Cooperative Localization in Wireless Networks

Yunlong Wang, Yang Cai, Yuan Shen
Tsinghua University, China

ABSTRACT

This chapter describes cooperative localization in wireless networks, where mobile nodes with unknown positions jointly infer their positions through measuring and exchanging information with each other. The technique of cooperation localization, efficiently even in harsh propagation environment, enables amounts of location-based services that rely on high-accuracy position information of mobile nodes. After a brief introduction of cooperative localization, the Cramer-Rao lower bound is given as a standard metric for performance. Then the information in the temporal and spatial domain is illustrated with geometrical interpretations. Two classes of cooperative localization algorithms, namely, centralized and distributed algorithms are presented to show the implementation of the cooperative localization in a wireless network. Then the performance of cooperative localization under non-line-of-sight condition is analyzed. Lastly, numerical results are given to illustrate the performance of cooperative localization algorithms.

Keywords: COOPERATIVE LOCALIZATION, WIRELESS NETWORK, HARSH PROPAGATION ENVIRONMENT, BAYESIAN INFERENCE, MAXIMUM LIKELIHOOD ESTIMATION, CRAMER RAO LOWER BOUND

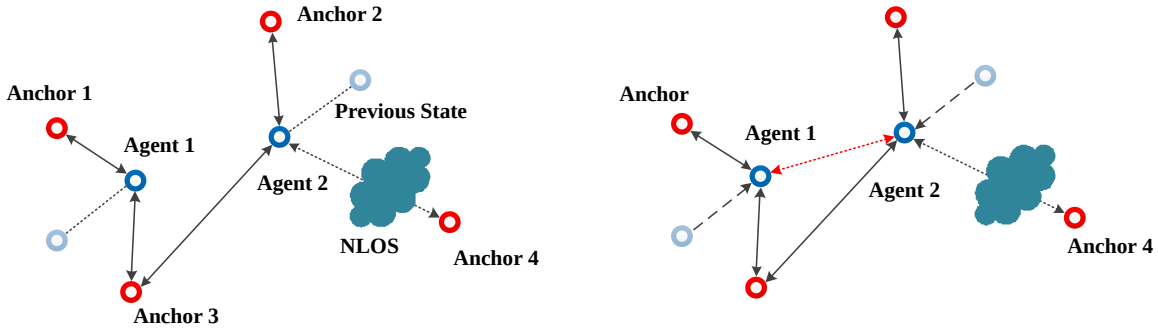
INTRODUCTION

The availability of absolute or relative position information is essential in many applications, such as localization services in cellular networks, search-and-rescue operations, asset tracking, blue force tracking, traffic monitoring and autonomous vehicles (Moe, 2010, Wang, 2017 & Cheng, 2017). Automatic localization of the nodes in a wireless network is a key enabling technology which, mainly because the availability of the node position is highly relevant to the value of the data it collects. With the integration of global positioning system (GPS) into cell phones, in conjunction with other common commercial signals, such as GSM, WiFi, Bluetooth and LED, the ubiquitous location-aware services with sub-meter accuracy are envisioned in the near future, bringing new opportunities for many other applications in commercial, public safety, and military sectors. Then, the wide range of potential applications will also motivate an increasing research interest in localization and tracking technologies for wireless networks.

Location-aware networks generally consist of two kinds of nodes: anchor nodes and agent nodes, where anchors have known positions while agents have unknown positions (Zhou, 2017). Figure 1 illustrates an example of comparison between traditional and cooperative localization/navigation in a simple network with two mobile agents and four static anchors. Conventionally, each agent determines its position based only on the range measurements to at least three anchors, e.g., Figure 1-(a), which is also known as triangulation. These techniques often fail to localize an agent (e.g., Agent 1) when there are insufficient neighboring anchors within the communication range, or fail to provide satisfactory localization accuracy (e.g., Agent 2) when the range measurements are subject to large errors due to non-line-of-sight (NLOS)

propagation. However, when the agents cooperate in both spatial and temporal domains, they are able to not only localize themselves with enough line-of-sight (LOS) signals, but also significantly improve the localization accuracy with more valid constraints e.g., Figure 1-(b).

Two common examples of anchor-based techniques include the GPS and beacon localization. In GPS, an agent determines its location based on the signals received from a constellation of GPS satellites. However, GPS does not operate well in harsh propagation environments, such as indoors or in urban canyons, since the signals suffer from severe attenuation when across obstacles and NLOS signals will degrade the localization accuracy severely. Beacon localization, on the other hand, relies on terrestrial anchors, such as WiFi access points or GSM base stations. In areas where network coverage is sparse, e.g., in emergency situations, localization errors can be unacceptably large. Therefore, there is an urgent need for localization systems that can achieve high accuracy in harsh propagation environments with limited infrastructure requirements.



(a) Before Cooperation *(b) After Cooperation*

Figure 1. A network with two mobile agents (blue circles) and four static anchors (red rings) in two time steps, in which light blue and dark blue circles denote the agents in the previous and current time step. Inter- and intra-node measurements are denoted by red and black arrows, respectively. Anchors 4 provides NLOS signals due to harsh propagation conditions while other three anchors provide LOS signals. (a) Before spatiotemporal cooperation. (b) After spatiotemporal cooperation.

If the agents with unknown locations can make measurements among themselves, such as relative ranges and relative angles, the additional information gained from these measurements between pairs of agents can enhance the accuracy and robustness of the localization system. We call this type of cooperation as spatial cooperation since the relative information in the space is utilized (Patwari, 2005). Moreover, if an agent can also incorporate its previous states and then the performance can be further improved. This type of cooperation can be viewed as temporal cooperation since the nodes' states at different time instants are utilized. The combination of these two types of cooperation in the wireless network enables the joint spatial and temporal cooperation, advocating a new paradigm for real-time localization and navigation in wireless networks, which is referred to as network localization (Shen, 2014). The joint spatial and temporal cooperation among network also improve the robustness of localization in harsh propagation environments by adding more information.

BACKGROUND

Cooperation among networks has been attracting great interests for localization due to its superior performance obtained by the sharing of information among spatial neighbors. Moreover, enabling the communications between agents circumvents the requirements of high-power transmission and, high-

density anchor deployment. In addition to spatial cooperation with neighbors, agents in the network can also benefit from the information obtained in different time instants, which is called temporal cooperation and can be utilized via filtering techniques in a process commonly known as tracking (Foxlin, 2005).

Although the relative information between nodes was utilized earlier, especially in the field of robotic control and monitoring, the cooperative localization in wireless network was first mentioned in (Savarese, 2001) and proposed in (Patwari, 2005). Based on different processing procedures, the cooperative localization algorithms can be categorized into two sets: Centralized algorithms and distributed algorithms (Patwari, 2005). The former requires a central processor to collect all the information of the network before calculating the agents' positions while the distributed algorithms only allow nodes to share information only with their neighbors.

The centralized methods are generally based on the non-Bayesian estimators. Least square and maximum likelihood (ML) are the two common criteria for implementing algorithms. A centralized maximum likelihood estimator (MLE) is proposed in (Patwari, 2003) to solve the relative location estimation in wireless networks. However, the algorithm involves a multivariable nonlinear optimization problem that makes it difficult to implement in practice. To alleviate this problem, the ML estimation was being relaxed to a centralized semi-definite programming (SDP) problem (Biswas & Liang, 2006) and a distributed second-order cone programming problem, respectively. With semidefinite relaxation (SDR), the network localization problem can be converted into a convex optimization program (Boyd, 2004), which can be solved efficiently. Another popular centralized algorithm is multidimensional scaling (MDS), which utilizes pairwise distances available from the received signal strength or time-of-arrival measurements (Leeuw, 1988 & Shang, 2004). While the MDS cost function is non-convex, it can still be solved approximately (Ramsay, 1982 & Zinnes, 1983). A generalized MDS method was proposed in (Chan, 2009) which utilizing the matrix transformation and subspace decomposition. It showed by replacing the transformation matrix, the classical MDS is obtained. Based on the connectivity of the network, MDS algorithms can also be developed to a distributed manner with partially connected network (Chan, 2009 & Chan, So, 2009).

Distributed cooperative localization algorithms generally belong to the Bayesian inference. In (Caceres2010, & Sottile, 2011), Bayesian estimators are utilized to localize the agents by treating the nodes' positions as random variables. The filtering techniques, such as the Kalman filter (KF), extended Kalman filter (EKF) and unscented Kalman filter (UKF) are proposed for cooperative localization in (Caceres, 2010). Both the EKF and UKF interpret the effects of neighbors' position uncertainties similarly to the range measurement errors, which lead to performance losses. A particle filter based cooperative localization is proposed in (Sottile, 2011) for hybrid GNSS-terrestrial environments, which is suitable for the nonlinear systems. However, it often suffers from a large computational complexity.

A class of novel methods based on the graph theory, which is referred to as message passing algorithms, were studied for the cooperative localization (Ihler 2005, Wymeersch 2009, & Savic, 2012). Particularly, a nonparametric belief propagation (NBP) method was proposed for self-localization in wireless networks (Ihler, 2005), which approximated the messages by particles with the corresponding weights. However, the large numbers of particles lead to a huge computational complexity and communication overhead. To reduce the communication overhead in NBP, a method using the approximations on the position beliefs by Gaussian mixtures and information censoring was proposed in (Savic, 2012). A cooperative localization method based on the sum-product algorithm and factor graph was proposed in (Wymeersch, 2009), which was known as the sum-product algorithm over a wireless network (SPAWN) algorithm. Unfortunately, the direct implementation of SPAWN still experiences the computational complexity and further approximation is necessary. In the literature of (Lien, 2012), the parametric and sample-based message passing algorithms of cooperative localization were compared. It indicated that parametric message

representation combined with simple censoring can yield excellent performance at a relatively low complexity.

In harsh propagation environments with the presence of obstacles, the NLOS biases will result in large localization errors if the a priori knowledge of NLOS distribution is unavailable (Shen, 2010). One of the effective ways to alleviate this problem is to identify the NLOS signals and then discard the corresponding measurements (if there are sufficient number of LOS paths) or reweight the contributions of different signals in the localization process (Marano, 2010, Wang, 2012). However, the paths may be identified incorrectly: i) the LOS signals are misidentified as NLOS (false alarm) and ii) NLOS signals are misidentified as LOS (missed detection) and both of them can significantly impact the localization accuracy (Wang, 2016). To cope with the NLOS condition, another way is to measure the propagation characteristics of the channel information and then localize the agent from a scattering model (Lv, 2015). The difficulty is to obtain an accurate model, and the model may change in a time varying environment which requires high computational resources. Some works were attempt to localizes with all NLOS and LOS measurements, and provide weighting or scaling to minimize the effects of the NLOS contributions (Abu, 2016). The map information can be also utilized for improving the localization accuracy in the indoor environment (Wang, 2017). Unlike non-cooperative localization, in cooperative localization the NLOS mitigation is more complex since the communication range and link between agents should be considered.

System Model

1. Network Model

The network considered in this chapter is with N anchors and M agents, where each of the agent is equipped with multiple sensors that can perform intra- and inter-node measurements (e.g., using inertial measurement unit (IMU) and range measurement unit (RMU), respectively) for the purpose of localization and navigation. The position of agents and anchors are denoted by $\mathbf{X}=[\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_M]^{\mathrm{T}}$ and $\mathbf{X}_{\mathrm{b}}=[\mathbf{x}_{M+1},\mathbf{x}_{M+2},\ldots,\mathbf{x}_{M+N}]^{\mathrm{T}}$. Using these intra- and inter-node measurements, represented by $\mathbf{z}=[\mathbf{z}_{\mathrm{s}},\mathbf{z}_{\mathrm{r}}]$, the agents can infer their positions $\mathbf{X}$. Let $M=\{1,2,\ldots,M\}$ and $N=\{M+1,M+2,\ldots,M+N\}$ be the sets of agents and anchors, respectively. The distance between node pairs (anchor-anchor, anchor-agent and agent-agent) is denoted by $\{d_{i,j}\}$, $i,j\in M\cup N\times M\cup N$. Since the nodes are with the constraint of communication range, we use the set $N(i)$ to denote the nodes set that node *i* can make measurements with $i\notin N(i)$ and $N(i)\subset\{1,\ldots,M+N\}$.

2. Signal Metrics

The time of arrival (TOA) is the time that a transmitted signal first arrives at a receiver in LOS environments. Suppose the transmitter and the receiver are synchronized, the receiver can determine the TOA, which is the time of transmission plus the time delay of propagation. The propagation induced time-delay between node *i* and node *j* is denoted by T_{ij}, is equal to the distance d_{ij}, divided by the velocity of propagation v_{p}. The propagation velocity for RF signals is the speed of light.

The angle of arrival (AOA) measurements provide information about the directions from neighboring nodes. There are two common ways to measure the AOA. The first method uses a node with antenna arrays whose locations are known, and estimates the AOA from the phase delays on different elements. The second

approach uses two antennas with different directions equipped on the node, and estimates the AOA from the ratio of their individual received signal strength (RSS) values.

The estimation accuracy of the propagation time at the receiver determines the localization performance directly in the time based approaches. It is mainly affected by additive noises and the multipath effects. Typically, the time delay can be estimated from highest peak of cross-correlation coefficients, but in the multipath channel, the first-arriving peak should be find as the delay estimation because the LOS signal may not be the strongest of the arriving signals. It is shown in (Cook, 2012) that the accuracy of TOA estimation satisfies

$$\mathrm{var}(\hat{t}) \geq \frac{1}{8\pi^2 b^2 \times \mathrm{SNR}} \tag{1}$$

where SNR is the signal to noise ratio, and b is the effective signal bandwidth defined by

$$b^2 = \frac{\int_{-\infty}^{\infty} f^2 |S(f)|^2 \, \mathrm{d}f}{\int_{-\infty}^{\infty} |S(f)|^2 \, \mathrm{d}f}. \tag{2}$$

By increasing the signal-to-noise ratio (SNR) and the effective signal bandwidth, the accuracy of TOA can be improved. For example, the time based techniques can achieve extremely accurate ranging estimates using ultra wideband (UWB) radios (Cassioli, 2002), which can transmit and receive signals with a very large bandwidth. Compared to the additive noise, the range error caused by the multipath effect is much larger. Since there is no guarantee that the LOS signal will be the strongest among all the arrival signals, the TOA is usually estimated by finding the first-arriving peak of the cross correlation. The higher temporal resolution can be achieved via increasing the signal bandwidth, since the peak width is inversely proportional to the signal bandwidth.

3. Graphic Model

Before introducing the cooperative localization, the following assumptions are given for clarity. These mild assumptions include movement model and measurement model which are common in many practical scenarios (Wymeersch, 2009).

Before we introduce the assumptions, several notations should be defined first. The state of agent i at time t is written as $\mathbf{x}_i^{(t)}$ and the state from time 0 to T is denoted by $\mathbf{x}^{(0:T)}$. Denote by $\mathrm{S}_{\rightarrow i}^{(t)}$ the set of nodes from which agent i may receive signals during time slot t, and $\mathrm{S}_{i\rightarrow}^{(t)}$ the set of agents that may receive a signal from agent i. At time slot t, the agent i may estimate self-metrics $z_{i,\mathrm{s}}^{(t)}$ (e.g., orientation or acceleration measurements), by intra-node measurements and signal metrics $z_{j\rightarrow i}^{(t)}$ (e.g., TOA or AOA measurements) by received signals. The collection of all relative signal metrics $\mathbf{z}_{\mathrm{r}}^{(t)}$ and local self-metrics $\mathbf{z}_{\mathrm{s}}^{(t)}$ of all agents is denoted by $\mathbf{z}^{(t)}$ at time t.

For the movement model, we assume that the agents move independently according to a memoryless walk:

$$p\left(\mathbf{x}^{(0:T)}\right) = p\left(\mathbf{x}^{(0)}\right)\prod_{t=1}^{T} p\left(\mathbf{x}^{(t)} \mid \mathbf{x}^{(t-1)}\right) = p\left(\mathbf{x}^{(0)}\right)\prod_{t=1}^{T}\left\{\prod_{i=1}^{M} p\left(\mathbf{x}_i^{(t)} \mid \mathbf{x}_i^{(t-1)}\right)\right\} \tag{3}$$

For the measurement model, we first assume that the intra-node measurements and the inter-node measurements are independent conditioned on the states

$$p\left(\mathbf{z}^{(1:T)} \mid \mathbf{x}^{(0:T)}\right) = p\left(\mathbf{z}_{\mathrm{s}}^{(1:T)} \mid \mathbf{x}^{(0:T)}\right) p\left(\mathbf{z}_{\mathrm{r}}^{(1:T)} \mid \mathbf{x}^{(0:T)}\right) \tag{4}$$

The intra-node measurements are mutually independent for different nodes and different time slots, and depend only on the current and previous location of the agent

$$p\left(\mathbf{z}_{\mathrm{s}}^{(1:T)} \mid \mathbf{x}^{(0:T)}\right)=\prod_{t=1}^{T} p\left(\mathbf{z}_{\mathrm{s}}^{(t)} \mid \mathbf{x}^{(t-1)}, \mathbf{x}^{(t)}\right)=\prod_{t=1}^{T}\left\{\prod_{i=1}^{M} p\left(z_{i,\mathrm{s}}^{(t)} \mid \mathbf{x}_{i}^{(t-1)}, \mathbf{x}_{i}^{(t)}\right)\right\} \tag{5}$$

The inter-node measurements at different time slots are independent, conditioned on the nodes' locations, and each of them depends only on the current locations of the two nodes involved:

$$p\left(\mathbf{z}_{\mathrm{r}}^{(1:T)} \mid \mathbf{x}^{(0:T)}\right)=\prod_{t=1}^{T} p\left(\mathbf{z}_{\mathrm{r}}^{(t)} \mid \mathbf{x}^{(t)}\right)=\prod_{t=1}^{T}\left\{\prod_{i=1}^{M} \prod_{j \in S_{\rightarrow i}^{(t)}} p\left(z_{j \rightarrow i}^{(t)} \mid \mathbf{x}_{i}^{(t)}, \mathbf{x}_{j}^{(t)}\right)\right\} \tag{6}$$

THE CRAMÉR-RAO BOUND

The Cramér-Rao bound (CRB) provides a lower bound on the variance achievable by any unbiased location estimator (Kay, 2013). It provides a guideline that how a best estimator can possibly to do and a measure of the estimator that we designed. Given an unbiased estimator $\hat{\boldsymbol{\theta}}$, it must satisfy that

$$\operatorname{var}\left(\hat{\boldsymbol{\theta}}\right) \geq \mathbf{J}\left(\boldsymbol{\theta}\right)^{-1} \tag{7}$$

and

$$\mathbf{J}\left(\boldsymbol{\theta}\right)=\mathrm{E}\left\{-\frac{\partial^{2} \ln f\left(\mathbf{x} \mid \boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}^{2}}\right\} \tag{8}$$

where $\mathbf{J}\left(\boldsymbol{\theta}\right)$ is the Fisher information matrix (FIM) of $\boldsymbol{\theta}$, var($\hat{\boldsymbol{\theta}}$) is the variance of the estimator and $f(\mathbf{x}|\boldsymbol{\theta})$ is the likelihood function of $\mathbf{x}$ condition on $\boldsymbol{\theta}$. The bound is also similar to the sensitivity analysis, applied to random measurements since the CRB is based on the curvature of the log-likelihood function. Generally, a sharp log-likelihood function performs superior to the broad one, resulting a lower bound. It should be noted that the CRB is limited to unbiased estimators. Such estimators provide coordinate estimates that, if averaged over enough realizations, are equal to the true coordinates. Although unbiased estimation is a desirable property, some bias might be tolerated to reduce variance; in such cases, the bound can be adapted (Hero III, 1996).

Consider a network with M agents with coordinates $\boldsymbol{\theta}=[\boldsymbol{\theta}_x, \boldsymbol{\theta}_y]$, where $\boldsymbol{\theta}_x=[x_1,\ldots,x_M]$ and $\boldsymbol{\theta}_y=[y_1,\ldots,y_M]$. There are N anchors and their coordinates are $[x_1,\ldots,x_N,y_1,\ldots,y_N]$. The pair-wise measurements are $\{z_{i,j}\}$ where $z_{i,j}$ is a measurement between devices *i* and *j*, $i,j \in M \cup N \times M \cup N$.

Table 1: Differences in CRB by measurements type

	Channel constant	Exponent	FIM
TOA	$\gamma=1/(v_p\sigma_T)^2$	*s*=0	$\mathbf{J}=\mathbf{J}_{\mathrm{TR}}$
RSS	$\gamma=(10n_p/\sigma_{\mathrm{dB}}\log 10)$	*s*=2	$\mathbf{J}=\mathbf{J}_{\mathrm{TR}}$
AOA	$\gamma=1/\sigma_\alpha^2$	*s*=2	$\mathbf{J}=\mathbf{J}_{\mathrm{A}}$

The measurements between node pairs can be TOA, AOA, RSS or connectivity. To derive the CRB of the cooperative localization, the FIM $\mathbf{J}$ is given as follows

$$\mathbf{J}_{\mathrm{TR}}=\begin{cases}\dfrac{\gamma}{d_{k,i}^{s}}\displaystyle\sum_{i\in N\ (k)}\mathbf{u}(\phi_{k,i})\mathbf{u}^{\mathrm{T}}(\phi_{k,i}), & k=l\\ -\dfrac{\gamma}{d_{k,i}^{s}}I_{N\ (k)}(l)\mathbf{u}(\phi_{k,i})\mathbf{u}^{\mathrm{T}}(\phi_{k,i}), & k\neq l\end{cases} \tag{9}$$

$$\mathbf{J}_{\mathrm{A}}=\begin{cases}\dfrac{\gamma}{d_{k,i}^{s}}\displaystyle\sum_{i\in N\ (k)}\mathbf{u}(\phi_{k,i}+\frac{\pi}{2})\mathbf{u}^{\mathrm{T}}(\phi_{k,i}+\frac{\pi}{2}), & k=l\\ -\dfrac{\gamma}{d_{k,i}^{s}}I_{N\ (k)}(l)\mathbf{u}(\phi_{k,i}+\frac{\pi}{2})\mathbf{u}^{\mathrm{T}}(\phi_{k,i}+\frac{\pi}{2}), & k\neq l\end{cases} \tag{10}$$

where $\mathbf{u}(\phi)=[\cos(\phi),\ \sin(\phi)]^{\mathrm{T}}$ is the unit vector with direction given by ϕ denoting the angle from one node to the other, $\mathbf{J}=\mathbf{J}_{\mathrm{TR}}$ denotes the FIM of TOA and RSS metrics, $\mathbf{J}=\mathbf{J}_{\mathrm{A}}$ denotes the FIM of AOA metric, γ is a channel constant and s is an exponent, both of which are functions of the measurements type and are given in the Table I. The indicator function $I_{N\ (k)}(l)$ represents the available measurements set with node *l*, i.e., $I_{N\ (k)}(l)=1$ when node $l\in N\ (k)$ can make measurements with node *k*, or 0 if not. Then CRB can be obtained by taking inversion of FIM, that is

$$\mathrm{CRB}_{\boldsymbol{\theta}}=\mathrm{trace}\left(\mathbf{J}^{-1}\right). \tag{11}$$

SPATIAL AND TEMPORAL COOPERATION

The concept of cooperation has been applied to wireless sensor networks (WSNs), where distributed nodes work together to draw a consensus about the environment or to make measurements with each other via spatial and temporal manner (Akyildiz, 2002). Similarly, network localization allows agents to communicate with each other in estimating their positions, offering additional benefits such as improved localization accuracy, resilience to system failure, increase in coverage, and reduction of cost per node (Wymeersch 2009, Patwari 2005, & Shen, 2010). Understanding the fundamentals of wireless cooperative localization is important, which not only provides a performance benchmark, but also guides algorithm and network design.

The accuracy of location estimation is inherently limited due to random phenomena such as noise, fading, shadowing, and multipath propagation. The fundamental limits of such accuracy have been derived by using the information inequality (Shen, 2010).

1. Spatial Cooperation

In a static network, or a dynamic network at a given time instant, only spatial cooperation among agents can be exploited. By using the notion of the FIM, the squared position error for agent *k* is bounded by the following inequality

$$\mathrm{E}\left\{\left\|\mathbf{x}_k-\hat{\mathbf{x}}_k\right\|^2\right\}\geq\mathrm{tr}\left\{\left[\mathbf{J}_{\mathrm{e}}(\mathbf{x})\right]_{\mathbf{x}_k}^{-1}\right\} \tag{12}$$

where $\mathbf{J}_{\mathrm{e}}(\mathbf{x})$ is the $2M\times 2M$ FIM for $\mathbf{x}$; $\mathrm{E}\{\square\}$ and $\mathrm{tr}\{\square\}$ are the expectation and trace operators, respectively, and $[\square]_{\mathbf{x}_k}$ denotes the square submatrix on the diagonal corresponding to $\mathbf{x}_k$. The right side of

Eq. (12) is referred to as the squared position error bound (SPEB). It can be shown that the FIM $\mathbf{J}_{\mathrm{e}}(\mathbf{x})$ is a sum of two parts, the localization information from anchors (shown as the block-diagonal matrices consisting of $\mathbf{K}_i, i=1,2,\ldots,M$) and that from agents' spatial cooperation (shown as the structured matrix consisting of $\mathbf{C}_{ij},\ i,j=1,2,\ldots,M$), shown as follows

$$\mathbf{J}_{\mathrm{e}}(\mathbf{x})=\begin{bmatrix} \mathbf{K}_1+\sum_{j\in M\ \backslash\{1\}}\mathbf{C}_{1j} & -\mathbf{C}_{12} & \cdots & -\mathbf{C}_{1M} \\ -\mathbf{C}_{12} & \mathbf{K}_2+\sum_{j\in M\ \backslash\{2\}}\mathbf{C}_{2j} & & -\mathbf{C}_{2M} \\ \vdots & & \ddots & \\ -\mathbf{C}_{1M} & -\mathbf{C}_{2M} & & \mathbf{K}_{N_a}+\sum_{j\in M\ \backslash\{M\}}\mathbf{C}_{Mj} \end{bmatrix} \tag{13}$$

The diagonal matrix in $\mathbf{J}_{\mathrm{e}}(\mathbf{x})$ is the summation of information from anchor $\mathbf{K}_i$ and cooperation from other agents $\mathbf{C}_{ij}$ $(i\neq j)$. The basic building blocks of the FIM $\mathbf{J}_{\mathrm{e}}(\mathbf{x})$ represent the localization information between pairs of nodes in the form $\lambda\mathbf{u}(\phi)\mathbf{u}^{\mathrm{T}}(\phi)$ where λ is a positive scalar that characterizes the ranging information intensity (RII) (Shen, 2010). The value of λ depends on the ranging technique as well as the power and bandwidth of the received waveform, multipath propagation, and prior statistical channel knowledge. In particular, the RII is proportional to the square of the effective bandwidth (Shen, 2010). Moreover, the localization information from anchors can be expressed in a canonical form as a weighted sum of "one-dimensional" information from individual anchors; while cooperation always improves the localization accuracy since it adds a positive semi-definite matrix to the FIM corresponding to non-cooperative localization. It was shown that the accuracy of localization is affected by two factors: the quality of point-to-point measurements, reflected by the expression of RII, and the network topology, reflected by the block structure of the matrix $\mathbf{J}_{\mathrm{e}}(\mathbf{x})$.

In the absence of cooperation, every $\mathbf{C}_{ij}$ corresponding to cooperation between nodes *i* and *j* becomes 0, resulting in a total FIM $\mathbf{J}_{\mathrm{e}}(\mathbf{x})$ with a block-diagonal structure. In such cases, localization information for different agents are uncorrelated, and the SPEB can be calculated using only local information at the agents.

2. Spatial-Temporal Cooperation

Building on the understanding of cooperative localization, we now discuss the case of cooperative navigation where agents in a dynamic network cooperate in both the space and time domains. For each time instant, the contribution of cooperation in space is similar to what we have seen in cooperative localization. In addition, another layer of cooperation in time, exploiting intra-node measurements and mobility (dynamic) models, yields new information for navigation. Such information at time *t* is characterized by $\mathbf{T}^{(t)}$ matrices in the total FIM $\mathbf{J}_{\mathrm{e}}(\mathbf{x}^{(1:T)})$ where $\mathbf{x}^{(1:T)}$ consists of the positions of all agents from time 1 to *T*. Consequently, the FIM for each agent at a given time instant can be given by

$$\mathbf{J}_{\mathrm{e}}(\mathbf{x}^{(1:T)})=\begin{bmatrix} \mathbf{S}^{(1)}+\mathbf{T}^{(2)} & -\mathbf{T}^{(2)} & & & \\ -\mathbf{T}^{(2)} & \mathbf{S}^{(2)}+\mathbf{T}^{(2)}+\mathbf{T}^{(3)} & & & \\ & & \ddots & & \\ & & & \mathbf{S}^{(T-1)}+\mathbf{T}^{(T-1)}+\mathbf{T}^{(T)} & -\mathbf{T}^{(T)} \\ & & & -\mathbf{T}^{(T)} & \mathbf{S}^{(T)}+\mathbf{T}^{(T)} \end{bmatrix} \tag{14}$$

where $\mathbf{S}^{(t)}$ represents the spatial cooperation information at time *t* and $\mathbf{S}^{(t)}$ has the same form of (13) and $\mathbf{T}^{(t)} = \mathrm{diag}\left\{\mathbf{T}_1^{(t)}, \mathbf{T}_2^{(t)}, \ldots, \mathbf{T}_M^{(t)}\right\}$ represents the temporal information.

Some observations can be drawn from the structure of the FIM for cooperative navigation (Moe, 2010). First, the total FIM consists of two major components: cooperation in space as well as in time. The former characterizes the localization information from inter-node measurements in the entire network at each time instant (shown as the diagonal $2M \times 2M$ block matrices), and the latter characterizes the information from intra-node measurements and mobility models at each individual agent (shown as components $\mathbf{T}^{(t)}$ outside the main block-diagonal). Second, since intra-node measurements and the mobility models for different agents are independent, the corresponding matrices $\mathbf{T}^{(t)}$ form a block diagonal matrix in the upper-right and lower-left quarter of the total FIM. Third, the $\mathbf{T}^{(t)}$ components can be viewed as the temporal link that connects localization information from spatial cooperation of the previous time instant to the current one. If the temporal link is not available (e.g., $\mathbf{T}^{(t)}$ matrices are zero), the total FIM is block diagonal, implying that position inference is independent from time to time. The structure of the FIM for cooperative navigation allows a recursive method to calculate the FIM at each time instant. This view also provides insights into the information evolution of spatial-temporal cooperation in cooperative navigation.

LOCALIZATION TECHNIQUES

The enabling technologies for network localization and navigation including network infrastructure, common signal metrics and error mitigation will be illustrated in this part.

1. Network Infrastructure

Although the GPS can provide localization accuracy on the order of meters in open outdoor environment, which has been widely used in practice, it may fail in several harsh propagation environments, such as in building and urban canyons. The blocking and NLOS signal will degrade the localization accuracy severely. Therefore, the terrestrial localization system is important in these GPS-challenged or even GPS-denied areas. Generally, these systems are based on cellular networks, radio frequency identification (RFID) networks, wireless local area networks (WLANs) and wireless networks and each of them has unique infrastructure requirements. In indoor environments, the cellular networks experience signal attenuation and cannot be used for localization in practice. WLANs typically suffer from the problems of obstruction, scattering and reflection from obstacles, furniture and people. Although RFIDs can provide the desired accuracy, it requires high node density to guarantee the stability. UWB technology offers high ranging accuracy in harsh propagation environments due to its ability to resolve multipath and penetrate obstacles (Gezici, 2005 & Dardari, 2009). It is envisioned that UWB-based localization will play an important role in future high-definition situation-aware and RFID networks. In particular, the IEEE 802.15.4a standard is the first to contemplate both communication and localization with high levels of availability and sub-meter accuracy. Moreover, since generally a wireless network is off-battery or difficult to recharge, the problem of energy consumption and allocation should be considered in network localization (Dai, 2017).

The network is based on intra- and inter-node measurements and the types of measurements available depend on the technologies utilized. The intra-node measurements include acceleration, angular velocity, Doppler shift and orientation etc., while inter-node measurements are waveforms, ranges and directions. For example, an IMU can support measures of distance and direction of the node itself while an agent with transceivers can infer the distance and direction to its neighbors based on signal metrics extracted form exchanged radio frequency waveforms.

2. Signal Metrics for Inter-Node Measurements

The classification of network localization can be categorized by the type of inter-measurements, such as range-based, direction-based and proximity-based. Among these measurements, the range-based methods are more suitable due to its high accuracy and lower complexity. Moreover, two common techniques are used to obtain range measurements: RSS-based and delay-based systems.

The distance is measured by the strength of the received signal, which follows the physical law that the signal experiences path loss due to the propagation in the medium. The measure of RSS can be based on theoretical or empirical models and the choice of mode strongly affects the ranging accuracy. A widely used mode is given by $P_{\mathrm{r}}(d) = P_0 - 10\gamma \log_{10}(d) + S$, where $P_{\mathrm{r}}(d)$ is the received power(dBm), P_0 is the receiver power (dBm) at 1m, the distance from the transmitter and the receiver is denoted by *d*, γ is the path loss exponent and *S* is the large-scale fading (shadowing) which is commonly modeled as a Gaussian random variable with zero mean and standard deviation σ_S. RSS is a signal metric that is easy to measure and generally cannot provide high-accuracy ranging, because it may suffer from the mismatch errors between distance and signal attenuation.

In a delay-based system, the range can be measured from the propagation delay, or time-of-flight $\tau_f = d / v$ where *d* is the distance between the nodes and *v* is the speed at which the waveform propagates in the medium. The process of time delay can be accomplished by one-way TOA, two-way TOA or time difference of arrival (TDOA). The factors that influence the time delay estimation are mainly noises, multipath effects, obstacles, interferences, and clock drifts. Generally, in a scenario with the dense multipath channel, the first path is often not the strongest which makes the TOA estimation challenging. The TDOA metric can be used to reduce the clock bias or in a non-cooperative localization scenario where the TOA cannot be measured directly. The method of ML-based time delay estimation can achieve the CRB at high SNR regimes, but the implementation of ML estimators usually requires sampling at the Nyquist sampling rate or higher, which is a challenging task for wideband and UWB signals.

In (Dardari, 2009), two kinds of TOA estimators based on energy detection and the effects of impairments are analyzed. Since the CRB is typically used in the high SNRs region and it is not accurate at low and moderate SNRs for TOA estimation. TOA estimation, like many other estimation problems, also can be characterized by the threshold effect with distinct SNR regions corresponding two different modes of operation. Particularly, at low SNRs (*a priori region*), measurements do not provide significant additional information, and the mean square error of the estimator is close to that obtained solely from the *a priori* information about TOA. At high SNRs (*asymptotic region*), the mean square error (MSE) of the estimator can be accurately described by the CRB. Furthermore, the region between these two extremes is called *transition* or *ambiguity* region where the performance is subject to ambiguities. Therefore, many other bounds like Ziv-Zakai bound (Chazan, Zakai & Ziv 1975) was proposed to characterize the performance by accounting both ambiguity effects and *a priori* information.

3. Ranging Error Mitigation in Harsh Propagation Environments

From the aforementioned analysis, the performance of network localization mainly determined by the following two factors: i) the geometric configuration of the system (i.e., the deployment of the anchors and agents), and ii) the quality of the measurements (waveform and ranges). Therefore, the design of such location-aware networks requires a clear understanding of these factors and their influences. For example, the obstacles may block partial or complete LOS, which leads to positively biased range estimates for time based ranging techniques. Discarding or refining unreliable range measurements can improve the

localization performance. Particularly, regardless of the specific range-based localization algorithm used, the following three-step procedure can be adopted: (i) calculate the preliminary position estimate $\hat{\mathbf{x}}^{(1)}$ from initial range estimates based on the measurements; (ii) correct range estimates based on $\hat{\mathbf{x}}^{(1)}$ by removing biases according to bias models; and (iii) estimate a refined position $\hat{\mathbf{x}}^{(2)}$ with the correct values.

Without environmental knowledge, ranging refinement can be made based on other information, such as the non-line-of-sight (NLOS) condition which is extracted from the received waveforms (Win, 2011). Amounts of techniques were proposed to identify NLOS conditions and mitigate their influence. The identification of signal obstruction is mainly accomplished by performing a likelihood ratio test on binary hypotheses. This requires the extraction of features that are significantly affected by the propagation conditions from the received waveform. For example, the NLOS identification technique in (Güvenc, 2008) is based on signal features such as root mean square delay-spread, Kurtosis, and mean excess delay. Other approaches include regression with support vector machines and Gaussian processes. The characterization of each possible link, by means of ranging and waveform measurements, enables the system designer to understand how to harness environmental knowledge, identify LOS and NLOS conditions, take advantage of cooperation, and choose cooperating nodes.

4. Localization and Navigation Algorithms

From the perspective of Bayesian, the localization and navigation algorithm aims to determine the posterior distribution $p(\mathbf{x}|\mathbf{z})$, also referred to its position belief. Once the posterior distribution is determined, point estimates can be accomplished with the minimum mean square error (MMSE) or maximum *a posteriori* (MAP) estimates, respectively. The primary tools to obtain the posterior distribution are Bayes' rule and marginalization, updating beliefs according to new measurements and reducing the dimension of inference problem, respectively.

In the traditional non-cooperative localization network, each agent can only determine its position from the neighbor anchors independently. The belief update and MMSE or MAP estimation with Bayes' rule can be calculated from the ML function modeled by a tractable statistical distribution. For a Gaussian distribution, the MMSE and MAP estimators for the agent's position are the solution of a weighted least squares (WLS), that is

$$\hat{\mathbf{x}}_{\mathrm{MMSE}} = \hat{\mathbf{x}}_{\mathrm{MAP}} = \arg\min_{\mathbf{x}} \sum_{j\in N} \frac{1}{\sigma_{\mathbf{z}_j}^2} \| \mathbf{z}_j - h\left(\mathbf{x}, \mathbf{x}_j\right) \|^2 + \frac{1}{\sigma_{\mathbf{x}}^2} \| \mathbf{x} - \hat{\mathbf{x}} \|^2 \tag{15}$$

where $\sigma_{\mathbf{z}_j}^2$ and $\sigma_{\mathbf{x}}^2$ are the variance of the noise in each measurement and variance of the agent's distribution, respectively, and $h(\mathbf{x}, \mathbf{x}_j)$ is a functional of the distance from anchor j to agent. When the likelihood function is a non-Gaussian distribution, the optimization process for MMSE and MAP may be complicated, in which case the WLS solution can be used as a tractable suboptimal solution. While each agent determines its belief based on the information obtained from its local measurements of anchors, the localization performance can be improved if agents incorporate the beliefs of neighbors (spatial cooperation) and the beliefs obtained in the past (temporal cooperation). The following parts will introduce the method of the spatial cooperation and temporal cooperation in algorithm designs.

5. Spatiotemporal Cooperation

In spatial cooperation, one agent will incorporate the beliefs from neighbor agents to determine its position and a centralized processor is required to update the joint belief of all agents by using all the measurements as well as the prior joint belief. Unfortunately, the position belief update and interference is a high

dimensional optimization that is difficult to handle. Since the position belief of all agents will transmit to the central processor and then communicate back after update, the large communication bandwidth and the resources are required, which is inefficient in practice. Moreover, this kind of centralized solutions relies on the stability of the central processor and this kind of network is not robust to failure. Therefore, the distributed algorithms in cooperative networks are attractive to tackle these problems. From Figure 2-(b) we can observe that the FIM does not have a block-diagonal structure due to the correlation between the inferred agents' positions. Thus, the distributed algorithms for cooperative localization cannot achieve the CRB.

A common method to implement distributed cooperative localization is by the loopy belief propagation (Wymeersch, 2009) which assumes the position beliefs are independent and propagate the beliefs of each agent through the cooperative network in an iterative manner by the sum-product algorithm. The Bayes' updates and marginalization are corresponding to the products and sums (integrals) respectively.

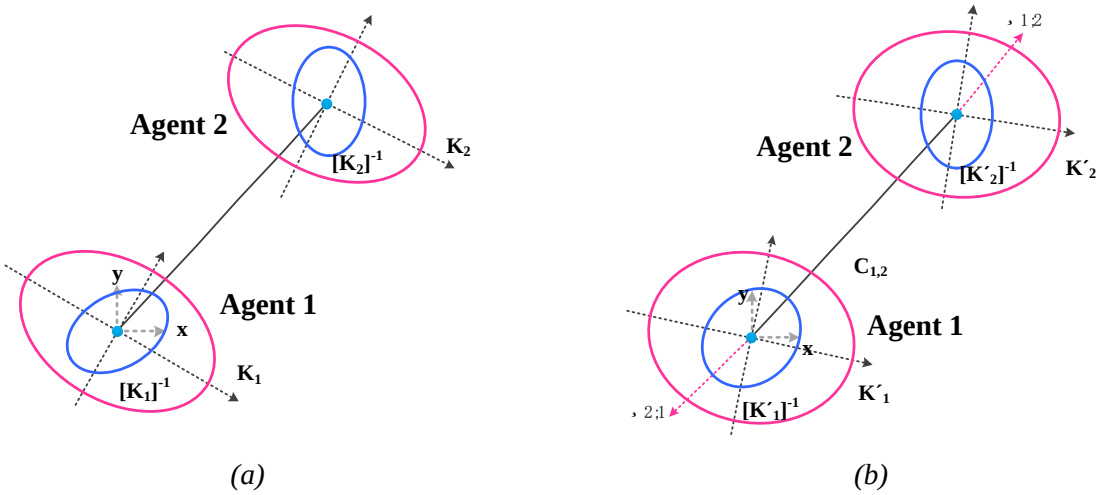


Figure 2. Localization information provided by spatial cooperation. The red and blue ellipses respectively represent the localization information and error.

To illustrate the benefits of the cooperative localization, an example of two agents with spatial cooperation is shown in Figure 2, describing the position errors before and after cooperation. The red and blue ellipses in Figure 2-(a) and Figure 2-(b) represent the agents' localization information and errors, respectively. As the two agents cooperate, the information ellipses increase and the error ellipses decrease, as predicted by the theoretical results. Correspondingly, the position beliefs of the two agents become more concentrated after cooperation.

In spatiotemporal cooperation, the agent's belief not only obtains from the neighbor agents, but also from the intra-node measurements and mobility models. The block-diagonal matrix $\mathbf{T}^{(t)}$ in (14) is the additional information to improve the network localization performance. Given $\mathbf{x}$ denotes the position states like positions and orientations, etc. The temporal cooperation can be accomplished using mobility and intra-node measurement models. The former statistically describes the evolution of the position states in time by $p(\mathbf{x}^{(t)} \mid \mathbf{x}^{(t-1)})$ and the relationship between intra-node measurements and the position states is $p(\mathbf{z}_s \mid \mathbf{x}^{(t)})$. Thus, the belief updates based on Bayes' rule and marginalization can be performed in the following two steps:

i) Prediction

$$p\left(\mathbf{x}^{(t)} \mid \mathbf{z}^{(1:t-1)}\right) = \int p\left(\mathbf{x}^{(t-1)} \mid \mathbf{z}^{(1:t-1)}\right) p\left(\mathbf{x}^{(t)} \mid \mathbf{x}^{(t-1)}\right) \mathrm{d}\mathbf{x}^{(t-1)} \tag{16}$$

ii) Correction:

$$p\left(\mathbf{x}^{(t)} \mid \mathbf{z}^{(1:t)}\right) \propto p\left(\mathbf{x}^{(t)} \mid \mathbf{z}^{(1:t-1)}\right) p\left(\mathbf{z}^{(t)} \mid \mathbf{x}^{(t)}\right) \tag{17}$$

where $p\left(\mathbf{z}^{(t)} \mid \mathbf{x}^{(t)}\right) = p\left(\mathbf{z}_{\mathrm{s}}^{(t)} \mid \mathbf{x}^{(t)}\right) p\left(\mathbf{z}_{\mathrm{r}}^{(t)} \mid \mathbf{x}^{(t)}\right)$.

If both measurements and mobility models are linear-Gaussian and the previous beliefs are also Gaussian, the updated beliefs at each time instant can be evaluated exactly by means of the Kalman filter. However, for nonlinear or non-Gaussian models in many scenarios, several approximation methods, such as, extended KF or unscented KF, to particle filters (PFs) based on sequential Monte Carlo techniques can be taken to improve the localization performance.

LOCALIZATION METHODS

1. Non-Bayesian Methods

In the non-Bayesian framework, the parameters to be estimated are treated as unknown, but deterministic variables. Two common estimators are LS estimator (Nguyen, 2015) and the ML estimator. The LS approach minimizes the unfitness between the true observations and the model induced results, while the ML approach maximizes the likelihood function of the observations given the concerned parameter (Poor, 1984). Both methods find their estimators by minimizing cost functions $\mathrm{C}^{(t)}(\mathbf{X})$ with respect to $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_M]$ at time t.

For LS estimator, the cost function is given by

$$\mathrm{C}_{\mathrm{LS}}^{(t)}(\mathbf{X}) = \sum_{i=1}^{M} \sum_{j \in \mathrm{S}_{\rightarrow i}^{(t)}} \left[z_{j \rightarrow i}^{(t)} - m(\mathbf{x}_i, \mathbf{x}_j) \right]^2 \tag{18}$$

where $m(\mathbf{x}_i, \mathbf{x}_j)$ is a suitable function based on the signal metrics. For example, in the case of ranging measurements, the function is the Euclid distance between the two nodes, which is $m(\mathbf{x}_i, \mathbf{x}_j) = \| \mathbf{x}_i - \mathbf{x}_j \|$.

For ML estimator, the cost function becomes

$$\mathrm{C}_{\mathrm{ML}}^{(t)}(\mathbf{X}) = -\log p(\mathbf{z}_{\mathrm{r}}^{(t)} \mid \mathbf{X}) = -\sum_{i=1}^{M} \sum_{j \in \mathrm{S}_{\rightarrow i}^{(t)}} \log p(z_{j \rightarrow i}^{(t)} \mid \mathbf{x}_i, \mathbf{x}_j) \tag{19}$$

where p is determined by the statistical model of the measurements. A common model is additive Gaussian noise $\mathrm{N}(0, \sigma^2)$, and in this case

$$p(z_{j \rightarrow i}^{(t)} \mid \mathbf{x}_i, \mathbf{x}_j) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left\{ -\frac{[z_{j \rightarrow i}^{(t)} - m(\mathbf{x}_i, \mathbf{x}_j)]^2}{2\sigma^2} \right\} \tag{20}$$

Note that the cost function can be summarized as $\mathrm{C}^{(t)}(\mathbf{X}) = \sum_{i=1}^{M} \sum_{j \in \mathrm{S}_{\rightarrow i}^{(t)}} c_{j \rightarrow i}(z_{j \rightarrow i}^{(t)}; \mathbf{x}_i, \mathbf{x}_j)$ for both LS estimator and the ML estimator aforementioned. To minimize this cost function, the gradient of the function should be set as a zero vector. This is equivalent to setting its partial derivatives with respect to all $\mathbf{x}_i$ $(i = 1, 2, \cdots, M)$ as zeros simultaneously. Specially, the derivative with respect to $\mathbf{x}_i$ is given by

$$\frac{\partial \mathrm{C}^{(t)}(\mathbf{X})}{\partial \mathbf{x}_i} = \sum_{j \in \mathrm{S}_{\rightarrow i}^{(t)}} \frac{\partial c_{j \rightarrow i}(z_{j \rightarrow i}^{(t)}; \mathbf{x}_i, \mathbf{x}_j)}{\partial \mathbf{x}_i} + \sum_{k \in \mathrm{S}_{i \rightarrow}^{(t)}} \frac{\partial c_{i \rightarrow k}(z_{i \rightarrow k}^{(t)}; \mathbf{x}_k, \mathbf{x}_i)}{\partial \mathbf{x}_i} = 0. \tag{21}$$

A close form solution for the above problem has not been found yet. In practice, the gradient decent based algorithms are usually applied to iteratively minimize the cost function $C^{(t)}(\mathbf{x})$.

2. Bayesian Methods

In the Bayesian framework, the parameter $\mathbf{x}$ is treated as a realization of a random variable $\mathbf{X}$ with an *a priori* distribution $p_{\mathbf{X}}(\mathbf{x})$. Two common estimators are MMSE estimator and MAP estimator. The MMSE approach minimizes the mean squared error of all possible realizations of the parameter, and the MAP approach finds the peak of the conditional probability of the concerned parameters given the observations (Poor, 1984). The key step of both methods is to find the posterior distribution.

The MMSE estimator has been proved to be the mean of the *a posteriori* distribution

$$\hat{\mathbf{x}}_{\mathrm{MMSE}} = \int \mathbf{x} \cdot p_{\mathbf{X}|\mathbf{Z}}(\mathbf{x}|\mathbf{z})\mathrm{d}\mathbf{x}. \tag{22}$$

The MAP estimator is the mode of the *a posteriori* distribution

$$\hat{\mathbf{x}}_{\mathrm{MAP}} = \arg\max_{\mathbf{x}} \mathrm{p}_{\mathbf{X}|\mathbf{Z}}(\mathbf{x}|\mathbf{z}). \tag{23}$$

Both estimators can be calculated by each component rather than the entire vector, which means that the estimation of the *k*-th component can be performed by the marginal distribution of the variable $\mathbf{X}_k$. With a little abuse of definition, the word marginal is used. For MMSE estimation, the marginal distribution of the *k*-th component is given by

$$p_{\mathbf{X}_k}(\mathbf{x}_k) = \int_{\sim\{\mathbf{x}_k\}} p_{\mathbf{X}}(\mathbf{x}) \tag{24}$$

and for MAP estimation, it is

$$p_{\mathbf{X}_k}(\mathbf{x}_k) = \max_{\sim\{\mathbf{x}_k\}} p_{\mathbf{X}}(\mathbf{x}) \tag{25}$$

where $\sim\{\mathbf{x}_k\}$ means applying a certain operation with respect to all variables but $\mathbf{x}_k$. The fast calculation of the marginal distribution of each component is essential to the realization of Bayesian methods. In the next section, a tool called factor graph will be introduced for this purpose.

CENTRALIZED ALGORITHM

If the measurements can be described well by a particular statistical model (e.g., Gaussian or log-normal), then the MLE can be derived and implemented. One reason that these estimators are used is that their variance asymptotically (as the SNR goes large) approaches the lower bound given by the CRB. However, there are two difficulties with this approach. i) Local maxima: an initial estimate close to the correct solution is required; ii) model dependency: If measurements deviate from the assumed model (or model parameters), the results are no longer guaranteed to be optimal. The method to tackle the local maxima problem is try to formulate the localization as a convex optimization problem. Convex constraints are utilized to estimate nodes' location in (Doherty, 2001).

More general constraints can be considered if SDP techniques are used (Biswas, 2003 & Luo, 2010). One difficulty that must be overcome in both techniques is their high computational complexity. Toward this end, a distributed SDP-based localization algorithm was presented in (Biswas, 2003 & Luo, 2010). MDS based algorithm is a well-known network localization algorithm by formulating localization from range measurements as an LS problem (Shang, Ruml 2003 & 2004) and the location estimation can be solved by eigen-decomposition in classical MDS. The MDS do not suffer the local maxima problem and can achieve the global convergence, which enables the MDS as an attracting method. The square distance matrix is the

kernel of MDS algorithms which contains the relative distance between all nodes in the network. According to the connectivity of the network (fully connected or partially connected), the MDS algorithms can be further divided into centralized or distributed version. By considering the sensor node uncertainties, a cooperative localization based on WLS was proposed in an asynchronous network (Rui, 2014). The unknown location and clock parameter are jointly estimated through WLS minimization. The least square cooperative localization was further investigated in a NLOS propagation environment and the theoretical bounds were derived to show that the Gaussian NOS bias is the worst case among all biases (Nguyen, 2015).

1. Subspace-based Algorithm

MDS is a field of study related to the search for a low-dimensional space (Cheung, 2005 & Tinsley, 2000) in which points in the space represent the objects such that the distances, or dissimilarities, between the points in the space matching.

MDS was first used in psychometrics and psychophysics and it can be seen as a set of data analysis techniques that display the structure of distance-like data as a geometrical picture (Borg, 1997). Generally, the MDS starts with one or more distance matrices (or similarity matrices) that are presumed to have been derived from points in a multidimensional space. It is usually used to find a placement of the points in a low dimensional space, usually two or three-dimensional, where the distances between points resemble the original similarities. MDS is often used as part of exploratory data analysis or information visualization. By visualizing objects as points in a low-dimensional space, the complexity in the original data matrix can often be reduced while preserving the essential information. MDS is related to principal component analysis, factor analysis, and cluster analysis. The approaches based on MDS have been applied in many fields, such as physical, biological, machine learning and computational chemistry, etc.

MDS based methods belong to deterministic localization and by applying the MDS to a fully connected network, the centralized MDS is developed. For a partially connected network, the distributed MDS can be developed. In this section, we will introduce the MDS localization method (Chan, 2009) in a cooperative network.

Consider a fully connected network with $M+N$ nodes in a plane (three-dimensional case can be extended straightforward). Let $\mathbf{x}_i=[x_i,y_i]^{\mathrm{T}}$ denotes the *i*-th node (*i*=1, 2, …, $M+N$). We repeat the positions of agents and anchors for completeness, $\mathbf{X}=[\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_M]^{\mathrm{T}}$ are agents while $\mathbf{X}_{\mathrm{b}}=[\mathbf{x}_{M+1},\mathbf{x}_{M+2},\ldots,\mathbf{x}_{M+N}]^{\mathrm{T}}$ are anchor nodes and our mission is to estimate $\mathbf{X}$ with inter-measurements from anchors and intra-measurements from agents.

Let $\mathbf{X}_{\mathrm{fs}}\in R^{(\mathrm{C}_{M+N}^2-\mathrm{C}_N^2)\times 2}$ where $\mathrm{C}_i^j=i!/\,j!(i-j)!$ includes all the differences of x_i-x_j and y_i-y_j, $i,j=1,2,\ldots,M+N$ and $i>j$ with at least one of the $\mathbf{x}_i$ and $\mathbf{x}_j$ being unknown. The matrix $\mathbf{X}_{\mathrm{fs}}$ has the following form

$$\mathbf{X}_{\mathrm{fs}}=\begin{bmatrix} x_1\mathbf{1}_N-\mathbf{s}_1^{\mathrm{x}} & y_1\mathbf{1}_N-\mathbf{s}_1^{\mathrm{y}} \\ x_2\mathbf{1}_{N+1}-\mathbf{s}_2^{\mathrm{x}} & y_2\mathbf{1}_{N+1}-\mathbf{s}_2^{\mathrm{y}} \\ \vdots & \vdots \\ x_M\mathbf{1}_{M+N-1}-\mathbf{s}_{N-1}^{\mathrm{x}} & y_M\mathbf{1}_{M+N-1}-\mathbf{s}_{N-1}^{\mathrm{y}} \end{bmatrix} \tag{26}$$

where $\mathbf{s}_i^{\mathrm{x}}=[x_{M+1}\ x_{M+2}\cdots x_{M+i}]^{\mathrm{T}}$, $\mathbf{s}_i^{\mathrm{y}}=[y_{M+1}\ y_{M+2}\cdots y_{M+i}]^{\mathrm{T}}$ and $\mathbf{1}_i$ denotes the *i*-dimensional vector with all 1. The matrix $\mathbf{X}_{\mathrm{fs}}$ can be expressed as an affine combination of matrices with coefficients x_i and y_i, as

$$\mathbf{X}_{\text{fs}} = \sum_{i=1}^{2M} \varphi_i \mathbf{X}_i + \mathbf{X}_0 \tag{27}$$

for $i = 1, 2, \ldots, M$, and φ_i is the element of the vector $\mathbf{\Phi}_{\text{fs}}$ which is defined as

$$\mathbf{\Phi}_{\text{fs}} = \left[x_1, y_1, x_2, y_2, \cdots x_M, y_M \right]^{\text{T}}. \tag{28}$$

The matrix $\mathbf{X}_0$ can be yielded by removing the coordinates of agents and the matrix $\mathbf{X}_i$ contains the signs of coordinates of anchors, i.e., $\mathbf{X}_{2i-1}$ and $\mathbf{X}_{2i}$ contain the signs of x_i and y_i for $i = 1, 2, \ldots, M$, respectively. The other entries of $\mathbf{X}_i$ are all zero when the corresponding element is absent.

$$\mathbf{F}_{\text{fs}} = \mathbf{X}_{\text{fs}} \mathbf{X}_{\text{fs}}^{\text{T}} \tag{29}$$

The matrix $\mathbf{X}_{\text{fs}}$ can be seen as full of column rank which is usually true in practice. Therefore, we have $\text{rank}(\mathbf{F}_{\text{fs}}) = \text{rank}(\mathbf{X}_{\text{fs}}) = 2$.

The matrix $\mathbf{F}_{\text{fs}}$ is constructed from the distance between node pairs (anchor-anchor, anchor-agent and agent-agent) $\{d_{i,j}\}$ where $d_{i,j}$ denotes the distance between the *i*-th node and *j*-th node. With the use of scalar dot product, we have

$$\left(\mathbf{x}_i - \mathbf{x}_j \right)^{\text{T}} \left(\mathbf{x}_m - \mathbf{x}_j \right) = \frac{1}{2} \left(d_{i,j}^2 + d_{j,m}^2 - d_{i,m}^2 \right) \tag{30}$$

Extending the formula above, yields

$$\begin{aligned} \left(\mathbf{x}_i - \mathbf{x}_j \right)^{\text{T}} \left(\mathbf{x}_m - \mathbf{x}_n \right) &= \left(\mathbf{x}_i - \mathbf{x}_j \right)^{\text{T}} \left(\mathbf{x}_m - \mathbf{x}_j - \left(\mathbf{x}_n - \mathbf{x}_j \right) \right) \\ &= \frac{1}{2} \left(d_{i,j}^2 + d_{j,m}^2 - d_{i,m}^2 - \left(d_{i,j}^2 + d_{j,\text{n}}^2 - d_{i,n}^2 \right) \right) \\ &= -\frac{1}{2} \left(d_{i,m}^2 - d_{i,n}^2 + d_{j,\text{n}}^2 - d_{j,m}^2 \right) \\ &= -\frac{1}{2} \left(\mathbf{e}_{M+N,i}^{\text{T}} \mathbf{R} \mathbf{e}_{M+N,m} - \mathbf{e}_{M+N,i}^{\text{T}} \mathbf{R} \mathbf{e}_{M+N,\text{n}} + \mathbf{e}_{M+N,j}^{\text{T}} \mathbf{R} \mathbf{e}_{M+N,n} - \mathbf{e}_{M+N,j}^{\text{T}} \mathbf{R} \mathbf{e}_{M+N,n} \right) \\ &= -\frac{1}{2} \left(\mathbf{e}_{M+N,i} - \mathbf{e}_{M+N,j} \right)^{\text{T}} \mathbf{R} \left(\mathbf{e}_{M+N,m} - \mathbf{e}_{M+N,n} \right) \end{aligned} \tag{31}$$

where $[\mathbf{R}]_{i,j} = d_{i,j}^2$ and $\mathbf{e}_{N,i}$ stands for the *i*-th column of the identity matrix with $M+N$ dimension. With the above formula, the matrix $\mathbf{F}_{\text{fs}}$ can be expressed as

$$\mathbf{F}_{\text{fs}} = -\frac{1}{2} \mathbf{K} \mathbf{R} \mathbf{K}^{\text{T}} \tag{32}$$

where $\mathbf{K} = [\mathbf{H}_1 \ \mathbf{H}_2 \ \cdots \ \mathbf{H}_M]^{\text{T}}$ and $\mathbf{H}_i = [-\mathbf{I}_{i-1} \ \ \mathbf{1}_{i-1} \ \ \mathbf{O}_{(i-1)\times(M+N-i)}]^{\text{T}}$ for $i = 1, 2, \cdots, M$. Square matrix $\mathbf{I}_K$ is an identity matrix with *K* dimension and $\mathbf{O}_{(i-1)\times(M+N-i)}$ denotes the $(i-1)\times(M+N-i)$-dimensional matrix with all zero.

Note that if matrix $\mathbf{K}$ is replaced by $\mathbf{J} = \mathbf{I}_{M+N} - \mathbf{1}_{M+N} \mathbf{1}_{M+N}^{\text{T}} / (M+N)$, then the matrix of $\mathbf{F}_{\text{fs}}$ is corresponding to the classical MDS (Shang, 2004 & Cheung, 2005). Since $\mathbf{F}_{\text{fs}}$ is a symmetric positive semidefinite and of rank 2, it can be decomposed with eigenvalue factorization as

$$\mathbf{F}_{\text{fs}} = \mathbf{U}_{\text{s}} \mathbf{\Lambda}_{\text{s}} \mathbf{U}_{\text{s}}^{\text{T}} + \mathbf{U}_{\text{n}} \mathbf{\Lambda}_{\text{n}} \mathbf{U}_{\text{n}}^{\text{T}} \tag{33}$$

where $\mathbf{U}_{\mathrm{s}} \in R^{(C_{M+N}^2 - C_N^2)\times 2}$ and $\mathbf{U}_{\mathrm{n}} \in R^{(C_{M+N}^2 - C_N^2)\times (C_{M+N}^2 - C_N^2 - 2)}$ are the signal and noise eigenvectors, respectively. $\mathbf{\Lambda}_{\mathrm{s}} = \mathrm{diag}(\lambda_1, \lambda_2)$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq 0$, and $\mathbf{\Lambda}_{\mathrm{n}} = \mathrm{diag}(\lambda_3, \lambda_4, \ldots, \lambda_{C_{M+N}^2 - C_N^2})$ where all entries are zeros. Note that $\mathbf{X}_{\mathrm{fs}}$ is orthogonal to the subspace spanned by $\mathbf{U}_{\mathrm{n}}$, then we have

$$\mathbf{U}_{\mathrm{n}}^{\mathrm{T}} \mathbf{X}_{\mathrm{fs}} = \mathbf{0}_{C_{M+N}^2 - C_N^2 - 2} \tag{34}$$

or

$$\sum_{i=1}^{2M} \varphi_i \mathbf{U}_{\mathrm{n}}^{\mathrm{T}} \mathbf{X}_i = -\mathbf{U}_{\mathrm{n}}^{\mathrm{T}} \mathbf{X}_0. \tag{35}$$

When the noise is present, the measured distance $\{r_{i,j}\}$ can be expressed by

$$r_{i,j} = d_{i,j} + q_{i,j} \tag{36}$$

where $q_{i,j} = q_{j,i}$ is the Gaussian noise with zero mean and variance $\sigma_{i,j}^2 > 0$. Then $\{d_{i,j}\}$ will be replaced with $\{r_{i,j}\}$ in (32) and (33) and the orthogonality between $\mathbf{X}_{\mathrm{fs}}$ and $\mathbf{U}_{\mathrm{n}}$ established approximately. Therefore, we have

$$\mathbf{A}_{\mathrm{fs}} \mathbf{\Phi}_{\mathrm{fs}} \approx \mathbf{b}_{\mathrm{fs}} \tag{37}$$

where $\mathbf{A}_{\mathrm{fs}} = [\mathrm{vec}(\mathbf{U}_{\mathrm{n}}^{\mathrm{T}} \mathbf{X}_1)\ \ \mathrm{vec}(\mathbf{U}_{\mathrm{n}}^{\mathrm{T}} \mathbf{X}_2)\ \cdots\ \mathrm{vec}(\mathbf{U}_{\mathrm{n}}^{\mathrm{T}} \mathbf{X}_{2M})]$ and $\mathbf{b}_{\mathrm{fs}} = -\mathrm{vec}(\mathbf{U}_{\mathrm{n}}^{\mathrm{T}} \mathbf{X}_0)$ with $\mathrm{vec}(\square)$ being the vectorization operator. Obviously, the equation above can be solved by the least square

$$\mathbf{\Phi}_{\mathrm{fs}} \approx \left(\mathbf{A}_{\mathrm{fs}}^{\mathrm{T}} \mathbf{A}_{\mathrm{fs}}\right)^{-1} \mathbf{A}_{\mathrm{fs}}^{\mathrm{T}} \mathbf{b}_{\mathrm{fs}} \tag{38}$$

Note that the above algorithm is formulated based on the fully connected network and if the network is partially connected, the distributed version can be deduced. More details can be found in (Chan, 2009).

After constructing the square distance matrix, the resolving process of MDS is based on the subspace principle which is in essence the same as the multiple signal classification (MUSIC) (Schmidt, 1986) algorithm for AOA estimation. However, the AOA estimation of MUSIC from the noise subspace requires the search of the spatial area of interest, while the MDS algorithm for localization based on the full-set subspace support a closed-form solution of (38).

By solving the equation of (38), all unknown coordinates of agents can be obtained and these results can be used to refine the measurements between agents, and therefore, the above algorithm can behave in an iterative manner to achieve better performance.

2. SDP-based Algorithm

Let $M \times N$ and $M \times N$ denote the set of agent-agent and agent-anchor edges, respectively. The measurements between two possible nodes are denoted by $\{d_{ik} : (i,k) \in M \times M\}$ (agent-agent) and $\{\bar{d}_{ik} : (i,k) \in M \times N\}$ (agent-anchor) are noise free, then we have

$$\begin{aligned} \left\|\mathbf{x}_i - \mathbf{x}_k\right\|^2 &= d_{ik}^2, \quad (i,k) \in M \times M, \\ \left\|\mathbf{x}_i - \mathbf{x}_k\right\|^2 &= \bar{d}_{ik}^2, \quad (i,k) \in N \times M. \end{aligned} \tag{39}$$

In general, the equations above are difficult to solve, as the quadratic constraints in it are nonconvex. The problem of determining the feasibility of (40) is NP-hard. However, one can derive a computationally efficient semidefinite relaxation of (40) as follows. First, we observe that

$$\left\|\mathbf{x}_i-\mathbf{x}_k\right\|^2=\mathbf{x}_i^{\mathrm{T}}\mathbf{x}_i-2\mathbf{x}_i^{\mathrm{T}}\mathbf{x}_k+\mathbf{x}_k^{\mathrm{T}}\mathbf{x}_k,\quad (i,k)\in M\ \times M\ .$$

In particular, we can find that $\|\mathbf{x}_i-\mathbf{x}_k\|^2$ is linear in the inner products $\mathbf{x}_i^{\mathrm{T}}\mathbf{x}_i$, $\mathbf{x}_i^{\mathrm{T}}\mathbf{x}_k$ and $\mathbf{x}_k^{\mathrm{T}}\mathbf{x}_k$. Hence, we can rewrite the equation above as

$$\left\|\mathbf{x}_i-\mathbf{x}_k\right\|^2=\left(\mathbf{e}_{M,i}-\mathbf{e}_{N_a,k}\right)^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\left(\mathbf{e}_{M,i}-\mathbf{e}_{M,k}\right)=\mathrm{trace}\left(\mathbf{E}_{ik}\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)\tag{40}$$

where $\mathbf{E}_{ik}=(\mathbf{e}_{M,i}-\mathbf{e}_{M,k})(\mathbf{e}_{M,i}-\mathbf{e}_{M,k})^{\mathrm{T}}$. Similarly, for the anchor-agent node pairs, we have

$$\left\|\mathbf{x}_i-\mathbf{x}_k\right\|^2=\mathbf{x}_i^{\mathrm{T}}\mathbf{x}_i-2\mathbf{x}_i^{\mathrm{T}}\mathbf{x}_k+\mathbf{x}_k^{\mathrm{T}}\mathbf{x}_k,\quad (i,k)\in N\ \times M\ .\tag{41}$$

and it can be transformed as the following form (Luo, 2010)

$$\left\|\mathbf{x}_i-\mathbf{x}_k\right\|^2=\begin{bmatrix}\mathbf{x}_i^{\mathrm{T}} & \mathbf{e}_{M,k}^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\mathbf{I}_2 & \mathbf{X}\\ \mathbf{X}^{\mathrm{T}} & \mathbf{X}^{\mathrm{T}}\mathbf{X}\end{bmatrix}\begin{bmatrix}\mathbf{x}_i\\ \mathbf{e}_{M,k}\end{bmatrix}=\mathrm{trace}\left(\bar{\mathbf{M}}_{ik}\mathbf{Z}\right),\quad (i,k)\in N\ \times M\tag{42}$$

where

$$\bar{\mathbf{M}}_{ik}=\begin{bmatrix}\mathbf{x}_i\\ \mathbf{e}_{M,k}\end{bmatrix}\begin{bmatrix}\mathbf{x}_i^{\mathrm{T}} & \mathbf{e}_{M,k}^{\mathrm{T}}\end{bmatrix},\quad (i,k)\in N\ \times M$$

and

$$\mathbf{Z}=\begin{bmatrix}\mathbf{I}_2 & \mathbf{X}\\ \mathbf{X}^{\mathrm{T}} & \mathbf{X}^{\mathrm{T}}\mathbf{X}\end{bmatrix}=\begin{bmatrix}\mathbf{I}_2\\ \mathbf{X}^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\mathbf{I}_2 & \mathbf{X}\end{bmatrix}.\tag{43}$$

Note that the matrix $\mathbf{Z}$ is a rank-2 positive semidefinite matrix with the upper left submatrix constrained to be an identity matrix. Moreover, using the Schur complement, it is not difficult to show that any rank-2 positive semidefinite matrix $\mathbf{Z}$ whose upper left sub-block is an identity matrix must have the form given by (43) for some $\mathbf{X}$. Moreover, by defining the matrix

$$\mathbf{M}_{ik}=\begin{bmatrix}\mathbf{O}_{2\times 2} & \mathbf{O}_{2\times n}\\ \mathbf{O}_{2\times n}^{\mathrm{T}} & \mathbf{E}_{ik}\end{bmatrix}\tag{44}$$

the cooperative localization can be formulated as

$$\begin{aligned}\text{Find}\quad & \mathbf{Z}\\ \text{s.t.}\quad & \mathrm{trace}\left(\mathbf{M}_{ik}\mathbf{Z}\right)=d_{ik}^2\ \ (i,k)\in M\ \times M\ ,\\ & \mathrm{trace}\left(\bar{\mathbf{M}}_{ik}\mathbf{Z}\right)=\bar{d}_{ik}^2\ \ (i,k)\in N\ \times M\ ,\\ & \mathbf{Z}_{1:2,1:2}=\mathbf{I}_2,\\ & \mathbf{Z}\succeq\mathbf{O},\\ & \mathrm{rank}\left(\mathbf{Z}\right)=2.\end{aligned}\tag{45}$$

By omitting the rank constraint causing non-convexity, the above formulation gives the SDR of cooperative localization. If we solve the SDR of (45), and obtain a rank-*r* solution $\mathbf{Z}$, then we can extract from a set of *r*-dimensional coordinates for the nodes such that those coordinates satisfy the distance constraints (So, 2007).

Note that the above discussions only focused on the case where the measured distances are noise-free. However, in practice, the measured distances are usually corrupted by noise (say, by an additive Gaussian

noise). In this case, there is no solution to satisfy the constraints in the above SDR, and a more general optimization problem should be set up. A MLE based method is introduced as follows (Biswas, 2006 & Liang, 2004).

Suppose there are some measurement errors between agent-agent and agent-anchor denoted by w_{ik} and $\bar{w}_{ik}$, respectively

$$\begin{aligned} d_{ik} &= \|\mathbf{x}_i - \mathbf{x}_k\| + w_{ik}, \quad (i,k) \in M \times M_{\text{a}}, \\ \bar{d}_{ik} &= \|\mathbf{x}_i - \mathbf{x}_k\| + \bar{w}_{ik}, \quad (i,k) \in N \times M . \end{aligned} \tag{46}$$

where we assume the noise $w_{ik} \sim N\left(0,\sigma_{ik}^2\right), (i,k) \in M \times M$, $\bar{w}_{ik} \sim N\left(0,\bar{\sigma}_{ik}^2\right), (i,k) \in N \times M$ and they are independent. Let the ML function p to estimate $\mathbf{X}$, using all distance measurements, be

$$\begin{aligned} p\left(\left(d_{ik},(i,k) \in N \times M ,\bar{d}_{ik},(i,k) \in M \times M \right),\mathbf{X}\right) = & \prod_{(i,k) \in M \times M} \frac{1}{2\pi^{\frac{1}{2}}\sigma_{ik}} \exp\left(-\frac{1}{2\sigma_{ik}^2}\left(d_{ik} - d\left(\mathbf{x}_i,\mathbf{x}_k\right)\right)^2\right) \\ & \prod_{(i,k) \in N \times M} \frac{1}{2\pi^{\frac{1}{2}}\bar{\sigma}_{ik}} \exp\left(-\frac{1}{2\bar{\sigma}_{ik}^2}\left(\bar{d}_{ik} - \bar{d}\left(\mathbf{x}_i,\mathbf{x}_k\right)\right)^2\right) \end{aligned} \tag{47}$$

where $d\left(\mathbf{x}_i,\mathbf{x}_k\right)$ and $\bar{d}\left(\mathbf{x}_i,\mathbf{x}_k\right)$ denote the true distance between $\mathbf{x}_i$ and $\mathbf{x}_k$ for $(i,k) \in M \times M$ and $(i,k) \in N \times M$, respectively. Then the MLE of the locations can be given by

$$\mathbf{X}_{\text{ML}} = \arg\max_{\mathbf{X}} p\left(\left(d_{ik},(i,k) \in M \times M ,\bar{d}_{ik},(i,k) \in N \times M \right),\mathbf{X}\right). \tag{48}$$

Then $\mathbf{X}_{\text{ML}}$ can be written explicitly as

$$\mathbf{X}_{\text{ML}} = \arg\min_{\mathbf{X}} \sum_{(i,k) \in M \times M} \frac{1}{2\sigma_{ik}^2}\left(d_{ik} - d\left(\mathbf{x}_i,\mathbf{x}_k\right)\right)^2 + \sum_{(i,k) \in N \times M} \frac{1}{2\bar{\sigma}_{ik}^2}\left(d_{ik} - d\left(\mathbf{x}_i,\mathbf{x}_k\right)\right)^2. \tag{49}$$

Note that the variances of the distance measurements are unknown, the following optimization formulation can solve the MLE problem

$$\begin{aligned} \min \quad & \sum_{(i,k) \in M \times M} \varepsilon_{ik}^2 + \sum_{(i,k) \in N \times M} \bar{\varepsilon}_{ik}^2 \\ \text{s.t.} \quad & \left(d_{ik} - \|\mathbf{x}_i - \mathbf{x}_k\|\right) = \varepsilon_{ik}, \quad \forall (i,k) \in M \times M \\ & \left(\bar{d}_{ik} - \|\mathbf{x}_i - \mathbf{x}_k\|\right) = \bar{\varepsilon}_{ik}, \quad \forall (i,k) \in N \times M \end{aligned} \tag{50}$$

and then can be relaxed as

$$\begin{aligned} \min \quad & \alpha \\ \text{s.t.} \quad & \text{trace}\left(\mathbf{M}_{ik}\mathbf{Z}\right) - d_{ik}^2 = \varepsilon_{ik}, \quad \forall (i,k) \in M \times M , \\ & \text{trace}\left(\bar{\mathbf{M}}_{ik}\mathbf{Z}\right) - \bar{d}_{ik}^2 = \bar{\varepsilon}_{ik}, \quad \forall (i,k) \in N \times M , \\ & \left(\sum_{(i,k) \in M \times M} \varepsilon_{ik}^2 + \sum_{(i,k) \in N \times M} \bar{\varepsilon}_{ik}^2\right)^{1/2} \le \alpha \\ & \mathbf{Z}_{1:2,1:2} = \mathbf{I}_2, \\ & \mathbf{Z} \succeq \mathbf{O}. \end{aligned} \tag{51}$$

Solving the SDR of (51), the coordinates of all agents can be obtained. If we restrict the communication range of the node, the optimization problem of (45) and (51) can be implemented by modifying the available set $M \times M$ and $N \times M$.

DISTRIBUTED ALGORITHMS

Although the centralized algorithm can achieve a better accuracy and reach the CRB asymptotically, a central processor is required to process the measurements (Zhou, 2016). When the central processor is unavailable, the distributed algorithm is necessary. Particularly, when a large network of nodes should be processed, the communication bandwidth and computational resources in the central processor become the bottleneck of the performance of the whole network. Moreover, the selection of the node paris for cooperation, refereed to network scheduling, is a vital part in a large network scenario (Wang, 2017). The network scheduling strategies affects the localization errors directly and it is crucial to schedule the nodes with an appropriate distributed manner. Generally, the distributed algorithms for cooperative localization generally fall into one of two categories:

i) Network multilateration: Each node estimates its multi-hop range to the nearest reference nodes. Note that finding the shortest path is readily distributed across the network. When each node has multiple range estimates to known positions, its coordinates are calculated locally via multilateration (Torrieri, 1984 & Caffery, 1999).
ii) Successive refinement: These algorithms try to find the optimum of a global cost function, e.g., LS, WLS, or ML criteria. After the self-localization process, each node will send the related information to its neighbors and neighbors will refine their locations again until convergence.

Typically, better statistical performance is achieved by successive refinement compared to network multilateration, but convergence issues must be addressed.

In this section, distributed localization algorithms are introduced. The study of distributed cooperative localization algorithms is essential, since they are more scalable than the centralized ones and can be applied to large localization networks. Four algorithms will be introduced, namely the cooperative LS and ML localization, and the SPAWN with MMSE and MAP.

1. Cooperative LS/ML Localization

A centralized method for LS/ML based localization has been introduced in (Mao, 2007). However, note that agent i is not aware of the measurements $z_{i\to k}^{(t)}$ $(k \in \mathrm{S}_{i\to}^{(t)})$. Thus in a distributed realization, the second term of the (18) is omitted, and only the first term is considered

$$\frac{\partial C_{\mathrm{dis}}^{(t)}(\mathbf{x})}{\partial \mathbf{x}_i} = \sum_{j\in \mathrm{S}_{\to i}^{(t)}} \frac{\partial c_{j\to i}(\mathbf{z}_{j\to i}^{(t)}; \mathbf{x}_i, \mathbf{x}_j)}{\partial \mathbf{x}_i} \tag{52}$$

Similarly, a gradient decent algorithm can be applied to achieve suboptimal solution.

The initial state can be chosen as the result of location estimation in time slot $(t-1)$

$$\hat{\mathbf{x}}_i^{(t,0)} = \hat{\mathbf{x}}_i^{(t-1,N_{\text{iter}})} \tag{53}$$

where N_{iter} represents the number of iterations at a given time instant. In each iteration, every agent broadcasts its current location estimation $\hat{\mathbf{x}}_i^{(t,l-1)}$ and receives the location estimations from its neighbors $\hat{\mathbf{x}}_j^{(t,l-1)}$ $(j \in \mathrm{S}_{\to i}^{(t)})$. The agent then updates it estimation as

$$\hat{\mathbf{x}}_i^{(t,l)} = \hat{\mathbf{x}}_i^{(t,l-1)} + \delta_i^{(t,l)} \sum_{j \in \mathrm{S}_{\to i}^{(t)}} \psi_{j \to i}^{(t,l-1)} \tag{54}$$

where for notation convenience, we introduce

$$\psi_{j \to i}^{(t,l-1)} \triangleq \left. \frac{\partial c_{j \to i}\left(z_{j \to i}^{(t)}; \mathbf{x}_i, \hat{\mathbf{x}}_j^{(t,l-1)}\right)}{\partial \mathbf{x}_i} \right|_{\mathbf{x}_i = \hat{\mathbf{x}}_i^{(t,l-1)}} \tag{55}$$

Step size $\delta_i^{(t,l)}$ controls the convergence speed and can be chosen by methods such as backtracking line search (Boyd, 2004). The updated result in the maximum number of iteration is set as the location estimation of each agent $\hat{\mathbf{x}}_i^{(t)} = \hat{\mathbf{x}}_i^{(t,N_{\text{iter}})}$.

2. Factor Graph

A factor graph $\mathrm{G}=(\mathrm{V},\mathrm{E})$ is a bipartite graph that describes the structure of function factorization (Kschischang, 2001), where V and E are respectively the sets of vertices and edges. Consider a high dimensional function g, with variables collected in the set $\mathrm{X} = \{\mathbf{x}_1, \cdots, \mathbf{x}_M\}$. Suppose it can be written as a product of several local functions, each having some subset as arguments

$$g(\mathbf{x}_1, \cdots, \mathbf{x}_M) = \prod_{j \in \mathrm{J}} f_j(\mathrm{X}_j) \tag{56}$$

where J is a discrete index set, X_j is a subset of X, and $f_j(\mathrm{X}_j)$ is a function having X_j as the argument. Then the factor graph corresponds to the factorization can be constructed by the following rules

i) In set V, there is a node for each variable $\mathbf{x}_i$ and a vertex for each local function f_j.

ii) In set E, there is an edge connecting node $\mathbf{x}_i$ with vertex f_j if and only if $\mathbf{x}_i \in \mathrm{X}_j$.

Based on the factor graph, the calculation of marginal distributions for each variable can be efficiently completed via the sum-product algorithm (Liu, 2017).

3. Sum Product Algorithm

Let $\mu_{a \to b}$ denotes the message passed from node a to node b in the operation of the sum-product algorithm, and let N_v be the set of neighbors of a given node v. It can be that the neighbors of a variable node are function vertices, and the neighbors of a function vertex are variable nodes in a factor graph. Thus there is one and only one variable node for each edge, and it is called the variable associated with the edge.

Before we introduce the information update rule, an operation named summary should to be defined. Summary over variable x means doing certain operation without x. In the context of MMSE estimation, the operation refers to integral; and it refers to maximization for MAP estimation. Without ambiguity, the nonstandard notation $\sum_{\sim\{x\}} f$ is used to denote summary over x.

- The Sum-Product Update Rule

The message sent from a node v on an edge e is the product of the local function at v (or the unit function if v is a variable node) with all messages received at v on edges other than e, summarized for the variables associated with e.

By applying the above results, an explicit presentation of the message is given by

- Variable to local function

$$\mu_{x\to f}(x)=\prod_{h\in \mathrm{N}_x\setminus\{f\}}\mu_{h\to x}(x) \tag{57}$$

- Local function to variable

$$\mu_{f\to x}(x)=\sum_{\sim\{x\}}\left(f(\mathrm{N}_f)\prod_{y\in \mathrm{N}_f\setminus\{x\}}\mu_{y\to f}(y)\right) \tag{58}$$

By the definition of the factor graph, the marginal distribution g_i of the variable x_i is the product of all messages directed toward node x_i and thus can be calculated. Equivalently, it can be calculated as the product of the messages that are passed in opposite directions over any single edge associated with the variable x_i.

The calculation of marginal distributions for more than one variable is of much interest in many problems. Such a task can be accomplished by applying the procedure above for all indices separately. However, many intermediate values can be reused. In a cycle-free finite graph, the computation for all indices simultaneously can be efficiently accomplished by starting from any leaf node, passing through all possible edges in both directions, and calculating the message. For factor graph with cycles, iterative algorithms are developed but the convergence is not guaranteed. For more details, we refer readers to (Loeliger, 2004).

4. SPAWN Algorithm

As previously mentioned, the factorization of $p(\mathbf{x}^{(0:T)}\,|\,\mathbf{z}^{(1:T)})$ is of great concern in Bayesian methods. By the previous assumptions, the conditional probability function $p(\mathbf{x}^{(0:T)}\,|\,\mathbf{z}^{(1:T)})$ can be factorized as

$$p\left(\mathbf{x}^{(0:T)}\,|\,\mathbf{z}^{(1:T)}\right)\propto p\left(\mathbf{x}^{(0)}\right)p\left(\mathbf{z}_{\mathrm{s}}^{(1:T)}\,|\,\mathbf{x}^{(0:T)}\right)p\left(\mathbf{z}_{\mathrm{r}}^{(1:T)}\,|\,\mathbf{x}^{(0:T)}\right) \tag{59}$$

Then we have

$$p\left(\mathbf{x}^{(0:T)}\,|\,\mathbf{z}^{(1:T)}\right)\propto p\left(\mathbf{x}^{(0)}\right)\prod_{t=1}^{T}\left\{\left[\prod_{i=1}^{M}p\left(\mathbf{x}_i^{(t)}\,|\,\mathbf{x}_i^{(t-1)}\right)p\left(z_{i,\mathrm{s}}^{(t)}\,|\,\mathbf{x}_i^{(t-1)},\mathbf{x}_i^{(t)}\right)\right]\times p\left(\mathbf{z}_{\mathrm{r}}^{(t)}\,|\,\mathbf{x}^{(t)}\right)\right\} \tag{60}$$

where the likelihood function of the relative measurements at time t is given by

$$p\left(\mathbf{z}_{\mathrm{r}}^{(t)}\,|\,\mathbf{x}^{(t)}\right)=\prod_{i=1}^{M}\prod_{j\in \mathrm{S}_{\to i}^{(t)}}p\left(z_{j\to i}^{(t)}\,|\,\mathbf{x}_i^{(t)},\mathbf{x}_j^{(t)}\right) \tag{61}$$

An example network and its factor graph are show in Figure 3.

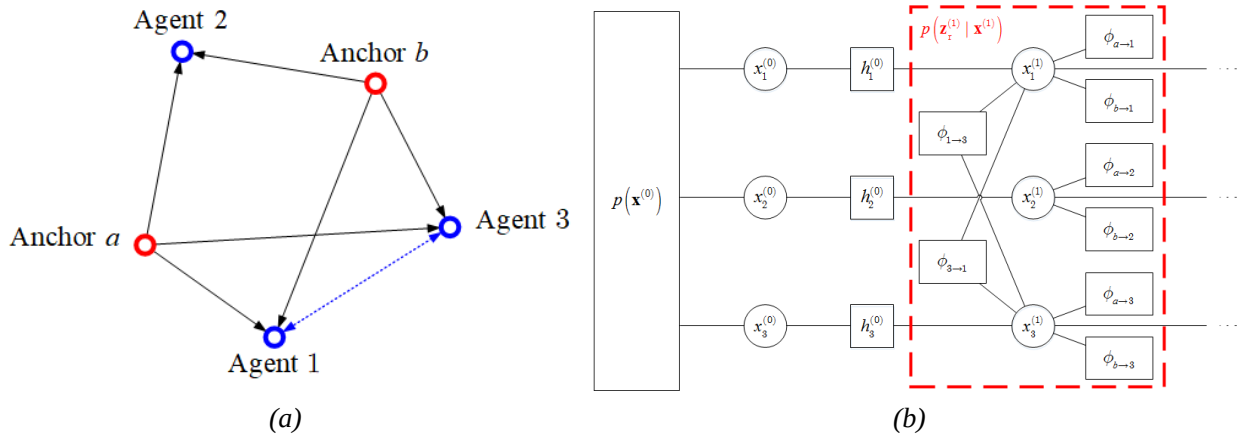


Figure 3. Each agent links to all anchors, and there is a cooperation link between Agent 1 and Agent 3. The red box represents the factor graph of the communication topology.

In mobile networks, the topology of the network varies from time to time. We can break the messages into intra-node messages and inter-messages. The former type is computed locally at each node, while the latter is passed via communication links between node pairs (Wymeersch, 2009).

Since the SPAWN algorithm is based on Bayesian filtering, the description of the SPAWN algorithm can be departed into the prediction phase and the correction phase.

In the prediction phase, each node predicts its current position based on the estimation in last time slot and the intra-node measurement, which is

$$\mu_{h_i^{(t-1)}\to X_i^{(t)}}\left(\mathbf{x}_i^{(t)}\right)\propto\int\underbrace{p\left(\mathbf{x}_i^{(t)}\mid\mathbf{x}_i^{(t-1)}\right)p\left(z_{i,s}^{(t)}\mid\mathbf{x}_i^{(t-1)},\mathbf{x}_i^{(t)}\right)}_{=h_i^{(t-1)}\left(\mathbf{x}_i^{(t-1)},\mathbf{x}_i^{(t)}\right)}\times\mu_{X_i^{(t-1)}\to h_i^{(t-1)}}\left(\mathbf{x}_i^{(t-1)}\right)\mathrm{d}\mathbf{x}_i^{(t-1)}\tag{62}$$

In the correction phase, we define message $b_{X_i^{(t)}}^{(l)}(\cdot)$ as the belief of agent i at the l-th iteration in time slot t, which is initialized by the estimation in the prediction phase

$$b_{X_i^{(t)}}^{(0)}(\cdot)=\mu_{h_i^{(t-1)}\to X_i^{(t)}}(\cdot).\tag{63}$$

In the following iteration, each node broadcast its belief, and receives the ones from its neighbors. The message passing from node j to agent i is given by

$$\mu_{\phi_{j\to i}\to X_i^{(t)}}^{(l)}\left(\mathbf{x}_i^{(t)}\right)\propto\int p\left(z_{j\to i}^{(t)}\mid\mathbf{x}_i^{(t)},\mathbf{x}_j^{(t)}\right)b_{X_j^{(t)}}^{(l-1)}\left(\mathbf{x}_j^{(t)}\right)\mathrm{d}\mathbf{x}_j^{(t)}\tag{64}$$

and the belief of agent i after this iteration is updated by

$$b_{X_i^{(t)}}^{(l)}\left(\mathbf{x}_i^{(t)}\right)\propto\mu_{h_i^{(t-1)}\to X_i^{(t)}}\left(\mathbf{x}_i^{(t)}\right)\prod_{j\in\mathrm{S}_{\to i}^{(t)}}\mu_{\phi_{j\to i}\to X_i^{(t)}}^{(l)}\left(\mathbf{x}_i^{(t)}\right).\tag{65}$$

At time t, the MMSE/MAP estimations can be determined at each agent by taking the mean or the mode of the final messages $b_{X_i^{(t)}}^{(N_{\mathrm{iter}})}(\cdot)$.

COOPERATIVE LOCALIZATION IN COMPLEX ENVIRONMENTS

As aforementioned, localization suffers accuracy degradation in harsh propagation environments, i.e., in the presence of obstacles, where the NLOS errors will result in large localization errors when *a priori*

knowledge of NLOS distribution is unavailable (Shen, 2010). To understand the theoretical performance of cooperative localization in wireless network is essential to the system design and algorithm development. The theoretical bounds of cooperative localization under NLOS conditions was analyzed in (Nguyen, 2015) and a distributed cooperative algorithm based on LS-Gaussian variational message passing was proposed. The localization problem based on a geometric interpretation was formulated as an implicit convex feasibility problem in which some of the sets depend on the agents' positions and apply a parallel projection onto convex sets approach to estimate the unknown agents' positions (Gholami, 2013). By taking the NLOS biases as nuisance parameters along with agents' locations, a simple and efficient SDP based localization was given in (Vaghefi, 2015).

In this subsection, the fundamental limits of cooperative localization under NLOS conditions will be briefly discussed. Note that the localization under NLOS conditions in a cooperative network is still not fully investigated and many open problems requires further investigation.

1. Cooperative Localization Model under NLOS Conditions

The network settings are repeated here for completeness. There are M agents and N anchors with positions $\mathbf{X}=[\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_M]^{\mathrm{T}}$ and $\mathbf{X}_{\mathrm{b}}=[\mathbf{x}_{M+1},\mathbf{x}_{M+2},\ldots,\mathbf{x}_{M+N}]^{\mathrm{T}}$, respectively. The distance estimate between two node i and j at n-th measurement under NLOS conditions can be model as (Van, 2016)

$$\hat{d}_{i,j}[k]=\left\|\mathbf{x}_i-\mathbf{x}_j\right\|+z_{i,j}[k]+b_{i,j}[k],i,j\in M\ \cup N\ \times M\ \cup N \tag{66}$$

where $z_{i,j}[k]\square\ N\ \left(0,\sigma_{i,j}^2\right)$ is Gaussian measurement noise; $b_{i,j}[k]$ is the bias term accounting for the NLOS effect with $\mathrm{E}\{b_{i,j}\}=\mu_{i,j}$ and $\mathrm{Var}\{b_{i,j}\}=\varsigma_{i,j}^2$. NLOS errors $b_{i,j}[k]$ are always positive and generally assumed to be much larger than the measurement noise. The sample mean of K (k=1, 2, …, K) ranging measurements is given as follows

$$\hat{d}_{i,j}=\frac{1}{K}\sum_{k=1}^{K}\hat{d}_{i,j}[k]=\left\|\mathbf{x}_i-\mathbf{x}_j\right\|+z_{i,j}+b_{i,j} \tag{67}$$

where $z_{i,j}=(1/K)\sum_{k=1}^{K}z_{i,j}[k]$ and $b_{i,j}=(1/K)\sum_{k=1}^{K}b_{i,j}[k]$. Note that the LOS biases $b_{i,j}[k]$ are independent and identical distributed random variables underlying the same known PDF with a set of unknown parameters.

The LS solution of agents' positions for cooperative localization can be written as

$$\mathbf{X}=\arg\min_{\mathbf{x}} f_{\mathrm{LS}}(\mathbf{x}) \tag{68}$$

with

$$f_{\mathrm{LS}}(\mathbf{x})=\sum_{i\in M}\sum_{\substack{j>i\\ j\in M\ \cup M}}\frac{c_{i,j}}{\sigma_{i,j}^2+\varsigma_{i,j}^2}\left(\hat{d}_{i,j}-\hat{b}_{i,j}-\left\|\mathbf{x}_i-\mathbf{x}_j\right\|\right)^2$$

where $\hat{b}_{i,j}$ is the estimate of $b_{i,j}$, $c_{i,j}$ is an indicator, such that $c_{i,j}=1$ if nodes i and j can make measurements and $c_{i,j}=0$ otherwise. When the bias estimate $\hat{b}_{i,j}$ is assumed to be unbiased and this LS solution achieves the CRB as $K\to\infty$ (Larsson, 2010).

Moreover, if we square the range estimate, the square-range LS (SR-LS) estimator is given by

$$\mathbf{X}=\arg\min_{\mathbf{x}} f_{\text{SR-LS}}(\mathbf{x}) \tag{69}$$

where

$$f_{\text{SR-LS}}(\mathbf{x}) = \sum_{i\in M} \sum_{\substack{j>i \\ j\in M \cup M}} \frac{c_{i,j}}{\sigma_{i,j}^2 + \varsigma_{i,j}^2} \left[(\hat{d}_{i,j} - \hat{b}_{i,j})^2 - \left\| \mathbf{x}_i - \mathbf{x}_j \right\|^2 \right]^2. \tag{70}$$

Note that the LS and SR-LS formulations are non-convex problem and it is hard to find an exact solution. The SR-LS localization of a non-cooperative case was investigated in (A. Beck, 2008) and its globally optimal solution can be obtained by generalized trust region sub-problem (GTRS) technique (Moré, 1993).

By introducing the weighting factors, the SR-WLS for cooperative localization can be obtained by

$$\mathbf{X} = \arg\min_{\mathbf{x}} f_{\text{SR-WLS}}(\mathbf{x}) \tag{71}$$

where

$$f_{\text{SR-WLS}}(\mathbf{x}) = \sum_{i\in M} \sum_{\substack{j>i \\ j\in M \cup M}} \left\{ \frac{c_{i,j}}{\sigma_{i,j}^2 + \varsigma_{i,j}^2} \frac{1}{(\hat{d}_{i,j} - \hat{b}_{i,j})^2} \times \left[(\hat{d}_{i,j} - \hat{b}_{i,j})^2 - \left\| \mathbf{x}_i - \mathbf{x}_j \right\|^2 \right]^2 \right\}.$$

It is shown that the LS and SR-WLS cooperative solutions are asymptotically efficient as $K \to \infty$ even in the presence of NLOS biases.

2. FIM of Cooperative Localization under NLOS Conditions

Let $\hat{\mathbf{d}}$ be the vector representing all ranging observations $\hat{d}_{i,j}$, the FIM can be calculated by (8) as

$$\mathbf{J}_{\text{e}}(\mathbf{x}) = \text{E}_{\hat{\mathbf{d}}} \left\{ - \begin{bmatrix} \frac{\partial^2 \ln f(\hat{\mathbf{d}};\mathbf{x})}{\partial \mathbf{x}_1 \partial \mathbf{x}_1^{\text{T}}} & \cdots & \frac{\partial^2 \ln f(\hat{\mathbf{d}};\mathbf{x})}{\partial \mathbf{x}_{N_{\text{a}}} \partial \mathbf{x}_1^{\text{T}}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \ln f(\hat{\mathbf{d}};\mathbf{x})}{\partial \mathbf{x}_1 \partial \mathbf{x}_{N_{\text{a}}}^{\text{T}}} & \cdots & \frac{\partial^2 \ln f(\hat{\mathbf{d}};\mathbf{x})}{\partial \mathbf{x}_{N_{\text{a}}} \partial \mathbf{x}_{N_{\text{a}}}^{\text{T}}} \end{bmatrix} \right\} \tag{72}$$

with

$$f(\hat{\mathbf{d}};\mathbf{x}) = \prod_{i\in M} \prod_{\substack{j>i \\ j\in M \cup M}} \left[p_{\hat{d}_{i,j}}(\hat{d}_{i,j}) \right]^{c_{i,j}}. \tag{73}$$

Note that $\hat{d}_{i,j} \,|\, b_{i,j} \sim N_1(b_{i,j} + \left\| \mathbf{x}_i - \mathbf{x}_j \right\|, \sigma_{i,j}^2 / K)$, then we have

$$\frac{\partial p_{\hat{d}_{i,j}}(x)}{\partial \mathbf{p}_i} = \frac{K}{\sigma_{i,\text{j}}^2} \frac{\mathbf{x}_i - \mathbf{x}_j}{\left\| \mathbf{x}_i - \mathbf{x}_j \right\|} \left[(x - \left\| \mathbf{x}_i - \mathbf{x}_j \right\|) p_{\hat{d}_{i,j}}(x) - \sqrt{\frac{K}{2\pi\sigma_{i,\text{j}}^2}} \text{E} \left\{ b_{i,j} \text{e}^{-\frac{K}{2\sigma_{i,\text{j}}^2}(x - b_{i,j} - \left\| \mathbf{x}_i - \mathbf{x}_j \right\|)^2} \right\} \right] \tag{74}$$

Let

$$V_{i,j} = \frac{\sqrt{K}}{\sigma_{i,j}^2}(z_{i,j} + b_{i,j}), \quad U_{i,j} = \frac{\sqrt{K}}{\sigma_{i,j}^2} b_{i,j}$$

be the normalized noise-plus-bias variable and the normalized NLOS bias of ranging $\hat{d}_{i,j}$, respectively. Then it follows the form of (74) that

$$\mathrm{E}_{\hat{d}_{i,j}}\left\{\frac{\partial^2 \ln p_{\hat{d}_{i,j}}(\hat{d}_{i,j})}{\partial \mathbf{p}_i \partial \mathbf{p}_j^{\mathrm{T}}}\right\} = \mathrm{E}_{\hat{d}_{i,j}}\left\{-\frac{\partial^2 \ln p_{\hat{d}_{i,j}}(\hat{d}_{i,j})}{\partial \mathbf{p}_i \partial \mathbf{p}_i^{\mathrm{T}}}\right\},\ \forall j \in M\ \cup M\ , j > i.$$

$$= \frac{N}{\sigma_{i,\mathrm{j}}^2} \frac{\left(\mathbf{x}_i - \mathbf{x}_j\right)\left(\mathbf{x}_i - \mathbf{x}_j\right)^{\mathrm{T}}}{\left\|\mathbf{x}_i - \mathbf{x}_j\right\|} \chi_{i,j}$$

where $\chi_{i,j}$ is called the LOS figure for the link between nodes *i* and *j*, is given by

$$\begin{aligned} \chi_{i,j} &= \mathrm{E}_{V_{i,j}}\left\{\left[V_{i,j} - \frac{\mathrm{E}_{U_{i,j}}\left\{U_{i,j}\mathrm{e}^{-\left(V_{i,j}-U_{i,j}\right)^2/2}\right\}}{\sqrt{2\pi}\, p_{V_{i,j}}\left(V_{i,j}\right)}\right]^2\right\} \\ &= 1 - \frac{\varsigma_{i,j}^2 + K\mu_{i,j}^2}{\sigma_{i,j}^2} + \int_{-\infty}^{+\infty} \frac{\left(\mathrm{E}_{U_{i,j}}\{U_{i,j}\mathrm{e}^{-\left(x-U_{i,j}\right)^2/2}\}\right)^2}{\sqrt{2\pi}\mathrm{E}_{U_{i,j}}\{e^{-\left(x-U_{i,j}\right)^2/2}\}}\mathrm{d}x. \end{aligned} \tag{75}$$

Then we obtain the (i, j) 2×2 submatrix of the FIM $\mathbf{J}_{\mathrm{e}}\left(\mathbf{x}\right)$ as

$$\mathrm{E}_{\hat{\mathbf{d}}}\left\{-\frac{\partial^2 \ln f\left(\hat{\mathbf{d}};\mathbf{x}\right)}{\partial \mathbf{x}_i \partial \mathbf{x}_j^{\mathrm{T}}}\right\} = \begin{cases} \dfrac{Kc_{i,j}}{\sigma_{i,j}^2} \dfrac{\left(\mathbf{x}_i - \mathbf{x}_j\right)\left(\mathbf{x}_i - \mathbf{x}_j\right)^{\mathrm{T}}}{\left\|\mathbf{x}_i - \mathbf{x}_j\right\|} \chi_{i,j} & j = i, \\ -\sum_{l\in M\ \cup M,\ l\neq i} \dfrac{Kc_{i,l}}{\sigma_{i,l}^2} \dfrac{\left(\mathbf{x}_i - \mathbf{x}_l\right)\left(\mathbf{x}_i - \mathbf{x}_l\right)^{\mathrm{T}}}{\left\|\mathbf{x}_i - \mathbf{x}_l\right\|} \chi_{i,j} & j \neq i. \end{cases} \tag{76}$$

Note that when all LOS figures are equal to 1, (76) reduces to the FIM expression for the LOS-only environment. The best and worst case FIMs can be given as follows.

Let $\chi_{i,j}^{\mathrm{G}}$ and $\chi_{i,j}^{\mathrm{LOS}}$ be the LOS figures for a Gaussian bias $b_{i,j} \sim N_1(\mu_{i,j}, \varsigma_{i,j}^2 / K)$ and a constant bias $b_{i,j}$ with $\varsigma_{i,j}^2 = 0$ (i.e., $b_{i,j} = \mu_{i,j}$ with probability 1), respectively. Then, we have

$$\frac{\sigma_{i,j}^2}{\sigma_{i,j}^2 + \varsigma_{i,j}^2} = \chi_{i,j}^{\mathrm{G}} \le \chi_{i,j} \le \chi_{i,j}^{\mathrm{LOS}} = 1 \tag{77}$$

and the FIM $\mathbf{J}_{\mathrm{e}}\left(\mathbf{x}\right)$ is bounded as

$$\mathbf{G}_{\mathrm{e}}\left(\mathbf{x}\right) \preceq \mathbf{J}_{\mathrm{e}}\left(\mathbf{x}\right) \preceq \mathbf{L}_{\mathrm{e}}\left(\mathbf{x}\right) \tag{78}$$

where $\mathbf{G}_{\mathrm{e}}\left(\mathbf{x}\right)$ represents the FIM corresponding to Gaussian biases for all links, i.e., $\chi_{i,j} = \chi_{i,j}^{\mathrm{G}}$ whereas $\mathbf{L}_{\mathrm{e}}\left(\mathbf{x}\right)$ is the FIM for full LOS links or equivalently constant biases for all links ($\chi_{i,j} = 1$).

The proof of above conclusion in (78) can be referred to (Nguyen, 2016) and it discloses that the NLOS biases always degrade the localization accuracy, and the Gaussian bias is the worst case among those with the same mean and variance.

NUMERICAL RESULTS

In this part, the simulations are presented for evaluating the performance gain of the cooperative localization algorithm. Both of centralized methods and distributed methods are given as follows.

1. Centralized Methods

Consider a 2-dimensional dynamic network within a square of $20 \times 20\,\mathrm{m}^2$, consisting 5 anchors and 4 agents. The location of anchors are [20, 20]m, [20, -20]m, [-20, 20]m, [-20, -20]m and [0 0]m while agents are random placed in a rectangle area bounded by $[\pm 15, \pm 15]$ m. The MDS and SDP based algorithms with cooperative and non-cooperative version are illustrated in Figure 4. The agents in the network are assumed to be fully connected with all anchors and agents.

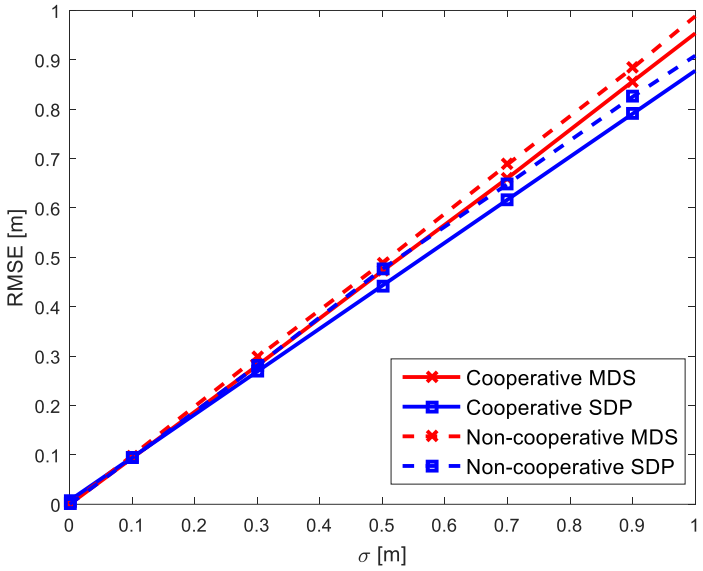


Figure 4. The RMSE comparison of MDS and SDP based non-cooperative and cooperative algorithms. The standard deviation of range measurement noise σ between anchor and agent is ranged from 0 to 1 meter, and the measurement noise between agent and agent is 0.1σ.

As shown in the Figure 4, the root mean squared error (RMSE) of MDS and SDP algorithms (cooperative and non-cooperative) showing the superior performance of cooperative algorithm due to the additional measurements of unknown agents. The performance comparison shows that the cooperative SDP performs superior to the MDS counterparts.

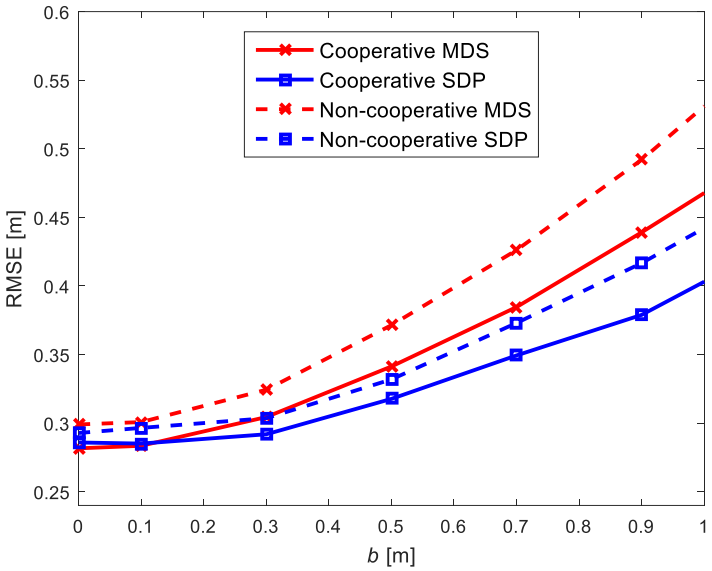


Figure 5. The RMSE comparison of MDS and SDP based non-cooperative and cooperative algorithms. The standard deviation of range measurement noise σ is 0.3 m, and the upper bound of NLOS errors are varying from 0 to 1 m. We randomly choose 1-3 anchors provide the NLOS measurements to each agent and the NLOS errors obey the uniform distribution between 0 to b meters.

Furthermore, we consider the localization performance comparison under NLOS propagation environments. The standard deviation of measurement noise is $\sigma = 0.3$ m. From Figure 5, we can observe that, under NLOS propagation, the cooperation between agents bring more information in localization which performs superior to the noncooperation methods. The performance gain is larger than the result shown in Figure 4.

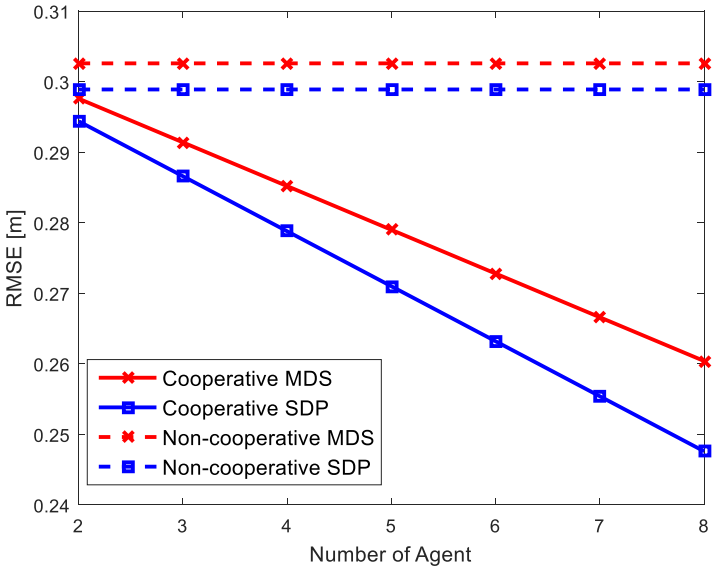


Figure 6. The cooperation gain in centralized methods. The standard deviation of measurement noise σ is 0.3 meters and the number of agents varying from two to eight.

The cooperation gain with the number of agents is shown in Figure 6 and we can observe that the cooperative SDP method has the superior performance when the number of agent increases. The cooperative SDP method also experiences a steeper slope which means the cooperation constrains are tighter than cooperative MDS method. On the other hand, the RMSE of non-cooperative methods almost keep constant during the all range since no more additional information is available.

2. Distributed Methods

Consider a 2-dimensional dynamic network within a square of $20 \times 20\ \mathrm{m}^2$, consisting five agents and four anchors placed at its corners. The agents move independently according to Gaussian random walk $\mathrm{N}(0\ \mathrm{m}, \mathbf{I}_{2\times 2}\ \mathrm{m}^2)$ reflecting in the area, and are equipped with IMUs whose measure errors are modeled as Gaussian variables $\mathrm{N}(0\ \mathrm{m}, 0.01\times\mathbf{I}_{2\times 2}\ \mathrm{m}^2)$. Assume that each agent is within the communication range of all nodes, and the maximum iteration number is set as $N_{\mathrm{iter}} = 20$.

The RMSE of location estimation increases as the variance of ranging measurements grows. The comparison of non-Bayesian and Bayesian methods under both non-cooperative and cooperative schemes are presented in Figure 7.

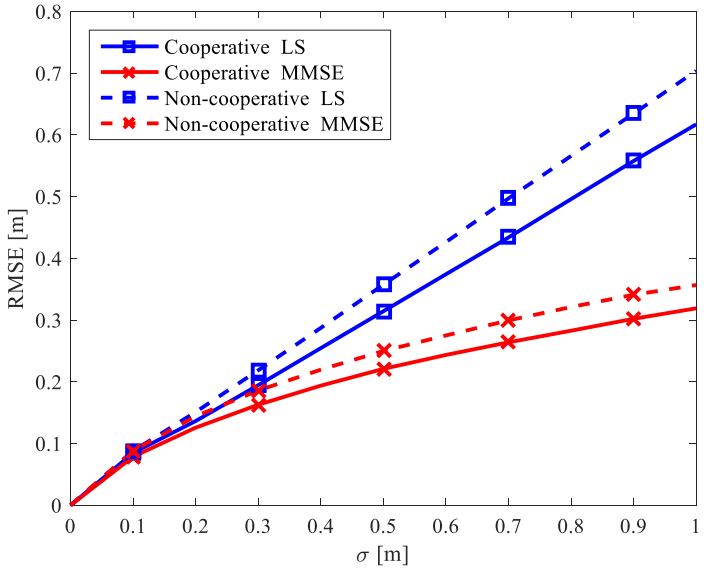


Figure 7. The comparison between Bayesian method (MMSE) and the non-Bayesian method (LS). The standard deviation of range measurement noise σ is ranged from 0 to 1 meter. It is shown that cooperation between agents helps to improve the localization accuracy.

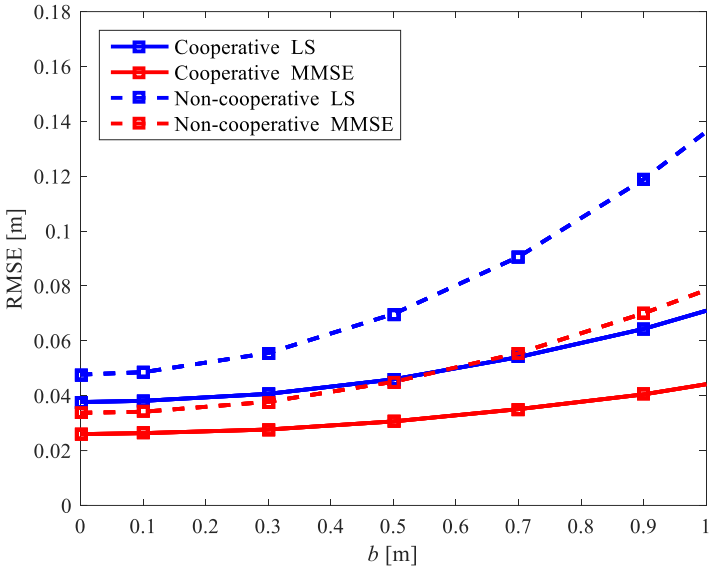


Figure 8. The comparison between Bayesian method (MMSE) and the non-Bayesian method (LS) under NLOS propagation environment. The standard deviation of measurement noise σ is 0.3 meters and the upper bound of NLOS errors are varying from 0 to 1 meter.

The curves in Figure 8 show the superior performance of distributed cooperative localization under NLOS propagation environment since the cooperation between agents are more robust to the NLOS errors. In Figure 8, one to three anchors are randomly chosen that suffer from the NLOS errors which is uniformly distributed between 0 to b meters. From Figure 7 and Figure 8, it is illustrated the cooperative LS is better than non-cooperative MMSE when NLOS errors becomes large.

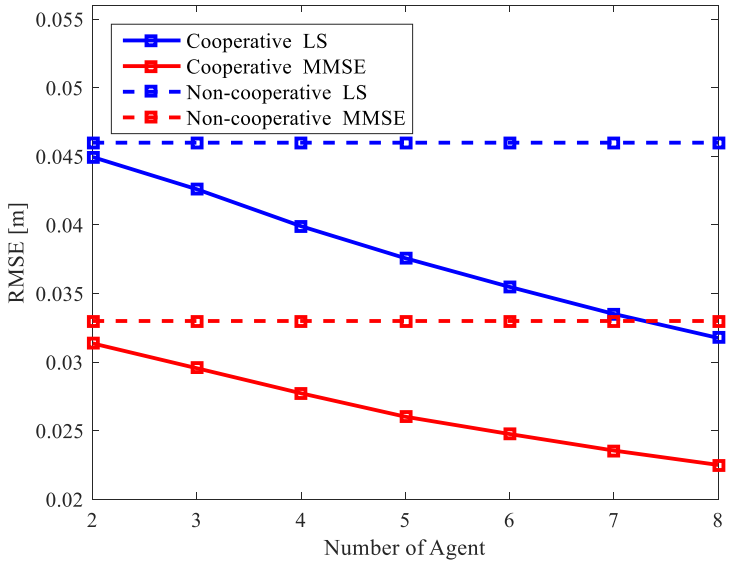


Figure 9. The cooperation gain in Bayesian and non-Bayesian cooperative algorithms. The standard deviation of measurement noise σ is 0.3 meters and the number of agents varying from two to eight.

In Figure 9, we can see that when the number of agents increases, the cooperative methods achieve higher localization accuracy due to more cooperation constraints between these agents. The non-cooperative methods do not benefit from the additional information and the localization accuracy remains constant.

CONCLUSION

This chapter introduces both the theory and algorithms of cooperative localization in wireless networks. There are two kinds of cooperation among nodes in a network, i.e., spatial cooperation between neighboring nodes and temporal cooperation from the previous states of a node itself. With the relative measurements between agents, cooperative localization can provide high-accuracy location estimates of mobile agents, which is important for many coming location-based services. The analysis based on the FIM demonstrates the superior performance of cooperative localization compared to their non-cooperative counterparts. In terms of algorithms, cooperative localization algorithms can be categorized into centralized and distributed ones, depending on whether centralized computation is required. The centralized algorithms can generally achieve higher localization accuracy, while distributed ones require lower communication overheads which is more suitable for practical usage. Then the analyses of cooperative localization under NLOS conditions show that the NLOS biases always degrade the localization accuracy, and the Gaussian bias is the worst case among all of the same mean and variance. Finally, the two types of localization algorithms are compared with numerical results, and the advantages of cooperative localization are also verified through these results.

ACKNOWLEDGEMENT

This work is supported by National Science Foundation of China under Grant No. 61501279 and 91638204.

REFERENCES

Akyildiz I. F. Su W, et al. (2002), A Survey on Sensor Networks, *IEEE Communications. Magzine*, 40(8), 102-114.

Abu-Shaban Z, Zhou X, Abhayapala T D. A Novel TOA-Based Mobile Localization Technique Under Mixed LOS/NLOS Conditions for Cellular Networks. *IEEE Transactions on Vehicular Technology*, 2016, 65(11):8841-8853.

A. Beck, P. Stoica, and J. Li, (2008). Exact and approximate solutions for source localization problems, *IEEE Transaction Signal Processing*, 56(5): 1770-1778.

Biswas P., Liang T.C., & Toh K.C., et al. (2006), Semidefinite programming approaches for sensor network localization with noisy distance measurements, *IEEE Transactions on Automation Science Eng*ineering, 3 (4), 360–371.

Borg I. & Groenen P. (1997), Modern Multidimensional Scaling, Theory and Applications New York: Springer-Verlag.

Biswas P. & Ye Y. (2003), Semidefinite programming for ad hoc wireless sensor network localization, Dept. of Computer Science, Stanford University, Stanford, CA, Tech. Rep.

Biswas P. & Ye Y. (2003), A distributed method for solving semidefinite programs arising from ad hoc wireless sensor network localization, Department. of Computer Science, Stanford Univ., Stanford, CA, Tech. Rep., Oct.

Biswas P., Lian T.-C., Wang T.-C., & Ye Y. (2006), Semidefinite programming based algorithms for sensor network localization, *ACM Transactions Sensor Networks*, 2(2), 188–220.

Chan F. K. W., So H. C., Ma W K. (2009), A Novel Subspace Approach for Cooperative Localization in Wireless Sensor Networks Using Range Measurements. *IEEE Transactions on Signal Processing*, 2009, 57(1), 260-269.

Cheung K. W. and So H. C. (2005), A Multidimensional Scaling Framework for Mobile Location Using Time-of-Arrival Measurements, *IEEE Truncations on Signal Processing*, 53(2).
Cook C. (2012), Radar signals: An introduction to theory and application. Elsevier.
Chazan D., Zakai M., & Ziv J. (1975), Improved Lower Bounds on Signal Parameter Estimation," *IEEE Transaction Information Theory*, 21(1), 90–93.
Cheng Z, Wang Y., Shen Y. (2017), Direct position determination of multiple targets via reduced-dimension beamspace, *Communications Workshops (ICC Workshops), 2017 IEEE International Conference*: 1030-1035.
Cassioli D., Win M. Z. and Molisch A. F., (2002), The ultra-wide bandwidth indoor channel: from statistical model to simulations," in *IEEE Journal on Selected Areas in Communications*, 20(6):1247-1257.
Caffery J. J. (1999), Wireless Location in CDMA Cellular Radio Systems. Norwell, MA: Kluwer.
Capkun S., Hamdi M., & Hubaux J.P. (2001), GPS-free positioning in mobile ad hoc networks, *in Proc. 34th IEEE Hawaii Int. Conf. System Sciences (HICSS-34)*, 34, 1:10..
Chan F. K. W. & So H. C. (2009), Efficient weighted multidimensional scaling for wireless sensor network localization. *IEEE Transactions on Signal Processing*, 57(11), 4548-4553.
Boyd S., Vandenberghe L. (2004), Convex Optimization, Cambridge University Press, New York.
Caceres M.A., Sottile F., & Garello R., et al. (2010), Hybrid GNSS-ToA localization and tracking via cooperative unscented Kalman filter, *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications Workshops*, 272–276.
Dardari D., Conti A., Ferner U., Giorgetti A. and Win M. Z., (2009), Ranging With Ultrawide Bandwidth Signals in Multipath Environments," in Proceedings of the IEEE, 97(2): 404-426.
Dai W.; Shen Y.; Win M. Z., (2017), A Computational Geometry Framework for Efficient Network Localization, *IEEE Transactions on Information Theory*, PP(99): 1-1.
Doherty L., Pister K. S. J., & Ghaoui L.E. (2001), Convex position estimation in wireless sensor networks, *in Proc. IEEE INFOCOM*, 3, 1655–1663.
Foxlin E. (2005), Pedestrian tracking with shoe-mounted inertial sensors, *IEEE Computer Graphics and Applications*, 25(6), 38-46.
Gezici S. et al. (2005), Localization via Ultra-Wideband Radios: A Look at Positioning Aspects for Future Sensor Networks, *IEEE Transaction Signal Processing*, 22(4), 70–84.
Gholami M R, Tetruashvili L, Strom E G, et al. (2013), Cooperative Wireless Sensor Network Positioning via Implicit Convex Feasibility. *IEEE Transactions on Signal Processing*, 61(23):5830-5840.
Hero III A.O., Fessler J.A., & Usman M., (1996), Exploring estimator bias-variance tradeoffs using the uniform CR bound, *IEEE Transaction. Signal Processing*, 44(8), 2026–2041.
Ihler A.T., Fisher J.W., & Moses R.L., et al. (2005), Nonparametric belief propagation for self-localization of sensor networks, *IEEE Journal Selected Area in Communications*, 23 (4), 809–819.
I. Güvenc et al. (2008), NLOS Identification and Weighted Least-Squares Localization for UWB Systems Using Multipath Channel Statistics, *EURASIP Journal Advance Signal Processing*, 36, 1–14.
Moré J. J. (1993), Generalizations of the trust region problem, Optimum. Methods Software, vol. 2, pp. 189–209, Feb.
Kay S. M. (2013), Fundamentals of statistical signal processing: Practical algorithm development. Pearson Education.
Kschischang F R, Frey B J, & Loeliger H. A. (2001), Factor graphs and the sum-product algorithm. *IEEE Transactions on information theory*, 47(2), 498-519.
Larsson E. G. and Danev D., (2010), Accuracy comparison of LS and squared range LS for source localization, *IEEE Transactions on Signal Processing*, 58(2):916-923.
Luo Z-Q., Ma W. k., So A. M-C., Ye Y. & Zhang S. (2010), Semidefinite Relaxation of Quadratic Optimization Problems, *IEEE Signal Processing Magazine*, 27(3), 20-34.
Loeliger H A.(2004), An introduction to factor graphs. *IEEE Signal Processing Magazine*, 21(1), 28-41.

Leeuw J. De (1988), Convergence of the majorization method for multidimensional scaling, *Journal of Classification*, 5(2), 163–180.
Liang T C, Wang T C, Ye Y. (2004), A Gradient Search Method to Round the Semidefinite Programming Relaxation Solution for Ad Hoc Wireless Sensor Network Localization. Sanford University.
Lien J, Ferner U J, Srichavengsup W, et al. (2012), A Comparison of Parametric and Sample-Based Message Representation in Cooperative Localization. *International Journal of Navigation and Observation*, 1687-5990.
Lv T, Gao H, Li X, et al. (2015) Space-Time Hierarchical-Graph Based Cooperative Localization in Wireless Sensor Networks. *IEEE Transactions on Signal Processing*, 2015, 64(2):322-334.
Modeling. San Diego, CA: Academic.Torrieri D. J. (1984), Statistical theory of passive location systems, *IEEE Transactions Aerospace Electronic System*, AES-20, (2) 183–198.
Mao G., Fidan B., & Anderson B. D. O. (2007), Wireless sensor network localization techniques. *Computer networks*, 51(10), 2529-2553.
Marano S, Gifford W M, Wymeersch H, et al. (2010) NLOS identification and mitigation for localization based on UWB experimental data. *IEEE Journal on Selected Areas in Communications*, 28(7):1026-1035.
Patwari N. et al. (2005), Locating the Nodes: Cooperative Localization in Wireless Sensor Networks, *IEEE Transactions Signal Processing*, 22(4), 54–69.
Patwari, N., Hero A.O., & Perkins M. et al. (2003), Relative location estimation in wireless sensor networks, *IEEE Transactions Signal Processing* 51 (8), 2137–2148.
Poor H V. (2013), An introduction to signal detection and estimatio. Springer Science & Business Media.
Savarese C., Rabaey J.M., & Beutel J. (2001), Locationing in distributed ad-hoc wireless sensor networks, *in Proc. IEEE ICASSP*, 2037–2040.
Ramsay J. O. (1982), Some statistical approaches to multidimensional scaling data, J. Roy. Statist. Soc., Series A, 285–312.
Rui L, Ho K C. (2014), Algebraic Solution for Joint Localization and Synchronization of Multiple Sensor Nodes in the Presence of Beacon Uncertainties, *IEEE Transactions Wireless Communications*, 13(9):5196-5210.
Shang Y, Rumi W, Zhang Y, et al (2004). Localization from connectivity in sensor networks. *IEEE Transactions on Parallel & Distributed Systems*, 15(11), 961-974.
Sottile F., Wymeersch H., & Caceres M.A., et al. (2011), Hybrid GNSS-terrestrial cooperative positioning based on particle filter, *IEEE Global Telecommunications Conference*, 1-5.
Savic V., Zazo S. (2012), Reducing communication overhead for cooperative localization using nonparametric belief propagation, *IEEE Wireless Communications Letter*, 1(4), 308–311.
Shen Y., Wymeersch H., & Win M. Z. (2010), Fundamental Limits of Wideband Localization -- Part II: Cooperative Networks, *IEEE Transaction Information Theory*, 56(10), 4981-5000.
Schmidt R. (1986), Multiple emitter location and signal parameter estimation. *IEEE transactions on antennas and propagation*, 34(3), 276-280.
Saxe J. B. (1979), Embeddability of weighted graphs in k –space is strongly NP– hard, *in Proc. 17th Annual. Allerton Conf. Communication, Control, and Computing*, 480-489
Shang Y., Ruml W., Zhang Y., & Fromherz M.P.J. (2003), Localization from mere connectivity, *in Proc. Mobihoc*, 03, 201–212.
Shang, Y., & Ruml, W. (2004). Improved MDS-based localization. *Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies* 4, 2640-2651.
So A. M-C. and Ye Y. (2007), Theory of semidefinite programming for sensor network localization, *Mathematic Programming, Series B*, 109(2), 367–384.
Tenenbaum J. B., Silva V. de, & Langford J. C. (2000), A global geometric framework for nonlinear dimensionality reduction, *Science*, 290, 2319–2323.
Tinsley H. E. A. & Brown S. D. (2000), Handbook of Applied Multivariate Statistics and Mathematical
Wymeersch H., Lien J., & Win M.Z. (2009), Cooperative localization in wireless networks, *Proc. IEEE 97* (2), 427–450.

Win Z. M., Conti A., Mazuelas S.., Shen Y., et al. (2011) Network localization and navigation via cooperation. *IEEE Communications Magazine*, 49(5): 56-62.
Liu Z.; Dai W.; Win M. Z., (2017), Mercury: An Infrastructure-free System for Network Localization and Navigation, *IEEE Transactions on Mobile Computing* , PP(99,):1-1.
Wang J, Shen Y. (2017), On the Discretization Schemes in Map-aided Indoor Localization. *IEEE Communications Letters*, PP(99):1-1.
Wang T.; Shen Y.; Conti A.; Win M. Z.,(2017), Network Navigation with Scheduling: Error Evolution, *IEEE Transactions on Information Theory* , PP(99): 1-1.
Wang Z, Zekavat S A. (2012), Omni-Directional Mobile NLOS Identification and Localization via Multiple Cooperative Nodes. *IEEE Transactions on Mobile Computing*, 11(12):2047-2059.
Wang Y, Wu Y. (2017), An Efficient Semidefinite Relaxation Algorithm for Moving Source Localization Using TDOA and FDOA Measurements. *IEEE Communications Letters*, 21(1): 80-83.
Wang G, So M. C., Li Y. (2016), Robust Convex Approximation Methods for TDOA-Based Localization Under NLOS Conditions. *IEEE Transactions on Signal Processing*, 2016, 64(13):3281-3296.
Shen Y. (2014), Network localization and navigation: theoretical framework, efficient operation, and security assurance, Cambridge, MA, Massachusetts Institute of Technology.
Vaghefi R M, Buehrer R M. (2015), Cooperative Localization in NLOS Environments Using Semidefinite Programming. *IEEE Communications Letters*, 2015, 19(8):1382-1385.
Nguyen T. V., Jeong Y., Shin H., Win M. Z. (2015), Least Square Cooperative Localization. *IEEE Transactions Vehicular Technology*, 64(4):1318-1330.
Zinnes J. L. & MacKay D. B. (1983), Probabilistic multidimensional scaling: Complete and incomplete data, *Psychometrika*, 48(1), 27–48.
Zhou B, Chen Q. (2016) On the Particle-Assisted Stochastic Search Mechanism in Wireless Cooperative Localization. *IEEE Transactions on Wireless Communications*, 15(7):4765-4777.
Zhou F, Shen Y. (2017) On the Outage Probability of Localization in Randomly Deployed Wireless Networks. *IEEE Communications Letters*, 21(4): 901-904.