Sensor Network Localization from Local Connectivity: Performance Analysis for the MDS-MAP Algorithm

Sewoong Oh and Andrea Montanari†
Electrical Engineering and Statistics Department
Stanford University, Stanford, CA 94304
swoh@stanford.edu, montanari@stanford.edu

Amin Karbasi
LTHC and LCAV
EPFL, Lausanne, Switzerland
amin.karbasi@epfl.ch

Abstract—Sensor localization from only connectivity information is a highly challenging problem. To this end, our result for the first time establishes an analytic bound on the performance of the popular MDS-MAP algorithm based on multidimensional scaling. For a network consisting of $n$ sensors positioned randomly on a unit square and a given radio range $r = o(1)$, we show that resulting error is bounded, decreasing at a rate that is inversely proportional to $r$, when only connectivity information is given. The same bound holds for the range-based model, when we have an approximate measurements for the distances, and the same algorithm can be applied without any modification.

I. INTRODUCTION

Localization plays an important role in wireless sensor networks when the positions of the nodes are not provided in advance. One way to acquire the positions is by equipping all the sensors with a global positioning system. Clearly, this adds considerable cost to the system. As an alternative, many research efforts have focused on creating algorithms that can derive positions of sensors based on basic local information such as proximity (which nodes are within communication range of each other) or local distances (pairwise distances between neighboring sensors).

The problem is solvable, meaning that it has a unique set of coordinates satisfying the given local information, only if there are enough constraints. Note that based on local information any solution is unique only up to rigid transformations (rotations, reflections and translations).

The simplest of such algorithms assumes that all pairwise distances are known. However, it is well known that in practical scenarios such information is unavailable and far away sensors (the ones outside the communication range) cannot obtain their pairwise distances. Many algorithms have been proposed to resolve this issue by using heuristic approximations to the missing distances where the success of them has mostly been measured experimentally.

Of the same interest and complementary in nature, however, are the theoretical guarantees associated with the performance of the existing methods. Such analytical bounds on the performance of localization algorithms can provide answers to practical questions: for example, how large should the radio range be in order to get the error within a threshold? With this motivation in mind, our work takes a forward step in this direction. In particular, our analysis focuses on providing a bound on the performance of the popular MDS-MAP [1] algorithm when applied to sensor localization from only connectivity information, which is a highly challenging problem. We prove for the first time that using MDS-MAP, we are able to localize sensors up to a bounded error in a connected network where most of pairwise distances are missing and only local connectivity information is given.

A. Related work

A great deal of work over the past few years has focused on the sensor localization problem. A detailed survey of this area is provided in [2].

In the case when all pairwise distances are known, the coordinates can be derived by using a classical method known as multidimensional scaling (MDS) [3], [4]. The underlying principle of the MDS is to convert distances into an inner product matrix, whose eigenvectors are the unknown coordinates. In practice, MDS tolerates errors gracefully due to the overdetermined nature of the solution. However, when most pairwise distances are missing, the problem of finding the unknown coordinates becomes more challenging. Two types of practical solutions to this problem have been proposed. The first group consists of algorithms who try first to estimate the missing entries of the distance matrix and then apply MDS to the reconstructed distance matrix to find the coordinates of the sensors. MDS-MAP introduced in [1] can be mentioned as a well known example of this class where it computes the shortest paths between all pairs of nodes in order to approximate the missing entries of the distance matrix. The algorithms in the second group mainly consider the sensor localization as a non-
convex optimization problem and directly estimate the coordinates of sensors. A famous example of this type is a relaxation to semidefinite programming (SDP) [5].

The performance of these practical algorithms are invariably measured through simulations and little is known about the theoretical analysis supporting their results. One exception is the work done in [6] where they use matrix completion methods as a mean to reconstruct the distance matrix. The main contribution of their paper is that they are able to provably localize the sensors up to a bounded error. However, their analysis is based on a number of strong assumptions. First, they assume that even far away sensors have non-zero probability of detecting their distances. Second, the algorithm explicitly requires the knowledge of detection probabilities between all pairs. Third, their theorem only works when the average degree of the network (i.e., the average number of nodes detected by each sensor) grows linearly with the number of sensors in the network.

Our work has a similar flavor as [6], namely we provide theoretical guarantee that backs up experimental results. We use shortest paths (the same way it is used in MDS-MAP) as our primary guess for the missing entries in the distance matrix and apply MDS to find the topology of the network. In contrast to [6], we require weaker assumptions for our result. More specifically, we assume that only neighboring sensors have information about each other and that only connectivity information is known. Furthermore, the knowledge of detection probabilities plays no role in our analysis or the algorithm. And last, in our analysis we assume that the average degree grows logarithmically with the number of sensors not linearly, which results in needing much less revealed entries in the distance matrix. In particular, the last condition is quite realistic: if the average degree grows any slower then the network is not even connected. Due to the fact that the shortest paths algorithm works for both range-free and range-aware cases, our analysis includes both and provides the first error bounds on the performance of MDS-MAP.

The organization of this paper is as follows. Section II describes the state of the art algorithm known as MDS-MAP and states the main result on the performance of MDS-MAP. In Section III, we provide detailed proof of the main theorem.

II. LOCALIZATION USING MULTIDIMENSIONAL SCALING

In this section, we first define the basic mathematical model for the sensors and the measurements and state preliminary facts on MDS, a technique to derive sensor configuration from pairwise distances [1]. We will then describe the algorithm and give a bound on the error between \(X\) and \(\hat{X}\), the estimate returned by the algorithm.

A. Model definition

For the sensor localization problem, we consider the random geometric graph model \(G(n, r) = (V, E, P)\) where \(V\) is a set of \(n\) vertices that correspond to \(n\) sensor nodes which are distributed uniformly at random within the \(d\)-dimensional hypercube \([0, 1]^d\), \(E \subseteq V \times V\) is a set of undirected edges that connect pairs of sensors which are close to each other, and \(P : E \to \mathbb{R}_+\) is a non-negative real-valued function. We consider the function \(P\) as a mapping from a pair of connected nodes \((i, j)\) to the approximate measurement for the distance between \(i\) and \(j\), which we call the proximity measurement. Let \(||\cdot||\) be the Euclidean norm on \(\mathbb{R}^d\) and let \(r\) be some positive parameter. Given a set of random positions of \(n\) sensors \(X = \{x_1, \ldots, x_n\}\), where \(x_i\) is the position of sensor \(i\) in \(d\)-dimensional space, a common model for the random geometric graph is the disc model where node \(i\) and \(j\) are connected if the Euclidean distance \(d_{i,j} = ||x_i - x_j||\) is less than \(r\). More formally,

\[(i, j) \in E \iff d_{i,j} \leq r.\]

To each edge \((i, j) \in E\), we associate the proximity measurement \(P_{i,j}\) between sensors \(i\) and \(j\), which is a function of the distance \(d_{i,j}\) and random noise. In an ideal case when our measurements are exact, we have \(P_{i,j} = d_{i,j}\). On the other extreme, when we are given only network connectivity information and no distance information, we have \(P_{i,j} = r\) for all \((i, j) \in E\).

The sensor localization algorithms can be classified into two different categories. In the first, only connectivity information is available (connectivity-based). In the second category, however, approximate measurement for the distance between neighboring sensors are available, perhaps with limited accuracy (range-based). For the connectivity-based model, which is alternatively also known as the range-free model, only the connectivity information is available. Formally,

\[P_{i,j} = r \quad \text{if} \quad (i, j) \in E.\]

For the range-based model, which is also known as the range-aware model, the distance measurement is available but may be corrupted by noise or have limited accuracy.

\[P_{i,j} = [d_{i,j} + z_{i,j}]_+ \quad \text{if} \quad (i, j) \in E,\]

where \(z_{i,j}\) models the measurement noise (in the noiseless case \(z_{i,j} = 0\), possibly a function of the distance \(d_{i,j}\), and \([a]_+ \equiv \max\{0, a\}\). Common examples are the additive Gaussian noise model, where the \(z_{i,j}\)’s are i.i.d. Gaussian random variables, and the multiplicative noise model, where \(P_{i,j} = [(1 + N_{i,j})d_{i,j}]_+\), for independent Gaussian random variables \(N_{i,j}\)’s.

In the following, let \(D\) denote the \(n \times n\) squared distance matrix where \(D_{ij} = d_{i,j}^2\) and \(X\) denote the
$n \times d$ position matrix where the $i$th row corresponds to $x_i$, the $d$-dimensional position vector of sensor $i$. Given the graph $G(n, r) = (V, E, P)$ with associated proximity measurements for each edges in $E$, the goal of sensor network localization is to find an embedding of the nodes that best fits the measured proximity between all the pairs in $E$. The notion of embedding, or configuration, will be made clear in the following sections.

The main contribution of this paper is that we provide, for the first time, the performance bound for the MDS-MAP algorithm when only the connectivity information is available as in the connectivity-based model. However, the algorithm can be readily applied to the range-based model without any modification. Further, given $G(V, E, P)$ from the range-based model we can easily produce $G'(V, E, P')$ where $P'_{ij} = r$ for all $(i, j) \in E$. This implies that the performance of MDS-MAP under the range-based model is also bounded by the same bound.

B. Multidimensional scaling (MDS)

MDS is a technique used in finding the configuration of objects in a low dimensional space such that the measured pairwise distances are preserved. If all the pairwise distances are measured without error then a naive application of MDS exactly recovers the configuration of sensors [1].

Algorithm : Classical Metric MDS [1]

**Input:** dimension $d$, estimated distance matrix $M$

1: Compute $(-1/2)LM\Sigma L$,
    where $L = I_n - (1/n)1_n1_n^T$;
2: Compute the best rank-$d$ approximation $U_d\Sigma_d U_d^T$
    of $(-1/2)LM\Sigma L$;
3: Return $MDS_d(M) \equiv U_d\Sigma_d^{1/2}$.

There are many types of MDS techniques, but, throughout this paper, whenever we say MDS we refer to the classical metric MDS, which is defined as follows. Let $L$ be an $n \times n$ symmetric matrix such that $L = I_n - (1/n)1_n1_n^T$, where $1_n \in \mathbb{R}^n$ is the all ones vector and $I_n$ is the $n \times n$ identity matrix. Let $MDS_d(D)$ denote the $n \times d$ matrix returned by MDS when applied to the squared distance matrix $D$. Then, in formula, given the singular value decomposition (SVD) of a symmetric and positive definite matrix $(-1/2)LD$L as $(-1/2)LD$L $\equiv U\Sigma U^T$, $MDS_d(D) \equiv U_d\Sigma_d^{1/2}$, where $U_d$ denotes the $n \times d$ left singular matrix corresponding to the $d$ largest singular values and $\Sigma_d$ denotes the $d \times d$ diagonal matrix with $d$ largest singular values in the diagonal. This is also known as the MDS-LOCALIZE algorithm in [6]. Note that since the columns of $U$ are orthogonal to $1_n$, by construction, it follow that $L \cdot MDS_d(D) = MDS_d(D)$.

It can be easily shown that when MDS is applied to the correct squared distance matrix without noise, the configuration of sensors is exactly recovered [6]. This follows from the following equality

$$-(1/2)LDL = LXX^TL.$$  

Note that only we get the configuration and not the absolute positions, in the sense that $MDS_d(D)$ is one version of infinitely many solutions that matches the distance measurements $D$. Intuitively, it is clear that the pairwise distances are invariant to a rigid transformation (a combination of rotation, reflection and translation) of the positions $X$, and therefore there are multiple incidents of $X$ that result in the same $D$. For future use, we introduce a formal definition of rigid transformation and related terms.

Denote by $O(d)$ the orthogonal group of $d \times d$ matrices. A set of sensor positions $Y \in \mathbb{R}^{n \times d}$ is a rigid transform of $X$, if there exists a $d$-dimensional shift vector $s$ and an orthogonal matrix $Q \in O(d)$ such that $Y = XQ + 1_n s^T$. $Y$ should be interpreted as a result of first rotating (and/or reflecting) sensors in position $X$ by $Q$ and then adding a shift by $s$. Similarly, when we say two position matrices $X$ and $Y$ are equal up to a rigid transformation, we mean that there exists a rotation $Q$ and a shift $s$ such that $Y = XQ + 1_n s^T$. Also, we say a function $f(X)$ is invariant under rigid transformation if and only if for all $X$ and $Y$ that are equal up to a rigid transformation we have $f(X) = f(Y)$. Under these definitions, it is clear that $D$ is invariant under rigid transformation, since for all $(i, j)$, $D_{ij} = ||x_i - x_j||^2 = ||(x_iQ + s^T) - (x_jQ + s^T)||^2$, for any $Q \in O(d)$ and $s \in \mathbb{R}^d$.

Note that when solving a localization problem, there are two possible outputs : a relative map and an absolute map. The task of finding a relative map is to find a configuration of sensors that have the same neighbor relationships as the underlying graph $G$. In the following we use the terms configuration, embedding, and relative map interchangeably. The task of finding an absolute map is to determine the absolute geographic coordinates of all sensors. In this paper, we are interested in presenting an algorithm MDS-MAP, that finds a configuration that best fits the proximity measurements, and providing an analytical bound on the error between the estimated configuration and the correct configuration.

While MDS works perfectly when $D$ is available, in practice not all proximity measurements are available because of the limited radio range $r$. The next section describes the MDS-MAP algorithm introduced in [1], where we apply the shortest paths algorithm on the graph $G = (V, E, P)$ to get a good estimate of the unknown proximity measurements.
Based on MDS, MDS-MAP consists of two steps:

**Algorithm:** MDS-MAP [1]

**Input:** dimension \(d\), graph \(G = (V, E, P)\)

1. Compute the shortest paths, and let \(\hat{D}\) be the squared shortest paths matrix;
2. Apply MDS to \(\hat{D}\), and let \(\hat{X}\) be the output.

The original MDS-MAP algorithm introduced in [1] has additional post-processing step to fit the given configuration to an absolute map using a small number (typically \(d + 1\)) of special nodes with known positions, which are called anchors. However, the focus of this paper is to give a bound on the error between the relative map found by the algorithm and the correct configuration. Hence, the last step, which does not improve the performance, is omitted here.

**Shortest paths.** The shortest path between node \(i\) and \(j\) in graph \(G = (V, E, P)\) is defined as a path between two nodes such that the sum of the proximity measures of its constituent edges is minimized. Let \(d_{ij}\) be the computed shortest path between node \(i\) and \(j\). Then, the squared shortest paths matrix \(\hat{D} \in \mathbb{R}^{n \times n}\) is defined as \(\hat{D}_{ij} = d_{ij}^2\) for \(i \neq j\), and 0 for \(i = j\).

Note that \(\hat{D}\) is well defined only if the graph \(G\) is connected. If \(G\) is not connected, there are multiple configurations resulting in the same observed proximity measures and global localization is not possible. In the unit square, assuming sensor positions are drawn uniformly at random as defined in the previous section, the graph is connected, with high probability, if \(\pi r^2 > (\log n + c_n)/n\) for \(c_n \rightarrow \infty\) [7]. A similar condition can be derived for generic \(d\)-dimensions as \(C_d r^d > (\log n + c_d)/n\), where \(C_d \leq \pi\) is a constant that depends on \(d\). Hence, we focus in the regime where \(r = (\alpha \log n/n)^{1/d}\) for some positive constant \(\alpha\).

As will be shown in Lemma III.1, the key observation of the shortest paths step is that the estimation is guaranteed to be arbitrarily close to the correct estimate (with high probability) for large enough radio range \(r\) and large enough \(n\). Moreover, the all-pairs shortest paths problem has an efficient algorithm whose complexity is \(O(n^2 \log n + n|E|)\) [8]. For \(r = (\alpha \log n/n)^{1/d}\) with constant \(\alpha\), the graph is sparse with \(|E| = O(n \log n)\), whence the complexity is \(O(n^2 \log n)\).

**Multidimensional scaling.** In step 2, we apply the MDS to \(\hat{D}\) to get a good estimate of \(X\). As explained in II-B, we compute \(\hat{X} = \text{MDS}_d(\hat{D})\). The main step in MDS is singular value decomposition of a dense matrix \(\hat{D}\), which has complexity of \(O(n^3)\).

**D. Main results**

Our main result establishes that MDS-MAP achieves an arbitrarily small error from a radio range \(r = (\alpha \log n/n)^{1/d}\) with a large enough constant \(\alpha\), when we have only the connectivity information as in the case of the connectivity-based model. The same bound holds immediately for the range-based model, when we have an approximate measurements for the distances, and the same algorithm can be applied without any modification.

Let \(\hat{X}\) denote an \(n \times d\) estimation for \(X\) with estimated position for node \(i\) in the \(i\)th row. Then, we need to define a metric for the distance between the original position matrix \(X\) and the estimation \(\hat{X}\) which is invariant under rigid transformation of \(X\) or \(\hat{X}\).

Define \(L = I_n - (1/n)1_n 1_n^T\) as in the MDS algorithm. \(L\) is an \(n \times n\) rank \(n - 1\) symmetric matrix which eliminates the contributions of the translation, in the sense that \(LX = L(X + 1s^T)\) for all \(s \in \mathbb{R}^d\). Note that \(L\) has the following nice properties: (1) \(LXX^T L\) is invariant under rigid transformation. (2) \(LXX^T L = L\hat{X}\hat{X}^T L\) implies that \(X\) and \(\hat{X}\) are equal up to a rigid transformation. This naturally defines following distance between \(X\) and \(\hat{X}\).

\[
d_1(X, \hat{X}) = (1/n) \left\| LXX^T L - L\hat{X}\hat{X}^T L \right\|_F, \tag{2}
\]

where \(\left\| \cdot \right\|_F\) denotes the Frobenius norm. Notice that the factor \((1/n)\) corresponds to the usual normalization by the number of entries in the summation, and \((1/\sqrt{d})\) corresponds to the maximum distance between any two nodes in the \([0, 1]^d\) hypercube. Indeed this distance is invariant to rigid transformation of \(X\) and \(\hat{X}\). Furthermore, \(d_1(X, \hat{X}) = 0\) implies that \(X\) and \(\hat{X}\) are equal up to a rigid transformation. Define

\[
r_0 \equiv \frac{24d^2}{\sqrt{10}} \left( \frac{\log n}{n} \right)^{\frac{3}{2}}. \tag{3}
\]

**Theorem II.1.** Assume \(n\) sensors distributed uniformly at random in the \([0, 1]^d\) hypercube, for a bounded dimension \(d = O(1)\). For a given radio range \(r\), a corresponding graph \(G(V, E, P)\) is defined according to the connectivity-based model. Then, with high probability, the distance between the estimate \(\hat{X}\) given by MDS-MAP and the correct position matrix \(X\) is bounded by

\[
d_1(X, \hat{X}) \leq \frac{r_0}{r} + O(r), \tag{4}
\]

for \(r > r_0\), where \(d_1(\cdot)\) is defined in Eq. (2) and \(r_0\) in Eq. (3).

The proof is provided in Section III. As described in the previous section, we are interested in the regime where \(r = (\alpha \log n/n)^{1/d}\) for some constant \(\alpha\). Given a small positive constant \(\delta\), this implies that MDS-MAP is guaranteed to produce estimated positions that satisfy \(d_1(X, \hat{X}) \leq \delta\) with a large enough constant \(\alpha\) and large enough \(n\).

In the following, whenever we write that a property \(A\) holds with high probability (w.h.p.), we mean that \(\mathbb{P}(A)\) actually occurs...
approaches 1 as the number of sensors \( n \) goes to infinity. Given a matrix \( A \in \mathbb{R}^{m \times n} \), the spectral norm of \( A \) is denoted by \( \|A\|_2 \), and the Frobenius norm is denoted by \( \|A\|_F \). For a vector \( a \in \mathbb{R}^n \), \( \|a\| \) denotes the Euclidean norm. Finally, we use \( C \) to denote generic constants that do not depend on \( n \).

### III. Proof of Theorem II.1

We start by bounding the distance \( d_1(X, \hat{X}) \), as defined in Eq. (2), in terms of \( D \) and \( \hat{D} \). Since \( L(XX^T - \hat{X}\hat{X}^T)L \) has rank at most \( 2d \), it follows that

\[
\|L(XX^T - \hat{X}\hat{X}^T)L\|_F \leq \sqrt{2d}\|L(XX^T - \hat{X}\hat{X}^T)L\|_2 , \tag{5}
\]

where we used the fact that, for any rank-\( 2d \) matrix \( A \) with singular values \( \{\sigma_1, \ldots, \sigma_{2d}\} \), \( \|A\|_F = \sqrt{\sum_{i=1}^{2d} \sigma_i^2} \leq \sqrt{2d}\|A\|_2 \). The spectral norm can be bounded in terms of \( D \) and \( \hat{D} \). Let \( M = -(1/2)LDDL \). Then,

\[
\|L(XX^T - \hat{X}\hat{X}^T)L\|_2 \leq \|LXX^TL - M\|_2 + \|M - \hat{X}\hat{X}^T\|_2 \\
\leq \frac{1}{2}\|L(-D + \hat{D})L\|_2 + \frac{1}{2}\|L(-\hat{D} + D)L\|_2 \\
\leq \|\hat{D} - D\|_2 , \tag{6}
\]

where in the first inequality we used the triangular inequality and the fact that \( X = LX \) as shown in Section II-B. In the second inequality we used Eq. (1) and the fact that

\[
\|M - \hat{X}\hat{X}^T\|_2 \leq \|M - A\|_2 , \tag{7}
\]

for any rank-\( d \) matrix \( A \). Since the rank of \( -(1/2)LDDL \) is \( d \), the second term follows by setting \( A = -(1/2)LDDL \). The inequality in (6) follows trivially from the observation that \( \|L\|_2 = 1 \).

Eq. (7) can be proved as follows. From the definition of \( \hat{X} = MDS_d(\hat{D}) \) in Section II-B, we know that \( \hat{X}\hat{X}^T \) is the best rank-\( d \)-approximation to \( M \). Hence, \( \hat{X}\hat{X}^T \) minimizes \( \|M - A\|_2 \) for any rank-\( d \) matrix \( A \), and this proves Eq. (7).

To bound \( \|\hat{D} - D\|_2 \), we use the following key result on the shortest paths. The main idea is that, for sensors with uniformly random positions, shortest paths provide arbitrarily close estimation to the correct distances for large enough radio range \( r \). Define

\[
\hat{r}_0(\beta) \equiv 24 d ((1 + \beta) \log n/n)^{1/d} . \tag{8}
\]

For simplicity we denote \( \hat{r}_0(0) \) by \( \hat{r}_0 \).

**Lemma III.1** (Bound on the shortest paths). Under the hypothesis of Theorem II.1, w.h.p., the shortest paths between all pairs of nodes in the graph \( G(V, E, P) \) are uniformly bounded by

\[
\hat{d}_{i,j}^2 \leq (1 + (\hat{r}_0/r))\hat{d}_{i,j}^2 + O(r) ,
\]

for \( r > r_0 \), where \( \hat{r}_0 \) is defined as in Eq. (8).

The proof of this lemma is provided in a longer version of this paper. Define an error matrix \( Z = \hat{D} - D \). Then by Lemma III.1, \( Z \) is element-wise bounded by \( 0 \leq Z \leq (\hat{r}_0/r)D + O(r) \). Since all the entries are non-negative, we can bound the spectral norm of \( Z \). Let \( u \) and \( v \) be the left and right singular vectors of \( Z \) respectively. Then by Perron-Frobenius theorem, \( u \) and \( v \) are non-negative. It follow that

\[
\|\hat{D} - D\|_2 = u^T Z v \\
\leq (\hat{r}_0/r)u^T D v + O(rn) \\
\leq (\hat{r}_0/r)\|D\|_2 + O(rn) . \tag{9}
\]

The first inequality follows from the element-wise bound on \( Z \) and the non-negativity of \( u \) and \( v \). It is simple to show that \( \|D\|_2 = O(n) \). Typically, we are interested in the region where \( r = o(1) \), which implies that the first term in Eq. (9) dominates the error.

Now we are left with the task of bounding \( \|D\|_2 \), using the following lemma, and the proof is provided in a longer version of this paper.

**Lemma III.2**. \( \|D\|_2 \leq (d/\sqrt{20})n + o(n) \), w.h.p.

Applying Lemma III.2 to Eq. (9) and substituting \( \|\hat{D} - D\|_2 \) in Eq. (6), we get the desired bound, and this finishes the proof of Theorem II.1.

**Acknowledgment**

We thank Rüdiger Urbanke for stimulating discussions on the subject of this paper.

**References**


