless mobility.

1. INTRODUCTION

Wireless localization and tracking has attracted tremendous interest from a wide range of sectors such as security, ecology, property control, and targeted marketing. Accurate location of targets can facilitate a number of location based services in these domains. For example, in perimeter surveillance a network of RF sensors can be used to estimate the location of intruders in the network. For retail stores, these services can be used to locate equipment or inventory in a warehouse or can advertise different products to users based on their location in the store. The problem of estimating target locations based on range information (e.g., RSS or time of arrival) has been an active area of research during the last decade.

Mulltilateration can be used to locate energy emitting targets based on the range information provided by the targets to multiple APs or other receiving sensors. When there are many users or when some of the locations of APs in range of the targets are not known, the problem of estimating target locations can be formulated as an Euclidean distance matrix completion problem. Previous approaches for solving the EDMCP use semi-definite programming methods [1] which are not scalable. In this paper, we propose to use the sparsity penalized distributed weighted multidimensional scaling (dwMDS) algorithm introduced in [2] to solve the EDMCP. Furthermore, we show how this general solution can be applied to geo-localize the unknown AP locations in the wireless network in addition to obtaining target coordinates. We conclude with a tracking illustration of our algorithm on real-time data sets from the WTD project at the UCSD campus.

2. PROBLEM FORMULATION

The problem of multi-emitter tracking can be represented as follows. Let $\{\mathbf{x}_i\}_{i=1}^N$ denote the unknown locations of N users in the wireless network. APs in the network measure RSS values from these users. Let $\{\mathbf{k}_i\}_{i=1}^M$ and $\{\mathbf{u}_i\}_{i=1}^P$ be the locations of the APs, where the first M AP locations are known and the last P locations are unknown. Denote the RSS measurements between target i and AP j as $\text{RSS}_{i,j}$. At any time, a target i is only in range of a small subset of the APs and hence the matrix of RSS values is only partially known. Our goal is to estimate the locations of the targets $\{\mathbf{x}_i\}$ and the locations of the unknown APs $\{\mathbf{u}_i\}$ given these RSS measurements. A pictorial representation of the problem is shown in Figure 1. The fundamental question is: given the knowledge of connectivity (edges) of only some edges of a Euclidean network, can one recover the connectivity of the entire network?

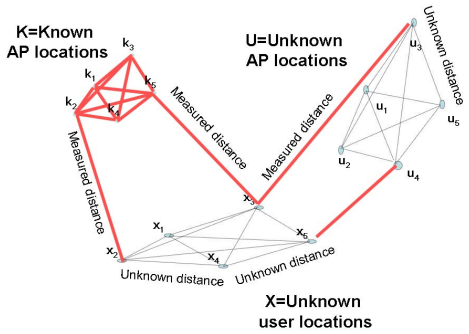

Fig. 1. Wireless users X and access points (AP) U are at unknown positions while access points K are at known positions. Given measured distances between X and K, X and U, and K and K the Euclidean matrix completion problem is to recover all pairwise distances between X, K and U.

3. PROPOSED SOLUTION

Denote the distance between emitter i and AP j as $d_{i,j}$. Here $i \in \{1, \ldots, N\}$ and j belongs to $\{1, \ldots, M, M+1, \ldots, M+P\}$ corresponding to the M known and P unknown APs. In many environments it can be shown that the RSS is approximately log-normal in its distribution [3], i.e., if the received RMS power $P_{i,j}$ is in milliWatts, then $\text{RSS}_{i,j} = 10\log_{10}(\text{P}_{i,j})$ is approximately Gaussian. Thus $\text{RSS}_{i,j}$ in dBm is typically modeled as

$$\text{RSS}_{i,j} \sim \mathcal{N}(\text{RSS}_0 - 10n_p \log\left(\frac{d_{i,j}}{d_0}\right), \sigma_0^2) \qquad (1)$$

This research was partially supported by NSF grant CCR-0325571.

where σ_0 is the standard deviation of the received power in dBm and RSS_0 is received power in dBm at a reference distance d_0. The constant n_p is referred to as the path-loss exponent and it depends on the environment. We use single observation maximum likelihood estimation to compute the range, $d_{i,j}$ from $\text{RSS}_{i,j}$. The maximum likelihood estimator of $d_{i,j}$ is given by

$$\delta_{i,j} = d_0 10^{\left((\text{RSS}_0 - \text{RSS}_{i,j})/10n_p\right)}. \quad (2)$$

The complete inter-point distance matrix $\mathbf{D}$, called the Euclidean distance matrix, is a symmetric matrix of the form

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_{kk} & \mathbf{D}_{kx} & \mathbf{D}_{ku} \\ \mathbf{D}_{xk} & \mathbf{D}_{xx} & \mathbf{D}_{xu} \\ \mathbf{D}_{uk} & \mathbf{D}_{ux} & \mathbf{D}_{uu} \end{pmatrix}, \quad (3)$$

where $\mathbf{D}_{kk}$ is the distance matrix between the known AP locations, $\mathbf{D}_{kx} = \mathbf{D}_{xk}^T$ is the distance matrix between the known AP locations and the users, and the rest of the sub-distance matrices are similarly defined. Among these sub-matrices, only $\mathbf{D}_{kk}$ is completely known. The distance matrices $\mathbf{D}_{kx}$ and $\mathbf{D}_{ux}$ are partially known since different users are within range of only a fraction of the known and unknown APs. For any matrix $\mathbf{D}$ define $\mathbf{D}^*$ as the matrix $\mathbf{D}$ with some of its entries deleted and define $*$ as the matrix $\mathbf{D}$ with all of its entries deleted. With this notation, if the exponent n_p is known then the incomplete Euclidean matrix of distances can be recovered from the noiseless RSS measurements

$$\mathbf{D}^* = \begin{pmatrix} \mathbf{D}_{kk} & \mathbf{D}_{kx}^* & * \\ \mathbf{D}_{xk}^* & * & \mathbf{D}_{xu}^* \\ * & \mathbf{D}_{ux}^* & * \end{pmatrix}. \quad (4)$$

The objective is to reconstruct the complete distance matrix $\mathbf{D}$ in (3) from the RSS measurements. Once $\mathbf{D}$ is recovered we can recover the user tracks and unknown AP locations. We consider two cases: (a) the exact completion problem, i.e., when the distances in $\mathbf{D}^*$ are available; (b) the approximate completion problem, i.e., when only a noise contaminated version of $\mathbf{D}^*$ is available.

3.1. Exact completion problem

In the noiseless case ($\sigma_0 = 0$) the problem of localizing the unknown APs and the targets can be formulated as an Euclidean distance matrix completion problem (EDMCP) [4]. This is a classical problem in geometry and can be stated as the problem of recovering the set of all pairwise distances between n points given only a subset of these distances. A solution exists when a sufficient number of entries of $\mathbf{D}$ are specified.

Specifically, let $\mathbf{A} = ((a_{i,j}))$ be an $N \times N$ partial distance matrix in $\mathbb{R}^m$. Let $G = (V, E)$ be an undirected graph with $V = 1, 2, \ldots, N$, $E = \{(i,j) \mid a_{i,j} \text{is specified}\}$, and whose specified entries are chordal.

Definition: A graph is *chordal* if each of its cycles of four or more vertices has a chord, which is an edge joining two nodes that are not adjacent in the cycle.

Theorem 3.1. *[4] Every partial distance matrix in $\mathbb{R}^m$, the graph of whose specified entries is chordal, admits a completion to a distance matrix in $\mathbb{R}^m$. The matrix completion is unique if and only if*

$$\mathbf{B} = \begin{pmatrix} 0 & \mathbf{e}^T \\ \mathbf{e} & \mathbf{A}(S) \end{pmatrix} \textit{ has rank } m+2 \textit{ for any } S \in \mathcal{S}, \quad (5)$$

where $\mathcal{S}$ is the collection of all minimal vertex separators of G, $\mathbf{e}$ is a column vector of ones and $\mathbf{A}(S)$ is a matrix formed by using the set of vertices in S.

A solution to the exact completion problem for an arbitrary known set of partial matrix entries is not generally in closed-form [5]. A special case for which a closed form solution does exist is given in the following result.

Theorem 3.2. *Let $\mathbf{A}$ be a $(N+M) \times (N+M)$ partial distance matrix with rank $m+2$ and the following structure:*

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}, \quad (6)$$

where $\mathbf{A}_{11}$ is $N \times N$, $\mathbf{A}_{12}$ is $N \times M$, and $\mathbf{A}_{22}$ is an $M \times M$ matrix. Given, $\mathbf{A}_{11}, \mathbf{A}_{12}$, there exists an unique Euclidean matrix completion to $\mathbf{A}$ given by $\mathbf{A}_{22} = \mathbf{A}_{21}\mathbf{A}_{11}^+\mathbf{A}_{12}$ if rank$(\mathbf{A}_{11}) = m+2$.

Proof. The set of nonadjacent vertices for the partially complete distance matrix $\mathbf{A}$ defined in (6) is given by $N_v = \{(i,j) \mid N+1 \leq i, j \leq N+M\}$. The corresponding collection of minimal vertex separators of the graph G is a singleton set $S = \{1, 2, \ldots, N\}$. This guarantees that the graph associated with the entries of $\mathbf{A}$ is chordal. As $N \geq m+2$, $\mathbf{B}$ in (5) has rank $m+2$ if $\mathbf{A}_{11}$ has rank $m+2$. Let $r = m+2$. From Theorem 3.1, there exists a unique solution to the exact completion problem. The eigendecomposition of $\mathbf{A}$ is given by

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T,$$

where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_r)$ and $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r]$ are the corresponding set of orthonormal eigenvectors. Let $\mathbf{V} = [\mathbf{V}_1^T \mathbf{V}_2^T]^T$, where $\mathbf{V}_1$ is $N \times r$ and $\mathbf{V}_2$ is $M \times r$. Then the sub-matrices can be written as

$$\begin{aligned} \mathbf{A}_{11} &= \mathbf{V}_1\mathbf{\Lambda}\mathbf{V}_1^T \\ \mathbf{A}_{12} &= \mathbf{V}_1\mathbf{\Lambda}\mathbf{V}_2^T \\ \mathbf{A}_{22} &= \mathbf{V}_2\mathbf{\Lambda}\mathbf{V}_2^T \end{aligned}$$

Since $\mathbf{A}_{11}$ has rank $m+2$, the pseudo inverse of $\mathbf{A}_{11}$ can be written as

$$\mathbf{A}_{11}^+ = \mathbf{V}_1(\mathbf{V}_1^T\mathbf{V}_1)^{-1}\mathbf{\Lambda}^{-1}(\mathbf{V}_1^T\mathbf{V}_1)^{-1}\mathbf{V}_1^T.$$

Then

$$\begin{aligned} &\mathbf{A}_{21}\mathbf{A}_{11}^+\mathbf{A}_{12} \\ &= \mathbf{V}_2\mathbf{\Lambda}\mathbf{V}_1^T \left(\mathbf{V}_1(\mathbf{V}_1^T\mathbf{V}_1)^{-1}\mathbf{\Lambda}^{-1}(\mathbf{V}_1^T\mathbf{V}_1)^{-1}\mathbf{V}_1^T\right)\mathbf{V}_1\mathbf{\Lambda}\mathbf{V}_2^T \\ &= \mathbf{V}_2\mathbf{\Lambda}\mathbf{V}_2^T = \mathbf{A}_{22}. \end{aligned}$$

□

Theorem 3.2 does not apply when all the points yielding the partial matrix $\mathbf{A}_{11}$ lie on a m-dimensional sphere such that rank$(\mathbf{A}_{11})$ is $m+1$. It is easy to verify that a solution exists for this case using Theorem 3.1, which is in fact unique, but Theorem 3.2 does not yield the optimal completion.

This theorem can be applied to completion of the partial matrix $\mathbf{D}^*$ defined in (4). Assume that rank$(\mathbf{D}_{kk}) \geq m+2$, which requires the number M of known AP positions to be no less than $m+2$. Then a two step procedure recovers $\mathbf{D}$ from $\mathbf{D}^*$ when $\mathbf{D}_{kx}^*$ and $\mathbf{D}_{ux}^*$ are completely known. First recover $\mathbf{D}_{xx}$ from $\mathbf{D}_{kk}$ and $\mathbf{D}_{kx}$ by applying Thm. 3.2 to the upper left $(M+N) \times (M+N)$ submatrix of $\mathbf{D}^*$. Then, plugging the solution of $\mathbf{D}_{xx}$ into its place in the upper left $(M+N) \times (M+N)$ submatrix of $\mathbf{D}^*$ solve for $\mathbf{D}_{uu}$ and $\mathbf{D}_{ku} = \mathbf{D}_{uk}^T$ together by a second application of the theorem. Thus in this case all unknown entries in $\mathbf{D}$ can be recovered from $\mathbf{D}^*$.

3.2. Approximate completion problem

When only noisy measurements $\mathbf{A} = \mathbf{D}^* + \mathbf{N}$ of the partial distance matrix $\mathbf{D}^*$ are available an approximation to the Euclidean distance matrix $\mathbf{D}$ can be obtained by formulating a nonlinear least squares problem. Let $\mathbf{W}$ be a symmetric weight matrix with nonnegative elements, e.g., $w_{i,j} = 1$ if $a_{i,j}$ is given and zero otherwise. Consider the Frobenius norm minimization

$$\min_{\mathbf{D}} \|\mathbf{W} \circ (\mathbf{A} - \mathbf{D})\|_F^2 \qquad \text{s.t} \qquad \mathbf{D} \in \mathcal{D}, \tag{7}$$

where $\circ$ denotes the Hadamard product and $\mathcal{D}$ is the convex cone of Euclidean distance matrices. The objective function can be rewritten as

$$\min_{\mathbf{D}} \sum_{i,j} w_{i,j} \left(a_{i,j} - d_{i,j}\right)^2, \qquad \text{s.t} \qquad \mathbf{D} \in \mathcal{D}. \tag{8}$$

A semi-definite programming solution to this problem is provided in [1] that does not scale well due to its high computational complexity in the number of locations $N + M + P$. A lower complexity iterative algorithm is the dwMDS algorithm of Costa [6] and it directly yields location estimates by minimizing the equivalent objective function over location vectors $\mathbf{z}_i$:

$$\min_{\mathbf{Z}} \sum_{i,j} w_{i,j} \left(a_{i,j} - \|\mathbf{z}_i - \mathbf{z}_j\|\right)^2, \tag{9}$$

where $\mathbf{Z} = [\mathbf{k}_1, \ldots, \mathbf{k}_M, \mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{u}_1 \ldots, \mathbf{u}_P]$ is the set of M known positions $\mathbf{k}_i$ of the APs, the set of N unknown user positions $\mathbf{x}_i$, and the set of P unknown AP positions $\mathbf{u}_j$. As in standard MDS, the solution to (9) is not unique since rotations and translations of the location matrix leave the objective function invariant. However, by constraining the positions $\mathbf{k}_i$ to be equal to the known AP positions (anchor nodes), as long as these known AP positions do not lie on a plane or a line, a unique solution can often be found [6].

The dwMDS algorithm was reformulated in [7], [2] for target tracking by introducing a sparsity penalty on changes in the user part ($\{\mathbf{x}_i\}$) of the solution to (9) over time. This sparse dwMDS algorithm is a distributed iterative procedure that can be implemented in a decentralized in-network manner. Furthermore, the associated weights $w_{i,j}$ can be selected to emphasize more accurate RSS measurements and de-emphasize others. We adopt a two-step version of the dwMDS procedure to successively localize the unknown users and APs that mimics the two step exact matrix completion procedure discussed above. First dwMDS is implemented to recover tracks $\{\mathbf{x}_i\}$ without using the RSSs measured with respect to the unknown AP's (U in Figure 1) then, substituting the estimated track locations into the partial distance matrix we rerun dwMDS to recover the unknown AP locations $\{\mathbf{u}_i\}$.

4. APPLICATION TO UCSD WTD DATA

4.1. UCSD wireless trace data

The wireless topology discovery (WTD) project[1] was undertaken by researchers at UCSD [8]. The project collects data on dynamic characteristics and user behavior in a real world wireless network. The primary objective of the WTD project was to test and develop reliable routing protocols in a geographically constrained wireless network. To collect data, 275 UCSD freshman were given HP Jordana PDAs which were equipped with symbol 802.11 compact flash cards and the WTD data collection software. The software recorded all

[1] http://sysnet.ucsd.edu/wtd/

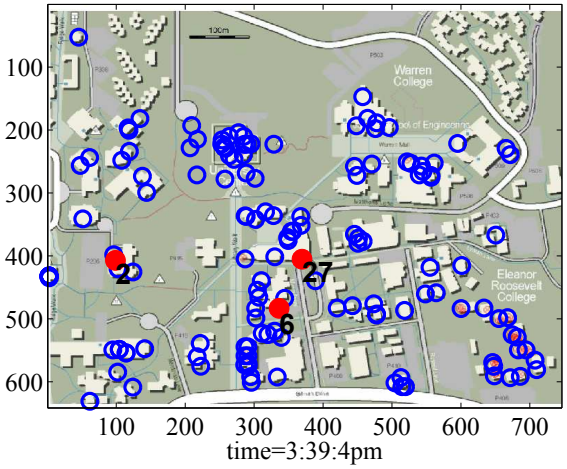

Fig. 2. Campus map showing the 200 known AP locations. Only the horizontal (x,y) part of the 3D coordinates (x,y,z) is shown. In addition to these locations, there were 100 APs at unknown locations. Also shown is a sample RSS data from a single user to APs. The sensed APs by the user are indicated in blocked red along with their RSS measurements.

APs sensed by the user every 20 seconds. The trace data were collected over a 11 week period and then transferred to a centralized database for analysis. The data indicated that around 300 APs were sensed over this time among which only 200 of them had knowledge about their locations.

The trace data collected consists of the following information: user identity, sample time, AP identity, RSS, and AC/battery power indicator. The coordinates of the known APs were also provided in the database. Our objective was to recover user trajectories over time and to reconstruct the network topology (locations of the unknown APs) using available data. The map of the known AP locations on the UCSD campus and the data samples at a particular time instant for a single user are shown in Fig. 2. The APs sensed by the user are shown as filled circles. The corresponding RSS values (2, 6, 27) are shown next to these APs.

To reconstruct the user trajectories, we need to estimate user locations over time based on measured RSS at the various access points. The RSS values provided in the database were 5-bit quantized values between 0 and 31. Since the mapping from these quantized values to the actual signal strength in dB is unknown, we needed to calibrate the quantized numbers to signal strength values. This was performed using an iterative least squares procedure that yielded the best linear fit between the quantized values and the corresponding RSS value in dB [9].

5. SIMULTANEOUS LOCALIZATION OF TARGETS AND APS

First, we consider all users that sense at least 4 known APs and one unknown AP. Using only the knowledge of the RSS values between the users and the known APs, we estimate the locations of the users in the network. We then use the set of user locations with the corresponding RSS values to the unknown APs to estimate the location of the unknown APs. As a validation of our two stage dwMDS al-

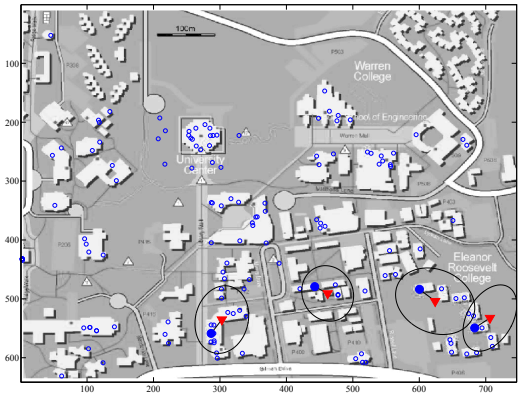

Fig. 3. Location estimates of three known APs that were made unknown to the two stage dwMDS algorithm and the corresponding uncertainty ellipses. Only horizontal coordinates of the 3D coordinate estimates are shown.

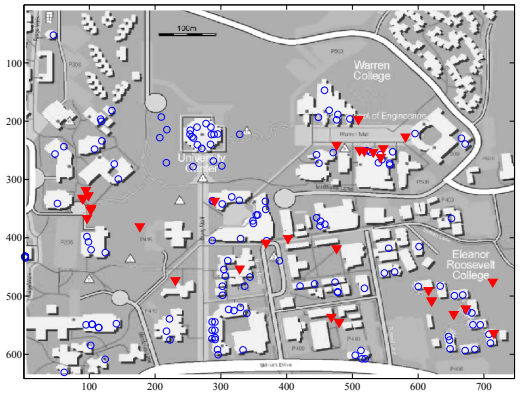

Fig. 4. Location estimates of the unknown APs are shown in red. The known AP locations are shown in blue.

gorithm for localization of targets and APs we performed the following experiment. We randomly choose a small set of known APs K and added them to the set U of unknown APs. We reconstruct the location of the APs in the augmented set U using different user trajectories in the sparsity penalized dwMDS algorithm of [2]. The knowledge of some of the locations in the augmented set U allows us to indirectly measure the RMS error of the reconstruction of the unknown AP locations (see Fig. 3).

In Fig. 3, the known AP locations in the augmented U set are indicated by filled circles. The mean estimates of the APs are shown as triangles. The black ellipse is the standard error ellipse for the unknown AP location. The error in the mean estimate of the AP location is roughly 30m. We would like to emphasize that these results were obtained despite the following limitations: inaccurate model (n_p may vary from one link to another), 5 bit quantization of the RSS, low sampling rate (one sample every 20 seconds), and measurement noise. The actual estimated locations of the unknown APs are shown in Fig. 4. The known AP locations are indicated as circles while the location estimates of the unknown APs are shown as triangles. These result indicate that dwMDS can perform robust geo-localization even in the absence of information about some of the links.

6. CONCLUSIONS

In this paper, we considered the problem of multitarget localization and tracking in a sensor (AP) network having some unknown sensor locations. For the case where there is no noise in the RSS measurements we formulated this problem as a Euclidean distance matrix completion problem (EDMCP) and proposed a two step exact completion algorithm to recover the complete pairwise distance matrix between all sensors and targets under some conditions on the rank of the distance matrix. The target positions and the unknown sensor (AP) locations can be recovered from the completed distance matrix up to a rotation and translation. We then showed how a two stage constrained dwMDS algorithm can be implemented to directly recover the unknown locations. The sparsity constrained dwMDS algorithm was illustrated for tracking wireless PDA users in the UCSD WTD data set.

We intend to explore the following possibilities to improve upon the localization performance in the UCSD WTD data set. (1) to perform multiple local fits to the RSS model rather than a single global fit. This would allow us to adaptively fit the log-normal model for each AP based on its relative location in the network. (2) to include side information, e.g., smoothness of user trajectories or topographic maps of the UCSD campus.

7. REFERENCES

[1] S. Al-Homidan and H. Wolkowicz, "Approximate and exact completion problems for Euclidean distance matrices using semidefinite programming," *Linear Algebra and its Applications*, vol. 406, pp. 109–141, 2005.

[2] R. Rangarajan, R. Raich, and A. O. Hero III, "Blind tracking using sparsity penalized multidimensional scaling," in *Proc. IEEE Workshop on Stat. Signal Processing*, Aug 2007, pp. 670–674.

[3] A. J. Coulson, A. G. Williamson, and R. G. Vaughan, "A statistical basis for lognormal shadowing effects in multipath fading channels," *IEEE Trans. on Veh. Tech.*, vol. 46, no. 4, pp. 494–502, 1998.

[4] M. Bakonyi and C. R. Johnson, "The Euclidean distance matrix completion problem," *SIAM J. Matrix. Anal. Appl.*, vol. 16, no. 2, pp. 646–654, April 1995.

[5] M. W. Trosset, "Distance matrix completion by numerical optimization," Rice University, Houston, Texas, Tech. Rep., 1997.

[6] J. Costa, N. Patwari, and A. O. Hero III, "Distributed multidimensional scaling with adaptive weighting for node localization in sensor networks," *ACM J. Sensor Networking*, vol. 2, no. 1, pp. 39–64, 2006.

[7] R. Rangarajan, R. Raich, and A. O. Hero III, "Sparsity penalized MDS for blind tracking in sensor networks," in *Advances in Sensor Networks*, V. Saligrama, Ed. NY: Springer, 2007.

[8] M. McNett and G. M. Voelker, "Access and mobility of wireless PDA users," *Mobile Computing Communications Review*, vol. 9, no. 2, pp. 40–55, Apr 2005.

[9] R. Rangarajan, "Resource constrained adaptive sensing," Ph.D. dissertation, University of Michigan, Ann Arbor, MI, July 2007.