REPUBLIC OF TÜRKİYE

AKDENIZ UNIVERSITY

COOPERATIVE LOCALIZATION USING PROBABILISTIC INFERENCE

Berk ERCİN

INSTITUTE OF NATURAL AND APPLIED SCIENCES

DEPARTMENT OF COMPUTER ENGINEERING

MASTER OF SCIENCE THESIS

JUNE 2024

ANTALYA

REPUBLIC OF TÜRKİYE
AKDENIZ UNIVERSITY

COOPERATIVE LOCALIZATION USING PROBABILISTIC INFERENCE

Berk ERCİN

INSTITUTE OF NATURAL AND APPLIED SCIENCES

DEPARTMENT OF COMPUTER ENGINEERING

MASTER OF SCIENCE THESIS

JUNE 2024

ANTALYA

REPUBLIC OF TÜRKİYE
AKDENIZ UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

COOPERATIVE LOCALIZATION USING PROBABILISTIC INFERENCE

Berk ERCİN

DEPARTMENT OF COMPUTER ENGINEERING

MASTER OF SCIENCE THESIS

This thesis was unanimously accepted by the jury on 14/06/2024.

Asst. Prof. Dr. Hüseyin Gökhan AKÇAY (Supervisor)
Asst. Prof. Dr. Mustafa Berkay YILMAZ
Asst. Prof. Dr. Mehmet KARAKOÇ

ABSTRACT

COOPERATIVE LOCALIZATION USING PROBABILISTIC INFERENCE

Berk ERCİN

MSc Thesis in Computer Engineering

Supervisor: Asst. Prof. Dr. Hüseyin Gökhan AKÇAY

June 2024; 66 pages

Accurate localization is crucial in today's interconnected world for enhancing the functionality and efficiency of various localization-based services. Traditional localization techniques such as triangulation, trilateration, and fingerprinting offer acceptable accuracy under ideal conditions but face significant challenges in complex environments due to factors like multipath interference, signal attenuation and limited connectivity. These challenges highlight the necessity for more robust and flexible localization methods. Cooperative localization, which utilizes collaborative information sharing among multiple targets and reference points, emerges as a promising solution to enhance localization accuracy and reliability, particularly in environments where traditional systems fall short.

This thesis adopts a probabilistic approach to address the limitations of traditional localization approaches through utilizing collaboration among agents. In this study, formulation of localization problem using a Markov Random Field graphical model is utilized and Nonparametric Belief Propagation to solve it is used. The methodology involves simulating Received Signal Strength measurements as input data, which are then used to approximate the belief distributions of target locations using Particle Filtering and Kernel Density Estimation. Through extensive simulations, the thesis demonstrates that probabilistic framework not only enhances the accuracy and robustness of localization but also adapts effectively to varying environmental conditions and noise levels. The results reveal that probabilistic inference methods significantly improve localization performance by effectively modeling uncertainties and incorporating prior knowledge, making them well-suited for dynamic and complex scenarios where traditional methods may fail.

This comprehensive exploration of cooperative localization using probabilistic inference provides valuable insights into its potential applications and effectiveness, offering a robust alternative to conventional localization techniques.

KEYWORDS: Belief Propagation, Cooperative Localization, Graphical Models, Probabilistic Inference, Received Signal Strength

COMMITTEE: Asst. Prof. Dr. Hüseyin Gökhan AKÇAY

Asst. Prof. Dr. Mustafa Berkay YILMAZ

Asst. Prof. Dr. Mehmet KARAKOÇ

ÖZET

OLASILIKSAL ÇIKARSAMA KULLANARAK İŞBİRLİKLİ KONUMLANDIRMA

Berk ERCİN

Yüksek Lisans Tezi, Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Hüseyin Gökhan AKÇAY

Haziran 2024; 66 sayfa

İsabetli konumlandırmalar, günümüzün birbirine bağlı dünyasında çeşitli lokalizasyon tabanlı hizmetlerin işlevselliğini ve verimliliğini arttırmak için son derece önemlidir. Üçgenleme, trilaterasyon ve parmak izi gibi geleneksel konumlandırma teknikleri ideal koşullar altında kabul edilebilir doğruluk sunarken, sinyallerin çok yolluluğu, sinyal zayıflaması ve sınırlı bağlantı mesafeleri gibi faktörler nedeniyle karmaşık ortamlarda zorluklarla karşılaşırlar. Bu zorluklar, daha güvenilir ve esnek konumsal tespit yöntemlerinin gerekliliğini ortaya koymaktadır. Birden fazla hedef ve referans noktası arasında iş birliğine dayalı bilgi paylaşımını kullanan iş birlikli lokalizasyon, özellikle geleneksel sistemlerin yetersiz kaldığı ortamlarda lokalizasyon doğruluğunu artırmak için umut verici bir çözüm olarak ortaya çıkmaktadır.

Bu tez, aktörler arasındaki iş birliğinden yararlanarak geleneksel lokalizasyon yaklaşımlarının kısıtlamalarını ele almak için olasılıksal bir yaklaşımı benimsemektedir. Bu çalışmada, Markov Rastgele Alan grafik modeli kullanılarak lokalizasyon problemi formüle edilmiş ve çözümü için Parametrik Olmayan Kanı Yayılımı kullanılmıştır. Metodoloji, alınan sinyal gücü ölçümlerinin girdi verisi olarak simüle edilmesini ve daha sonra Parçacık Filtreleme ve Çekirdek Yoğunluğu Tahmini kullanılarak hedef konumların güven dağılımlarının yaklaşık olarak hesaplanmasını içermektedir. Kapsamlı simülasyonlar aracılığıyla tez, olasılıksal çerçevenin yalnızca konumlandırmaların doğruluğunu ve güvenliğini artırmakla kalmayıp aynı zamanda değişen çevresel koşullara ve gürültü seviyelerine etkili bir şekilde uyum sağladığını göstermektedir. Sonuçlar, olasılıksal çıkarım yöntemlerinin belirsizlikleri etkili bir şekilde modelleyerek ve önceki bilgileri dahil ederek lokalizasyon performansını önemli ölçüde artırdığını ve karmaşık senaryolar için oldukça uygun olduğunu ortaya koymaktadır.

Olasılıksal çıkarım kullanan iş birliğine dayalı konumlandırmaların bu kapsamlı araştırması, geleneksel yerelleştirme tekniklerine verimli bir alternatif sunarak potansiyel uygulamaları ve etkinliği hakkında değerli bilgiler sağlar.

ANAHTAR KELİMELER: İşbirliğine Dayalı Konumlandırma, Kanı Yayılımı, Olasılıksal Çıkarsama, Sinyal Gücü, Grafik Modelleri,

JÜRİ: Dr. Öğr. Üyesi Hüseyin Gökhan AKÇAY

Dr. Öğr. Üyesi Mustafa Berkay YILMAZ

Dr. Öğr. Üyesi Mehmet Karakoç

ACKNOWLEDGEMENTS

I would like to extend my deepest gratitude to my supervisor, Asst. Prof. Dr. Hüseyin Gökhan Akçay, for his unwavering support, guidance, and motivation throughout this study. His insights and expertise have been invaluable to the success of this thesis.

I am also profoundly grateful to the other esteemed members of my jury, Asst. Prof. Dr. Mustafa Berkay Yılmaz and Asst. Prof. Dr. Mehmet Karakoç, for their valuable feedback and constructive criticism, which have significantly enhanced the quality of this work.

A special thank you to my Research Assistant friends at Akdeniz University's Computer Engineering department. Your support, collaboration, and shared knowledge have greatly contributed to this journey.

A heartfelt thank you to my friends, the "Babalar" – Ata Varol Koçman, Görkem Geyik, Tolga Kendirli, and Bertuğ Beşkardaş. Your never-ending support, laughter, and camaraderie have been a constant source of joy and motivation. I couldn't have done this without you.

Finally, I wish to express my heartfelt love and appreciation to my mother, Semra Civan. Your endless support, encouragement, and love have been my constant source of strength and inspiration. This achievement would not have been possible without you.

LIST OF CONTENTS

TEXT OF OATH

I declare that this study "Cooperative Localization Using Probabilistic Inference", which I present as master thesis, is in accordance with the academic rules and ethical conduct. I also declare that I cited and referenced all material and results that are not original to this work.

14/06/2024

Berk ERCİN

SYMBOLS AND ABBREVIATIONS

Abbreviations

AoA	: Angle of Arrival
BN	: Bayesian Network
BP	: Belief Propagation
CL	: Cooperative Localization
CRLB	: Cramer-Rao Lower Bound
FG	: Factor Graph
FIM	: Fisher Information Matrix
IMU	: Inertial Measurement Unit
KDE	: Kernel Density Estimation
LBS	: Location-based Services
LS	: Least-Squares
MDS	: Multidimensional Scaling
MIS	: Mixture Importance Sampling
MRF	: Markov Random Field
NBP	: Nonparametric Belief Propagation
PDF	: Probability Density Function
PF	: Particle Filtering
RF	: Radio Frequency
RMSE	: Root Mean Square Error
RSS	: Received Signal Strength
ToA	: Time of Arrival
TLT	: Traditional Localization Techniques
UWB	: Ultrawide Bandwidth
SPAWN	: Sum-Product Algorithm over a Wireless Network

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

In our increasingly interconnected world, precise location determination is paramount for delivering personalized content and services. Whether it's mobile devices, Internet of Things (IoT) sensors, or other connected elements, accurate and timely position estimations provide invaluable insights into operational environments. These insights enable timely responses to various situations, such as navigating urban environments or coordinating disaster response (Ali, Hur, and Park 2019).

Traditional Localization Techniques (TLTs), often employed by Location-based Services (LBS), offer sufficient accuracy under ideal conditions. However, factors like multipath interference, signal attenuation, dynamic obstacles, and limited connectivity can significantly impact localization results (Rathnayake et al. 2023). To overcome these challenges researchers have proposed Cooperative Localization (CL), a collaborative paradigm where targets share information with neighboring unlocalized targets and utilize known reference points, often referred to as anchors, to enhance the accuracy, robustness, and reliability of positioning. This approach is beneficial for LBS applications ranging from navigation systems, social networks, advertising, emergency services, entertainment, and military operations (Sadoun and Al-Bayari 2007).

TLTs employ Radio Frequency (RF) signals, leveraging known geographical landmarks such as base stations to derive geometric positioning through distance or angle measurements relative to these landmarks (Bai, Mulvenna, and Bond 2023).

A widely used TLT is triangulation where the geographic coordinates of targets are determined by measuring the angles from anchors. For example, with two base stations at fixed locations, the angles of the signal from the target point to each base station are measured. As shown in Figure 1.1, the intersection of lines drawn at these angles from the base stations estimates the target's location, forming a triangle with the target as one of its vertices (Wang et al. 2013).

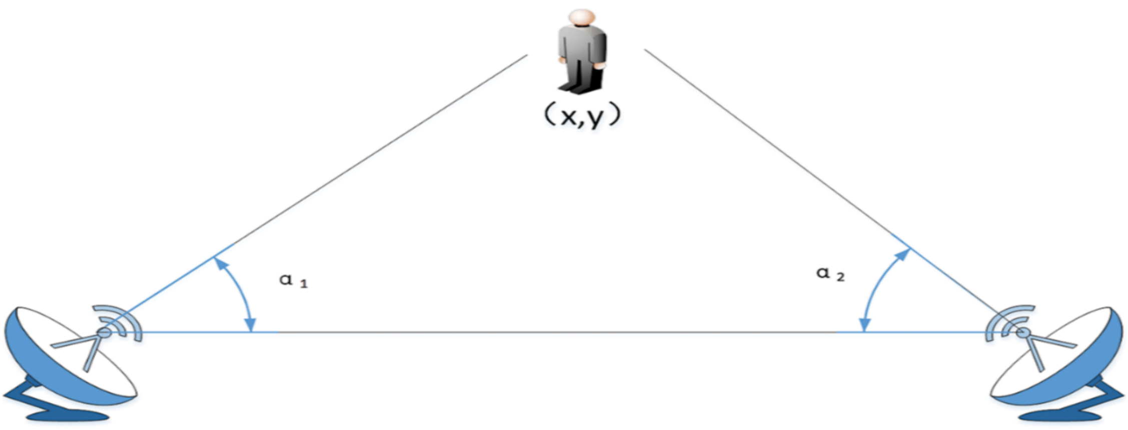


Figure 1.1. Illustration of angle tracking using two satellite dishes to determine the position of an object through triangulation

While triangulation is known for its precise geographic positioning capabilities, it also presents several challenges. First and foremost, triangulation necessitates the presence of pre-existing antenna arrays strategically placed at specific locations, which might not be feasible in areas with limited infrastructure or in remote regions. In indoor environments, where physical barriers such as walls and furniture are prevalent, enclosed spaces can obstruct signal propagation, leading to multipath effects where signals reflect off surfaces, causing significant localization errors (Wang et al. 2013). These multipath effects can create multiple signal paths that confuse the receiver, making it difficult to accurately pinpoint the original signal source.

Additionally, triangulation's accuracy heavily relies on precise angle measurements and the exact orientation of the antenna arrays. Even minor deviations in these measurements, such as slight misalignments or errors in angle calculation, can result in substantial errors in localization, as illustrated in Figure 1.2 (Peng and Sichitiu 2006). This dependency on precise measurements makes triangulation particularly sensitive to any form of measurement inaccuracy or environmental change. Therefore, despite its effectiveness and potential for high accuracy, triangulation requires careful consideration of these factors to ensure the reliability and robustness of the localization results.

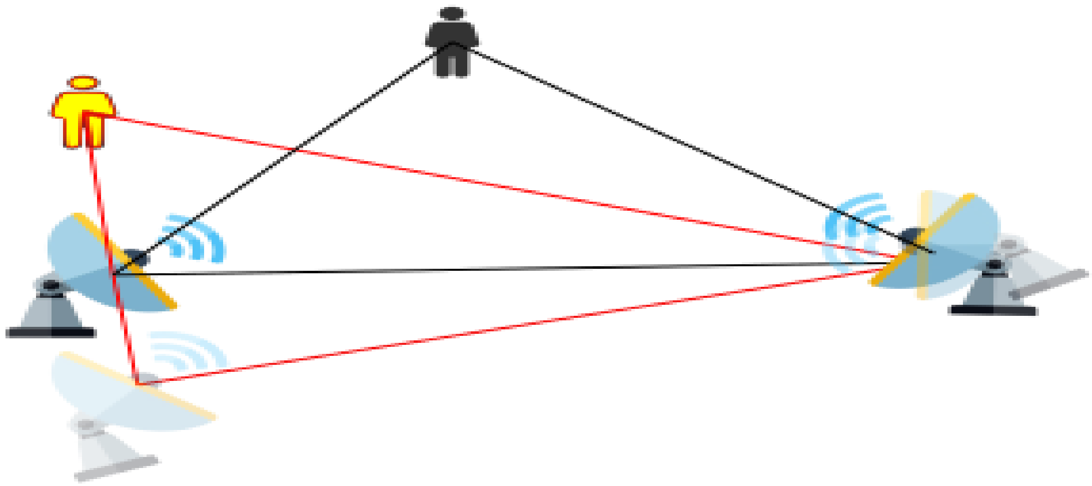


Figure 1.2. Slight deviations in the angle measurements can cause significant localization errors

Trilateration is another widely utilized TLT that determines the position of a target by measuring distances from multiple anchors. Unlike triangulation, which relies on angle measurements, trilateration calculates the location based on the known distances from at least three anchors (Yang and Liu 2010). This method is particularly effective in systems like GPS, where satellites act as anchors, and the intersection can be achieved using the Time of Arrival (ToA) of signals, where the distance is inferred from the signal's travel time. Alternatively, Received Signal Strength (RSS) can be used to estimate distance based on the attenuation of signal power over distance (Bai et al. 2023).

Trilateration's reliance on distance measurements makes it less susceptible to errors caused by angle inaccuracies, and it can provide robust localization in a variety of environments. The flexibility in distance measurement techniques makes trilateration a preferred method for applications.

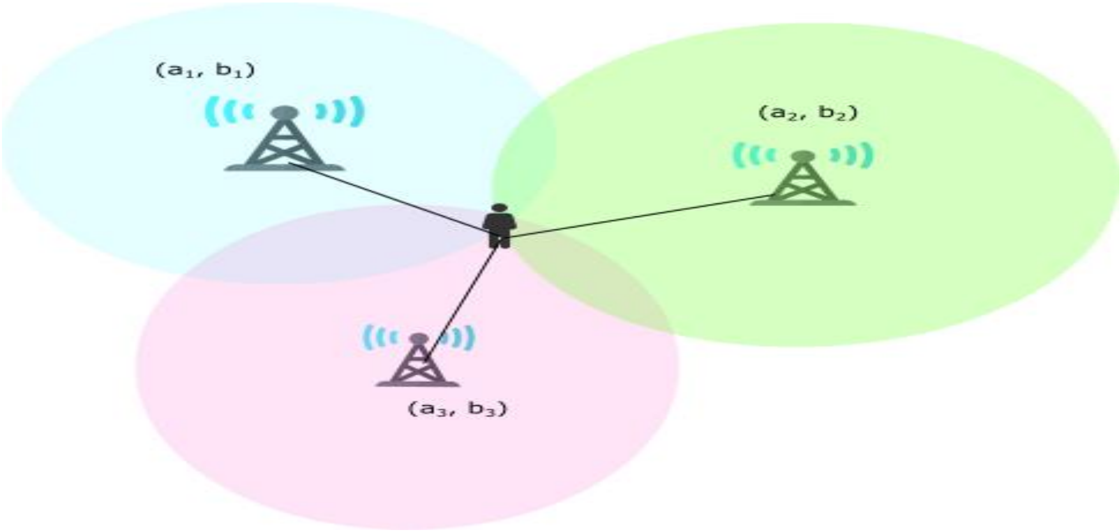


Figure 1.3. Illustration of intersection point of three cell towers to determine the position of an object through trilateration

Trilateration, while effective under certain conditions, has several notable limitations that can impact its reliability in localization tasks. One of the primary challenges is its dependence on the precise measurement of distances from at least three fixed anchors. As illustrated in Figure 1.4, when there are not enough anchor points, localization becomes ambiguous. The requirement for line-of-sight communication between the target and anchors may not always be feasible, particularly in indoor environments, leading to a lack of usable anchors. This method also struggles in scenarios where the anchors are not optimally positioned, resulting in poor geometric dilution of precision. This can lead to significant positional errors, especially when the anchors form narrow angles with respect to the target (Zafari, Gkelias, and Leung 2019). Furthermore, trilateration's reliance on fixed infrastructure limits its flexibility and scalability in dynamic or ad-hoc networks, where the positions of anchors may not be readily available or may frequently change. These limitations highlight the need for more robust and flexible localization techniques.

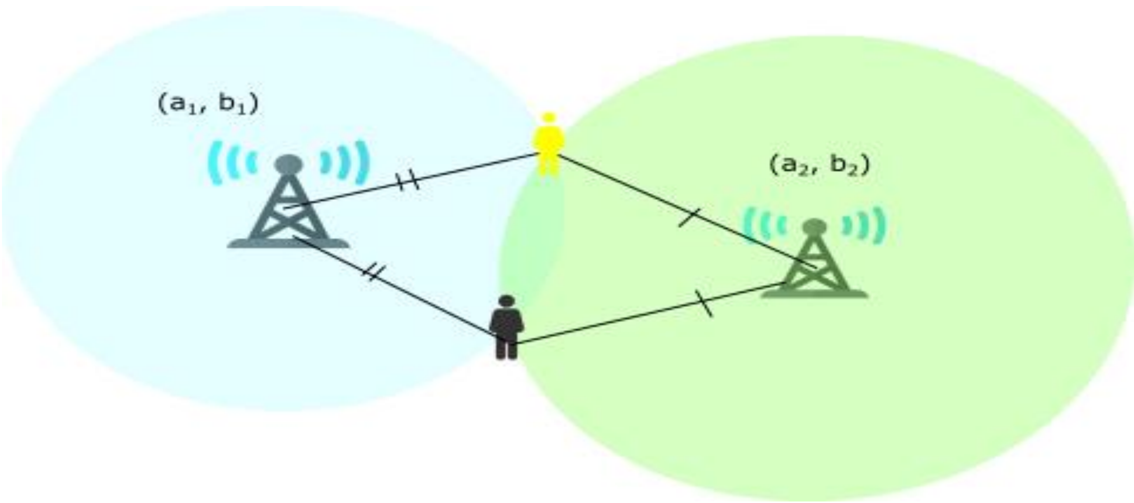


Figure 1.4. Trilateration with limited number of known positions can cause ambiguities

While previously mentioned TLT work well in outside environments, they lack applicability due to signal interference from surfaces and multipathing for indoor positioning. Fingerprinting is a widely used TLT for indoor positioning that leverages existing infrastructure within an environment. It utilizes hardware such as Bluetooth sensors or Wi-Fi access points to measure the RSS from target devices. The fundamental principle behind fingerprinting is the inverse relationship RSS and distance; as the distance between the transmitting and receiving devices increases, the RSS correspondingly decreases. This relationship allows for the inference of location information based on the RSS values obtained from the target devices (Vo and De 2016).

The offline phase of fingerprinting involves a meticulous process of data collection. During this phase, RSS measurements are taken from the environment where the hardware, such as sensors or access points, have known, fixed locations. The precision of the fingerprinting system is highly dependent on both the volume and the quality of collected data. To facilitate the localization of target points, the environment is segmented into numerous small sections, often squares. For each section, RSS measurements are gathered from the access points and subsequently stored in a database. This preparatory state is crucial as it establishes the foundational dataset against which online measurements will be compared.

In contrast, the online phase is where the actual localization occurs. In this phase, the RSS measurements obtained from the target devices are compared with the pre-collected data in the database. The system analyzes the patterns within the measurements to identify the most probable location for the targets. This comparison is based on the RSS measurements from the access points, and the system determines the location that best matches the observed RSS pattern. The technique of fingerprinting, also referred to as database correlation or correspondence, thus relies on the interplay between the offline collection of location-specific RSS data and the online analysis of real-time RSS measurements to accurately localize target devices within an indoor environment.

RSS fingerprinting, while a popular TLT for indoor localization, has several limitations that can affect its accuracy and reliability. One significant challenge is the variability of RSS measurements due to environmental factors such as multipath propagation, signal interference, and changes in the physical environment (e.g., moving people, furniture rearrangements). These factors can cause significant fluctuations in RSS, leading to inconsistent localization results. Additionally, the initial phase of creating the fingerprint database is labor-intensive and time-consuming, as it requires extensive data collection at numerous points throughout the environment. The precision of the localization is highly dependent on the density and quality of the collected fingerprint data. Over time, the fingerprint database can become outdated due to changes in the environment, necessitating periodic updates to maintain accuracy. Moreover, the performance of RSS fingerprinting can be significantly affected by the placement and number of access points, which may not be optimally distributed in all indoor settings (Wen et al. 2015).

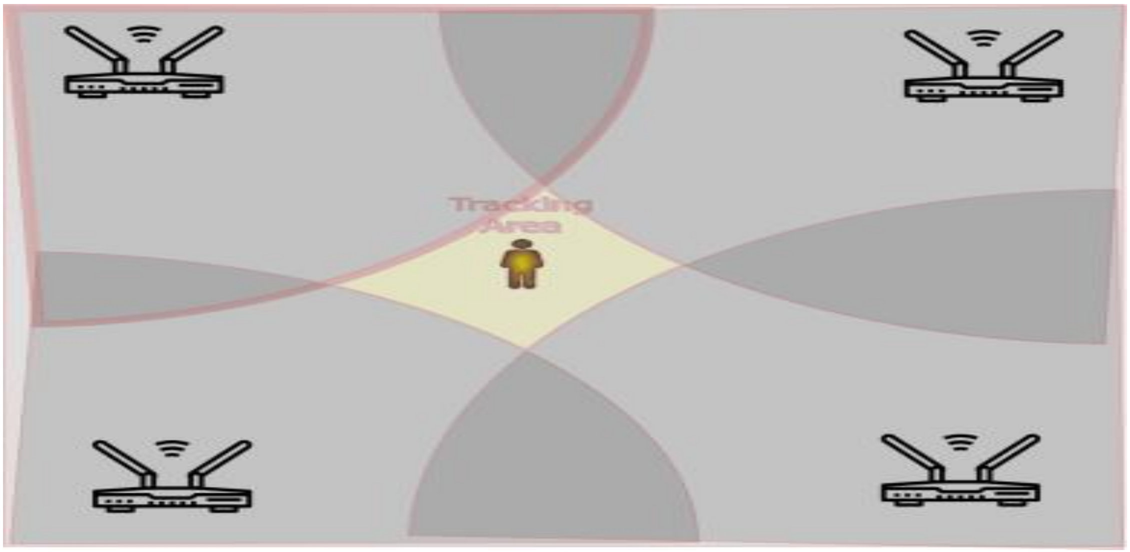


Figure 1.5. RSS fingerprinting relies on connectivity of known points in the environment, causing dead spaces within tracking area

Many TLT are widely used in LBS applications, including those previously mentioned. However, these techniques often require extensive preprocessing steps and environment-specific data collection, or they rely heavily on preexisting hardware installations, which can be costly (Farid, Nordin, and Ismail 2013). While these methods can provide accurate location information under ideal conditions, they can produce erroneous results in complex environments, such as those lacking sufficient number of anchors or with limited communication range. A major drawback of TLT is their focus on positioning individual targets, whereas positioning a group of targets relative to predetermined locations can be more advantageous. The inherent noise in sensor measurements affects each target individually, potentially distorting the group's overall topology. CL addresses this issue by leveraging the collective knowledge of multiple targets, providing a more reliable alternative in situations where TLTs may be inadequate or inaccurate.

CL, also referred to as network localization, or relative positioning, utilizes multiple targets to collaborate to enhance the accuracy and reliability of their location estimates. For localization tasks, it is crucial to have enough anchors in the localization frame. However, due to challenging environments, either through complex structure or limited infrastructure, availability of anchors may be limited. In such scenarios it may still be possible to provide reliable position information to targets through utilizing other targets that they have communication to. This way, they can collectively harness the connectivity within each other to improve positioning. This collaborative aspect of CL can lead to reliable positioning by leveraging the collective intelligence of the network of targets (Patwari et al. 2005).

Given the limitations of TLT and the inherent challenges in positioning within complex environments, it is essential to explore alternative approaches that can provide more robust and flexible solutions. While deterministic methods in CL, such as Multidimensional Scaling (MDS) and iterative Least-Squares (LS), offer fast solutions for localization, they often fail to account for the uncertainties and noise present in real-world environments (Fan et al. 2023; Shang et al. 2003). These methods typically provide point estimates without incorporating the statistical nature of the measurements, leading to potential inaccuracies and reduced reliability in dynamic or noisy settings.

In contrast, probabilistic approaches offer significant advantages by explicitly modeling uncertainties and leveraging prior knowledge to improve localization accuracy. Probabilistic methods use likelihoods of measured distances and known prior distributions to obtain posterior probability density function (PDFs), representing the beliefs about target positions. This allows for the incorporation of measurement noise and environmental variability into the localization process. By treating target positions as random variables and using graphical models such as Markov Random Fields (MRF), Bayesian Networks (BN), or Factor Graphs (FG), probabilistic approaches enable more accurate and reliable inference through techniques like belief propagation (BP) (Sudderth et al. 2003).

Despite the computational challenges associated with exact inference in continuous domains, nonparametric methods such as Particle Filtering (PF) provide practical solutions. PF approximates the beliefs using a set of particles, enabling efficient message passing in nonparametric BP (NBP) (Ihler et al. 2005). This approach allows for the estimation of complex, non-Gaussian probability distributions, making it particularly suited for environments where traditional parametric methods fall short.

In this thesis, we adopt a probabilistic perspective to address the problem of CL. By utilizing RSS values obtained from simulations as measurements, we model the problem as an MRF graphical model. NBP is employed to solve the localization problem, combining PF and Kernel Density Estimation (KDE) to approximate beliefs. This probabilistic approach not only enhances the accuracy and robustness of localization but also provides a flexible framework for adapting to varying environmental conditions and noise levels. Through extensive simulations, we demonstrate the effectiveness of this method and analyze the impact of various factors on localization performance.

The following sections of this thesis explore CL using probabilistic inference methods. The literature review introduces CL and its advantages over non-CL positioning approaches, followed by a taxonomy categorizing CL techniques into centralized vs. distributed, deterministic vs. probabilistic, distance-based vs. distance-free and relative vs. static localization approaches. Various measurement types, such as RSS and ToA, and their evaluation using Cramer-Rao Lower Bound (CRLB) are discussed. The section also explores how to formulate localization as a probabilistic inference problem, highlighting the efficiency of graphical models for scalable solutions. The BP algorithm is reviewed, focusing on effective approximations for marginalizing posterior probabilities. The review concludes with an overview of related works.

In the material and method section, the process of obtaining distance measurements between targets and their utilization is explained. The section details the representation of CL as Bayesian inference, providing a formulation and exploring how NBP can approximate marginals of posterior distributions. Next, we provide results from extensive simulations and explore scenarios in which CL can effectively obtain localization estimates. Finally, we conclude with the conclusion section.

2. LITERATURE REVIEW

In this section, we review the existing literature on CL and provide a comprehensive background from a probabilistic perspective. We begin by presenting the fundamentals of CL and its distinct advantages over non-CL techniques. Following this, we explore various categorizations of CL methods, providing a taxonomy. We then discuss the different types of measurements that can be employed in CL, with a focus on RSS, highlighting their respective strengths and limitations. Moving forward, we explain how the CL problem can be framed as a probabilistic inference problem using graphical models, such as MRFs. We also delve into advanced probabilistic methods, including BP, PF, mixture importance sampling (MIS), and KDE, illustrating how these techniques enhance localization accuracy and robustness. Lastly, we review the related works in CL.

2.1. Cooperative Localization

CL is a method that enables multiple devices within a network to collaboratively determine their positions by sharing information. This approach becomes crucial in environments where TLTs, such as GPS, are impractical due to factors like signal obstruction, limited communication range, or absence of sufficient reference points (Munoz et al. 2009). In CL, nodes utilize relative measurements from their neighbors, such as distance or angle data, to collectively estimate their positions with greater accuracy and reliability, thus making it ideal for applications such as indoor navigation, disaster response, and autonomous coordination.

Non-CL TLTs relying on direct communication with anchors like GPS satellites or base stations, can be highly accurate in open, unobstructed environments but often fail in areas with limited infrastructure or significant physical obstructions. In contrast, CL leverages input from multiple network agents, including non-localized ones, enhancing localization even in non-line-of-sight to anchors. This collaboration allows devices to adapt and maintain accurate localization in limited scenarios (Wymeersch, Lien, and Win 2009). In Figure 2.1a, a scenario with non-CL is presented. Lack of anchors within the communication range can lead to ambiguities in position estimates. Figure 2.1b, on the other hand, highlights how CL can provide more reliable position estimates by utilizing shared information among target nodes.

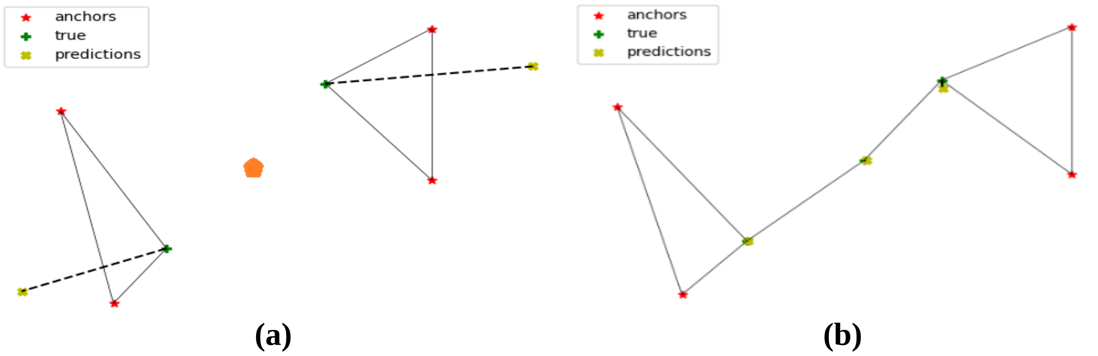


Figure 2.1. Comparison of collaborative and non-collaborative localization; **a)** non-collaborative localization may have failure to localize targets; **b)** collaborative localization allows localization of target points without any reference points

The main advantage of CL is its robustness and adaptability in diverse environments. By incorporating data from multiple agents, CL reduces the effect of individual measurement errors caused by sensors and overall uncertainty in position estimates. This collaborative approach also enhances scalability and lowers dependency on fixed infrastructure, making it cost-effective and practical in scenarios where extensive infrastructure is impractical. Furthermore, CL improves network localization coverage and accuracy, particularly when anchors have poor geometry, by leveraging relative distances among target devices, thus eliminating ambiguities and significantly enhancing performance (Patwari et al. 2005).

2.2. Taxonomy of Cooperative Localization

As described earlier, CL is regarded as a collaborative approach for estimating positions of unlocalized targets through utilizing information obtained from not only anchors, but also other unlocalized agents within the communication vicinity. There are various methods that can be used for this purpose, and each method can be categorized according to their considerations about various aspects of localization problem. Understanding these categorizations can aid researchers in finding the most suitable method for their specific application requirements and to their objectives.

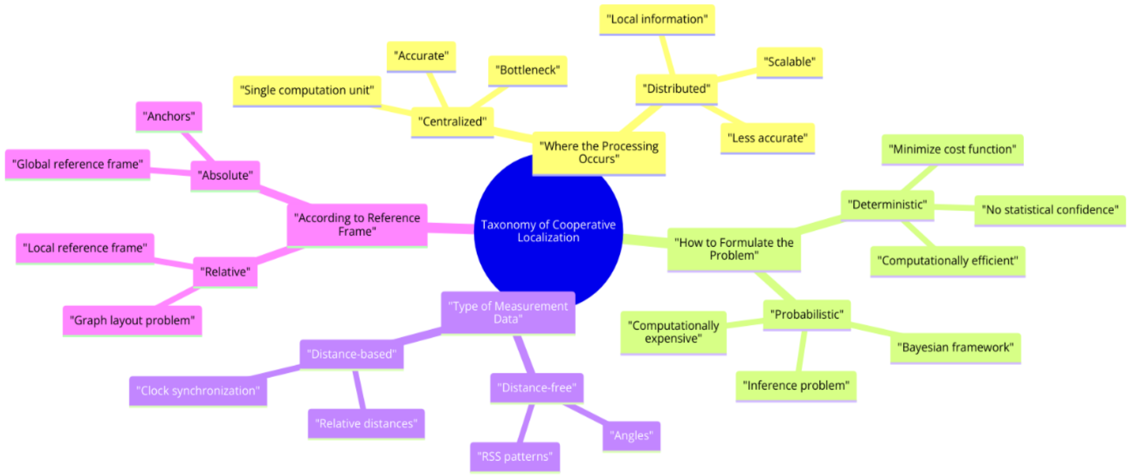


Figure 2.2. Taxonomy of cooperative localization approaches

2.2.1. Where the processing occurs

Localization algorithms can be categorized based on the location where data processing occurs. Centralized approaches gather all the information from the network of devices into a single computation unit (Shang et al. 2003; Vaghefi and Buehrer 2017). Having all the context of the network improves localization performance, therefore centralized approaches often produce more accurate results compared to distributed approaches (Buehrer, Wymeersch, and Vaghefi 2018). However, the central unit must process a large volume of data, which can lead the system into a bottleneck, causing slower response times. While there are many factors affecting the volume of data to be processed, like various measurements obtained from sensors and number of nodes in the network, the latter one is the main contributing factors to the complexity of the system. Such systems require efficient solvers to provide estimations in a timely manner.

Additionally, environmental factors may cause disconnections between nodes of the network and the central computation unit. In such cases, efficient routing algorithms must be implemented.

Distributed approaches, on the other hand, only utilize local information to carry out localization process. In these networks, each node processes its own data and exchanges this information with neighboring nodes. This decentralized approach reduces the computational burden on any one unit and prevents a bottleneck. Not requiring connecting to a central unit, allows this system to be easily scalable (Mensing and Nielsen 2010). However, these systems are less accurate due to their limited context and their performance relies on the number of neighbors they have.

2.2.2. How to formulate the problem

The aim of deterministic approaches is to obtain the best estimate for targets by minimizing a cost function related to localization error. These methods are computationally efficient, and their easy implementation makes them suitable for well-defined and stable environments. However, deterministic approaches do not provide any statistical confidence about the estimates. While it is possible to calculate residual error, this measure alone is not sufficient for the complex nature of CL tasks.

In contrast, probabilistic approaches formulate the localization task as an inference problem, computing marginal probability distributions for each target node in the network. This statistical representation can capture the uncertainties of the estimates. With each new observation, whether from measurements or connectivity information, the underlying Bayesian framework continuously updates these distributions. Although these approaches are computationally expensive compared to deterministic approaches, they can increase the reliability of localization process by fusing multiple observations (Gostar et al. 2021).

2.2.3. Type of measurement data

Localization algorithms can be categorized according to whether they use distance-based or distance-free measurements. A simple but effective approach is to use the relative distances of targets to reference points or to other targets in their vicinity to estimate their location. Approximate distances between agents can be obtained by measuring signal characteristics such as time of flight or signal strength at the receiver. Such measurements provide an easily quantifiable way to represent differences between agents' locations. Various localization algorithms can then minimize a cost function or calculate the likelihood of observing these distances. However, clock synchronization inaccuracies or multipath propagation and interference on observed signals can adversely affect distance approximations, leading to erroneous estimates (Zafari et al. 2019).

Another way that localization algorithms utilize signals is calculating the angles between agents with antenna arrays or matching the observed RSS patterns. Such approaches are known as distance-free methods. Instead of relying on distances they can use existing hardware to obtain angles and apply triangulation, or through using machine learning models, find the most likely position that the target resides in with fingerprinting (Vo and De 2016). Apart from signals, if the initial states of the agents are known, the Inertial Measurement Unit (IMU) can also be used to estimate the positions in dynamic systems. They provide acceleration and angular velocity measurements to estimate the position of the agent in the next state. However, sensor noise in such systems is additive, causing significant deviations over time, therefore requiring frequent synchronization (Liu et al. 2017). Using visual cues involves recognizing fixed landmarks in the environment to perform localization. Such approaches are often used together with IMUs. However, such methods require costly image processing steps therefore, computationally not feasible for many applications. Although simplistic, connectivity information can also be used to collaboratively localize targets. Such approaches only utilize whether they can observe any communication from other agents or not. These methods only provide coarse position estimates however, they are computationally inexpensive. Although distance-free approaches offer simplicity and computational efficiency, they often suffer from lower accuracy and reliability compared to distance-based methods.

2.2.4. According to reference frame

CL methods can be categorized based on their reference frame; absolute or relative. Absolute approaches involve estimation to be made within the context of global reference frame, such as GPS coordinates. These methods rely on anchors with known global positions. Anchors share their positions with targets within their vicinity and often the localization process starts from the neighbors of anchors. Through utilizing distances to anchors, targets can estimate their locations. Estimating positions requires at least three anchors for 2D and four anchors for 3D localization. By utilizing relative distances to other targets, even if a target does not have an anchor within its communication range, it can leverage the information from neighboring targets that are connected to anchors. Although such approaches can provide accurate localization estimates, they are heavily reliant on infrastructure of the environment. Additionally, geometry of anchors regarding to the localization environment plays crucial role in for accuracy of estimates (Rathnayake et al. 2023). Therefore, some researchers entirely focus on determining ideal anchor placement strategies.

Even though anchors can help eliminate measurement noise by comparing observations and their known locations, they are not necessary for localization process. Such approaches which do not utilize any anchors are known as relative localization. In these systems, targets estimate their positions based on relative measurements observed between themselves. When targets are represented as nodes in a graph, this approach is also known as graph layout problem. The performance of such approaches depends on the quality of measurements, number of targets, and their density. An anchor can be introduced into the network when available, allowing the network to shift to the anchor’s global reference frame. To obtain unambiguous position estimates, more anchors are necessary. Without enough anchors, localization will result in arbitrary rotation, reflection, and scale. To register the local reference frame into global one, registration algorithms can be used (Fan et al. 2023).

Table 2.1. Taxonomy of reference works

Works	*Processing*	*Formulation*	*Measurement Type*	*Reference Frame*
(Wymeersch et al. 2009)	Distributed	Probabilistic	Distance-based	Relative
(Mensing and Nielsen 2010)	Centralized	Probabilistic	Distance-based	Absolute
(Gostar et al. 2021)	Centralized	Probabilistic	Distance-free	Relative
(Carrillo-Arce et al. 2013)	Distributed	Probabilistic	Distance-based	Relative
(Vaghefi and Buehrer 2015)	Centralized	Deterministic	Distance-based	Absolute
(Gholami et al. 2013)	Distributed	Deterministic	Distance-based	Relative
(Zaidi et al. 2016)	Distributed	Deterministic	Distance-free	Relative
(Darakeh, Mohammad-Khani, and Azmi 2018)	Distributed	Deterministic	Distance-free	Relative
(Shang et al. 2003)	Centralized	Deterministic	Distance-free	Absolute
(Ihler et al. 2005)	Distributed	Probabilistic	Distance-based	Relative
(Fascista et al. 2018)	Distributed	Probabilistic	Distance-free	Absolute

2.3. Measurements

To perform localization, a measure of dissimilarity is required that characterizes the positions of the targets, so that estimation algorithms can perform optimization on objective functions or inference on graphical models. This measure of dissimilarity can be as simple as binary connectivity information. Such approaches are referred to as range-free methods.

One such method estimates the centroid of neighboring anchors for positioning (Blumenthal et al. 2007). Another method is the DV-HOP algorithm, which calculates the distance between nodes based on hop count and averages the distance per hop (Li, Zhang, and Xiande 2009). Distances, in this case, are approximated according to distance between anchor nodes. MDS is a statistical technique that transforms connectivity information into a spatial configuration of targets in a lower-dimensional space. First, all the targets share their connectivity information to form a neighborhood matrix, then applying dimensionality reduction with eigenvalue decomposition, rough estimates for relative positions can be obtained (Shang et al. 2003).

Such methods are especially dependent on the network geometry and slight deviations in connectivity information are enough to adversely affect localization. When the communication range is limited, such that only nearby targets are visible, neighborhood information can provide rough estimates with arbitrary rotation and scale. However, due to varying communication range capabilities due to heterogeneity of targets, potentially resulting in some devices having either all other devices as neighbors or, they may have no neighbor at all, so there is no information to utilize in the network. Without any quantifiable measurement that represents the dissimilarities of the network elements, localizations can be highly erroneous.

$$d_{ij} = \| x_i - x_j \| + v_{ij} \tag{2.1}$$

On the other hand, range-based approaches estimate the distance between targets based on some measurement with noise to provide representative dissimilarity information. Using such quantitative measurement approaches increases network localization accuracy compared to range-free approaches. The choice of measurements used for this purpose significantly affects the performance of the localization process.

The travel time of signals is often known; therefore, it is straightforward to estimate the distance between any two points. However, this requires precise clock synchronization between transmitter and receiver; any synchronization error can lead to significant errors due to signals traveling near the speed of light. Another approach is to utilize the loss of energy over the distance that the signal travels as an indicator of distance between receiver and transmitter. However, multipath effects, obstacles in the environment and interference from other signal sources can disrupt signal power and cause fluctuations.

These concerns about distance-based measurements can be attributed to variability of sensor noise. Therefore, unless noise mitigation actions are taken, the performance of the system will be limited by the quality of considered measurements. Based on the quality of measurements a theoretical lower bound, such as the Cramer-Rao Lower Bound (CRLB), can be calculated to act as a benchmark for evaluating the localization algorithm. In this section we will review commonly used distance-based measurements for CL and explain how the CRLB can be used to assess the performance of the system.

2.3.1. Time of arrival

ToA is a widely used measurement technique in localization systems. The fundamental principle of ToA is to determine the distance between a transmitter and a receiver by measuring the time it takes for a signal to travel from the transmitter to receiver. Given the known speed of signal propagation, the distance can be calculated as:

$$d_{ij} = \mathrm{c} \, . \, t_{ij} + v_{ij} \tag{2.2}$$

Where d_{ij} is the distance between nodes $\underline{i}$ and j, c is the speed of the signal and t_{ij} is the measured time differences between sending and receiving times of the package.

The primary benefit of ToA is its potential for high accuracy, especially in environments with minimal obstacles and clear line-of-sight conditions. ToA particularly effective in outdoor environments. However, ToA requires precise clock synchronization, any clock discrepancies can lead to substantial errors in distance estimation. Additionally, ToA measurements are susceptible to multipath propagation, where signals reflect off surfaces and arrive at the receiver at different times (Chen, Yao, and Peng 2020).

2.3.2. Received signal strength

RSS is one of the most popular measurement modalities that is used for localization problems. Basic principle of RSS is that signal power diminishes as it propagates so the distance between any two devices has inverse relationship between signal strength and distance (Farahani 2008). The main benefit of using RSS is that it can utilize the preexisting infrastructure of the environment such as Wi-Fi access points installed or Bluetooth connectivity strength between targets. Distance estimation from RSS is straightforward however, it depends greatly on the path loss exponent, in practice it ranges from 2 to 4, based on the complexity of the environment. As the complexity increases path loss exponent increases as well. There exists methods for estimating the path loss exponent along with locations of target devices (Jin et al. 2020).

Assuming two devices i and j are connected in a network, able to send or receive signals from one another, RSS measurements can be modeled as:

$$s_{ij} = P_i - 10\alpha \log_{10} \frac{d_{ij}}{d_0} + v_{ij} \tag{2.3}$$

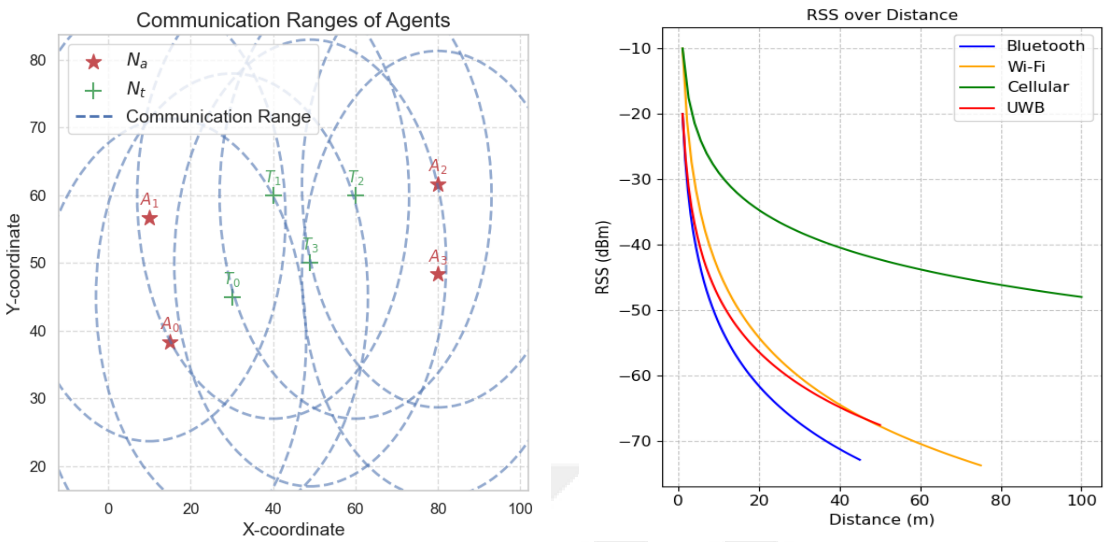
Where s_{ij} is the signal strength by agent i from agent j, P_i is the transmitted power, α is the path loss exponent, d_{ij} is the distance between nodes i and j, v_{ij} is the measurement noise, usually log-normal modeled by $v_{ij} \sim N(0, \sigma^2_{ij})$ (Mao, Anderson, and Fidan 2007).

Table 2.2. RSS dBm ranges of signal technologies

Quality / Technology	Excellent	Good	Fair	Poor
Bluetooth	-30 to -50	-51 to -60	-61 to -70	-71 to -90
Wi-Fi	-30 to -50	-51 to -60	-61 to -70	-71 to -80
Cellular	-50 to -70	-71 to -85	-86 to -100	-101 to -110
UWB	-30 to -50	-51 to -60	-61 to -70	-71 to -80

RF signal technologies are affected with various parameters such as path loss exponent, transmitter power and pathloss signal deviation. Although values provided in

Table 2.2 gives rough estimates, an arbitrary cut off dBm can be selected (Miao et al. 2016).



(a) **(b)**

Figure 2.3. Communication ranges of various signals; **a)** connectivity regions of agents; **b)** distance change over the RSS measurements

Despite its benefits, RSS-based localization faces several challenges. The accuracy of RSS measurements is highly dependent on the environment, particularly the path loss exponent, which can vary significantly in different settings. Factors such as multipath propagation, signal interference, and physical obstacles can cause fluctuations in the received signal strength, leading to errors in distance estimation.

2.3.3. Cramer-Rao lower bound

The CRLB provides a framework for evaluating the accuracy of localization algorithms by establishing the minimum possible variance that any unbiased estimator can achieve. This variance serves as a benchmark to assess the accuracy of estimations based on the measurements taken (Patwari et al. 2005). The statistical nature of the measurement model can influence the variance. The amount of noise presence is another leading factor to higher variance. Scarcity of observations typically means less information, therefore higher variance. CRLB can be calculated without having any prior knowledge about the localization method, but only statistical models of the measurements and observations are required. Suppose we have set of measurements X and a parameter vector θ for positions. The log-likelihood of measurement model relates the observations to the parameters in the form $\log L(\theta; X)$. Fisher Information Matrix (FIM) is then expected value of the negative Hessian of the log-likelihood function.

$$I(\theta) = -E\left[\frac{\partial^2}{\partial\theta\,\partial\theta^T} \log L(\theta; X)\right] \tag{2.4}$$

The lower bound is derived from FIM, which measures how much information the observations provide about estimated positions of targets. The diagonal elements of the inverse FIM indicate the lower bound on the variance of the estimates.

$$\mathrm{CRLB}(\theta_i) \geq [I(\theta)^{-1}]_{ii} \tag{2.5}$$

Equation (2.5) illustrates theoretical minimum variance, setting a benchmark for the accuracy of localization estimates. CRLB provides a lower bound on the variance, serving as a measure of best achievable performance given the noise characteristics and the measurement model.

2.4. Problem Formulation

Formulating the problem for estimating the positions of targets involves defining a mathematical framework to process available information in a structured way. One approach is to formulate a deterministic system, in which all states are assumed to be known and uncertainties can be neglected or accounted for minimal error. Deterministic methods typically either solve a series of equations to achieve a closed-form solution or minimize a cost function to find the global minimum (Shang et al. 2003; Zhang et al. 2016). These approaches can deliver fast and accurate results when the measurements are dependable, and the environment is stable. To adapt to dynamic changes, Kalman filtering can be used to estimate the state dynamics however, they require system to be linearized (Bahr, Walter, and Leonard 2009).

$$\hat{x}_{LS} = \mathrm{argmin}\left||z - f(x)|\right|^2 \tag{2.6}$$

However, noise models in complex environments are often non-Gaussian and deterministic solutions do not provide metrics for uncertainties about estimates. In contrast, probabilistic approaches inherently provide a measure of uncertainty due to their nature of using probability distributions to represent location estimates and update beliefs continuously based on new measurements introduced. They can use extended Kalman filtering and PF to deal with nonlinear and non-Gaussian systems (Djuric et al. 2003).

$$\hat{x}_{MMSE} = \int p_x(x|z)dx \tag{2.7}$$

Although probabilistic approaches are robust to non-Gaussian noise, inconsistencies, or when the nodes have poor geometry, they are computationally intensive, often requiring significant resources to handle the large state spaces and continuous domains (Nagaty et al. 2015). Therefore, such methods are generally used alongside distributed approaches to provide efficient solutions and scalability. In this thesis we will focus on probabilistic approaches over deterministic ones, because they provide flexible framework applicable to changing environments and robust to non-Gaussian noise.

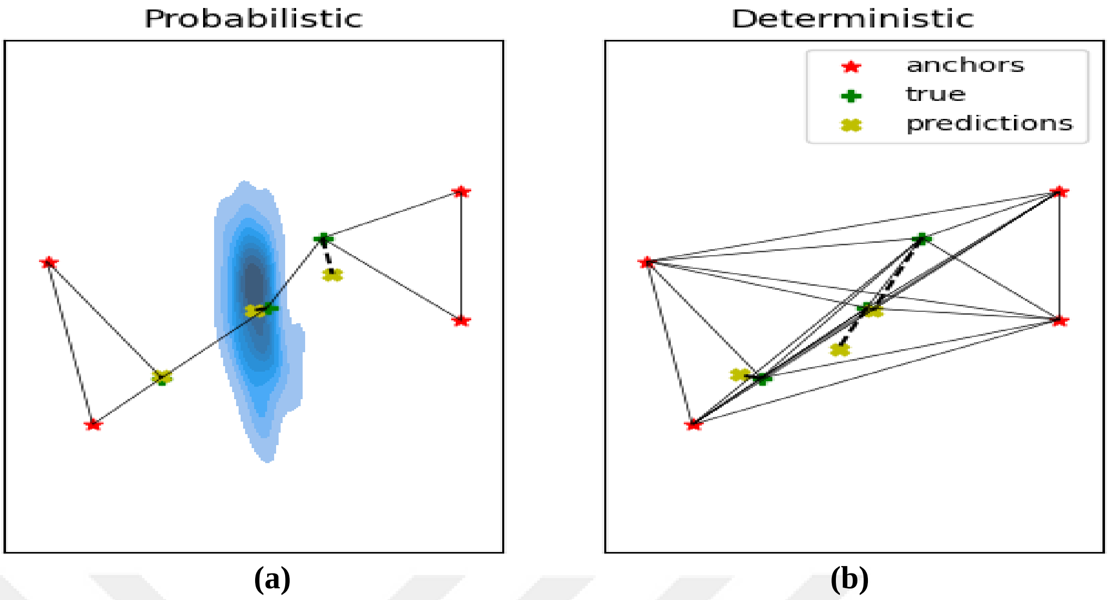


Figure 2.4. a) Belief representation of a single agent shown with blue distribution; **b)** results obtained with MDS localization technique

2.4.1. Probabilistic approach to localization

Probabilistic approaches can offer a robust framework for handling uncertainties of the position estimates. Although these methods do not correct uncertainties, they provide an efficient way to represent them. Consequently, provide foresight into how the model could be refined to reduce uncertainties. Bayesian inference is the central mechanism that systematically updates beliefs about position estimates based on new measurements and prior knowledge (Wymeersch et al. 2009). In this framework positions of nodes are treated as random variables, and the goal is to estimate the posterior distribution of these variables given the observations. Let $\mathbf{X} = \{x_1, x_2, \ldots, x_N\}$ denote the set of all node positions in the network and $\mathbf{Z}$ set of all measurements, including all pairwise information, the joint posterior distribution of the entire network can be expressed as:

$$\mathrm{p}(\boldsymbol{X} \mid \boldsymbol{Z}) \propto p(\boldsymbol{Z} \mid \boldsymbol{X})\mathrm{p}(\boldsymbol{X}) \tag{2.8}$$

Where $\mathrm{p}(\boldsymbol{Z} \mid \boldsymbol{X})$ is the likelihood of the observations given the positions, $\mathrm{p}(\boldsymbol{X})$ is the prior distribution of the positions. Through likelihoods, Bayesian inference incorporates uncertainties into estimations and any prior knowledge, for example locations of anchors, can be introduced before any measurements are observed (Savic and Zazo 2009). With each new measurement from sensors becoming available, this probabilistic inference approach iteratively improves estimates. Representative power of likelihoods can also be used to model non-Gaussian noise, which often is the case in real-world applications due to multipath effects, signal interference from other sources or obstacles in the environment. This can be achieved with techniques such as KDE or mixture models (Ihler and Thomas 2005).

While probabilistic framework offers many advantages, it also has limitations such as computational complexity of calculating the posterior distribution under continuous domains, which requires costly integrals. Exact inference is challenging to achieve under such scenarios. Therefore, approximate solutions such as PF can be used. PF represents the posterior distribution with a set of discrete particles. Therefore, these approximations can be used instead of explicit integrals. Another limitation is the potential convergence to local optima, and this can be mitigated by using efficient sampling techniques such as MIS to guide inference process.

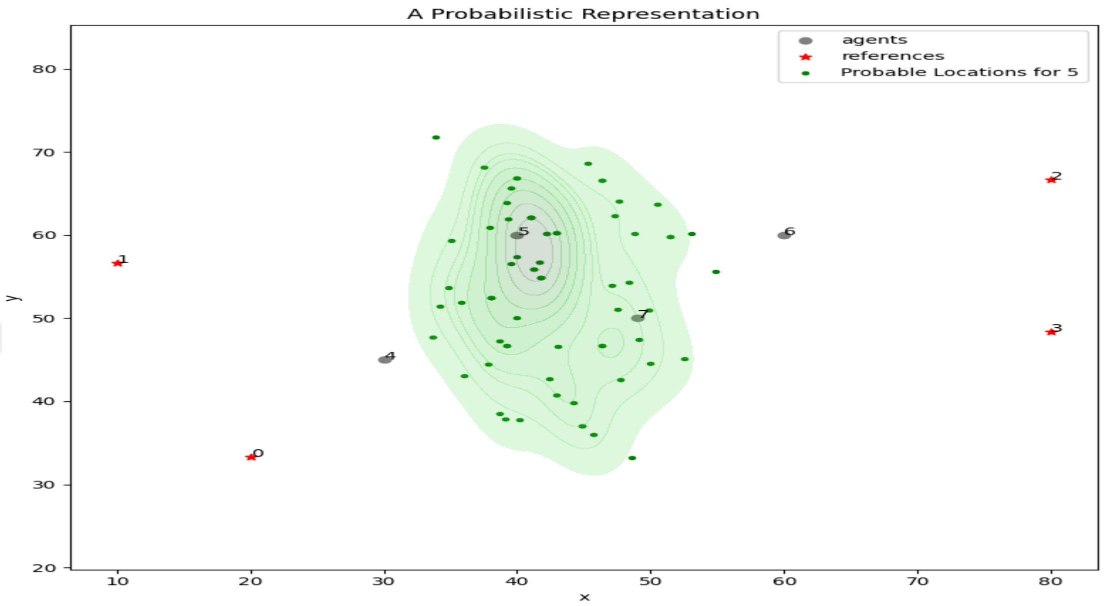


Figure 2.5. Probabilistic representation of belief of a single agent with particle-based distribution

In summary, probabilistic inference can provide a flexible framework for dealing with uncertainties by utilizing prior knowledge and likelihood models to represent non-Gaussian noise. With various techniques it becomes feasible to estimate positions under continuous domain.

2.5. Graphical Models

Representing probability distributions for each target is a difficult task. After all, there is no guarantee that every agent of the network will have the same parametric form, and modelling distributions for each of them requires a considerable amount of time. Enter graphical models; a powerful mathematical framework that can simplify the complex nature of probabilistic localization, which can also be used to visualize the relationships between targets. These models represent the problem as a graph with targets as random variables and connections between them as edges that encode dependencies. Such models allow implementation of efficient inference algorithms for calculating joint probabilities (Ihler and Thomas 2005).

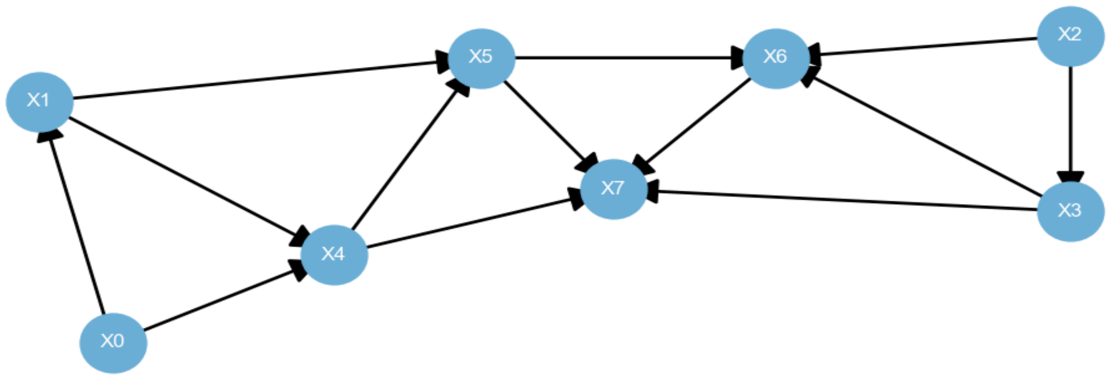


Figure 2.6. A representative Bayesian network to model hierarchical relationships

One type of graphical model is the BN, which is a directed acyclic graph where edges represent conditional dependencies between nodes. Each node in a BN is associated with a conditional probability distribution that reflects the influence of its parent nodes. While BNs are ideal for modeling hierarchical structures, they are less suitable for CL scenarios, where groups of targets are localized without a clear parent-child relationship. Additionally, the acyclic nature of BNs requires a well-defined ordering of nodes, whereas in CL applications, cycles are common.

$$\mathrm{p}(\boldsymbol{X}) = \prod_{i=1}^{n} p\big(x_i \mid \mathrm{Parents}(x_i) \big) \tag{2.9}$$

Overall probability of all variables in the network can be computed by multiplying the conditional probabilities of each variable X_i, given its parent variables. Although this factorization allows local computations simplifying complex relationships, it also implies a directional computation flow, often is not the case for localization.

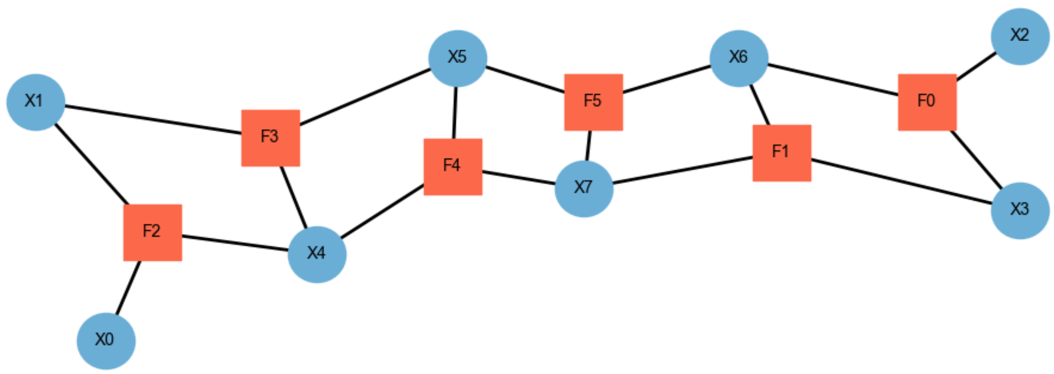


Figure 2.7. A representative factor graph to model dependencies between variables

FG, on the other hand, does not need a hierarchical structure instead, they represent factorization of a joint probability distribution into products of local functions. These factorizations lead to a flexible structure that is scalable and allows for distributed computation of joint probabilities. FG are bipartite graphs consisting of two sets of nodes, random variables and factor nodes that capture the dependencies between subsets of these variables (Wymeersch et al. 2009). However, FGs require explicit representations for each factor function of the graph, which increases complexity. Often, localization tasks involve changes to graph, modelling local computations with factorizations for dynamically changing graphs are challenging.

$$\mathrm{p}(\boldsymbol{X}) \propto \prod_{i=1}^{n} f_i(x_{i1}, x_{i2}, \ldots) \tag{2.10}$$

Each factor f_i represents a specific interaction or constraint among the variables it connects, which simplifies the complex joint distribution into manageable parts. However, explicit representation of each factor function can increase complexity, and often depends on the environmental aspects of the localization problem.

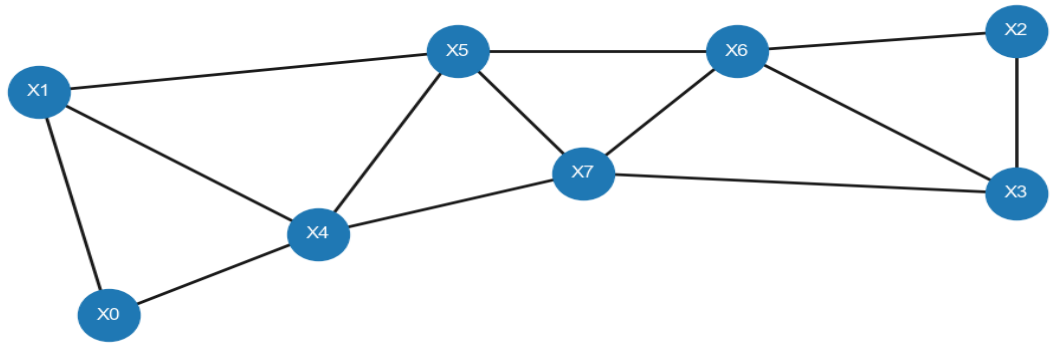


Figure 2.8. A representative Markov random field to model factorization of in terms of local computations

Another graphical model that can be used to quantify joint probability distribution is MRF. This graphical model represents each target with a random variable and edges between them enforce conditional independence of all other variables given its immediate neighbors. This is also known as Markov property, which allows local computations without the need for global information. Representing the localization problem with MRF helps develop distributed methods. Another benefit it provides is that this graphical model is easily scalable, making them ideal for large-scale problems (Ihler et al. 2005). Like FGs, MRFs can also be represented in terms of factorized functions of clique potentials:

$$\mathrm{p}(\boldsymbol{X}) \propto \prod_{Clique\ \mathcal{C}} \psi_{\mathcal{C}}(X_{\mathcal{C}}) \tag{2.11}$$

Where $\mathcal{C}$ denotes the set of cliques in the graph, X_C are the variables in clique, ψ_C are the clique potentials. A clique in an undirected graph is a subset of nodes such that every two distinct nodes in the subset are adjacent to each other.

Using graphical models to represent probabilistic problems provides an efficient way to model the CL solutions that can account for uncertainties regarding various measurements through collaboration. Another benefit of these models is that they distribute the computational burden of the CL through using a simplifying conditional independence assumption, thus providing a scalable model.

2.6. Belief Propagation

Probabilistic inference methods can be applied on graphical models to update each node's belief about its own state using message passing algorithms. One such algorithm that allows for efficient calculation of marginal distributions of variables instead of joint probabilities of complex distributions is the BP algorithm, also known as Sum-Product algorithm (Sudderth et al. 2003). BP allows each node in the graphical model to transmit messages to their neighbors that expresses their opinion about the state of their neighbors, and each recipient node updates its belief about its own state based on the aggregation of messages received and its own local information. This process allows each node to iteratively refine its marginal distribution of its state, thereby propagating local information throughout the graph. Breaking down the global computation into series of local ones allows this distributed approach to handle large-scale problems and provide easily scalable structure.

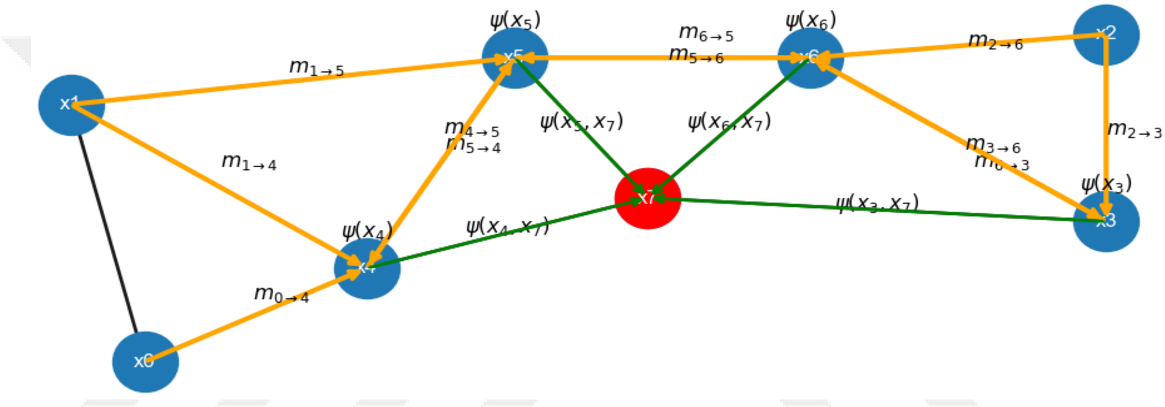


Figure 2.9. Passing messages along connected edges and calculating beliefs about positions

BP consists of two steps: message computation and belief update. Messages passed in BP are functions that represent the influence of one node's state on another. The calculation of messages in BP involves each node computing messages for its neighbors based on the incoming messages from all other neighboring nodes. A message $m_{i\to j}$ from node i to its neighbor j can be computed given a state x_j by summing over its all-possible states x_i. This calculation involves the product of local potential $\psi_{ij}(x_i, x_j)$ that defines the compatibility between nodes i and j, and the messages $m_{k\to i}$ received by node i from all other neighboring nodes k. For localization tasks summation becomes an integration due to state spaces being a continuous domain.

$$m_{i\to j}(x_j) = \int_{x_i} \psi_{ij}(x_i, x_j) \prod_{k\in N(i)\setminus j} m_{k\to i}(x_i)\, dx_i \tag{2.12}$$

After receiving messages from all neighboring nodes, each node updates its belief about its own state by combining these messages. The belief of a node i, b_i is calculated as the product of all incoming messages from its neighbors and the local evidence. This aggregation of messages ensures that the node incorporates all available information from its neighbors to refine its estimate of its own state. This collaboration of nodes increases localization performance.

$$b_i(x_i) \propto \psi_i(x_i) \prod_{j \in N(i)} m_{j \to i}(x_i) \tag{2.13}$$

In graphical models with loops, BP repeatedly considers the influence of a node's opinion on messages multiple times (see Equation 2.12), as demonstrated in Figure 2.9. This loopy behavior on graphs increases uncertainty, resulting in inaccurate estimates. To address this behavior, iterative formulation of messages can be employed for computing outgoing messages by accounting for incoming messages from previous iteration (Murphy, Weiss, and Jordan 2013).

$$m_{i \to j}^{n}(x_j) \propto \int_{x_i} \psi_{ij}(x_i, x_j) \frac{b_i^{n-1}(x_i)}{m_{j \to i}^{n-1}(x_i)} \tag{2.14}$$

The denominator $m_{j \to i}^{n-1}$ accounts for the message sent from node j to node i in the previous iteration. This formulation iteratively refines the beliefs by incorporating the most recent information from neighboring nodes, ensuring convergence towards a stable set of marginal distributions.

BP is a powerful and efficient algorithm for performing probabilistic inference in graphical models, particularly useful in the context of CL. Decomposition of global computations into localized message passing and belief updates, BP facilitates scalable and distributed computation, making it well-suited for large-scale and dynamic environments. The ability to iteratively refine marginal distributions utilizing collaboration among neighbors allows for accurate state estimation even in complex, uncertain scenarios.

2.7. Particle Filtering

BP is a powerful probabilistic inference method for discrete domains. However, when calculating messages in continuous state spaces, an integral over all possible states for each neighbor is required therefore, marginalizing probability distributions for exact inference tasks is intractable. Although there exist efficient methods for discretizing the state space, the computational burden for large environments can be prohibitive. This is where PF can offer a viable alternative for handling continuous states in complex environments. PF represents the belief of a random variable using a set of discrete particles, also known as samples from distribution for approximate inference tasks.

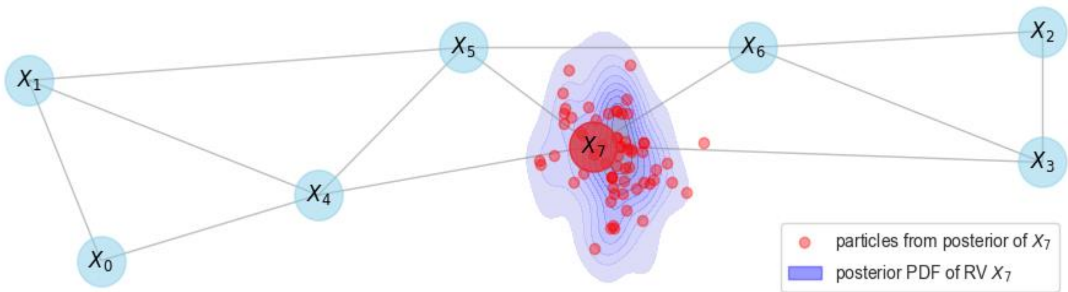


Figure 2.10. Beliefs about positions are represented using sampled particles from distribution

One of the main strengths of PF is its ability to represent and approximate non-Gaussian probability distributions. Traditional methods, such as the Kalman Filter, assume linearity and Gaussian noise, which limits their applicability in real-world scenarios where these assumptions do not hold. PF, on the other hand, does not make such assumptions and can thus handle complex, multi-modal, and non-linear distributions effectively.

$$\int f(x)p(x)\ dx \approx \frac{1}{M}\sum_{i=1}^{M} f\left(\boldsymbol{X}_i^p\right) \tag{2.15}$$

For an arbitrary function $f(x)$ and a true distribution $p(x)$ can be approximated with M particles $\boldsymbol{X}_i^p$ sampled from distribution $p(x)$. This is also known as sequential Monte Carlo integration and it is fundamental to PF, as it allows the method to approximate continuous probability distributions using discrete samples. However, PF can suffer from particle degeneracy, where few particles dominate the characteristics of distribution.

2.7.1. Kernel density estimation

While Monte Carlo integration approximates the true density distribution using sampled particles, the limited number of samples restricts the evaluation to specific states, thereby reducing the representative power of the approximation. Although limited particles can represent a nonparametric distribution function for discrete scenarios, continuous states require an approximate density function for effective approximation. Applying a kernel function to each sampled particles of distribution allows smoothing of evaluations with possible states onto nearby region. A common kernel that can be used for this smoothing operation is the Gaussian kernel. With KDE, a nonparametric density can be represented with mixture of Gaussians.

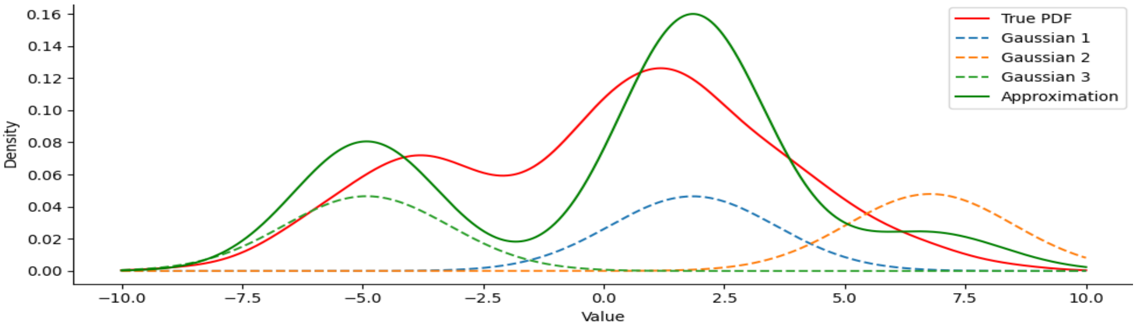


Figure 2.11. KDE estimate using multiple Gaussians to approximate a distribution

Particles sampled from distribution will be sampled from most likely places. As the number of particles increases, the representative power of the model will increase as well. Given a set of M particles $\{x_1, x_2, \ldots, x_N\}$, each representing a possible state in the continuous domain, the KDE $\hat{f}(x)$ at a state x is:

$$\hat{f}(x) = \frac{1}{Nh}\sum_{i=1}^{M} K\left(\frac{x - x_i}{h}\right) \tag{2.16}$$

The parameter h, which determines the smoothness of the density estimate is known as the bandwidth parameter. Using a lower h increases sensitivity so that model captures more details, however it can be easily affected by the noise also can be referred to as overfitting, on the other hand, a larger h would result in smoother estimates but a too large h may cause missing important features which is known as underfitting. Using the standard deviation of the particles and number of samples, bandwidth parameters can be automatically determined. Additionally, cross-validation techniques can be employed to optimize the bandwidth parameter by minimizing a cost function between the estimated and true distribution.

$$\int f(x)p(x)\,dx \approx \sum_{i=1}^{M} W_i^p f(X_i^p)K(x - X_i^p) \tag{2.17}$$

Depending on function f and the distribution p, approximations can be obtained for marginals for each random variable like Figure 2.10, or proposal distributions like Figure 2.12(a) with equation (2.16). KDE offers a flexible and effective means to approximate complex distributions without the need for parametric assumptions.

2.7.2. Mixture importance sampling

The goal of PF is to approximate the true distribution $p_i(x)$ of a random variable using a set of samples $\boldsymbol{X}_i^p$ However, directly sampling from $p_i(x)$ may not be feasible as it might not be parametrically representable, though evaluating with some samples possible. Instead, samples can be drawn from a proposal distribution $q_i(x)$ that best captures the characteristics of the true distribution. This $q_i(x)$ can be a mixture of proposal distributions, each designed to capture different aspect of the true distribution.

$$q_i(x) = \sum_{j \in N(i)} w_{j,i} q_{j,i}(x) \tag{2.18}$$

In graphical models, proposal distributions $q_{j,i}(x)$ are generally derived from pairwise compatibility functions $\psi_{ji}(x_j, x_i)$ and $w_{j,i}$ represents the importance of the specific proposal distribution. Importance factors for proposals often depend on the prior knowledge about each variable, for homogeneous problems that includes one type of variable, weights of proposals are equal.

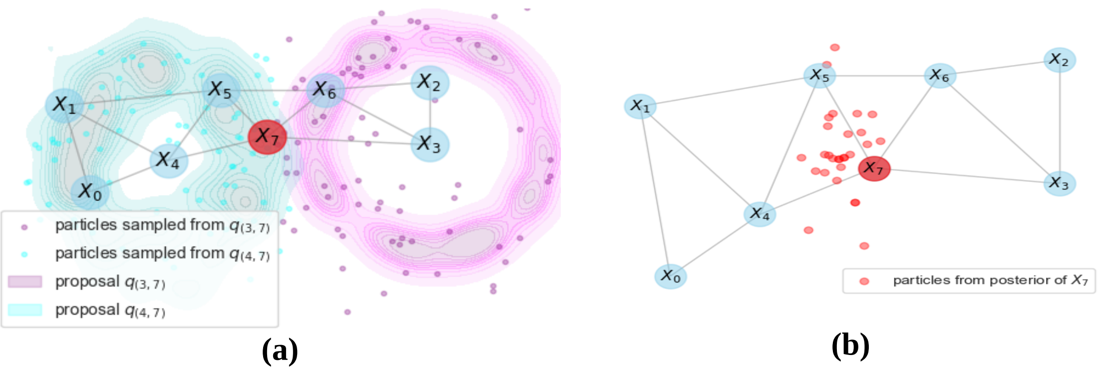


Figure 2.12. Messages are represented in terms of mixtures of importances; **a)** sampling candidate particles from proposals of messages; **b)** resampling with replacement stage

Drawing M samples $\boldsymbol{X_i^q}$ from each proposal and calculating their weights allows for resampling M samples that best represent $p_i(x)$. Resampling process mitigates the particle degeneracy problem, where few particles with very high weights dominate the particle set, leading to loss of diversity. Redistributing particles allows BP to focus on more likely regions of the state space, improving approximation of $p_i(x)$.

$$W_i^q = \frac{p_i(X_i^q)}{q_i(X_i^q)} \tag{2.19}$$

MIS enhances the particle filtering process by utilizing multiple proposal distributions, where each proposal captures different characteristics regarding target distribution, allowing collaborative inference to achieve higher accuracy.

2.8. Related Works

Cooperative Localization (CL) aims to improve accuracy and reliability of positioning systems under complex scenarios where traditional localization techniques suffer due to limited context. To address the uncertainties inherent in real-world positioning applications, advances in probabilistic approaches such as graphical models, particle filtering and belief propagation, have provided a reliable framework.

One of the initial work that considers pair-wise range estimates is a relative localization method proposed by (Patwari, O'Dea, and Wang 2001). Authors seek an effective way for localization without requiring a global reference frame due to expensive requirements and impracticalities of traditional localization techniques in dense environments. Instead, in their proposed peer-to-peer system, range estimates between neighboring devices are obtained from RSS measurements and a few reference points are used for determining positions. They use a simple Maximum Likelihood (ML) estimator to combine information from multiple pairs. However, their proposed approach is centralized, therefore scalability of their system is limited with the efficiency of information broadcasting algorithm. Additionally, they consider gaussian distribution, often not the case for localization.

(Howard, Mataric, and Sukhatme 2003) introduce a fully relative localization method for mobile robot teams that does not necessitate knowledge of any reference points. In their proposed approach, robots measure their relative poses to nearby robots and broadcast this information to generate an ego-centric estimate using a Bayesian framework and particle filtering. By employing particle filtering, they can represent non-Gaussian distributions. Additionally, authors use a dependency tree to prevent circular updates within the team. Although their proposed approach is decentralized and requires minimal communication, making it suitable for dynamic environments, without efficient approximations to densities, particle filtering will only provide rough estimates. This necessitates a large number of particles, which can be computationally costly.

(Sudderth et al. 2003) proposed a novel NBP algorithm extending particle filtering to arbitrarily structured graphical models. In their work, authors consider solving a computer vision problem in which they reconstruct hidden facial features, components of occluded faces such as nose or eyes. This method is crucial for applications involving high-dimensionality, continuous, non-Gaussian distributions. NBP iteratively updates kernel-based approximations to true distributions, allowing the use of a variety of potential functions. By leveraging local message updates and regularizing kernels, NBP efficiently combines information from neighboring nodes, overcoming the limitations of simpler Markov chains. Through using kernels, they can effectively represent the underlying uncertainty. A drawback of this approach is the computational complexity and updating these nonparametric representations, which can be mitigated by efficient sampling techniques like Gibbs sampling. In their work, authors applied NBP to infer components of occluded face features and showed that such features can be represented with continuous, non-Gaussian distributions.

On the other hand, (Ihler et al. 2005) considers self-localization within a sensor network by applying NBP to an MRF graph. In their work, they utilize conditional independence characteristic and develop a distributed CL mechanism, in which each random variable represents a sensor device and its marginal probability as a non-Gaussian region where the sensor could be positioned. Their approach leverages the Bayesian framework to iteratively update each variable's belief about its position based on their observations from neighboring sensors. Authors show that NBP is robust and suitable for large-scale CL problems.

In their seminal work, (Patwari et al. 2005) presents how collaboration among sensors can help improve positioning compared to TLT. They describe statistical models for measurements widely used for localization problems such as Angle of Arrival (AoA), ToA, and RSS. In their work, authors demonstrate how a theoretical benchmark can be established to evaluate effectiveness of localization algorithms. This is achieved by using statistical models to analyze the measurements obtained, without considering specific algorithm implementations.

Another influential work for CL on sensor networks is (Wymeersch et al. 2009). They consider ultrawide bandwidth (UWB) technology to achieve precise localization. Authors have developed the SPAWN (Sum-Product Algorithm over a Wireless Network), which is a fully distributed approach that collaboratively determines the positions of network elements. Similar to NBP of (Ihler et al. 2005), SPAWN also uses probabilistic inference to continuously update the beliefs about positions of nodes by exchanging information with neighbors. However, this work differs in its choice for graphical model; they use FGs to model dependencies between network elements. Additionally, NBP approach uses particles to represent messages passed, however SPAWN computes the integral over states. SPAWN is more efficient for smaller networks, on the other hand, NBP approach is more suitable for more complex, non-Gaussian distributions which allows flexibility in representing uncertainties.

The work of (Savic and Zazo 2009) builds on NBP for graphical models by introducing a method that constrains the sampling area of particles using boxes where anchor nodes' ranges overlap. In the traditional PF approach, a large number of particles must be sampled from the entire deployment area, which can diminish the representation power when the localization environment is large. Constraining the sampling area reduces the number of particles for effective localization, leading to an increase in computation efficiency. Particles sampled during either the initial stage or in belief update stage outside these bounding boxes are disregarded, allowing the algorithm to focus on most likely areas.

3. MATERIAL AND METHOD

In this thesis, the goal is to explore how collaboration among agents can help improve the localization performance of multiple agents. A probabilistic approach is chosen because it manages the uncertainties and its ability to work with incomplete or flawed information. Specifically, the graphical model approach by (Ihler et al. 2005) is adopted for CL task. This methodology leverages probabilistic inference to achieve robust and efficient localization through using a variant of BP with PF to handle non-Gaussian assumptions and nonparametric representations. Conditional independence assumption on neighboring agents allows for a distributed solution that is easily scalable. Additionally, the benefits of introducing a sampling constraint on reference points by building bounding boxes for agents to improve sampling quality and performance is explored (Savic and Zazo 2009).

The objective is to estimate position of N agents scattered in some environment with N_a anchors and N_t targets. Positions of anchors are known, and anchors do not necessarily have any information regarding any other anchor outside of their communication vicinity. Additionally, all agents are assumed to be able to broadcast a signal to their vicinity and sense the signals of others.

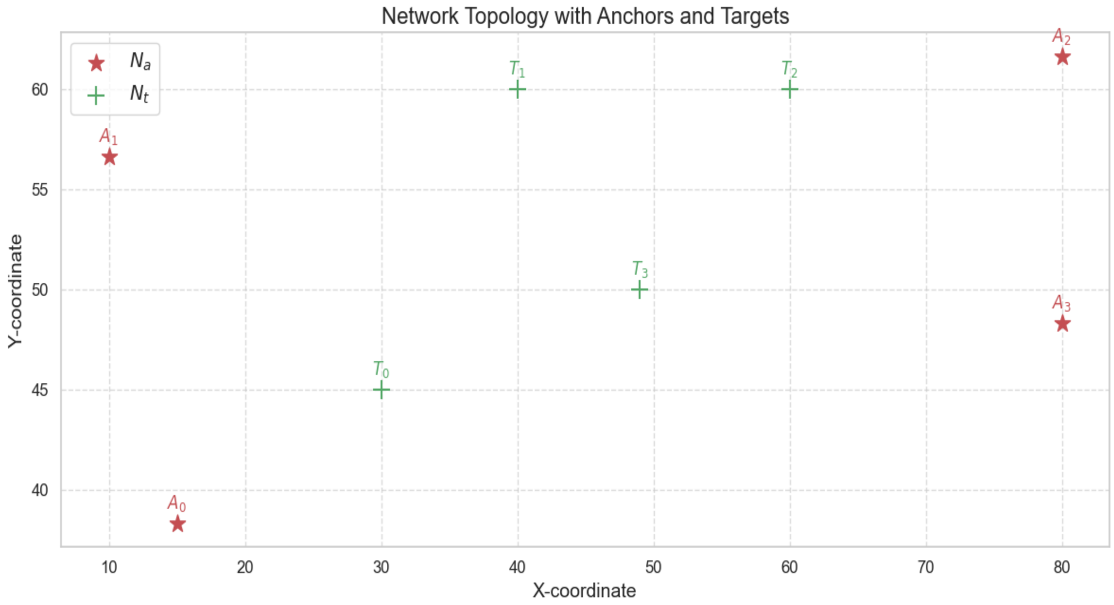


Figure 3.1. Agents distributed in a region with known points and targets

First, we investigate how agents can obtain approximate distance estimates through RSS and improve the context of individual agents by sharing observations with their neighbors. Next, we will formulate localization problem according to Bayesian inference and detail how prior knowledge, likelihood of states and detection probabilities contribute. Following this, graphical model representation of the problem formulation will be explored to utilize local potential functions for a distributed localization. Finally, we will employ NBP with PF to handle non-Gaussian assumption and approximate distributions with mixtures of gaussian kernels using KDE. Lastly, how MIS can be used as an efficient sampling technique to update beliefs about positions.

3.1. Estimating Distances

For localization tasks estimating the distances between agents that can communicate provides a quantifiable metric for determining their relative positions. From equation (2.3) it is straightforward to derive relative distances between agents d_{ij} given RSS measurements obtained from one of the agents s_{ij}:

$$d_{ij} = 10d_0 \frac{P_i - s_{ij} + n_{ij}}{10\alpha} \tag{3.1}$$

Distance between agent d_{ij} and d_{ji} may show differences, additionally, it is possible that while agent i observes j, the reverse may not be true. For convenience, we will accept both observations are the same. Although, it is straightforward to obtain a common distance by taking average of estimated distances. Obtaining d_{ij} allows to form a relationship between the positions of agents x_i and x_j as shown in equation (2.1).

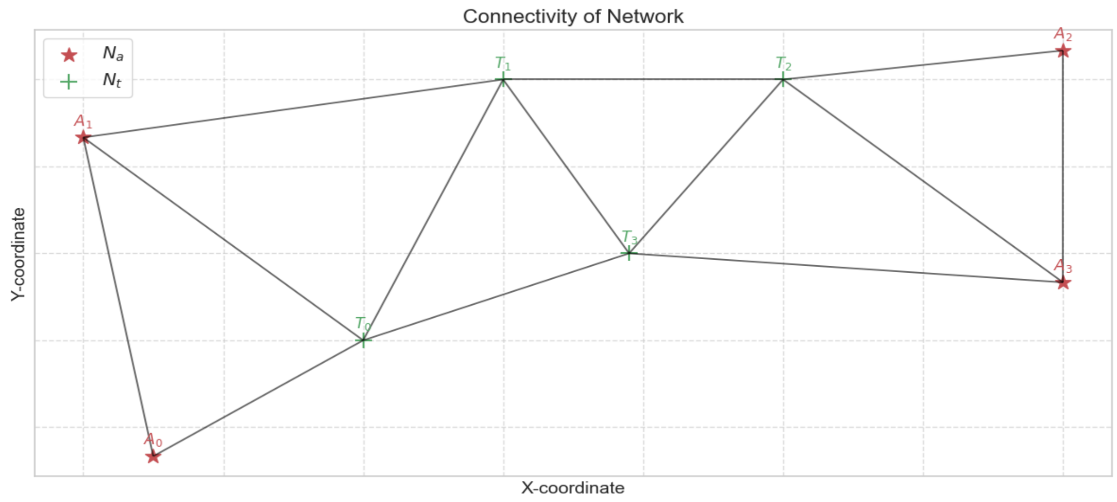


Figure 3.2. 1-Hop connectivity of agents

One-hop distances referred to the direct distances between agents that can communicate directly. If agents can observe each other, they are considered immediate neighbors. The neighbors of an agent i will be referred to as $N(i)$. The one-hop distances of the network of agents are the distances each agent i has calculated to their neighbors $N(i)$, c_{ij} is the connectivity variable that states whether agents observe their neighbors j directly or not.

$$\forall i \in N, \quad \forall j \in N(i), \quad d_{ij} \text{ as defined in (3.1)}, \quad c_{ij} = 1 \tag{3.2}$$

Based on the sensing technology used to estimate the distances from RSS network topology can differ. In this thesis, agents are assumed to be able to sense and exchange information within a limited range in an environment. Unfortunately, there are not readily available and commonly recognized datasets for localization problems, therefore, throughout this thesis RSS measurements are simulated according to Bluetooth and Wi-Fi technologies with various parameters.

3.1.1. Two-hop distances

While one-hop distances provide direct observations and are sufficient for localization, two-hop distances account for indirect relationship through intermediate agents. If an agent i cannot directly observe agent k, but can communicate with agent $j \in N(i)$, which in turn can communicate with agent $k \in N(j)$, then the distance between i and k can be inferred through j as:

$$\forall \mathrm{i} \in \mathrm{N}, \forall \mathrm{j} \in \mathrm{N(i)}, \forall \mathrm{k} \in \mathrm{N(j)} \setminus \mathrm{N(i)}: \quad \mathrm{d_{ik}} = \min_{\mathrm{j} \in \mathrm{N(i)}} \left(\mathrm{d_{ij}} + \mathrm{d_{jk}}\right), \quad \mathrm{c_{ik}} = 0 \tag{3.3}$$

While one-hop distances represent how close agents are, in contrast, two-hop distances highlight the necessary separation between agents, implying that they should maintain a certain distance from each other. One can conceptualize this network as a spring system; one-hop connections act as forces pulling agents closer together, while two-hop distances exert a repulsive force, pushing agents away as these agents are not meant to be in close vicinity.

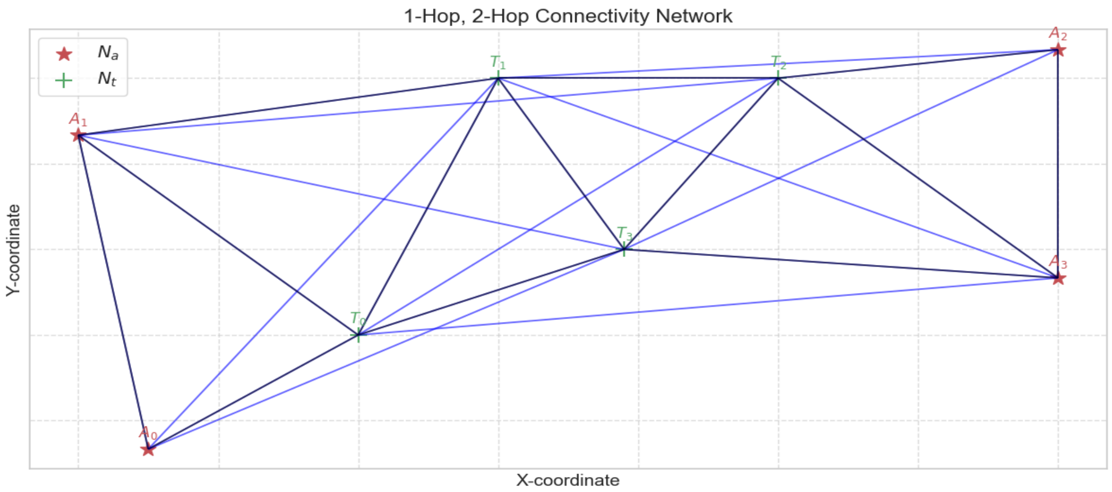


Figure 3.3. 2-Hop connectivity of agents

Utilizing both sets of distances helps achieve a balanced localization framework, ensuring that position estimates are not solely based on direct observations but also incorporate broader network structure. This dual approach mitigates the risk of skewed positioning by accounting for indirect relationships, leading to more accurate estimates.

Additionally, relying solely on direct measurements during the initial particle sampling stage may result in targets lacking connections to anchors. Consequently, rather than sampling from a defined smaller region, sampling must then be conducted across the entire deployment area, reducing the representative power of individual particles. Being able to utilize prior knowledge obtained from anchors through two-hop distances enhances overall reliability and precision of the localization process.

3.1.2. N-hop distances

The concept of obtaining two-hop distances can be generalized to n-hop distances. In this approach, each time agents share their distances, each of them obtains approximate distances about their n-hop neighbors. However, this information distribution step requires sequential distance update scheduling. Depending on the depth n, each agent i in clique $\mathcal{C}^n$, obtains distance information from $N_{c=1}(i)$ concurrently and each $\mathcal{C}$ can perform information distribution in parallel. However, depending on the density of $\mathcal{C}^n$ distribution can be time consuming, therefore some $\mathcal{C}$ needs to wait for new information to become available for the next nth-hop.

$$\forall i \in N, \forall k \in \left(\bigcup_{n=2} \mathcal{N}^n(i)\right) \setminus \mathcal{N}^{n-1}(i): d_{ik} = \min_{p \in \mathcal{P}_{ik}} \left(\sum_{(a,b)\in p} d_{ab}\right), c_{ik} = 0 \quad (3.4)$$

Here $\mathcal{N}^n$(i) is the set of agents that are within n-hop of agent i. k is any agent that is reachable after n-hop and p ∈ P_{ik} is a path in the set of all paths P_{ik} from $\mathcal{N}^{n-1}$(i) to $\mathcal{N}^n$(i) of i to k.

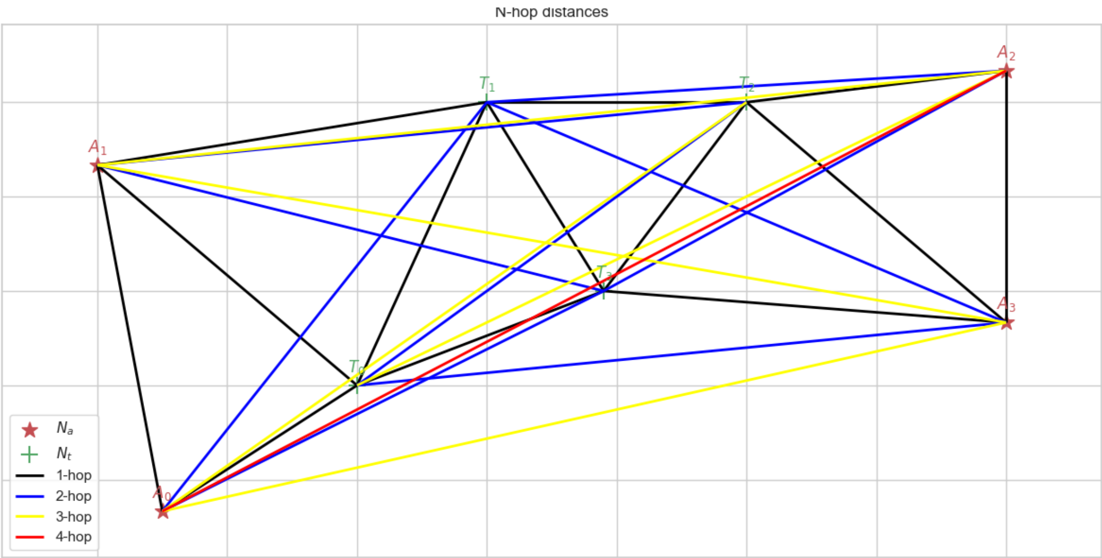


Figure 3.4. N-hop neighborhoods

In equation (3.4), c_{ik} is again set to zero, indicating that these distances are not obtained from immediate neighbors. Given that our objective is to achieve distributed localization, we assume that all agents are conditionally independent of each other with $c_{ik} = 0$, given their immediate neighbors.

For all $c_{ik} = 1$ and adequately large n, each agent has pseudo fully connected network, corresponds to centralized solution for each agent separately. Therefore, in addition to increasing uncertainties, the system also becomes computationally infeasible due to large number message computations.

Assumption of agents within some arbitrary n-hop distance are immediate neighbors ($c_{ik} = 1$) greatly degrades localization precision due to the stacking uncertainties; this degradation is multiplicative for naive BP because belief update phase of the algorithm requires product of all incoming messages (see equation (2.13)) and additive for PF-based BP uncertainties because using MIS allows fusing mixtures of gaussians based on their weights (see equation (2.18)).

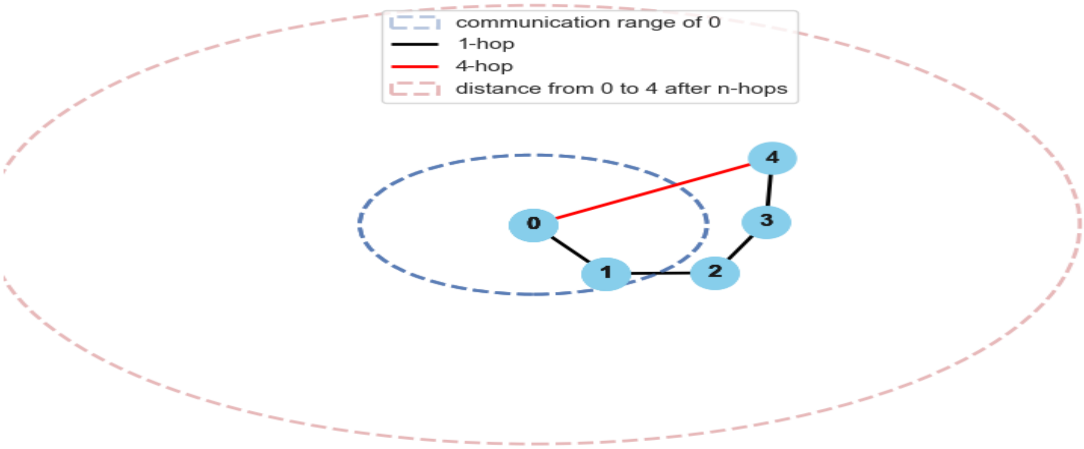


Figure 3.5. Distance approximation over multi-hop neighborhoods

The main contributing factor to increasing uncertainties, in addition to sensor noise, is the limitation of distance approximation over n-hops because $d_{ij} + d_{jk}$ often does not lie on a straight line. One important consideration is that there is no knowledge about angles between agents; therefore, distances are only approximated through n-hops. From Figure 3.5, belief about the distance of agent 0 for agent 4's location is on the red line. However, agent 4 is much closer. On the other hand, when n-hop neighbors are considered as indirect ($c_{ik} = 0$), messages over such neighbors can be considered as suggestions, as opposed to beliefs.

3.2. Localization Formulation

In CL, the primary objective is to estimate the positions of multiple agents through using Bayesian framework. This probabilistic approach allows to utilize pairwise connectivity, relationship and prior knowledge about the locations of agents. The problem can be formulated as the joint probability of the network of agents. By applying Bayesian inference, various sources of information are integrated, and inherent uncertainties are effectively managed.

$$p(\boldsymbol{X}, \boldsymbol{C}, \boldsymbol{D}) = \prod_{(i,\, j)} p\left(c_i \middle| x_i,\, x_j\right) \prod_{(i,\, j):c_{ij}=1} p\left(d_{ij} \middle| x_i,\, x_j\right) \prod_{i} p_i(x_i) \tag{3.5}$$

Equation (3.5) represent the probabilistic formulation of the localization problem, capturing the joint probability distribution over the positions of agents $\boldsymbol{X}$, the connectivity variables $\boldsymbol{C}$ between agents, and the approximate distance measurements $\boldsymbol{D}$. The goal is to estimate the maximum a posteriori (MAP) location for $\forall \mathrm{i} \in \mathrm{N}_{\mathrm{t}}, x_i \in \boldsymbol{X}$ from the information obtained relative to neighbors $\boldsymbol{D}$. MAP for a network to estimate positions jointly is different from individual MAP estimates. While the previous one corresponds to a centralized approach, the latter one allows to distribute the task to local computations, hence provides a more scalable option.

3.2.1. Prior knowledge

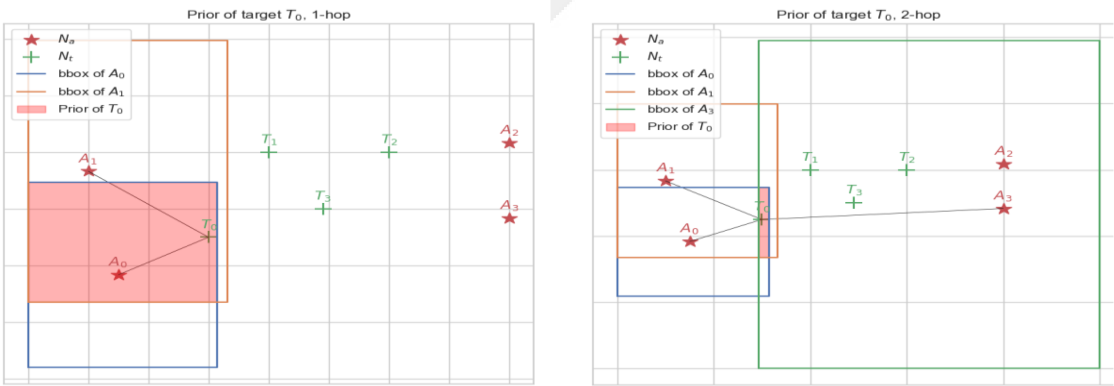
The existence of anchors with known positions can be used for localization according to some absolute reference frame. Although localization can be performed without relying on such information is also possible and it's known as the relative localization, the estimates for the agents under such circumstances will involve arbitrary scaling, rotation and translation. Instead, utilizing anchors allows one to take advantage of existing prior knowledge, leading to more accurate and consistent estimation results. Each target agent can leverage the known positions of anchors within its communication range to obtain a priori about its location. This approach mitigates the ambiguities associated with relative localization.

For all $i \in N_t$, and for all $j \in N_a^n(i)$, the bounding box for a target agent i can be defined using the distance d_{ij} to its n-hop anchor neighbors. Let x_j represent the coordinates of anchor j. Then, the bounding box for the target agent i is given by:

$$min_i = \min_{j \in N_a^n(i)} \{x_j + d_{ij}\} \tag{3.6}$$

$$max_i = \max_{j \in N_a^n(i)} \{x_j - d_{ij}\} \tag{3.7}$$

Here, min_i and max_i represent the bounds of the bounding box for the target agent i, ensuring the smallest area is used to represent the prior knowledge about the position of the agent. If the deployment area's boundaries are also known, they can be shared from anchors to further limit the bounded boxes to the deployment area, or based on statistical models that represents the effect of signal attenuation, various boundaries that depicting rooms, or furnishments.



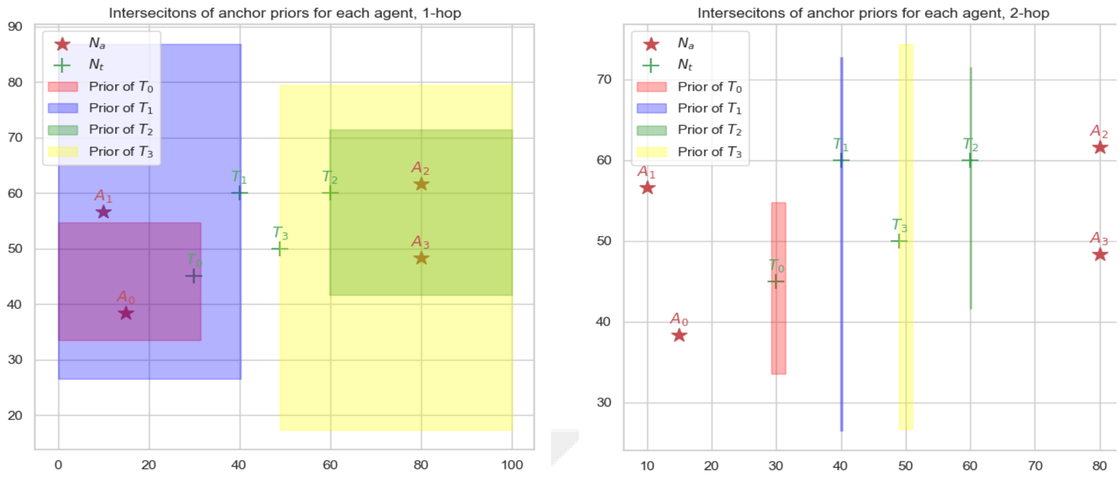
(a) **(b)**

Figure 3.6. Obtaining prior knowledge about positions from known anchor points; **a)** 1-hop priors obtained from immediate anchors; **b)** 2-hop priors obtained from anchors

A prior probability distribution function, given the bounding box defined by min_i and max_i, can be defined as a uniform distribution over the region specified by these bounds. The prior probability function $p_i(x_i)$ for the position x_i of the target agent i is:

$$p_i(x_i) = \begin{cases} \frac{1}{V_i}, & if\ min_i \le x_i \le max_i \\ 0\,, & otherwise \end{cases} \tag{3.8}$$

V_i is the volume of the bounded box defined by the intervals $[min_i, max_i]$. This ensures that the prior probability distribution is uniform over the defined bounded box region and zero everywhere else. If other contributing factors such as environmental obstacles exist, this distribution can either be a mixture of multiple uniforms or a specific statistical distribution.



(a) **(b)**

Figure 3.7. Prior probability regions of all target agents; **a)** prior regions of targets through 1-hop connectivity; **b)** prior regions of targets through 2-hop connectivity

Although localization can be performed without any priors, utilizing the existing information can significantly improve the positioning estimates. Since anchors plays an important role for determining priors, there are works that also considers anchor placement strategies in the field (Zhang, Cao, and Liu 2017). During the PF stage, particles can be sampled from these distributions to represent initial states of targets, granting efficiency by reducing the number of particles needed.

3.2.2. Likelihood of states

In the context of CL, the likelihood of the states represents the probability of observing a specific distance measurement given possible positions that each agent can be located in. In each iteration of the algorithm, each set of states represent the beliefs about the agents' positions, expressed as probability distribution. In the first iteration, this corresponds to the prior distribution and in the PF case each particle represents a state for its' agent. The likelihood of the states quantifies how well the observed data fits the assumed model of the agents' positions.

For agents $i, j \in N$, the likelihood function $p(d_{ij}|x_i, x_j)$ evaluates the probability of measuring a distance d_{ij} given the states x_i and x_j. Observations are subject to noise, which is typically assumed as Gaussian and the formulation of likelihood function incorporates this noise model p_n, enabling a probabilistic framework to handle uncertainties.

$$\mathrm{p}(d_{ij}|x_i, x_j) = p_n(d_{ij} - \|x_i - x_j\|), \quad n_{ij} \sim p_n \text{ (noise model)} \tag{3.9}$$

In an ideal situation, distribution of the likelihood function resembles a ring-like in 2D and a bowl-like shape in 3D around evaluating agent with distance d_{ij} as its' radius.

Each state of the evaluated agent is assessed using this likelihood function, determining the probability of observing the given distance measurement. States with the highest probabilities are the candidates for representing the belief of evaluated agent.

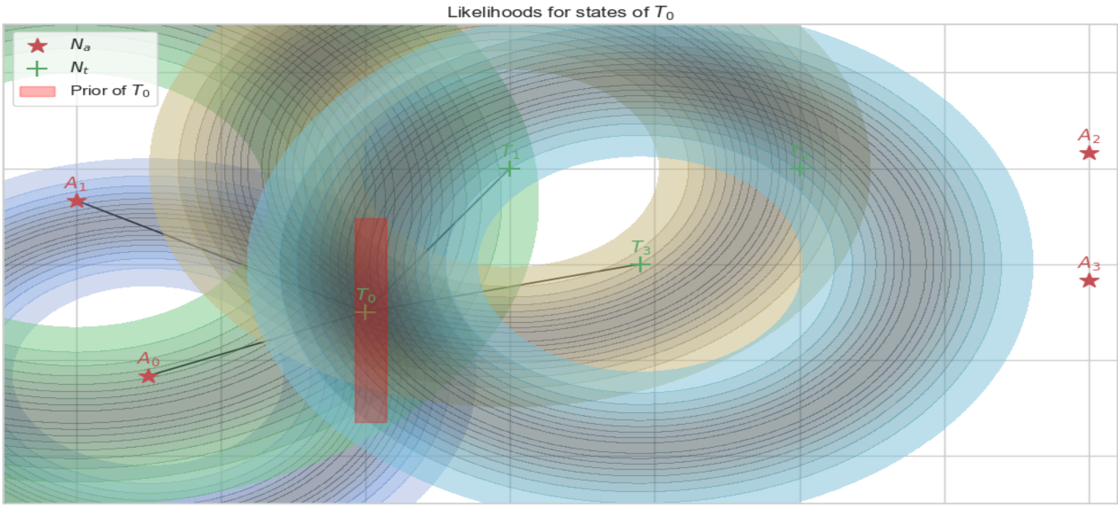


Figure 3.8. Likelihoods of observing the given states of neighboring targets with known distances

The likelihood of the states provides a critical measure for evaluating the positions of agents based on observed distance measurements. These likelihoods are the essential part of an agent's opinion about the position of its neighbor and such opinions correspond to messages passed on by BP.

3.2.3. Probability of connectivity

Probability of detection or connectivity provides insights into which agents should or should not be within their communication range. As the signals propagate through space, due to attenuations they lose power, resulting in less reliable measurements. Without an approach that accounts for this behavior incorrect estimates would have the same impact on localization process as with the correct ones, hence increasing the uncertainty. Probability of connectivity statistically models this behavior and reduces the affect from low confidence neighbors. A simple yet effective model has been proposed by (Moses and Patterson 2002)

For all agents $i \in N$, and their one hop neighbors $j \in N(i)$, probability of their connectivity decreases exponentially with squared distance with a threshold radius R.

$$c_{ij} = 1: \quad P_c(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2R^2}\right) \tag{3.10}$$

This statistical model evaluates the states of an agent that passes its message to a neighbor, hence providing a confidence metric on the messages passed. With the connectivity distribution centered around that neighbor, states closer to the neighbor will have messages with greater confidence. This can be conceptualized as a pulling force on springs, resulting in agents that have high connectivity to each other being estimated to be closer together.

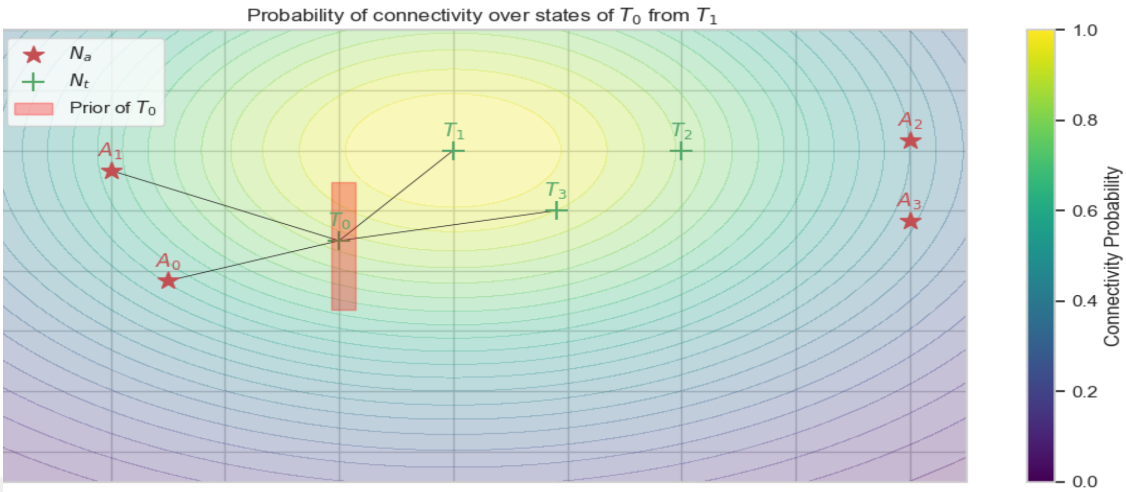


Figure 3.9. Probability of connectivity of targets whose are within each other's communication radius

Although this concept can be extended to multi-hop neighbors, uncertainty regarding distance measurements as demonstrated in Figure 3.5, adversely affects localization results, disrupting estimates. However, a similar approach can be considered for multi-hop neighbors. Instead of providing confidence about opinions based on how reliable measurements are, confidence about suggestions from multi-hop neighbors according to lack of connectivity can be used. This is simply $1 - P_c(x_i, x_j)$.

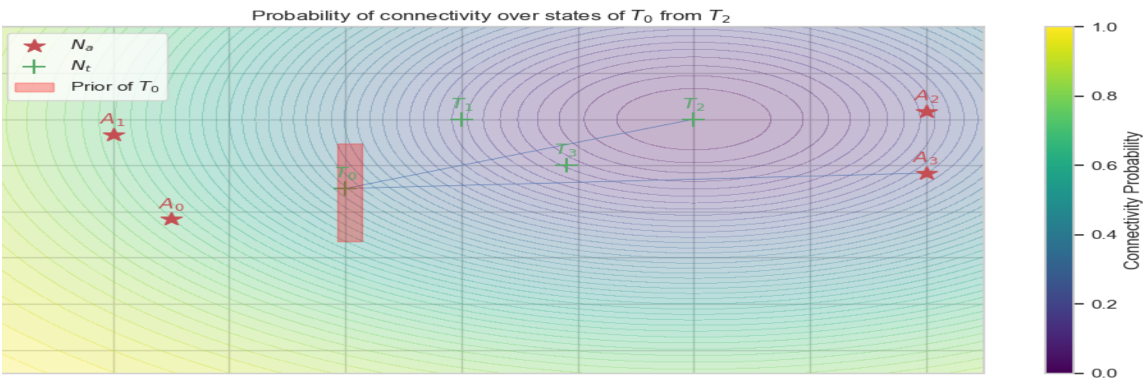


Figure 3.10. Probabilities of disconnectivity of 2-hop neighbors

For multi-hop neighbors, since agents can only observe each other through an intermediate agent, their confidence increases with their lack of connectivity. Utilizing this lack of connectivity increases the precision of localization because these agents provide useful information regarding which agents should not be close to each other. Because for distributed localization approaches due to conditional independence assumption, agents only rely on their immediate neighbors, this may lead to estimating incompatible agents together. This time, connectivity within multi-hop agents acts as a repulsive force in the spring system, pushing agents away from each other. This simple idea can drastically improve estimates for crowded scenarios.

3.3. Graph Representation

Graphical models provide a structured approach to represent the probabilistic relationships among variables. In this work, we use an MRF to model the CL problem. An MRF consists of an undirected graph $G = (V, E)$ where V represents the set of nodes, and E represents the set of edges indicating conditional dependencies between nodes.

The Prior knowledge about the positions of the nodes is incorporated as unary potentials in the MRF. Each node i in V has an associated prior probability distribution $p_i(x_i)$ representing initial belief about its position.

$$\psi_i(x_i) = p_i(x_i) \tag{3.11}$$

The likelihood of the states and the probability of connectivity is together represented as pairwise potential functions. These functions quantify the compatibility between the states of connected nodes based on observed distances. Additionally, probability of connectivity influences the form of the pairwise potential functions. This probability depends on factors such as signal strength and environmental conditions, and it determines whether a direct connection between the nodes exists. For nodes i and j with an observed distance d_{ij}, the pairwise potential $\psi_{ij}(x_i, x_j)$ is defined as:

$$\psi_{ij}(x_i, x_j) = \begin{cases} P_c(x_i, x_j)p(d_{ij}|x_i, x_j), & if\ c_{ij} = 1 \\ 1 - P_c(x_i, x_j), & otherwise \end{cases} \tag{3.12}$$

The overall joint posterior likelihood over the node positions and connectivity is given by the product of all unary and pairwise potentials:

$$p(\boldsymbol{X}, \boldsymbol{C}, \boldsymbol{D}) \propto \prod_{i \in V} \psi_i(x_i) \prod_{(i,j) \in E} \psi_{ij}(x_i, x_j) \tag{3.13}$$

This formulation allows to systematically incorporate prior knowledge, observed data, and connectivity probabilities into a coherent probabilistic framework for CL. This posterior probability density function then can be marginalized for each target agent to obtain beliefs about positions as their distribution and their means to be represented as their estimates.

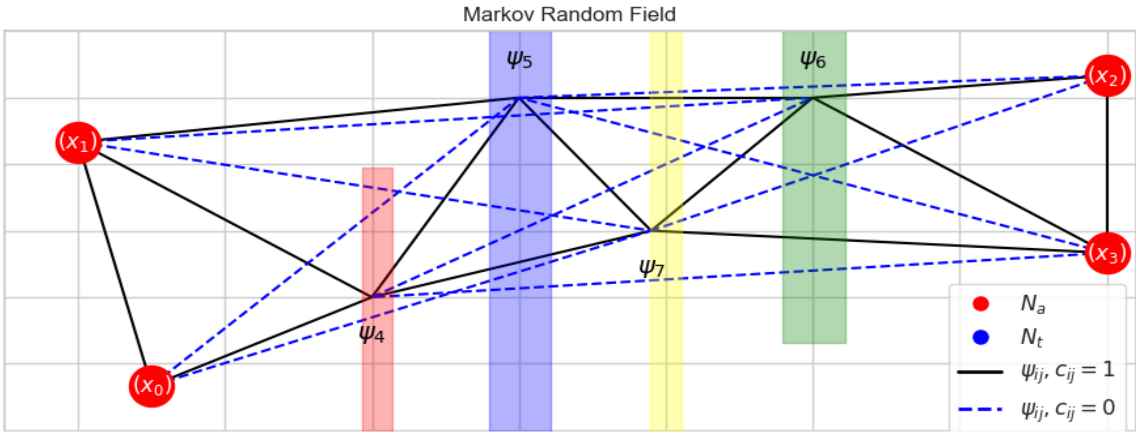


Figure 3.11. Complete MRF representation with pairwise potential functions and priors

3.4. Nonparametric Belief Propagation

Marginalization of each random variable's posterior distribution can be achieved using a probabilistic inference method known as NBP (Sudderth et al. 2003). In this approach NBP takes advantage of particle-based representation as opposed to parametric representation for distributions to handle complex, non-Gaussian forms, which is often the case for localization problems. Additionally, computation of messages that involves integral over continuous states can be efficiently approximated with set of particles. In each iteration of the algorithm particles sampled from proposal distributions are weighted and resampled to obtain posterior beliefs. In this section, we will outline the iterative process of NBP, the necessity for particle-based representation, and the methods used for message approximation and belief calculation.

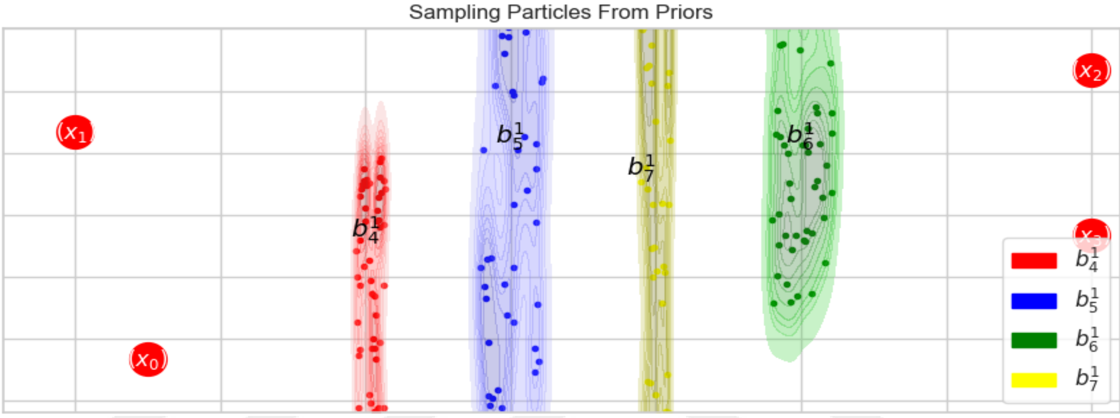


Figure 3.12. Sampling particles from initial beliefs

NBP operates iteratively to refine the estimates of each node's states within the MRF. In the first iteration of the algorithm $\forall i \in N_t$, M particles X_i^1 are drawn from initial beliefs $b_i^1 = p_i$, and all messages are set to uninformative states $m^1{}_{i\to j} = 1$. Each particle is assigned a weight $W_i^1 = p_i(X_i^1)$. The initial belief $b_i^1(\mathrm{x_i})$ is represented nonparametrically with an approximation using KDE as follows:

$$b_i^1(\mathrm{x_i}) \approx \sum_{p=1}^{M} W_i^{1,p} K\left(x_i - X_i^{1,p}\right) \tag{3.14}$$

In this context, $\mathrm{x_i}$ is the state at which nodes can be positioned in the density being estimated, and K is the kernel function typically a Gaussian. The summation reflects the contribution of each particle to the density at $\mathrm{x_i}$, weighted by W_i^1.

3.4.1. Message approximation

In NBP, messages between nodes represent the influence of one node's state on another's. For naïve BP, messages $m_{i\to j}^n$ are calculated as integral over all states of the sender node with pairwise compatibility function ψ_{ij}. Obtaining the marginalized distribution for continuous state space is intractable, because along with evaluating all states of the sender node, receiver node's all states must be considered, and this process must be repeated for all incoming messages for $i \in \mathrm{N(j)}$. Instead, M sampled and weighted particles $\{X_i^{n-1}, W_i^{n-1}\}$ from the belief b_i^{n-1}, where n is the current iteration, can be used to approximate each outgoing message as mixture of Gaussians from i to $j \in N(i), c_{ij} = 1$.

Although there is no information regarding the direction of node j relative to node i, distance measurement d_{ij} can be used to infer how far node j is from node i. Since particles X_i^{n-1} represents the position for node i, we can sample M angles θ uniformly from $[0,2\pi)$ and use these angles to scatter X_i^{n-1} with distance d_{ij}, allows to obtain particles $X_{i\to j}^n$ of message $m_{i\to j}^n$.

$$X_{i\to j}^n = X_i^{n-1} + \left(d_{ij} + \mathrm{n_{ij}}\right) \cdot (\cos(\theta)\sin(\theta)), \qquad n_{ij} \sim p_n(noise\ model) \tag{3.15}$$

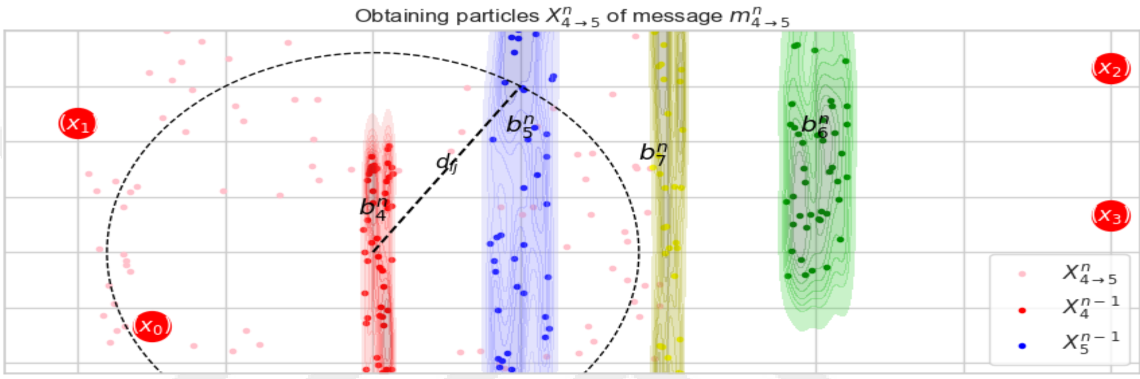


Figure 3.13. Spreading particles in a random direction to approximate messages

After obtaining particles $X_{i\to j}^n$ of message $m_{i\to j}^n$, each particle is weighted using the probability of connectivity function $P_c\left(X_{i\to j}^n, \bar{X}_j^{n-1}\right)$, which considers both particles of the message and the estimated position $\bar{X}_j^{n-1}$ of node j from previous iteration.

$$W_{i\to j}^n = P_c\left(X_{i\to j}^n, \bar{X}_j^{n-1}\right)\frac{W_i^{n-1}}{m_{j\to i}^{n-1}(X_i^{n-1})} \tag{3.16}$$

Denominator term adjust the weights W_i^{n-1} of the particle $X_{i\to j}^n$ from the previous iteration by the inverse of the message $m_{j\to i}^{n-1}$, ensuring that the weights for the particles of the message reflects the current connectivity and past information.

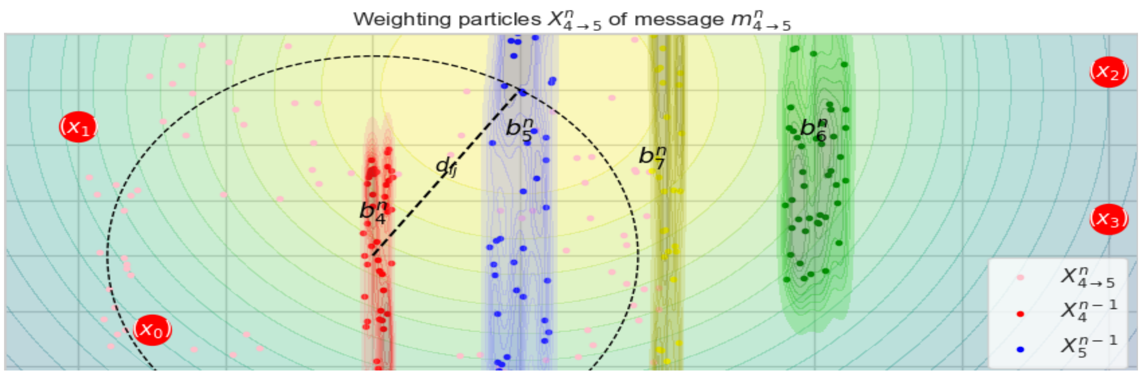


Figure 3.14. Weighting particles of messages obtained to approximate messages

Spreading particles of the node i allows to utilize the observed distance and from a mixture of Gaussian with weighted particles to approximate to the message $m_{i \to j}^{n}$. According to Equation 3.14, each particle of node j can be evaluated under this nonparametric distribution to obtain an opinion of i's particles about the position of particles of j.

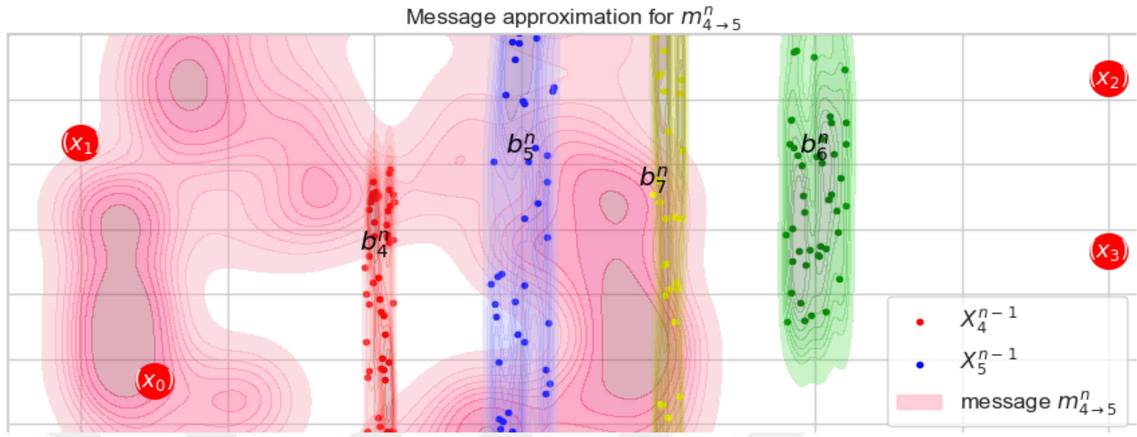


Figure 3.15. Message approximation

Scattering particles of i with uniformly sampled θ angles from $[0,2\pi)$ enables messages to be approximated without any directionality information. Although densities can be efficiently approximated with PF, to effectively represent the distributions, large number of particles are needed. However, from Figure 3.13, it can be seen that majority of particles for $m_{i \to j}^{n}$ will have little impact on the approximation. Instead, particles drawn from the marginals of previous iteration can be utilized to obtain a directionality about the nodes:

$$\theta_{ij}^{n-1} = \arctan\left(X_i^{n-1} - X_j^{n-1}\right),\ \mathrm{n} > 1 \tag{3.17}$$

Using such θ_{ij}^{n-1} angles enable to build a KDE, and to draw sample angles $\theta_{m_{i \to j}^{n}}$ to approximate the directionality and shift particles X_i^{n-1} to more focused region instead of randomly scattering.

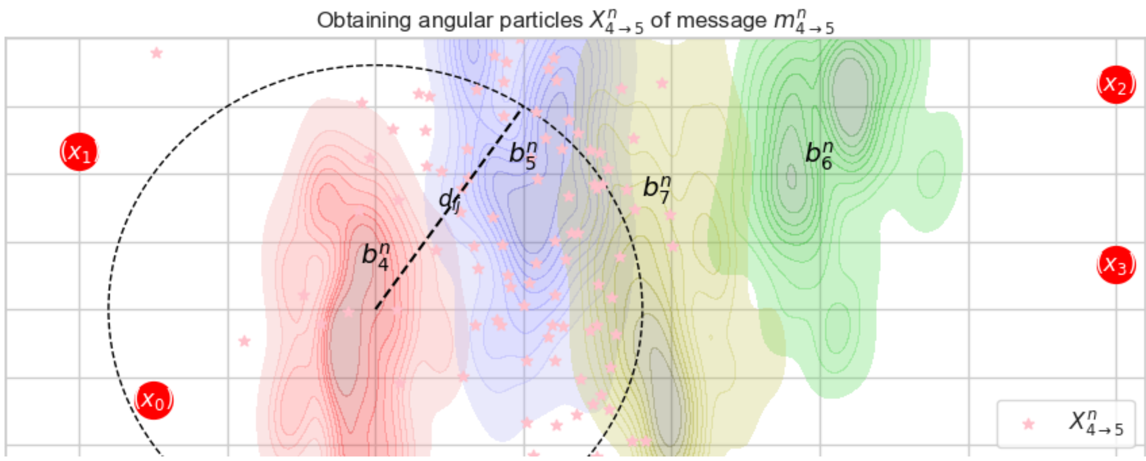


Figure 3.16. Angular particles of message

Utilizing the previous iteration's information about particles allows to represent messages effectively, therefore less particles can be used to efficiently approximate messages.

Approximations for two-hop messages are relatively straightforward. While distance estimates over intermediate neighborhoods can introduce uncertainties due to the nondirectional nature of distances, two-hop messages are crucial for maintaining balance in the network. These messages act as a counterforce, ensuring that nodes influenced by the pulling force of immediate neighbors do not get too close if they are beyond their direct communication range. In a distributed approach, nodes may lack information about those outside their one-hop neighborhoods, potentially leading to overestimation of distances. Two-hop messages help correct these estimates by pushing nodes apart when necessary, maintaining a more accurate and stable network configuration. Two-hop messages $m_{i\to k}^{n}$ is analytically calculated using the equation:

$$m_{i\to k}^{n}(X_k^{n-1}) = 1 - \sum W_i^{n-1} P_c(X_i^{n-1} - \bar{X}_k^{n-1}) \tag{3.18}$$

For each $\mathrm{k} \in \mathrm{N(i)}, c_{ik} = 0$, the equation captures the influence of node i on node k via an intermediate node j. The term $P_c(X_i^{n-1} - \bar{X}_k^{n-1})$ represents the probability of connectivity between the particles X_i^{n-1} and the estimated position $\bar{X}_k^{n-1}$ of node k. By summing the weighted connectivity probabilities W_i^{n-1} of all particles, obtaining a measure of the cumulative influence. Subtracting from 1 ensures that the message acts as a balancing factor, pushing nodes apart if they are not within their direct communication range.

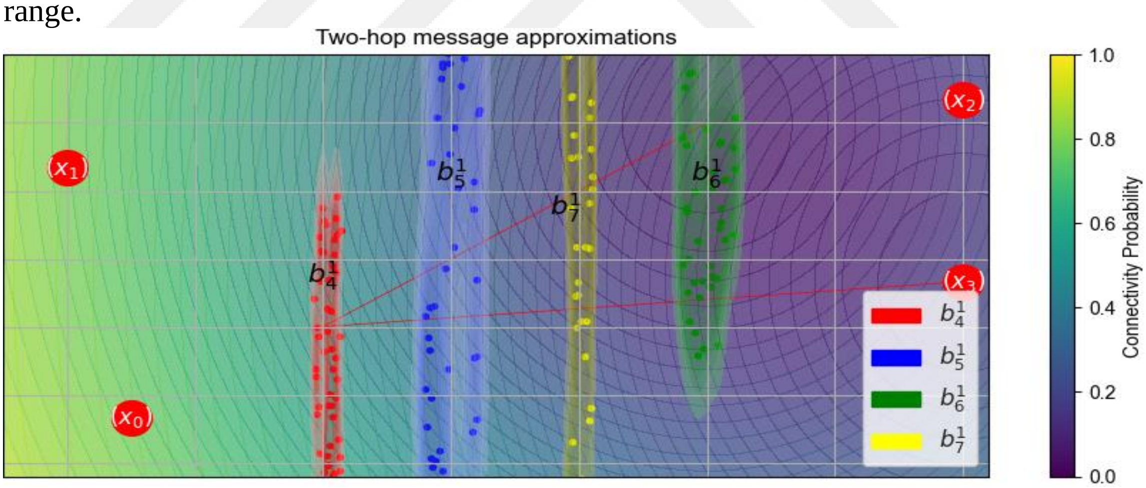


Figure 3.17. Two-hop message approximations

3.4.2. Belief update

Belief updates in NBP involve integrating information from incoming messages to refine the estimates of each node's state. Each node's belief is represented as a set of weighted particles, which are updated iteratively as new messages are received from neighboring nodes. In the belief update step, the belief b_i^n at node i, involves $\prod_{j\in N(i)} m_{j\to i}^{n}(x_i)$ from neighboring nodes, each represented as a mixture of Gaussians (see Equation 2.13). This process combines the influence of all neighbors to refine the belief at node i.

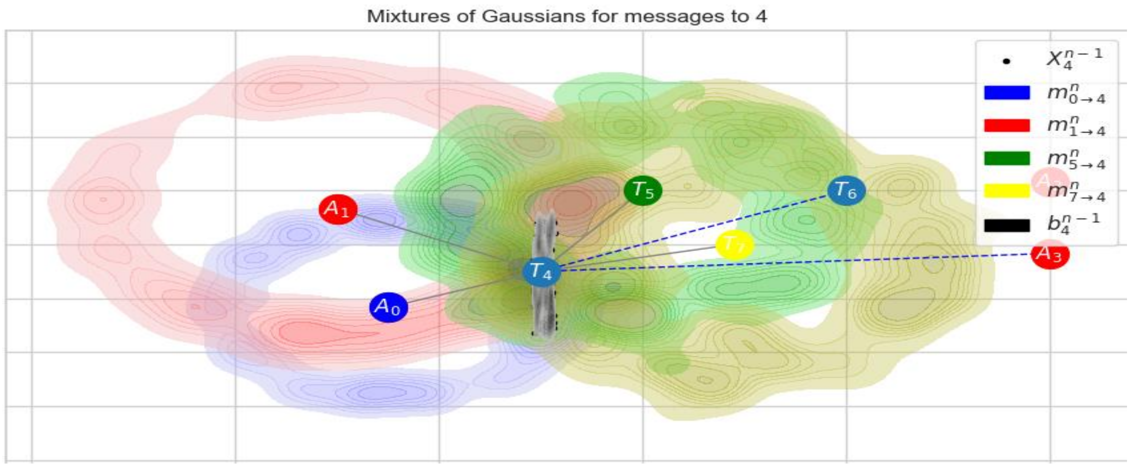
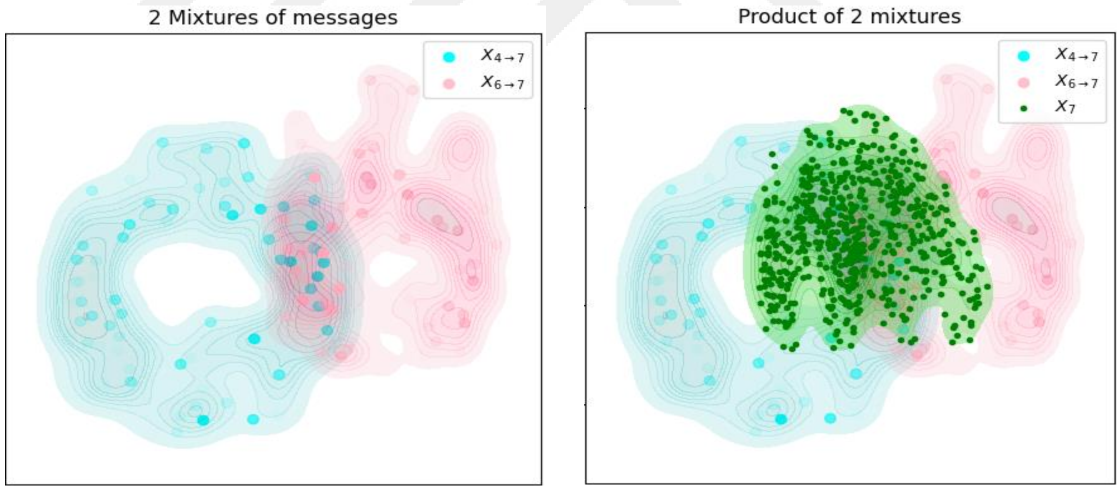


Figure 3.18. All incoming messages to a single node

However, the product of multiple mixtures of Gaussians results in an exponential increase in the number of components (particles), which quickly becomes computationally infeasible. When computing the product of multiple mixtures of Gaussians, each mixture is represented by a set of Gaussian components. The product of two mixtures results in a new mixture whose number of components is the product of the numbers of components in the original mixtures. This exponential growth continues with each additional mixture involved in the product.



(a) **(b)**

Figure 3.19. Product of mixtures of messages; **a)** sampling particles from incoming messages; **b)** naïve product of mixtures represented with particles

Therefore, instead of computing b_i^n beliefs, such marginals can be approximated by sampling particles from proposal distribution $\mathrm{q}_\mathrm{i}^\mathrm{n}$, resulting in approximation for belief $\tilde{b}_i^n$. The proposal distribution $\mathrm{q}_\mathrm{i}^\mathrm{n}$ designed to approximate the true distribution b_i^n which can be specifically formulated for the domain in which NBP is applied on. Often, a suitable proposal for the $\tilde{b}_i^n$ is the belief from previous iteration $\tilde{b}_i^{n-1}$. However, this approach is vulnerable to inefficiencies when the prior p_i is not informative enough. In such cases, the sampled particles may spread too broadly, leading to low representative power. This can result in particles within local maxima to dominate the distribution, potentially causing erroneous results due to poor coverage of the true state space.

For CL implementation, this proposal can be the influence of neighboring nodes. Therefore, proposal of i, q_i^n can be represented with mixtures of multiple proposals $q_{j\rightarrow i}^n$ where $j \in N(i), c_{ij} = 1$. In this sense, each proposal $q_{j\rightarrow i}^n$ represents the influence of node j on node i, consequently, such distributions are messages $m_{j\rightarrow i}^n$.

$$q_i^n(X_i^{n,J\rightarrow i}) = \sum_{j\in N(i)\wedge c_{ij}=1} m_{j\rightarrow i}^n(X_i^{n,J\rightarrow i}) \quad (3.19)$$

According to Equation 2.18, proposal q_i^n can be written as sum over all incoming messages. In the equation $X_i^{n,J\rightarrow i}$ represents union of all particles sampled from messages $m_{j\rightarrow i}^n$ from all immediate neighbors $j \in N(i)$ and $c_{ij} = 1$. Specifically, sampling $\frac{kM}{N(i)}$ particles from each incoming message, here k is the sampling parameter that controls how much particles from each message should be sampled.

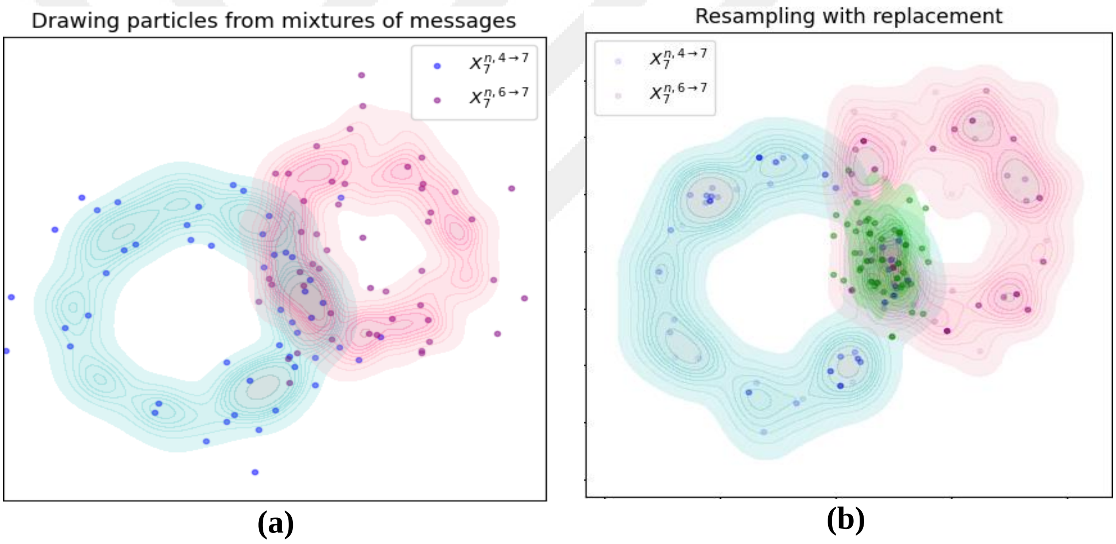


Figure 3.20. a) Drawing particles from each mixtures of messages; **b)** resampling with replacement according to their weights

Furthermore, each of kM particles are evaluated with the product of all incoming messages $m_{j\rightarrow i}^n$ and through using Equation 2.19 particles are weighted for resampling with replacement of M particles according to their weights, resulting in most probable particles to remain for representing the belief $\tilde{b}_i^n$.

$$W_i^{n,J\rightarrow i} = \frac{\tilde{b}_i^n(X_i^{n,J\rightarrow i})}{q_i^n(X_i^{n,J\rightarrow i})} \quad (3.20)$$

After obtaining $\{X_i^{n,J\rightarrow i}, W_i^{n,J\rightarrow i}\}$ for all i, Kullback-Leibler divergence is calculated to measure the change between consecutive iterations. When the divergence drops below a threshold for all nodes the algorithm is halted. On the other hand, nodes which have low divergence can also be used as pseudo anchors (Savic and Zazo 2013).

$$\mathrm{KL}_i^n = \sum_{p}^{M} W_i^{n,p} \log \frac{W_i^{n,p}}{M_i^{n-1}\left(X_i^{n,p}\right)} \tag{3.21}$$

Finally, for each target node's belief can be approximated and estimations about locations can be made from weighted sum of particles:

$$\forall i \in N_t, b_i^n(\mathrm{x_i}) \approx \sum_{p=1}^{M} W_i^{n,p} K\left(x_i - X_i^{n,p}\right) \tag{3.22}$$

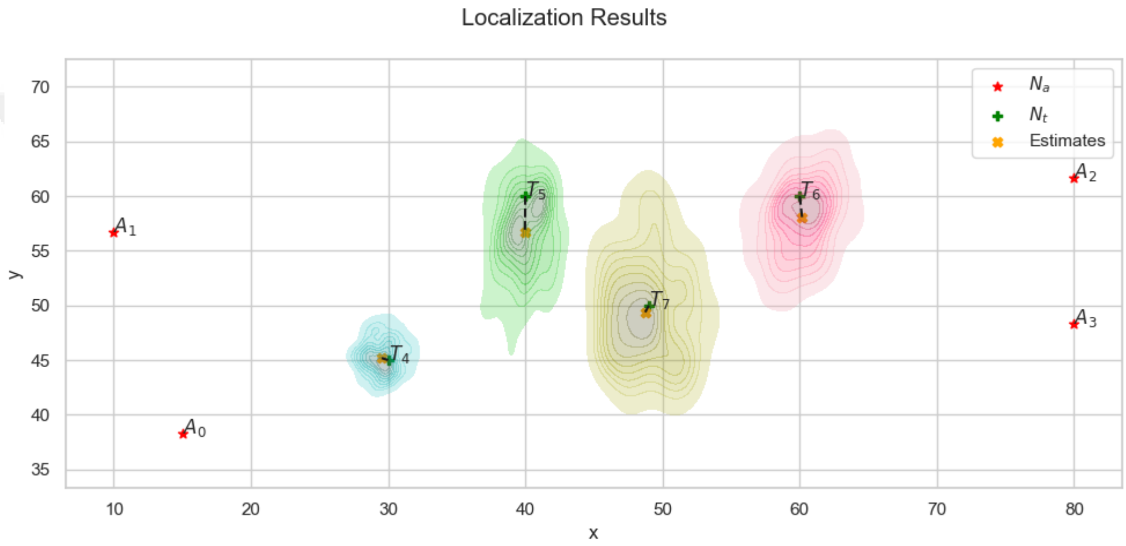


Figure 3.21. Location estimates along with beliefs about nodes

4. RESULTS AND DISCUSSION

In this section, the theoretical framework and methodologies described in previous section are performed on synthetically generated network of agents and the performance of the localization system under various scenarios are analyzed and reported. First, experimental setup is described, explaining the simulation environment used to perform experiments, how synthetic RSS measurements are generated to obtain distance approximations, parameters and evaluation metrics are presented. Following this, results obtained from various simulations considering the effect of n-hop connectivity, bounding boxes formed from direct and indirect anchors, and anchor placement strategies are explored. Finally, strengths and limitations are discussed and potential applications in real-world scenarios are considered.

The CL system was implemented in *Python*, utilizing *NumPy* for computations and vectorization of custom functions such as creating bounding boxes and sampling particles from each prior, along with iterative NBP implementation to approximate messages and computing marginalized distributions to efficiently compute messages and beliefs. *SciPy* library is used for statistical operations such as forming KDEs with particles and weights to represent continuous nonparametric distributions. For visualizations, *Matplotlib* and *Seaborn* libraries are used. To identify the bottlenecks, a profiling tool for dynamic program analysis *cProfile* is used. Since approximating messages and beliefs of each agent can be achieved in parallel, to asynchronously perform executions, thread pooling feature of *Python* is utilized for optimization of runtime performance.

4.1. Experimental Setup

The experiments were conducted on a machine equipped with an Intel i7-12700K processor and 16GB RAM running on Windows 11 Pro. The software setup included *Python 3.10* along with necessary libraries, the implementation utilized *Python's* built-in *cProfile* module for performance profiling and *concurrent* module for parallel execution of tasks. *VS Code* editor is used for programming environment and *Git* for version control.

For simulations, multiple synthetic datasets for various scenarios were generated to represent networks of agents. Various parameters are utilized to generate synthetic datasets are given in Table 4.1.

Table 4.1. Parameters of synthetic datasets

Symbol	Parameter	Typical Values
N	Number of agents	(8 - 200)
m	Half of one side of deployment area in meters	(50 - 250)
A	Deployment area agents will be localized in	$[-m, m, -m, m]$
a	Number of anchors with known positions	(2 - 16)
p	Number of particles for each agent	(20 - 200)
r	Connectivity radius for each agent in meters	(12 - 48)
i	Number of iterations algorithm runs for	(5 - 20)
k	Determines number of batches to sample p particles	(1 - 5)
n	Number of hops to consider indirect neighbors	1, 2, 3
c_n	Number of hops to consider as direct neighbors	1, 2

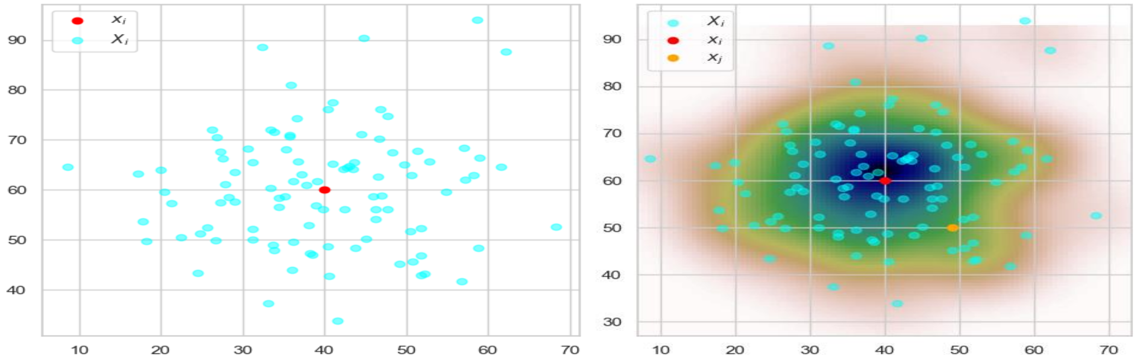
Some of the key parameters for CL task is N that specifies the density of the network along with A. The denser the network, the more neighbors to obtain messages from, hence leading to more reliable estimates. Another important parameter is the a, number of anchors which are used to determine prior knowledge about each target agent in the network. Each belief is approximated with particles p, higher the number of particles, leading to better representation of the probability distribution. Using k increases the representation power of each passed messages, however, increasing the number of batches heavily affects computational complexity. Utilizing n hops help estimations to avoid unrelated locations. Parameters p, i, k, n, c_n are algorithm parameters and they can be adjusted to fit to scenario while N, m, a, r defines the scenario.

For randomly generated networks consisting of N agents, where a of them are anchors, the agents are uniformly scattered within the A, and each agent has a connectivity radius r. To obtain distance estimates between each agent, Equation 2.1 is used to obtain the initial distances through analytical computation. The statistical model for RSS measurements can be approximated using Equation 2.3 which simulates RSS measurements based on these distances, which is referred to as forward filtering. Following this, Equation 3.1 is employed to backward filter these simulated measurements, therefore, obtaining distance estimates for the agents.

To avoid analytical computation through Equation 2.1, one can simulate the RSS measurements using a probabilistic approach. This involves sampling p particles X_i around x_i with a Gaussian distribution. Each particle's weight is then calculated using the probability of connectivity $P_c(x_i, X_i)$. Using the set of particles and corresponding weights $\{\mathrm{X_i}, \mathrm{W_i}\}$ for node i $\in N$, a KDE estimate similar to Equation 3.14 can be formed to represent the connectivity density. By evaluating this density at x_j, the likelihood of detection p_d can be obtained.

$$\hat{d}_{ij} = P_c^{-1}(1 - p_d) + v_{ij} \tag{4.1}$$

These distances can then be applied with forward and backward filtering to obtain simulated distance estimates between agents. This approach allows the use of any distance-based statistical measurement model to produce simulated measurement estimates.



(a) (b)

Figure 4.1. a) Sampling particles from around target x_i; **b)** forming KDE distribution with sampled particles

Although, this approach is parametric due to Gaussian assumption, even though it uses nonparametric KDE, any non-Gaussian distribution can be incorporated by adjusting the particle sampling distribution accordingly, allowing for greater flexibility and adaptability in modeling diverse network conditions.

Along with parameters of the network scenario, parameters representing conditions of the localization environment and agent characteristics are presented in Table 4.2.

Table 4.2. Parameters of measurement model

Symbol	Parameter	Typical Values
α	Path loss exponent	3
Pi	Transmission power of agent i	20 dBm
d_0	A reference distance	1

4.1.1. Evaluation metrics

In this thesis, deployment area is assumed to be homogeneous, and all agents are the same. Therefore, these parameters are set to their usual values and used the same throughout the results section. α parameter characterizes how quickly the signal attenuates as it propagates through the environment and typical values range from 2 to 4 depending on the environment. P_i is the power level at which agent i sends out signals. d_0 is a predefined distance at which the path loss exponent is known or assumed, often set to 1.

In this section, we evaluate the performance of localization method using Root Mean Square Error (RMSE), CRLB as the lowest achievable benchmark given the measurement model, uncertainties obtained from individual covariance matrices and Procrustes analysis.

RMSE is a commonly used measure of the differences between the true positions and the estimated positions of the nodes in the network. RMSE provides a direct measure of the accuracy of the localization algorithm.

$$\text{RMSE} = \sqrt{\frac{1}{N_t}\sum_{i\in N_t}\|x_i - \hat{x}_i\|^2} \tag{4.2}$$

Although RMSE provides a measure of accuracy of the algorithm, does not provide a context for how much more the algorithm can improve localization estimates. Therefore, CRLB can be used as a theoretical benchmark according to lower bound on the minimum achievable variance of the measurements.

$$CRLB = \sqrt{\mathrm{Tr}(I_{ii}^{-1})} \tag{4.3}$$

Here I is the FIM matrix formed from observations. Although each node will have its own CRLB, by taking the trace of the inverse FIM a representative CRLB for the network can be obtained to use it as a benchmark to evaluate the performance of the localization algorithm.

Another evaluation metric is the uncertainties of the estimates. Uncertainty refers to the degree of confidence in the position estimates of each node. During MIS step of the belief update stage, particles are sampled from proposal distribution q_i and resampled with replacement according to their weights. These weights are derived from incoming messages that represent the opinions of neighboring nodes about the position of the particles. If the particles have approximately equal weights, the entropy of the distribution will be high, indicating that the particles have low representative power suggesting low confidence in the position estimates.

$$C_i = \frac{\sum_{p=1}^{M} W_i^p \left(X_i^p - \hat{x}_i\right)\left(X_i^p - \hat{x}_i\right)^{\top}}{\sum_{p=1}^{M} W_i^p} \quad (4.4)$$

To quantify this uncertainty, the covariance matrix C_i of node i can be used, which is based on its particles $\{X_i, W_i\}$ and the estimate $\hat{x}_i$. The uncertainty u_i of the node is represented as $\sqrt{\mathrm{Tr}(C_i)}$ provides a measure of the total spread of the particle distribution around the estimate, therefore, it captures the combined standard deviation of the node's belief, reflecting overall uncertainty in the position estimate.

While other metrics provide a measure of accuracy by considering the distances between true and estimated positions or their reliability, Procrustes similarity focuses on geometric similarity, this metric complements other evaluation metrics by offering insight into overall alignment between sets of positions. Procrustes similarity ranges between 0 and 1, higher means more similar.

The evaluation of CL approach implemented in this thesis using RMSE, CRLB, uncertainties, and Procrustes analysis provides a comprehensive overview for the algorithm. RMSE offers a direct measure of accuracy, while CRLB serves as a theoretical benchmark indicating the potential for improvement. The covariance matrices quantify the uncertainty of estimates, providing insights into the reliability of the position estimates. Procrustes analysis with its focus on geometric similarity, complements these metrics by evaluating the overall alignment between position sets. Together, these metrics enable through evaluation of the effectiveness in various deployment scenarios.

4.2. Results

To investigate various approaches used in our work for improving localization quality, we first examine the localization estimates without incorporating multi-hop connectivity, prior knowledge, or any anchor placement strategies. Experiments are performed with parameters presented in Table 4.3, and each simulation is repeated 100 times to obtain reliable results.

Utilzing varying parameters allows to discover reliability of CL approach under various conditions. First, exploring how various degrees of n-hop connectivity affects localization results, comparing with the most primitive results and gradually introducing improvement techniques to demonstrate applicability under various situations. Secondly, we demonstrate that in scenarios where prior knowledge exists, such information can play a crucial role in improving estimates. Lastly, effective anchor deployments are considered for various network layouts.

Table 4.3. Parameters for simulations

Parameters	Typical Values
N	125
m	150
a	7
p	125
r	33
i	5
k	4
c_n	1

The comparison between the true and estimated graphs without multi-hop connectivity, prior knowledge or anchor placement strategy provides insight into the baseline performance of the localization algorithm. In this initial analysis, the nodes rely solely on immediate neighbors and raw measurements to estimate their positions. The true graph represents the ideal connectivity and positioning of nodes, while the estimated graph shows the outcome of the localization process under given constraints. This comparison highlights the inherent challenges in achieving accurate localization in the absence of additional information, setting reference points for evaluating subsequent improvements.

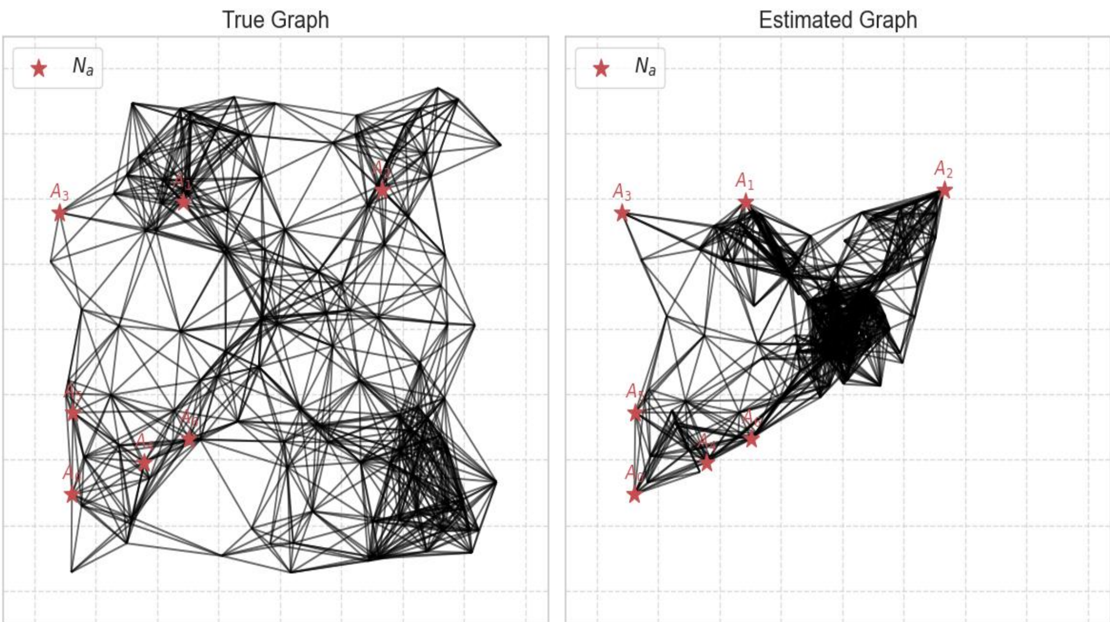
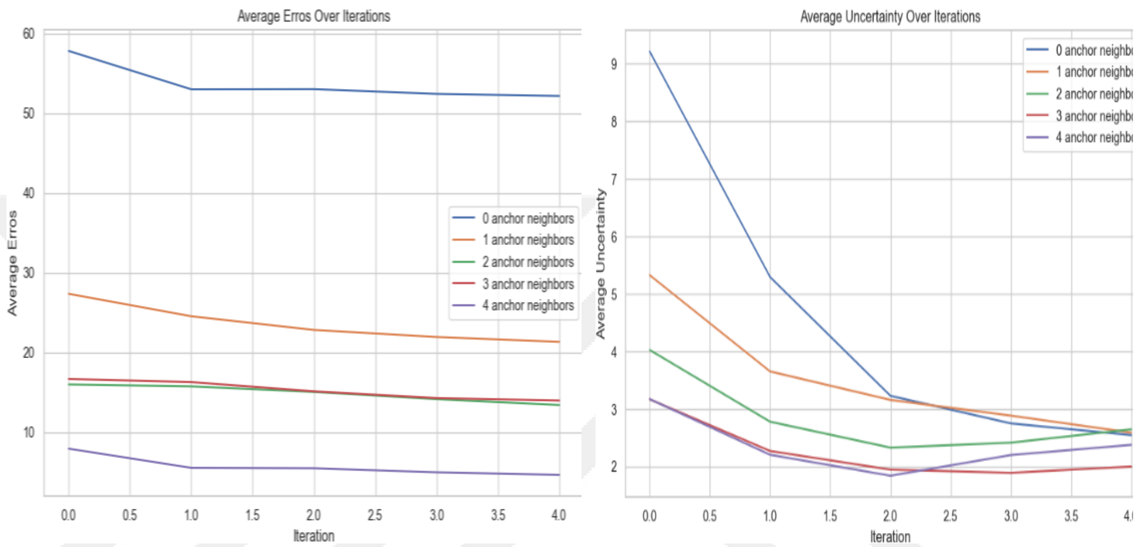


Figure 4.2. Comparison of graphs

Examining the graphs of Figure 4.2 reveals notable discrepancies between the true and estimated positions. The absences of multi-hop connectivity results in unrelated agents being localized in proximity, thereby significantly altering the neighborhood structure. Despite the lack of prior knowledge, anchor nodes provide reference points that assist in more accurately positioning the neighbors of anchors. Conversely, the accuracy of position estimates for nodes without anchor neighbors diminishes as the distance from anchors increases due to stacking uncertainties.

Measurement noise inherent in RSS-based distance estimation leads to greater variability in distances, increasing the uncertainty of the position estimates. High connectivity leads to more neighbors collaborating, therefore a potential decrease in the entropy, lowering the uncertainties.



(a) (b)

Figure 4.3. Investigation of errors and uncertainties with various numbers of anchor connectivity degrees; a) average errors over iterations; b) uncertainty change

From Figure 4.3, it is evident that having anchor neighbors greatly reduces both the errors and the uncertainties about the estimates. Therefore, it is necessary to utilize efficient anchor placement strategies.

Further investigation of the estimated positions for each agent and their true positions provided in Figure 4.4 reveals that although the estimations are highly erroneuos, agents are estimated within close proximity to their true neighbors. Since NBP can be applied as a distributed localization system without global context, agents might be significantly displaced from their original positions. However, they maintain similar relative distances to their neighbors.

Without global context, agents cannot determine that they should not be close to certain nodes. Consequently, agents estimated to be close to non-neighbor nodes do not share measurements due to the lack of connectivity, resulting in no knowledge of the closely estimated non-neighbor nodes. Strategic placement of anchors and utilizing multi-hop distances can provide each node with prior knowledge about its position, thereby reducing the likelihood of highly erroneous localization estimates.

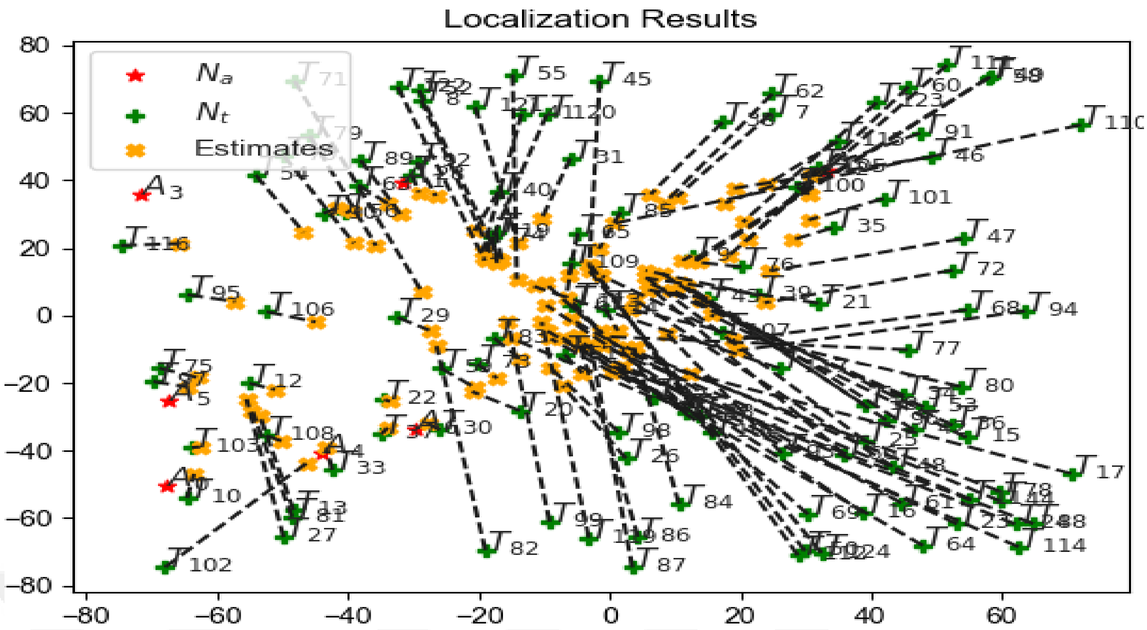


Figure 4.4. Localization results

To further understand the impact of neighborhood connectivity on localization accuracy, relationship between the number of neighbors each agent has, and the relative distances are investigated.

$$\Delta_i = \frac{1}{|N_i|} \sum_{j \in N_i, c_{ij}=1} \frac{|d_{ij} - \hat{d}_{ij}|}{\sum_{k \in N_i} d_{ik}} \tag{4.5}$$

Normalizing the distance difference between the original layout and the estimated layout under the original connectivity, over the sum of total neighborhood distances for multiple experiments yields that normalized distance difference decreases as the number of neighbors increases, despite number of measurements affected by environmental noise increases. This shows that collaboration among neighbors helps improve the relative localization performance.

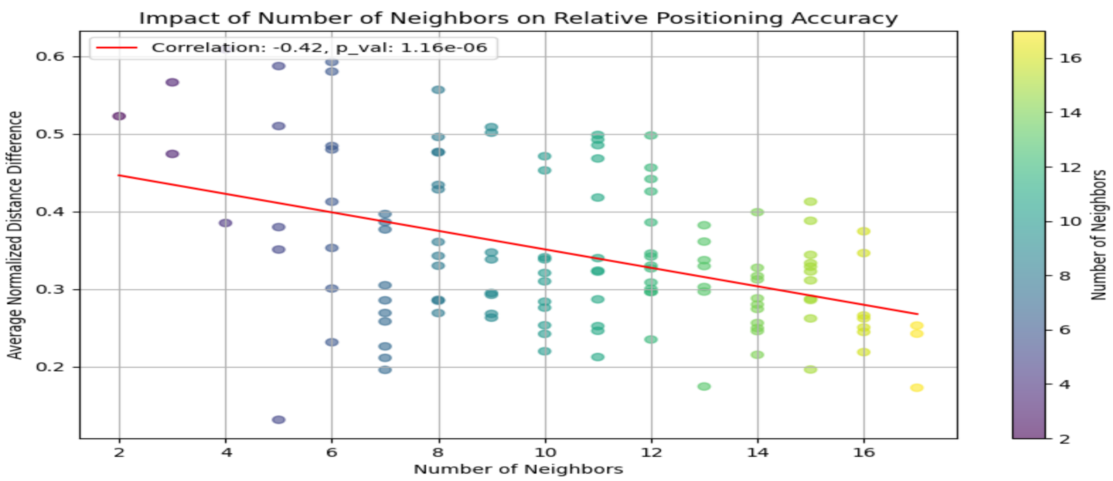


Figure 4.5. Impact of collaboration on relative localization

Figure 4.5 demonstrates a clear inverse relationship between the number of neighbors an agent has and the average normalized distance difference, indicating improved relative positioning accuracy with increased number of neighbors. Negative correlation and highly significant p-value confirms collaboration among increasing neighborhoods improves relative distances of estimates.

4.2.1. N-hop connectivity

When considering 1-hop connectivity, each node relies solely on information from its direct neighbors. This limited scope can result in higher localization errors, especially in sparse networks. By extending the connectivity to multiple hops, nodes can aggregate information from a broader context of the network, thereby improving the accuracy of the localization estimates. Following experiments have been conducted with various degrees of multi-hop connectivity.

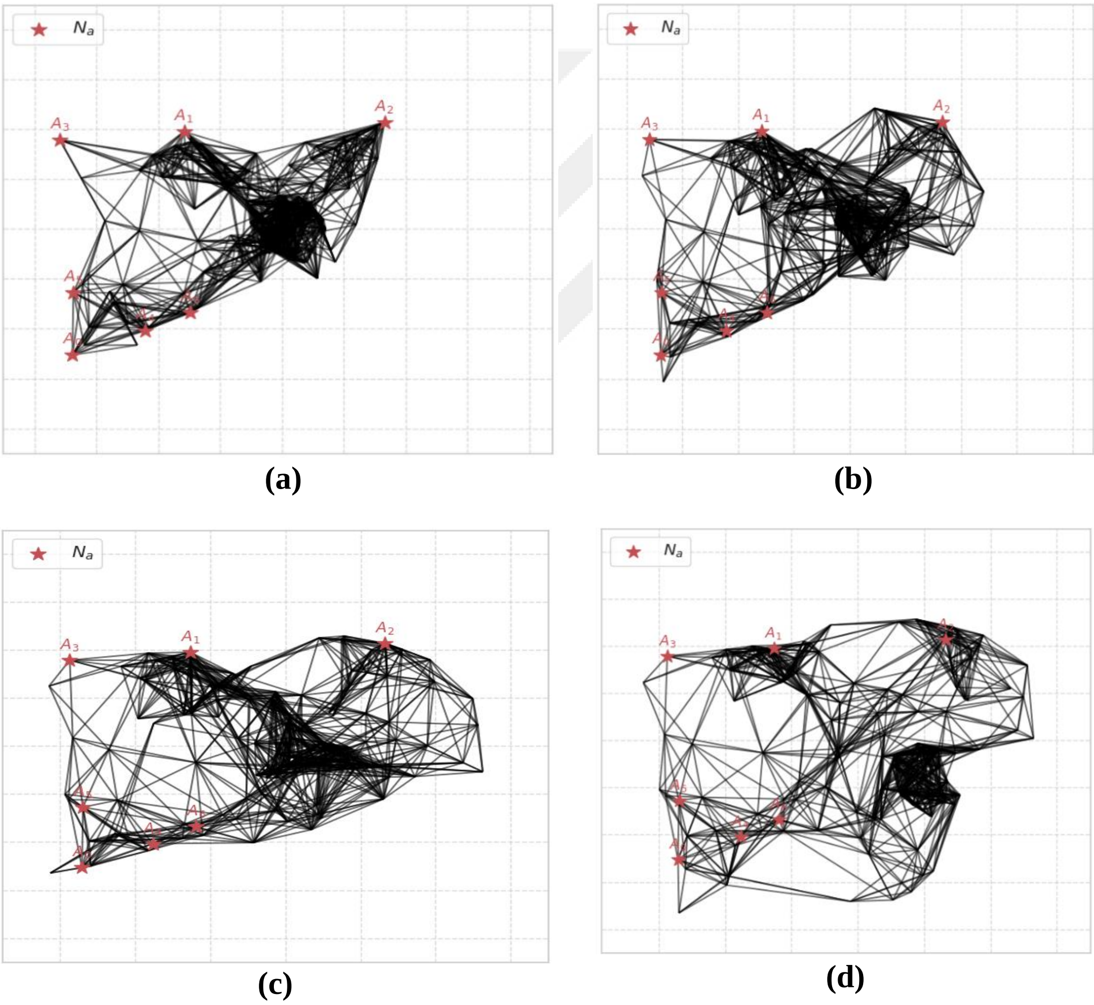
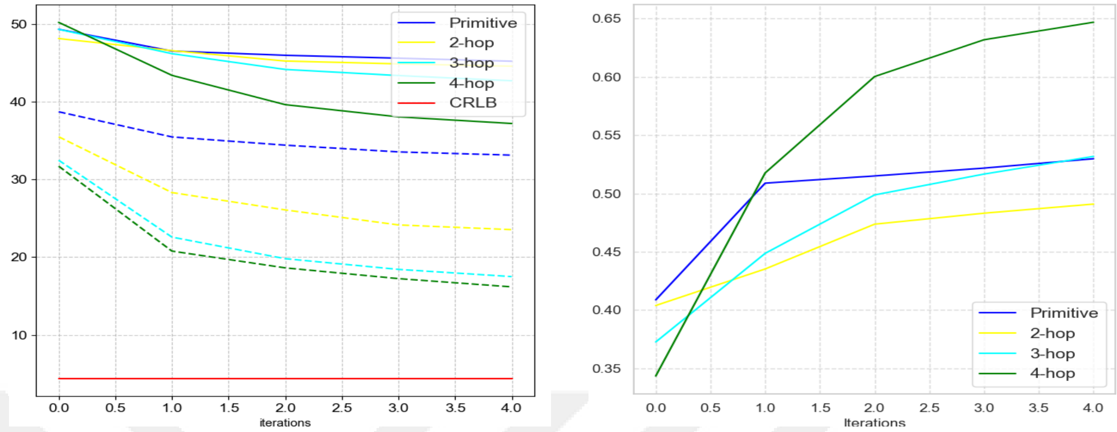


Figure 4.6. Various n-hop connectivity; **a)** 1-hop; **b)** 2-hop; **c)** 3-hop; **d)** 4-hop

As the multi-hop connectivity increases, the estimated graph layout more closely approximates to the true form. This improvement arises from the broader network context available to each node. By incorporating indirect measurements, nodes can better refine their estimates and understand which nodes should not be in proximity. While immediate neighborhoods provide information on the relative closeness of nodes, multi-hop distances offer insights about which nodes should be further apart. This additional information acts as a balancing force, ensuring a more accurate and well-distributed network layout.



(a) **(b)**

Figure 4.7. Errors and similarities; **a)** displays the RMSE with continuous lines and median error with dashed lines, accounting for outliers with large errors; **b)** Procrustes similarity of various multi-hop connectivity networks

Figure 4.7 illustrates the impact of multi-hop connectivity on localization performance and layout similarity. As the number of hops increases, both localization accuracy and layout similarity improve. In Figure 4.7(a), the RMSE and median error metrics are shown. While RMSE is sensitive to outliers with large errors, median error provides an additional measure of how closely the estimated layouts resemble the true network layout.

Utilizing additional constraints from multi-hop neighborhoods allows estimates to avoid regions where localization errors are likely. This broader context helps nodes adjust their positions by considering not only their immediate neighbors but also the extended network structure. As the number of hops increases, the nodes gain more information, reducing the probability of significant deviations from their true positions.

Consequently, utilizing multi-hop neighborhoods improves localization performance through increasing collaboration among target nodes to eliminate erroneous estimates. However, in a distributed framework, assumption of each clique performing localization independently restricts the obtainability of global context regarding the entire deployment area, hence, it also plays as a limiting factor. Although this characteristic may come up as an hinderance, it helps obtain accurate network layout due to eliminating ambiguous rotations occurring from not using enough anchor nodes.

One limitation of utilizing multi-hop connectivity is that it requires efficient broadcasting to share obtained information to neighborhoods. As the number of hops increases, computational complexity also increases, because the cliques regarding each neighborhood's population also increases, hence, this necessitates the analytical computation of messages over non-immediate neighborhoods. This approach does not require computation of outgoing messages.

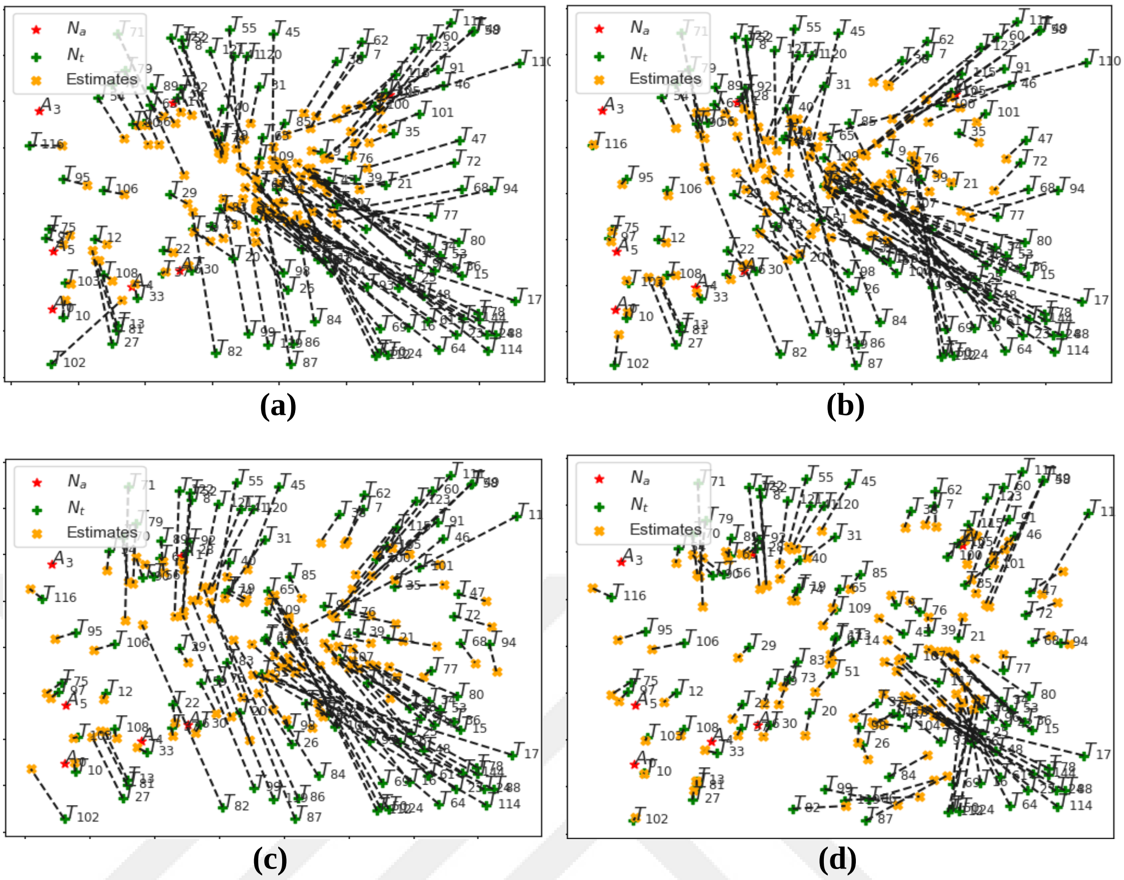


Figure 4.8. Localization results for various n-hop connectivity; **a)** 1-hop; **b)** 2-hop; **c)** 3-hop; **d)** 4-hop

However, using multi-hop connectivity feature is not without its limitation. One significant drawback is the increased computational complexity associated with incorporating additional hops. As the number of hops increases, each iteration of the localization algorithm takes significantly longer time.

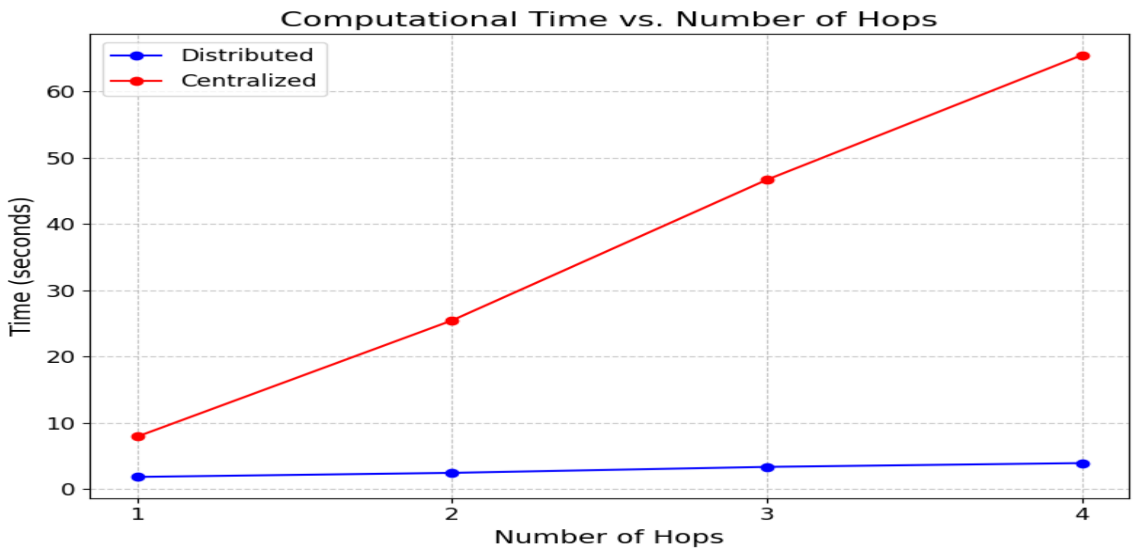


Figure 4.9. Computational time of number of hops per iteration

This increased computational burden is due to the NBP requiring consideration of larger sets of possible states for each neighbor, consequently their particles. As the neighborhood grows, leading to an increase in the number of computations needed. Each additional hop introduces more nodes into the neighborhood, requiring algorithm to process more information.

4.2.2. Effect of priors

Utilizing prior knowledge, such as the positions of anchors, significantly improves the accuracy and reliability of position estimates. This approach mitigates the ambiguities associated with relative localization by providing absolute reference frames that guide the estimation process. Without prior knowledge, particles representing the beliefs about node positions are sampled from the entire deployment area in the first iteration, necessitating a large number of particles to effectively represent the posteriors. This increases computational complexity and can lead to poor approximations if insufficient number of particles are used. Incorporating prior knowledge helps refine initial estimates, reduces overall uncertainty, and decreases the number of particles needed to represent distributions effectively.

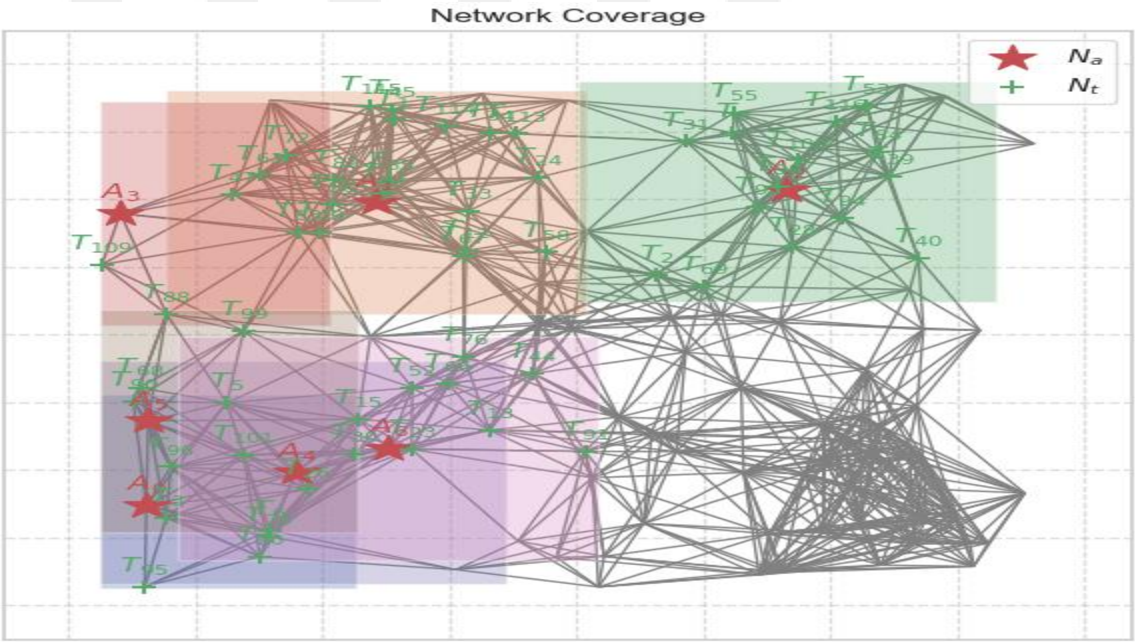
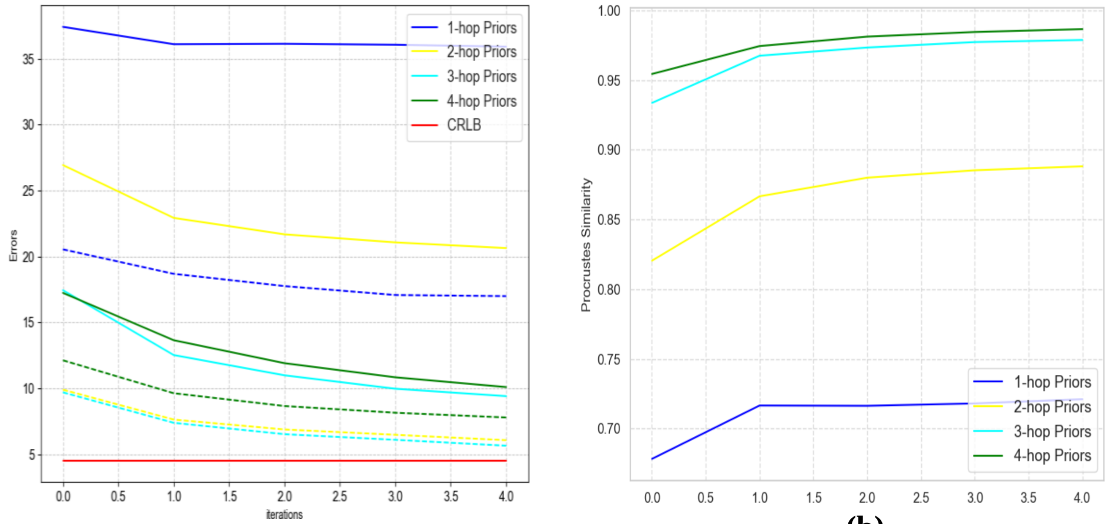


Figure 4.10. Network coverage by anchors of true graph

By leveraging the known positions of anchor nodes, bounding boxes can be created from their intersections, providing a prior belief about each node's position and thus narrowing the search space. In each iteration of the NBP, particles sampled from proposal distributions can be filtered using these priors. Additionally, utilizing multi-hop distances allows target nodes without immediate anchors to obtain pseudo priors. For nodes within the neighborhoods of anchors, increasing the number of anchors can further refine their priors, enhancing the precision of the estimates and improving localization accuracy.



(a) **(b)**

Figure 4.11. Errors and similarities over various n-hop connectivity's with prior knowledges obtained through anchors; **a)** RMSE and median error over various n-hops; **b)** Procrustes similarity to true network layout under various connectivity

Figure 4.11 presents RMSE and median error on various multi-hop connectivity situations, along with network layout similarity. Comparing with the errors and similarities provided in Figure 4.7, it is evident that utilizing prior knowledge significantly improves localization estimates. For each result in the figure, experiments are conducted with 100 different trials with various initializations. One interesting point to note is that, despite having less connectivity, median error on 2-hop with priors shows promising results. Although having a rather high RMSE, the comparable median error with 3 and 4 hop connectivity indicates that erroneous estimates are less frequent, but when they occur, they tend to be more severe.

Although, under 2, 3, and 4-hop connectivity median error approaches CRLB benchmark, discrepancy between RMSE results indicates that while majority of networks have plausible localization estimates, group, or groups of nodes have rather large errors causing network layout to have differences. In Figure 4.12, targets whose have multiple anchors within their vicinity show promising results, however due to limited connectivity to anchors, such targets have erroneous estimates. This is mainly due to if there is no anchor within connectivity of targets, their prior depends upon the deployment region, which can be rather large compared to number of particles used to represent each of their probability distributions. Considering priors as a uniform distribution led to sampling particles of such targets from a large area, hence, causing representative power of each particle to decrease significantly, increasing entropy for passed messages. This increase in the entropy causes selection of candidate particles away from true positions of targets. However, although such groups of targets may have been estimated based on relative distances to each other, they are together, estimated from the wrong region of the deployment area. Hence, causing groups of targets to be estimated together, but at wrong regions. Therefore, as the n-hop connectivity increases, targets obtain information regarding which agents that they should have distance to, leading network layout similarity to increase.

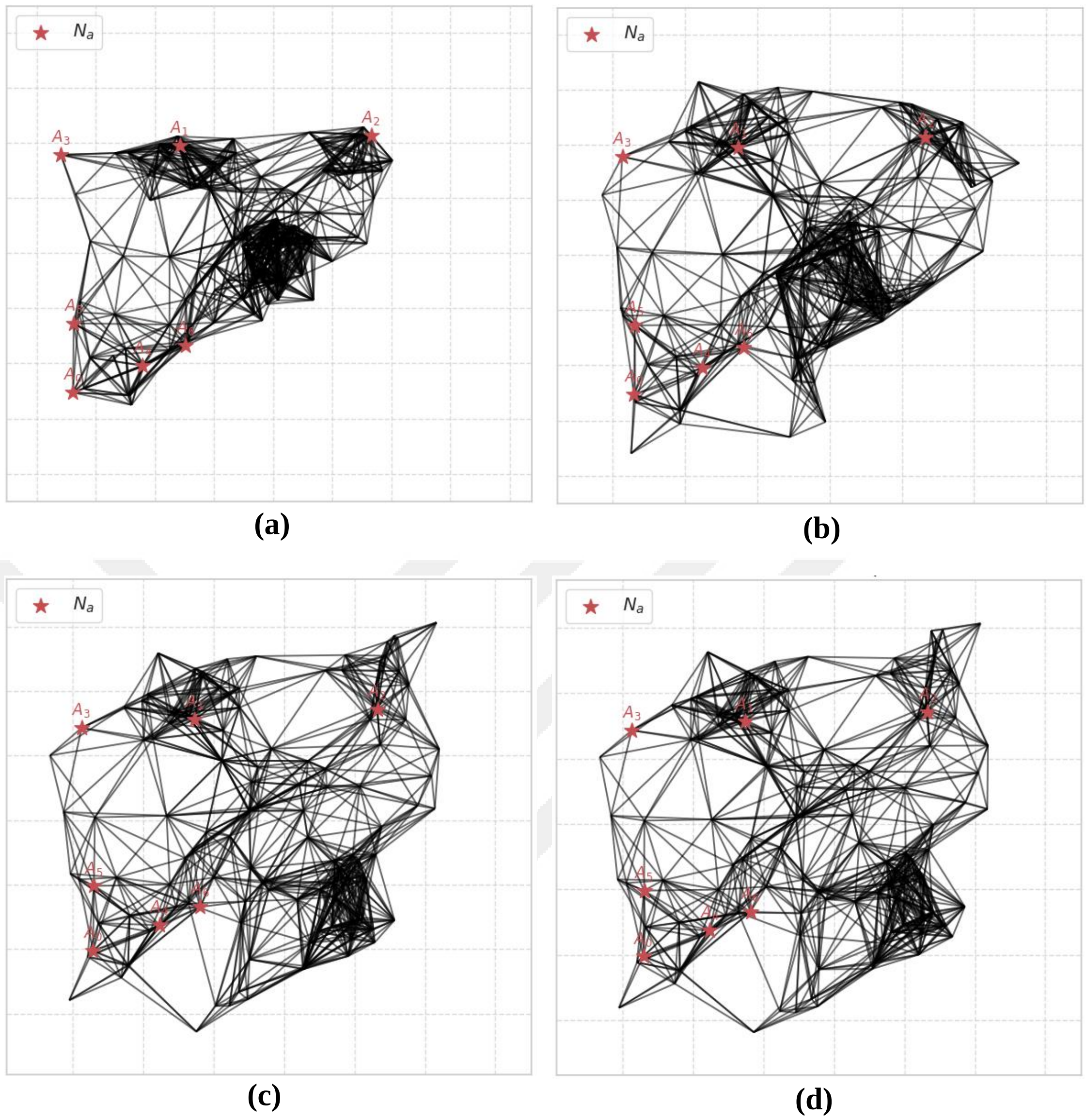


Figure 4.12. Estimated network layouts under various connectivity with priors obtained from anchor points; **a)** 1-hop; **b)** 2-hop; **c)** 3-hop; **d)** 4-hop

Further investigation of network layouts presented in Figure 4.12(b) suggests that 2-hop priors are effective in providing reliable localization under normal conditions but may be more susceptible to outliers in extreme cases compared to higher-hops priors. While 2-hop priors achieve sufficient accuracy in regions with at least one anchor, as seen in the upper-right and upper-middle parts of the network, the lack of connectivity to anchors in the local region adversely affects the estimates, such as the lower-right section of the network.

Although accurate localization typically requires at least three anchors within the vicinity to eliminate ambiguity of flip, rotation, and scale, CL techniques can achieve sufficient estimates with only one anchor within the vicinity. This suggests that with robust anchor placements, 2-hop connectivity is sufficient for estimations without requiring multiple anchors to overlap in the region, leading to cost effective localization solutions. Limiting connectivity to 2-hop also improves computational burden which can be seen from Figure 4.9 and allows for plausible real-time localizations.

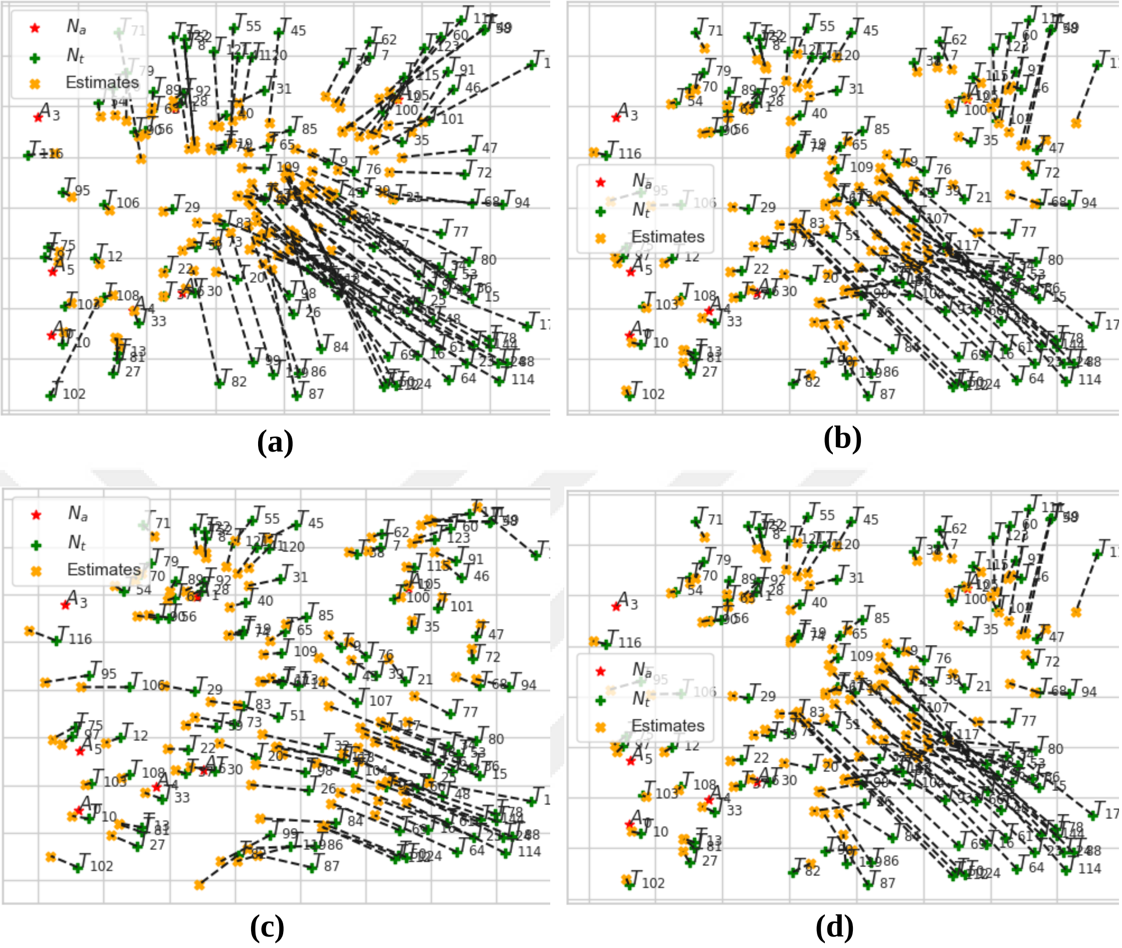


Figure 4.13. Localization results with priors; **a)** 1-hop; **b)** 2-hop; **c)** 3-hop; **d)** 4-hop

Moreover, considering Figure 3.6 true positions of nodes often falls near the edge of their priors. Therefore, under noisy measurements priors are misaligned with the true positions, forcing algorithm to make estimates away from their intended position and causing erroneous estimates. This can be seen from Figure 4.13, for some nodes, even 1-hop connectivity provides better results than 3 and 4-hops. As the number of hops increases, connectivity to anchors creates smaller regions for priors, therefore limiting the flexibility and probabilistic nature of the NBP.

In conclusion, prior knowledge of the positions can significantly enhance the localization estimates; however, with a drawback of increasing computational burden, along with erroneous estimates stemming from overly restricted priors obtained from 3+hops and due to inherent measurement noises. Therefore, efficient anchor placement strategies can be utilized to reduce the number of hops required for accurate localization and cost-effective solutions.

4.2.3. Anchor placement

Instead of randomly scattering anchors, strategically placing within the deployment area improves network coverage and enhances localization reliability and accuracy. Having at least one anchor in the vicinity significantly reduces ambiguities of estimations. Although some targets may have large priors with only single anchor, utilizing multi-hop connectivity allows to obtain optimal prior knowledge. When choosing anchor positions, it is essential to consider covering the maximum area with the minimum number of anchors to reduce deployment costs. Multi-hop distances enable targets to obtain rough estimates of anchor distances. The number of anchors chosen often depends on the environment of localizations and characteristics of signal propagation. In our experiments we deployed anchors in a hexagonal shape to cover maximum area most efficiently. With this shape, the distances between neighboring anchors are uniform, enabling messages from anchors to effectively represent the network's geometry.

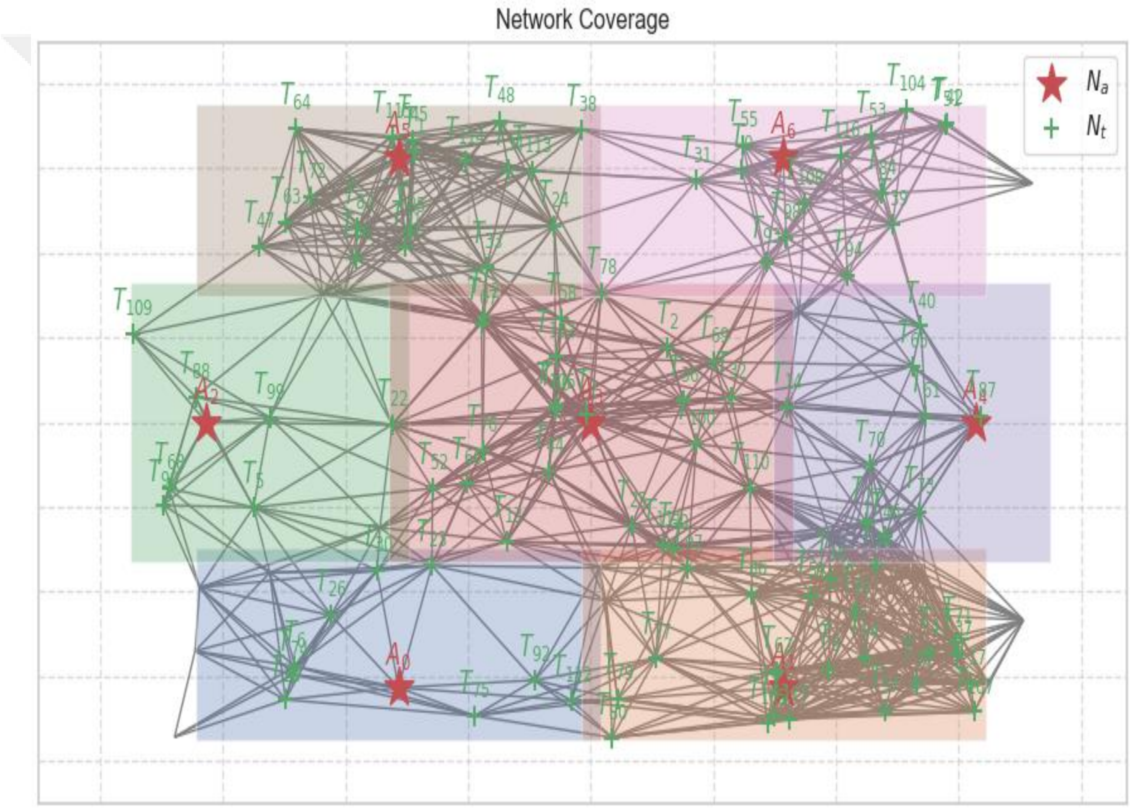
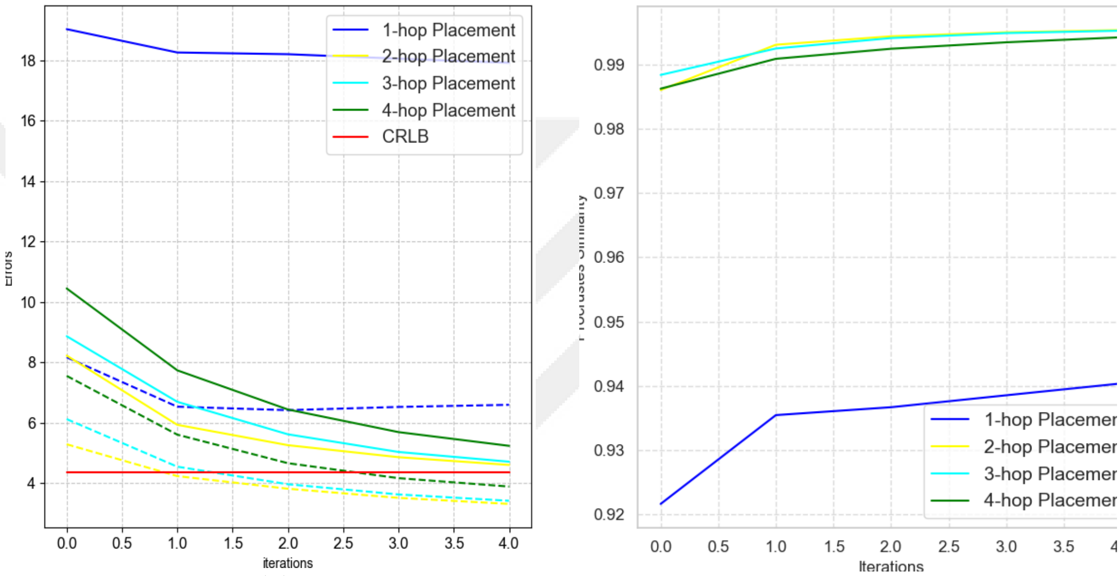


Figure 4.14. Anchor placement strategy

In this structure, given the connectivity ranges are plausible, some targets will be in the intersections of multiple anchors. In the first iteration of NBP, since there is no directionality information, each node's particles approximate messages that represent beliefs about target nodes will be spread around with random angles uniformly from $[0, 2\pi)$. Although, messages of anchors will be each node's particles from the known positions of anchors, due to no known direction's representation of opinions about the target's position estimates may have limited certainties. Therefore, following the second iteration, such targets with multiple immediate anchor neighbors will likely have low uncertainties about positions because starting from second iteration, angular information obtained from previous iteration may be utilized to focus each particle in a specific region for calculating messages for such targets.

Having obtained multiple messages from anchors, such targets are likely to be estimated close to their true positions. According to KL-divergence differences of consequent iterations, if such targets fall below a certain threshold they can also be used as pseudo anchors to provide prior knowledge. Although as discussed in the previous section, due to measurement noise, further limitations of the possible estimation region can adversely affect the estimates. In any case, obtaining accurate position estimations allows considering target in question as localized early in the iterations, leading to decrease in the number of outgoing messages needed to be approximated from neighbors of such targets, consequently, reducing the computational burden on the clique.



(a) (b)

Figure 4.15. a) Errors; **b)** similarities using both priors obtained from anchors and various n-hop connectivity

In Figure 4.15 results of various multi-hop connectivity with strategic anchor placements show the importance of utilizing localization improvement approaches together, significantly improving the estimates. Although collaboration among targets allows to utilize information effectively from wider network context, due to limitation of rough distance approximates over multi-hop neighbors can adversely affect localization performance. Using effective anchor placements, even with limited number of anchors eliminates the necessity for multi-hop distance approximates, therefore, along with improving accuracy of the estimates reduce the computational burden on each node. This results in 2-hop connectivity to surpass 3+hops for localization, and it is sufficient to approach CRLB limit with only limited number of iterations.

Considering the effect of anchor placements on network layout over multi-hop connectivity scenarios, we find that there is no significant improvement beyond 2-hop connectivity. This is due to connectivity radius r is sufficient to cover most of the network layout. Having observed the effect of anchor placement improvement under various multi-hop connectivity with sufficient network coverage in Figure 4.16, it also necessary to investigate the effect of poor network coverage with anchor placement strategy to draw conclusions about effectiveness of such approach.

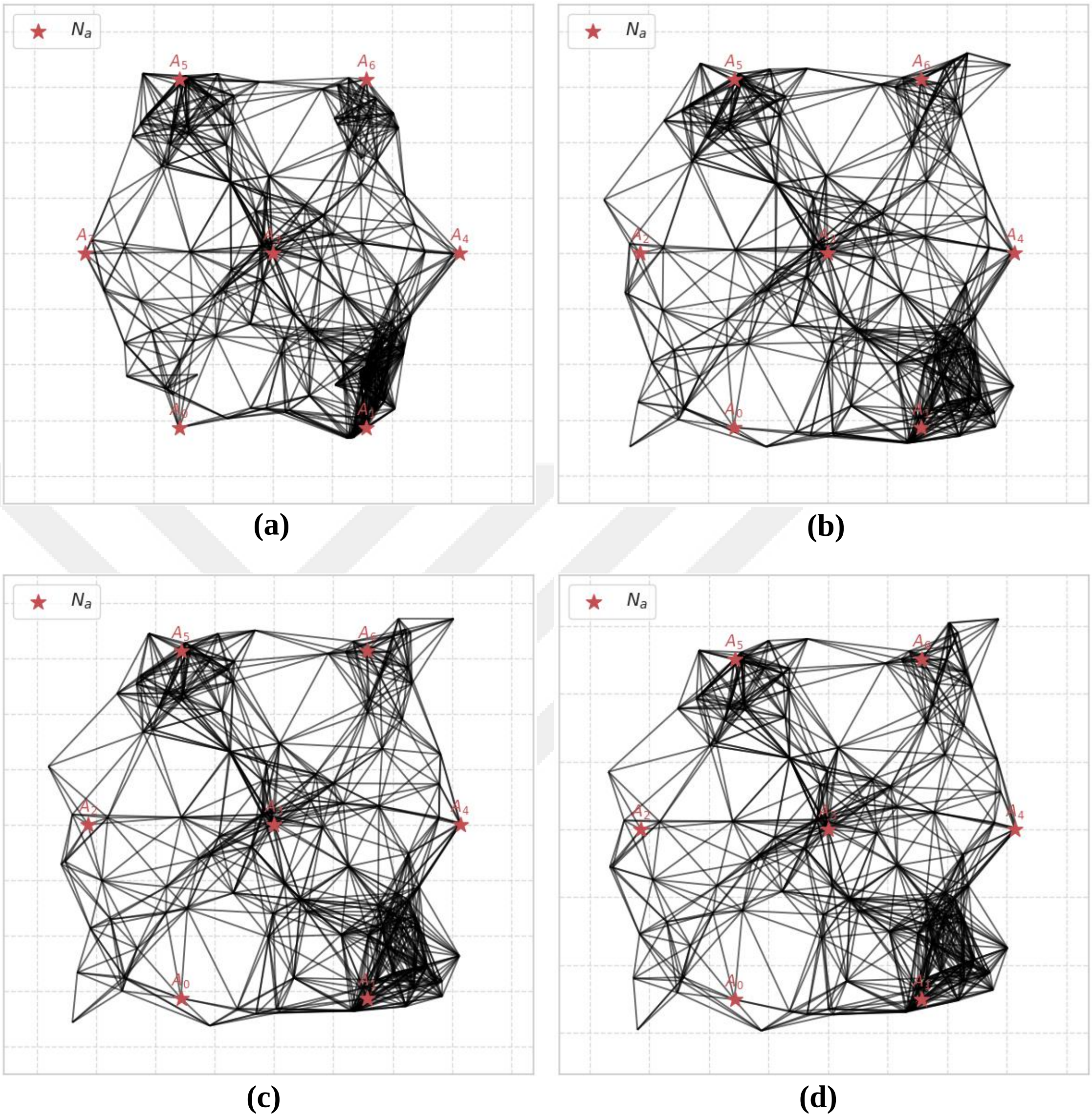
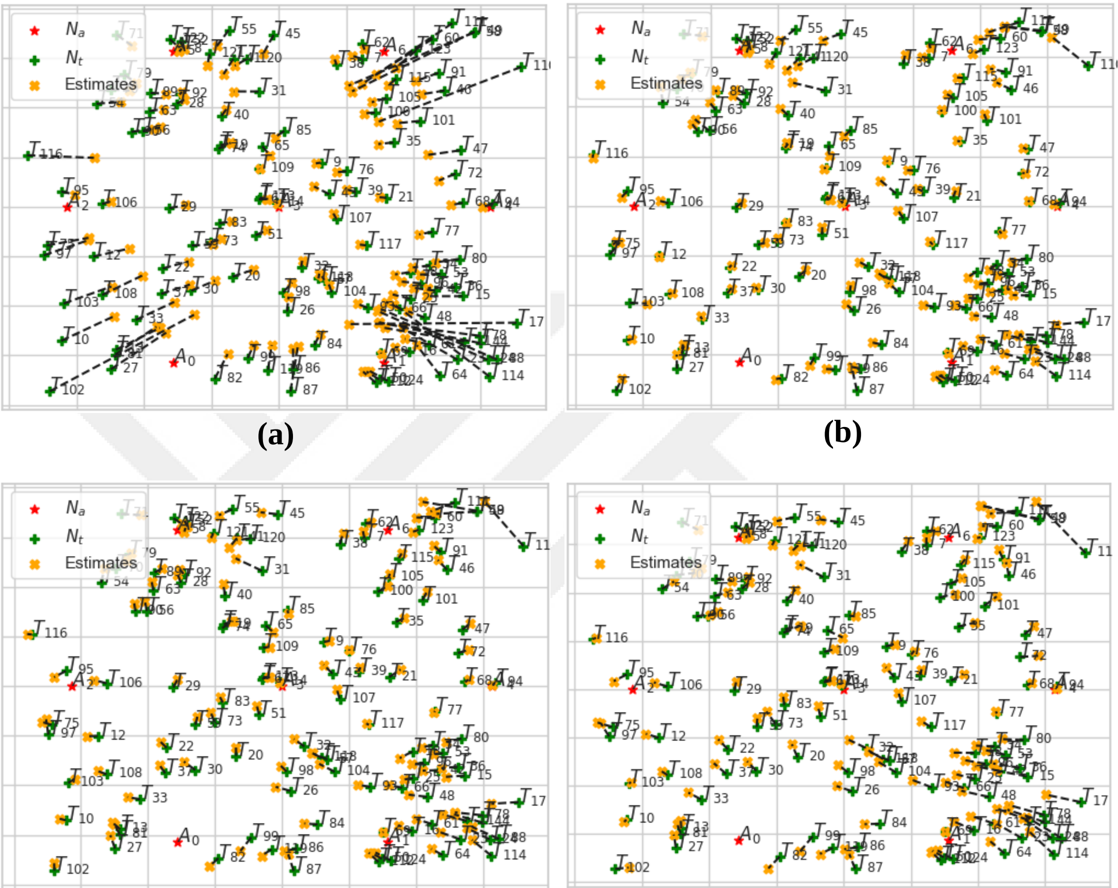


Figure 4.16. Network layouts with various n-hop connectivity using anchor placements; **a)** 1-hop; **b)** 2-hop; **c)** 3-hop; **d)** 4-hop

Network coverage of anchor nodes are shown in Figure 4.14. When given communication radius is enough to cover the network, priors obtained through 2-hop distance measurements from anchors provide sufficient knowledge about the position of targets. However, investigation of individual estimation errors in Figure 4.17 yields that priors obtained from large number of anchors for 2+hop connectivity restricts the estimation region, hindering the probabilistic nature of the NBP. This hindrance not only affects the belief update stage, it also affects message approximations leading to deviations about the belief about positions of neighboring targets.

From Figure 3.7, it is observed that true positions of nodes often lie near the edges of their bounded boxes. Given the inherent noise in sensor measurements, excessively restricting beliefs with multi-hop priors can cause the true positions to fall outside these bounded boxes due to the noise. This results in erroneous position estimates that cannot be easily corrected due to the overly restricted beliefs imposed by the priors. Consequently, the probabilistic flexibility of the NBP is compromised, leading to persistent inaccuracies in the localization process. Therefore, choosing the optimal number of anchors when there is insufficient communication radius for effective network coverage is important for network localization.



(a) **(b)** **(c)** **(d)**

Figure 4.17. Localization results with utilizing priors along with anchor placement strategy for various n-hop connectivity; **a)** 1-hop; **b)** 2-hop; **c)** 3-hop; **d)** 4-hop

Although there are ways to overcome excessive number of anchor placements, such as sequential update scheduling proposed by Ihler et al. this requires a single computation unit to decide the flow of computations, leading to a centralized approach and can hinder the responsiveness of the system. In the sequential update scheduling, each iteration of the NBP algorithm starts from targets with the greatest number of immediate anchors and based on the shortest distances to anchors targets are ordered. Since in this thesis, we consider fully distributed CL solutions, therefore, we consider the effect of anchor placement in terms of network coverage.

4.3. Discussion

In this section, discussions regarding the implications of the results presented in subsection 4.2. The focus is on analyzing the impact of multi-hop connectivity, prior knowledge, and anchor placement strategies on the accuracy and reliability of CL system. By examining these factors, we aim to derive insights that could guide future improvements and practical applications of the system.

The results clearly indicate that increasing the degree of multi-hop connectivity significantly enhances localization accuracy. As demonstrated, 2-hop connectivity substantially reduces both the errors and uncertainties compared to 1-hop connectivity. This improvement is due to the increased number of neighboring nodes contributing to the localization process, indicates that as the collaboration among nodes increases, both the accuracy of estimates and reliability of network layout increases as well.

However, it is also observed that beyond 2-hop connectivity, the improvements plateau and in some cases even degrade. This phenomenon is attributed to the compounding of errors over longer distances and the increased computational burden, which can introduce additional uncertainties. Therefore, while multi-hop connectivity is beneficial, its effectiveness is maximized according to the network coverage of the anchor's connectivity regions.

Incorporating prior knowledge, such as the positions of anchors, plays a crucial role in improving localization performance. The use of priors helps narrow down the search space for the nodes' positions, thus reducing the overall uncertainty and the number of particles needed for effective representation. This is evident from the lower RMSE and more accurate position estimates when prior knowledge is utilized compared to the scenarios without any prior information.

However, there are limitations associated with the use of priors. In scenarios with high measurement of noise or misaligned priors, the algorithm can be misled, resulting in erroneous estimates. This issue is particularly pronounced when the priors are overly restrictive, as seen in the cases with 3-hop and 4-hop connectivity, where flexibility of NBP is compromised.

The strategy for placing anchors significantly influences the localization accuracy and efficiency. The results show that strategic anchor placement, such as in hexagonal pattern, ensures uniform coverage and effective utilization of the network's geometry. This approach minimizes the deployment cost by reducing the number of anchors needed while maintaining high localization accuracy.

Furthermore, the results indicate that having at least one anchor within vicinity of a node significantly reduces ambiguities in the position estimates. Multi-hop distances allow nodes to derive rough estimates of anchor distances, which, when combined with efficient placement strategies, optimize the localization process. However, excessive reliance on multi-hop anchors can lead to overly restricted beliefs, adversely affecting the NBP's probabilistic nature and leading to persistent inaccuracies.

The findings highlight the importance of balancing multi-hop connectivity, prior knowledge, and strategic anchor placement to optimize CL system's performance.

5. CONCLUSION

Localization is vital in our interconnected world for providing personalized services and insights into various operational environments. CL aims to overcome the limitations of TLTs by enabling targets to share information with neighboring nodes and utilize known reference points. This collaborative aspect enhances the accuracy, robustness and reliability of positioning.

In this study, we have leveraged several methodologies to achieve a robust and efficient CL system. We adopted a probabilistic approach, specifically utilizing NBP with PF (Ihler et al. 2005), to manage uncertainties and work with incomplete or flawed information. This method allows for a distributed solution that is easily scalable by assuming conditional independence among neighboring agents. The graphical model approach we employed is based on MRF, which represents the CL problem by modeling the probabilistic relationships among variables. In this model, prior knowledge about the positions of nodes is incorporated as unary potentials, while pairwise potential functions represent the likelihood of states and the probability of connectivity based on observed distances. To enhance the efficiency and quality of sampling, we introduced constraints on the sampling area by building bounding boxes for agents (Savic and Zazo 2009). This approach helps improve sampling performance and localization accuracy by focusing on the most likely areas.

Our results indicate that increasing multi-hop connectivity up to sufficient network coverage reduces localization errors and uncertainties by leveraging contributions from a greater number of neighboring nodes, effectively increasing collaboration between neighbors. Utilizing prior knowledge of anchor positions proved crucial in enhancing localization accuracy by narrowing the search space, though its effectiveness depends on the precision of the priors. Additionally, our investigation revealed that a strategic anchor placement ensures uniform network coverage and efficient use of the network’s geometry, minimizing deployment costs while maintaining high accuracy. Having at least one anchor near a node significantly reduces positional ambiguities, optimizing the localization process.

Throughout this study we have simulated the networks to achieve collaborative localization, as a future work comprehensive measurement campaign can be conducted in real-world environments to validate and further refine the methodologies, allowing for practical implementation of CL system and assessing robustness under different conditions and constraints. Another future consideration is the development of an optimized belief update schedule, which would prioritize target nodes with the highest number of nearby anchors, thereby enhancing the efficiency and accuracy of the localization process in sub iterations of the algorithm, effectively utilizing most confident messages from nearby targets.

6. REFERENCES

Ali, Muhammad Usman, Soojung Hur, and Yongwan Park. 2019. “Wi-Fi-Based Effortless Indoor Positioning System Using IoT Sensors.” *Sensors* 19(7):1496. doi: 10.3390/s19071496.

Bahr, Alexander, Matthew R. Walter, and John J. Leonard. 2009. “Consistent Cooperative Localization.” Pp. 3415–22 in *2009 IEEE International Conference on Robotics and Automation*.

Bai, Lu, Maurice D. Mulvenna, and Raymond R. Bond. 2023. “Indoor Localization Using Trilateration and Location Fingerprinting Methods.” Pp. 71–99 in *Machine Learning for Indoor Localization and Navigation*, edited by S. Tiku and S. Pasricha. Cham: Springer International Publishing.

Blumenthal, Jan, Ralf Grossmann, Frank Golatowski, and Dirk Timmermann. 2007. “Weighted Centroid Localization in Zigbee-Based Sensor Networks.” Pp. 1–6 in *2007 IEEE International Symposium on Intelligent Signal Processing*.

Buehrer, R. Michael, Henk Wymeersch, and Reza Monir Vaghefi. 2018. “Collaborative Sensor Network Localization: Algorithms and Practical Issues.” *Proceedings of the IEEE* 106(6):1089–1114. doi: 10.1109/JPROC.2018.2829439.

Carrillo-Arce, Luis C., Esha D. Nerurkar, José L. Gordillo, and Stergios I. Roumeliotis. 2013. “Decentralized Multi-Robot Cooperative Localization Using Covariance Intersection.” Pp. 1412–17 in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*.

Chen, Yuanpeng, Zhiqiang Yao, and Zheng Peng. 2020. “A Novel Method for Asynchronous Time-of-Arrival-Based Source Localization: Algorithms, Performance and Complexity.” *Sensors* 20(12):3466. doi: 10.3390/s20123466.

Darakeh, Fatemeh, Gholam-Reza Mohammad-Khani, and Paeiz Azmi. 2018. “CRWSNP: Cooperative Range-Free Wireless Sensor Network Positioning Algorithm.” *Wireless Networks* 24(8):2881–97. doi: 10.1007/s11276-017-1505-2.

Djuric, P. M., J. H. Kotecha, Jianqui Zhang, Yufei Huang, T. Ghirmai, M. F. Bugallo, and J. Miguez. 2003. “Particle Filtering.” *IEEE Signal Processing Magazine* 20(5):19–38. doi: 10.1109/MSP.2003.1236770.

Fan, Ying Sheng, Fei Meng, Guo Jun Li, and Xiao Gang Qi. 2023. “A Survey and Evaluation of Iterative Optimization Algorithms for Cooperative Localization.” *Physical Communication* 60:102165. doi: 10.1016/j.phycom.2023.102165.

Farahani, Shahin. 2008. “Chapter 7 - Location Estimation Methods.” Pp. 225–46 in *ZigBee Wireless Networks and Transceivers*, edited by S. Farahani. Burlington: Newnes.

Farid, Zahid, Rosdiadee Nordin, and Mahamod Ismail. 2013. “Recent Advances in Wireless Indoor Localization Techniques and System.” *Journal of Computer Networks and Communications* 2013:e185138. doi: 10.1155/2013/185138.

Fascista, Alessio, Giovanni Ciccarese, Angelo Coluccia, and Giuseppe Ricci. 2018. “Angle of Arrival-Based Cooperative Positioning for Smart Vehicles.” *IEEE*

Transactions on Intelligent Transportation Systems 19(9):2880–92. doi: 10.1109/TITS.2017.2769488.

Gholami, Mohammad Reza, Luba Tetruashvili, Erik G. Ström, and Yair Censor. 2013. "Cooperative Wireless Sensor Network Positioning via Implicit Convex Feasibility." *IEEE Transactions on Signal Processing* 61(23):5830–40. doi: 10.1109/TSP.2013.2279770.

Gostar, Amirali Khodadadian, Tharindu Rathnayake, Ruwan Tennakoon, Alireza Bab-Hadiashar, Giorgio Battistelli, Luigi Chisci, and Reza Hoseinnezhad. 2021. "Centralized Cooperative Sensor Fusion for Dynamic Sensor Network With Limited Field-of-View via Labeled Multi-Bernoulli Filter." *IEEE Transactions on Signal Processing* 69:878–91. doi: 10.1109/TSP.2020.3048595.

Howard, A., M. J. Mataric, and G. S. Sukhatme. 2003. "Putting the 'I' in 'Team': An Ego-Centric Approach to Cooperative Localization." Pp. 868–74 vol.1 in *2003 IEEE International Conference on Robotics and Automation (Cat. No.03CH37422)*. Vol. 1.

Ihler, A. T., J. W. Fisher, R. L. Moses, and A. S. Willsky. 2005. "Nonparametric Belief Propagation for Self-Localization of Sensor Networks." *IEEE Journal on Selected Areas in Communications* 23(4):809–19. doi: 10.1109/JSAC.2005.843548.

Ihler, and Alexander Thomas. 2005. "Inference in Sensor Networks : Graphical Models and Particle Methods."

Jin, Di, Feng Yin, Carsten Fritsche, Fredrik Gustafsson, and Abdelhak M. Zoubir. 2020. "Bayesian Cooperative Localization Using Received Signal Strength With Unknown Path Loss Exponent: Message Passing Approaches." *IEEE Transactions on Signal Processing* 68:1120–35. doi: 10.1109/TSP.2020.2969048.

Li, Jian, Jianmin Zhang, and Liu Xiande. 2009. "A Weighted DV-Hop Localization Scheme for Wireless Sensor Networks." Pp. 269–72 in *2009 International Conference on Scalable Computing and Communications; Eighth International Conference on Embedded Computing*. Dalian, China: IEEE.

Liu, Ran, Chau Yuen, Tri-Nhut Do, Dewei Jiao, Xiang Liu, and U.-Xuan Tan. 2017. "Cooperative Relative Positioning of Mobile Users by Fusing IMU Inertial and UWB Ranging Information." Pp. 5623–29 in *2017 IEEE International Conference on Robotics and Automation (ICRA)*.

Mao, Guoqiang, Brian D. O. Anderson, and Barış Fidan. 2007. "Path Loss Exponent Estimation for Wireless Sensor Network Localization." *Computer Networks* 51(10):2467–83. doi: 10.1016/j.comnet.2006.11.007.

Mensing, Christian, and Jimmy Jessen Nielsen. 2010. "Centralized Cooperative Positioning and Tracking with Realistic Communications Constraints." Pp. 215–23 in *Navigation and Communication 2010 7th Workshop on Positioning*.

Miao, Guowang, Jens Zander, Ki Won Sung, and Slimane Ben Slimane. 2016. *Fundamentals of Mobile Data Networks*. Cambridge: Cambridge University Press.

Moses, Randolph L., and Robert Patterson. 2002. "Self-Calibration of Sensor Networks." Pp. 108–19 in *Unattended Ground Sensor Technologies and Applications IV*. Vol. 4743. SPIE.

Munoz, David, Frantz Bouchereau Lara, Cesar Vargas, and Rogerio Enriquez-Caldera. 2009. *Position Location Techniques and Applications*. Academic Press.

Murphy, Kevin, Yair Weiss, and Michael I. Jordan. 2013. "Loopy Belief Propagation for Approximate Inference: An Empirical Study."

Nagaty, Amr, Carl Thibault, Michael Trentini, and Howard Li. 2015. "Probabilistic Cooperative Target Localization." *IEEE Transactions on Automation Science and Engineering* 12(3):786–94. doi: 10.1109/TASE.2015.2424865.

Patwari, N., J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal. 2005. "Locating the Nodes: Cooperative Localization in Wireless Sensor Networks." *IEEE Signal Processing Magazine* 22(4):54–69. doi: 10.1109/MSP.2005.1458287.

Patwari, N., R. J. O'Dea, and Yanwei Wang. 2001. "Relative Location in Wireless Networks." Pp. 1149–53 vol.2 in *IEEE VTS 53rd Vehicular Technology Conference, Spring 2001. Proceedings (Cat. No.01CH37202)*. Vol. 2.

Peng, Rong, and Mihail L. Sichitiu. 2006. "Angle of Arrival Localization for Wireless Sensor Networks." Pp. 374–82 in *2006 3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks*. Vol. 1.

Rathnayake, R. M. M. R., Madduma Wellalage Pasan Maduranga, Valmik Tilwari, and Maheshi B. Dissanayake. 2023. "RSSI and Machine Learning-Based Indoor Localization Systems for Smart Cities." *Eng* 4(2):1468–94. doi: 10.3390/eng4020085.

Sadoun, Balqies, and Omar Al-Bayari. 2007. "Location Based Services Using Geographical Information Systems." *Computer Communications* 30(16):3154–60. doi: 10.1016/j.comcom.2007.05.059.

Savic, Vladimir, and Santiago Zazo. 2009. "Nonparametric Boxed Belief Propagation for Localization in Wireless Sensor Networks." Pp. 520–25 in *2009 Third International Conference on Sensor Technologies and Applications*.

Savic, Vladimir, and Santiago Zazo. 2013. "Cooperative Localization in Mobile Networks Using Nonparametric Variants of Belief Propagation." *Ad Hoc Networks* 11(1):138–50. doi: 10.1016/j.adhoc.2012.04.012.

Shang, Yi, Wheeler Ruml, Ying Zhang, and Markus P. J. Fromherz. 2003. "Localization from Mere Connectivity." Pp. 201–12 in *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, *MobiHoc '03*. New York, NY, USA: Association for Computing Machinery.

Sudderth, E. B., A. T. Ihler, W. T. Freeman, and A. S. Willsky. 2003. "Nonparametric Belief Propagation." P. I–I in *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.* Vol. 1.

Vaghefi, Reza Monir, and R. Michael Buehrer. 2015. "Cooperative Localization in NLOS Environments Using Semidefinite Programming." *IEEE Communications Letters* 19(8):1382–85. doi: 10.1109/LCOMM.2015.2442580.

Vaghefi, Reza Monir, and R. Michael Buehrer. 2017. “Cooperative Source Node Tracking in Non-Line-of-Sight Environments.” *IEEE Transactions on Mobile Computing* 16(5):1287–99. doi: 10.1109/TMC.2016.2591540.

Vo, Quoc Duy, and Pradipta De. 2016. “A Survey of Fingerprint-Based Outdoor Localization.” *IEEE Communications Surveys & Tutorials* 18(1):491–506. doi: 10.1109/COMST.2015.2448632.

Wang, Yapeng, Xu Yang, Yutian Zhao, Yue Liu, and Laurie Cuthbert. 2013. “Bluetooth Positioning Using RSSI and Triangulation Methods.” Pp. 837–42 in *2013 IEEE 10th Consumer Communications and Networking Conference (CCNC)*.

Wen, Yutian, Xiaohua Tian, Xinbing Wang, and Songwu Lu. 2015. “Fundamental Limits of RSS Fingerprinting Based Indoor Localization.” Pp. 2479–87 in *2015 IEEE Conference on Computer Communications (INFOCOM)*. Kowloon, Hong Kong: IEEE.

Wymeersch, Henk, Jaime Lien, and Moe Z. Win. 2009. “Cooperative Localization in Wireless Networks.” *Proceedings of the IEEE* 97(2):427–50. doi: 10.1109/JPROC.2008.2008853.

Yang, Zheng, and Yunhao Liu. 2010. “Quality of Trilateration: Confidence-Based Iterative Localization.” *IEEE Transactions on Parallel and Distributed Systems* 21(5):631–40. doi: 10.1109/TPDS.2009.90.

Zafari, Faheem, Athanasios Gkelias, and Kin K. Leung. 2019. “A Survey of Indoor Localization Systems and Technologies.” *IEEE Communications Surveys & Tutorials* 21(3):2568–99. doi: 10.1109/COMST.2019.2911558.

Zaidi, Slim, Ahmad El Assaf, Sofiène Affes, and Nahi Kandil. 2016. “Accurate Range-Free Localization in Multi-Hop Wireless Sensor Networks.” *IEEE Transactions on Communications* 64(9):3886–3900. doi: 10.1109/TCOMM.2016.2590436.

Zhang, Ping, Along Cao, and Tao Liu. 2017. “Bound Analysis for Anchor Selection in Cooperative Localization.” Pp. 1–10 in *Industrial IoT Technologies and Applications*, edited by F. Chen and Y. Luo. Cham: Springer International Publishing.

Zhang, Tingting, Andreas F. Molisch, Yuan Shen, Qinyu Zhang, Hao Feng, and Moe Z. Win. 2016. “Joint Power and Bandwidth Allocation in Wireless Cooperative Localization Networks.” *IEEE Transactions on Wireless Communications* 15(10):6527–40. doi: 10.1109/TWC.2016.2580504.

CURRICULUM VITAE

Berk ERCİN

EDUCATION

Master of Science 2021-2024	Akdeniz University Institute of Natural and Applied Sciences, Department of Computer Engineering, Antalya
Bachelor of Science 2014-2020	Kastamonu University Faculty of Engineering and Architecture, Computer Engineering, Kastamonu, Türkiye

WORK EXPERIENCE

Research Assistant 2022-Present	Akdeniz University Faculty of Engineering, Computer Engineering, Antalya, Türkiye

PUBLICATIONS

1- Ercin, B. & Karacı, A. (2021, December). Real-Time Identification from Gait Features Using Cascade Voting Method. In Applied Computer Systems, vol.26, no.2, 2021, pp.164-172. https://doi.org/10.2478/acss-2021-0020