

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO CEARÁ
PRÓ-REITORIA DE PESQUISA, PÓS-GRADUAÇÃO E INOVAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Daniel Jean Rodrigues Vasconcelos

A Probabilistic Approach to the Multilateration Problem

Fortaleza – CE

2 de novembro de 2019

Daniel Jean Rodrigues Vasconcelos

A Probabilistic Approach to the Multilateration Problem

Dissertação submetida ao Programa de Pós-Graduação em Ciência da Computação do Instituto Federal de Educação, Ciência e Tecnologia do Ceará, como requisito parcial para obtenção do grau de Mestre em Ciência da Computação.

Linha de Pesquisa: Inteligência Artificial

Supervisor: Prof. Dr. Amauri H. Souza Junior

Co-supervisor: Prof. Dr. Francesco Corona

Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação (PRPI)

Programa de Pós-Graduação em Ciência da Computação (PPGCC)

Fortaleza – CE

2 de novembro de 2019



Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)
Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação (PRPI)
Programa de Pós-Graduação em Ciência da Computação (PPGCC)

Daniel Jean Rodrigues Vasconcelos

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Ciência da Computação, sendo aprovada pela Coordenação do Programa de Pós-Graduação em Ciência da Computação do Instituto Federal de Educação, Ciência e Tecnologia do Ceará e pela banca examinadora:

Supervisor: Prof. Dr. Amauri H. Souza Junior
Instituto Federal de Educação, Ciência e Tecnologia do
Ceará (IFCE)

Co-supervisor: Prof. Dr. Francesco Corona
Universidade Federal do Ceará (UFC)

Prof. Dr. Ajalmar Rêgo da Rocha Neto
Instituto Federal de Educação, Ciência e Tecnologia do
Ceará (IFCE)

Prof. Dr. César Lincon Cavalcante Mattos
Universidade Federal do Ceará (UFC)

Fortaleza – CE
2 de novembro de 2019

*Este trabalho é dedicado aos que,
direta ou indiretamente, o tornaram possível:
Deus, minha esposa, meus pais, professores e amigos.*

Acknowledgements

Special gratitude to the Giver of life, God, to Whom be all the glory and praise forever and ever.

Extraordinary gratitude to supervisor Amauri and co-supervisor Francesco, who provided helpful and relevant insights for this research.

Main acknowledgments to professors and colleagues at IFCE who supported and give valuable advice regarding the research and other areas, as well.

*“To raise new questions, new possibilities, to regard old problems from a new angle,
requires creative imagination and marks real advance in science.”*
(Albert Einstein)

Resumo

Esta dissertação, utilizando uma abordagem probabilística, tem por objetivo oferecer uma alternativa válida para resolver o problema de Multilateração. Esse problema específico consiste em localizar um ponto de interesse $\mathbf{q}$ dados pontos de referência e medições destes pontos para o ponto de interesse. Na literatura existem métodos determinísticos e probabilísticos que solucionam o problema supracitado. Em geral a abordagem determinística soluciona o problema de multilateração via minimização de erros residuais; métodos probabilísticos usam estimadores de verossimilhança, onde, em geral, os ruídos das distâncias observadas são normalmente distribuídos. Neste trabalho nós utilizamos uma completa abordagem probabilística em termos das incertezas que podem ser atribuídas ao ponto de interesse e às distâncias medidas. Nesse sentido apresentamos um método que possui como saída a distribuição posterior da variável latente $\mathbf{q}$ (o ponto objetivo) dado o conjunto de distâncias observadas tomadas entre os pontos de referência e o ponto de interesse. Para criar o método fazemos uso de distribuições específicas para matematicamente e corretamente modelar as distâncias observadas no seu contexto correto; a saber, medições exclusivamente positivas. Nós também avaliamos como estas distribuições se comportam em casos específicos para que pesquisadores possam escolher de acordo com a aplicação em questão. Existem casos onde estimações pontuais são necessárias para cumprir determinada tarefa. Para tais casos discutimos a utilização dos valores estatísticos de média e moda da distribuição resultante para a obtenção de estimativas pontuais. Validamos o trabalho apresentado através de testes em dados sintéticos e reais. Dados sintéticos são artificialmente gerados a partir da abordagem discutida neste trabalho. Os dados reais são oriundos de um problema de localização que usa estações Bluetooth para estimar a posição de um celular (o ponto de interesse, nesse caso). Em última análise nós mostramos que o método proposto é compatível, ou até melhor em alguns casos, que métodos clássicos. O resultado final desta pesquisa é o método, resultados e figuras que mostram como o mesmo se comporta. Nós também disponibilizamos o código-fonte para utilização por pesquisadores.

Keywords: Método probabilístico. Problema de Multilateração. Distribuições. Métodos de aproximação.

Abstract

This dissertation aims to offer a viable alternative for solving the Multilateration problem using a probabilistic approach. This specific problem consists in localizing a query point $\mathbf{q}$ given reference points and measurements from these points to the point of interest. There exist in the literature deterministic and probabilistic methods for solving the aforementioned problem. In general, the deterministic approach tries to solve the multilateration problem via minimization of the error residue; probabilistic methods use likelihood estimators where distance observations are assumed, in general, to have a normally distributed noise. In this research, however, we take the approach of fully describing the multilateration problem in terms of the uncertainty that can be ascribed to the target point and distance measurements. In that regard we present a method that outputs the posterior distribution of the latent variable $\mathbf{q}$ (the target point) given the set of random distance observations that are taken between reference points and the target point. To create the method we make use of specific distributions to mathematically and correctly model distance observations in the proper context; namely, exclusively positive distance observations. We also evaluate how these distributions are expected to behave in specific cases, so researchers in general have options to choose from as their application demands. There are cases where point estimations are required to fulfill a given task. For such cases we also discuss using statistical values of mean and mode of the resulting posterior distribution to provide point estimations of the target point. We validate the present work by testing it against synthetic and real-world datasets. The synthetic dataset is artificially generated following the reasoning presented throughout the research; whereas, the real-world dataset comes from a localization problem that uses Bluetooth beacons to estimate the position of a cellphone (the target point, in this case). In the final analysis we showed that the proposed method performs on par, or even better in some cases, than other classical solutions. The final outcome of this research is, in addition to the method, results and plots originated from the method. We also make available the source code for researchers to use.

Keywords: Probabilistic method. Multilateration problem. Bayesian approach. Positive supported distributions. Approximation methods.

List of Figures

Figure 1 – Locating the target point $\mathbf{q}$ given reference points and distance measurements.	27
Figure 2 – multilateration: Geometric perspective.	31
Figure 3 – Multilateration: Probabilistic perspective.	31
Figure 4 – The Normal distribution.	38
Figure 5 – The Uniform distribution.	39
Figure 6 – The Gamma distribution.	40
Figure 7 – Multilateration: Probabilistic graphical model.	48
Figure 8 – The χ^2 distribution behavior.	51
Figure 9 – Gamma distribution for $\nu = 1$	52
Figure 10 – Gamma distribution for $\nu = 1$ and $t = 3.2$	53
Figure 11 – Probabilistic Multilateration: Left column, $\log p(\mathbf{q})$; Middle column, $\log \sum_{k=1}^K p(\delta^{(k)} \mathbf{q})$, and; Right column, $\log [p(\mathbf{q}) \prod_{k=1}^K p(\delta^{(k)} \mathbf{q})]$, for $D = 2$ and $K \in \{2, 4, 16, 32\}$	60
Figure 12 – Room where dataset was generated.	61

List of Tables

Table 1	– Parameters for the Gamma distribution and their variations	53
Table 2	– Probabilistic Multilateration Method Results.	58
Table 3	– Sample table of measurements from beacons to each location	62
Table 4	– Table of errors from estimates to the target point $\mathbf{q}^{(i)}$	62
Table 5	– Comparison between results from related works and the proposed method	63
Table 6	– Parameters for the Gamma distribution and their variations	73
Table 7	– Gradient for optimization functions with $\delta^{(k)}$ centered at the mean. . . .	77

List of Algorithms

1	Metropolis-Hastings	43
2	Proposed Method	55

List of abbreviations and acronyms

BC	Before Christ
CDF	Cumulative Density Function
IS	Importance Sampling
MAP	Maximum A Posteriori
MH	Metropolis-Hastings
MLM	Minimal Learning Machine
PDF	Probability Density Function
PMF	Probabilistic Mass Function
RV	Random Variable

List of symbols

$p(x)$	Probability associated with assignment $X = x$
$\mathbf{q}$	Query point, point of interest, position of target point
$\{\delta^{(k)}\}_{k=1}^K$	Set of distance observations
$\{\mathbf{r}^{(k)}\}_{k=1}^K$	Set of reference points
2D	Two dimensions
3D	Three dimensions
X	Representative Random Variable
$X = x$	Assignment of Random Variable X
$\mathbf{X}$	Set of Random Variables X_i

Contents

1	INTRODUCTION	27
1.1	Motivation	29
1.2	Problem Description	30
1.3	General and Specific Goals	32
1.4	Structure of the Dissertation	32
2	BACKGROUND	35
2.1	Random Variables and their Characterization	35
2.1.1	Expectation, Variance	37
2.1.2	Mode and Kurtosis	37
2.2	Some Probability Density Functions	38
2.2.1	Normal Distribution	38
2.2.2	Uniform Distribution	39
2.2.3	Gamma Distribution	39
2.3	Bayesian Reasoning	41
2.4	Maximum A Posteriori	42
2.5	Bayes Estimator via Sampling Methods	42
2.5.1	Importance Sampling Method	42
2.5.2	Metropolis-Hastings Method	43
2.5.3	Numeric Integration for Evidence Computation	44
2.6	Closing Remarks	44
3	PROBABILISTIC MULTILATERATION	47
3.1	Prior Distribution	48
3.2	Likelihood Function	49
3.3	Inference	54
3.3.1	Computation of MAP Estimates	54
3.3.2	Computation of Bayes estimate	55
3.4	Closing Remarks	55
4	EXPERIMENTS	57
4.1	Validating the Proposed Method	57
4.2	Validation via real-world application Dataset	59
5	CONCLUSION	65

BIBLIOGRAPHY 67

APPENDIX 71

**APPENDIX A – COMPLETE LIST FOR GAMMA DISTRIBUTION
PARAMETERS 73**

ANNEX 75

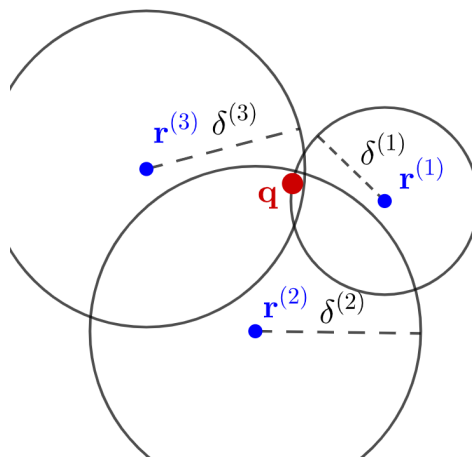
**ANNEX A – PARTIAL DERIVATIVES FOR THE GAMMA DIS-
TRIBUTION 77**

A.1 Gradient for measurements centered at the mean 77

1 Introduction

In this research we study the well-known problem of inferring the position of a target point in an unbounded D -dimensional Euclidean space, starting from observations of its distance from a set of other points whose locations are assumed to be fixed and known. In Figure 1 we show a system where distances $\{\delta^{(1)}, \delta^{(2)}, \delta^{(3)}\}$ are taken from the set of fixed reference points $\{\mathbf{r}^{(1)}, \mathbf{r}^{(2)}, \mathbf{r}^{(3)}\}$.

Figure 1 – Locating the target point $\mathbf{q}$ given reference points and distance measurements.



Source: Author.

The task is recurrent in many engineering areas ranging from localization in self-configuring wireless networks (PINEDA-BRISE et al., 2015) and airport surveillance (MANTILLA-GAVIRIA et al., 2015) to locating a node in sensor networks (SAVVIDES; PARK; SRIVASTAVA, 2003) or locating an aircraft from antenna measurements (NAGANAWA; MIYAZAKI; TAJIMA, 2017). Nonetheless, they all share the same interest.

Many mathematical techniques for calculating the position of a point of interest have been developed. One particular well-known method is Triangulation, which dates to the 6th century BC. Such an approach relies on receiving measurements of length and angle between the point of interest and two reference points. This schema forms a triangle that can be easily solved for finding the point of interest (STAFF, 2014). Although primitive, Triangulation is still used today in many applications, ranging from rocketry and gun direction of weapons to surveying and navigation (VOGEL, 2002).

Another familiar technique is Trilateration, which finds the position of the desired point using the measured distances between it and reference points. The optimal solution is found using geometry analysis of circles, spheres or triangles (STAFF, 2014). One of

the key differences to Triangulation is that Trilateration does not need measurements of angles to find the point of interest. Similarly, Trilateration has been applied to navigation and surveying, including global positioning systems (GPS) (BAJAJ; RANAWEERA; AGRAWAL, 2002).

A equally well-known method is multilateration, which gives the location of the point of interest using the distance¹ between known reference points. multilateration methods are found in Wide Area multilateration environments (HE, 2017), location of a set of nodes in a mobile ad hoc network (PINEDA-BRISE et al., 2015) and in many other applications.

This technique differs from Trilateration by generalizing for D -dimensional systems. In addition, the original formulation of the multilateration technique made use of hyperbolic curves instead of circles and spheres as in the Trilateration technique (SMITH; HULSTROM, 2005). One strong characteristic of multilateration is that estimation accuracy highly dependent on having multiple measurements from points of reference.

Formally, a possible approach to solve the multilateration problem is

$$\hat{\mathbf{q}} = \arg \min_{\mathbf{q}} \sum_k \left(\|\mathbf{q} - \mathbf{r}^{(k)}\|_2 - \delta^{(k)} \right)^2, \quad (1.1)$$

where $\|\cdot\|_2$ is the Euclidean vector norm and denotes the computed distance from $\mathbf{q}$ to $\mathbf{r}^{(k)}$ and $\delta^{(k)}$ is the observed distance measurement that might come from devices such as antennas, beacons or sensors. In Eq. (1.1) we show the minimization of a non-linear Least-Squares (LS) problem and due to its nonlinear nature, a closed form for it is only found in very specific scenarios. Such cases are largely discussed in Zekavat e Buehrer (2011), Abel e Smith (1987) and Beck, Stoica e Li (2008). In addition, there are stochastic methods to approximately solve this problem, which are discussed by Yao et al. (1998), Fang (1990), Norrdine (2012), Chan e Ho (1994) and Stefański (2009).

Nonetheless, these methods have confidence in receiving error-free, or known error measurements. Naturally, if distance observations without errors, fluctuations or contaminated data were given, locating the query point would be trivial, as described by Norrdine (2012). However, in real world applications, one might find useful to have a probabilistic framework to account for the uncertainty on distance measurements, which are directly related to computing the point of interest. Moreover, a probabilistic framework makes easier updating the previous result upon arrival of new distance measurements, which means also, less use of memory resources.

In the literature we found a vast corpus of multilateration algorithms to solve the localization problem for many applications. In order to be acquainted with multilateration

¹ There exists other types of measurement used in the literature but this research focuses on measurements related to distances.

methods and with this specific type of localization problem we refer the reader to Mantilla-Gaviria et al. (2015), which classifies multilateration methods in two main categories, namely: open-form and closed-form. In open-form algorithms we have the probabilistic approach; as for the closed-form methods we encounter numeric and algebraic solutions. Another interesting work is the thesis by Gaviria (2013), which provides formal definitions for multilateration systems. He also covers many algorithms and techniques for such systems. In both works the authors make considerations on possible applications for multilateration systems and provide an overview on techniques for signal processing for computing distance observations between reference points and the point of interest.

1.1 Motivation

Due to noise and uncertainty in general on observed distance measurements there has been some research on solving the multilateration problem using a probabilistic schema. PINEDA-BRISE et al. (2015) describe a method that uses this probabilistic approach for location estimation in mobile ad hoc network. They use distance estimates from the node of interest to three or more references. However, they assume independence between distance estimates. To assess their method they make use of five protocols, two of them being modified with this probabilistic schema. Their results were promising when dealing with 2D problems. For 3D situations, however, only one of the standard algorithms obtained good results.

Ramadurai e Sichertiu (2003) describes a position estimation algorithm for wireless sensor network in the presence of range measurement inaccuracy using a probabilistic foundation. They considered a gaussian error in the distance measurements, this assumption was then used to compute the probability that a unknown node would be in a given position. Comparing to Pineda's method their work presented better results because it was used as an indoors localization solution, which is expected to have less error on the distances than in outdoor localization. For more related solutions, we refer to Helen et al. (2001), Bahl, Padmanabhan e Balachandran (2000), Paschalidis, Li e Guo (2009) and Helen et al. (2001).

As aforementioned, probabilistic multilateration can be used to solve navigation problems, which includes also aeronautical surveillance (GALATI et al., 2005). Naganawa, Miyazaki e Tajima (2017) present a probability-of-detection modeling. This is attained by connecting together in the same model the main factors that affect probability of detection, such as radio propagation, interference, receiver, etc. Their work tries to model the error in measurements based on the source from receivers and emitters. They use filtering models to develop their research. The modeling framework was validated with evaluations and implementations, which assist them to derive valuable insights and future opportunities for

improvement. For more related works, we refer to Gaviria (2013) and Smith e Hulstrom (2005).

In the main works cited above the authors used maximum likelihood estimators to find an estimation for the target point and use the normal distribution to model uncertainty on the distance observations.

Finally, probabilistic solutions for multilateration problems have been extensively studied. However, to the extent of our knowledge, none of the works used the probabilistic reasoning as presented in this research; nor did they assume that error measurements could also occur from other distribution that is not gaussian. One of the motivations for this research is to tackle the former by presenting a method using the inverse probability principle to model the posterior distribution over $\mathbf{q}$, which in turn is used to offer a possible solution via point estimation. As for the latter, this method is intended to work with many distributions not to model the error of measurements but the distances themselves. Thus, the result of this research can be applied in multilateration problems that go beyond 2D or 3D situations.

1.2 Problem Description

In this work, we focus on the variant of the multilateration problem in which the geometry of the space is Cartesian and the position of the reference points is assumed to be fixed and known exactly. We let $\mathbf{q} = [q_1, q_2, \dots, q_D]^T \in \mathbb{R}^D$ be the target point of unknown coordinates and we let $\{\mathbf{r}^{(k)}\}_{k=1}^K$ with also $\mathbf{r}^{(k)} \in \mathbb{R}^D$ indicate the set of K reference points of known coordinates $\{r_d^{(k)}\}$, again with $d = 1, \dots, D$. We are given a set of distance observations $\mathcal{D} = \{\delta^{(k)}\}_{k=1}^K$ between the target $\mathbf{q}$ and each reference point $\mathbf{r}^{(k)}$. The goal is to estimate the position $\mathbf{q}$. Distances are assumed to be Euclidean, so that $\delta^{(k)} \approx \|\mathbf{q} - \mathbf{r}^{(k)}\|_2$. Geometrically, multilateration is formulated as the problem of locating the target point somewhere at the intersection of K D -dimensional balls centred at the reference points and radius $\delta^{(k)}$, under the assumption that $K \geq D + 1$ distances are observed and that such an intersection exists. Had distance measurements been exact, such an intersection would always exist and would unambiguously reduce to a single point, at $\mathbf{q}$. In Figure 2 geometric multilateration is illustrated for the case $D = 2$ and $K = 3$, firstly when distance measurements are inexact, Figure 2(a), and then when they are exact, Figure 2(b).

From a probabilistic perspective, the target point $\mathbf{q}$ is an unobserved random D -vector that we want to infer, starting from some prior belief $p(\mathbf{q})$ on its location and a set $\mathcal{D}$ of distance observations from the reference points. As the position of the reference points is fixed and known, we let each distance $\delta^{(k)}$ to the target to be a non-negative random variable with some specified probability density distribution $p(\delta^{(k)}|\mathbf{q})$, so that

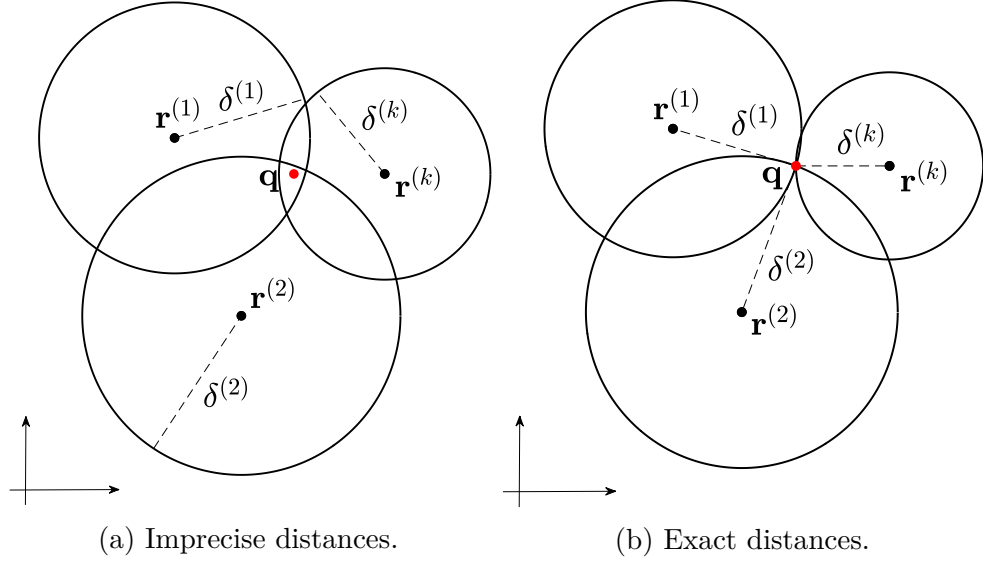


Figure 2 – multilateration: Geometric perspective.

the parametric form of the density function is tied to the latent location $\mathbf{q}$. Moreover, we assume that each distance is observed independently of the other ones, so that the likelihood factorizes as $p(\mathcal{D}|\mathbf{q}) = p(\delta^{(1)}, \delta^{(2)}, \dots, \delta^{(K)}|\mathbf{q}) = \prod_{k=1}^K p(\delta^{(k)}|\mathbf{q})$. The probabilistic model for this form of multilateration is thus given by the joint probability density function $p(\mathcal{D}, \mathbf{q})$ over latent and observed variables. Inference, as plain application of the inverse probability principle, thus corresponds to determining the posterior probability density function of the target given the distances; that is, $p(\mathbf{q}|\mathcal{D}) = p(\mathbf{q})p(\mathcal{D}|\mathbf{q})/p(\mathcal{D})$. As the marginal probability density of the data, or marginal likelihood, $p(\mathcal{D}) = \int_{\mathbf{q}} p(\mathbf{q})p(\mathcal{D}|\mathbf{q})$, does not depend on $\mathbf{q}$, it may be treated as a normalization constant. In Figure 3, probabilistic multilateration is illustrated again for $D = 2$ and $K = 3$. A prior density on the target position is stipulated before any data is observed and, then, the prior is coupled to the likelihood function $p(\delta^{(1)}, \delta^{(2)}, \delta^{(3)}|\mathbf{q}) = \prod_{k=1}^3 p(\delta^{(k)}|\mathbf{q})$ from the three observations. The resulting posterior $p(\mathbf{q}|\delta^{(1)}, \delta^{(2)}, \delta^{(3)})$ is the density of $\mathbf{q}$, given the data.

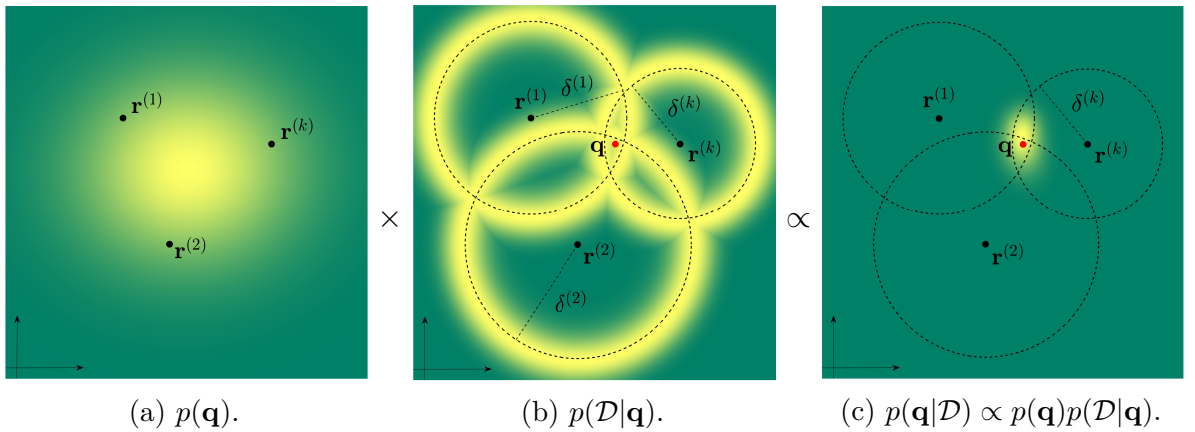


Figure 3 – Multilateration: Probabilistic perspective.

1.3 General and Specific Goals

This research focuses on postulating a novel probabilistic method to compute the target position $\mathbf{q}$ given reference points and distances. That is accomplished by modeling the posterior distribution over possible positions of $\mathbf{q}$. Point estimates (mode and mean) are computed as candidate solutions for the multilateration problem. In addition, it is paramount to test and validate the results of this method in artificial and real-world datasets to evaluate its performance.

The specific goals are:

- to propose a method that gives the distribution over possible values of $\mathbf{q}$;
- to develop a model that uses mean and mode of $p(\mathbf{q})$ as possible solutions for the multilateration problem;
- to construct a model that accounts for distances that are assumed to be distributed following a probabilistic function;
- to validate the model with artificial datasets;
- to validate the model with a real-world problem dataset.

1.4 Structure of the Dissertation

For a better flowing and understanding of this research, this document is divided into five chapters: Introduction (Chapter 1), Background (Chapter 2), Probabilistic Multilateration (Chapter 3), Experiments (Chapter 4) and Conclusion (Chapter 5). In addition, the Appendix section holds secondary topics, tables and plots.

Following the introductory chapter, the theory used to achieve the proposed method is expounded in the Background Chapter. It reviews, amongst other topics, probabilistic distributions. In addition, the chapter shows important concepts on sampling methods that are necessary for computing the evidence, mean and mode of posterior distributions. Moreover, model selection is described as means of choosing the best distribution for explaining the input data. In summary, this chapter substantiates and give support for the development of a valid and reliable probabilistic method for the multilateration problem.

The third chapter contains the main aspects of the proposed method that solves multilateration problem. The chapter describes the probabilistic framework for each model to find the point of interest. It also shows how parameters for each distribution are computed and how the method gives its estimation for the point of interest.

The results are shown in the fourth chapter (Experiments). This particular chapter sheds light on the two types of dataset used to validate the method: Artificial and real-world applications. The Artificial dataset is created to verify how the method will behave in a controlled setting. The second type of dataset comes from real-world measurements to

gauge the method's results. Finally, in this chapter it is also presented tables of results and figures for both types of datasets.

The last chapter (Conclusion), finishes this document with important remarks about the method and briefly discusses challenges and solutions encountered throughout this research.

2 Background

At the core of this work is the notion of probability to explain data. More specifically, the distance measurements between the reference points and the point of interest are assumed to follow a probabilistic distribution. This premise accounts for uncertainty or even errors in measurements or calibration of instruments. This line of thought allow methods in general to make good decisions even with incomplete or ambiguous information; in other words, probability theory grants a consistent structure for quantifying and manipulating uncertainty in pattern recognition (LEHMANN; CASELLA, 2006).

This chapter describes important topics that underlies the conception of the proposed method to solve the Multilateration problem using a probabilistic framework. This chapter encompasses probabilistic distributions, Bayesian reasoning, numeric computations for finding evidence, mean and mode of distributions.

2.1 Random Variables and their Characterization

A Random Variable (RV) is a variable that can take on different values randomly. A RV is often linked with a probability distribution to specify how likely each possible state of the RV is; otherwise, it would only specify the states that are possible, but not quantify their likelihood of happening (GOODFELLOW; BENGIO; COURVILLE, 2016).

There are two major conditions of existence that govern probability. Given any Random Variable X ,

$$\forall x \in \text{dom}(X), p(x) \geq 0 \text{ and,} \quad (2.1)$$

$$\int_{\text{dom}(X)} p(x) dx = 1. \quad (2.2)$$

In addition to Eq. (2.1) and (2.2), the domain of function $p(\cdot)$ must be all possible values of X . Notice, that these conditions refer to univariate functions and univariate RVs. The assignment $X \sim p(x)$ represents that X follows the distribution given by $p(x)$. That is, the RV X is characterized by the Probability Density Function (PDF) $p(x)$. By contrast, discrete RVs have *Probability Mass Functions* (PMFs) to characterize their assigned values.

From Eq. (2.1) and (2.2) it can be derived other rules and concepts, as discussed below. For the formal proofs of these concepts and rules the more interested reader should refer to Ramsey (2016).

The first concept is *joint* probability. Given RVs X and Y , $p(x, y)$ is the probability density function that describes the probability of each RV falling in a particular range of

values specified for those variables. When a joint probability is given and the interest is on only one of them, the *marginal* density can be obtained by integration, as

$$p(x) = \int_{\text{dom}(Y)} p(x, y) dy. \quad (2.3)$$

In words, the RV Y is marginalized out yielding only $p(x)$.

Another concept that often appears in problems involving probability is called *conditional* probability. It accounts for the probability of some event happening given that another event has happened. That is, $p(y|x)$ is the conditional probability density function of Y given that $X = x$, where x represents a given interval. This PDF is computed as

$$p(y|x) = \frac{p(x, y)}{p(x)}, \quad (2.4)$$

provided that $p(x) > 0$.

The two concepts of joint and conditional probability are useful tools to assess whether two RVs X and Y are *independent*, *conditionally independent*, or not. More specifically, X and Y are said to be independent if they can be expressed as

$$p(x, y) = p(x)p(y), \forall x, y. \quad (2.5)$$

Moreover, they are considered to be conditionally independent given a third RV Z if

$$p(x, y | z) = p(x | z)p(y | z), \forall x, y, z \quad (2.6)$$

holds. Notice that for both Eq. (2.5) and (2.6) it is required that independence conditions hold for all possible values of the RV.

It is natural to think that problems using a probabilistic framework present themselves with more than two RVs. In fact, given n RVs X , their joint probability can be factorized as

$$p(x^{(1)}, \dots, x^{(n)}) = \prod_{i=1}^n p(x^{(i)} | x^{(1)}, \dots, x^{(i-1)}). \quad (2.7)$$

This is known as the *chain rule* or the *product rule*. For the sake of illustration, applying Eq. (2.7) on three RVs yields

$$\begin{aligned} p(x, y, z) &= p(x | y, z)p(y, z) \\ &= p(x | y, z)p(y | z)p(z), \end{aligned}$$

where $p(y, z) = p(y|z)p(z)$; in addition, this is only one of the possible factorizations. This result is advantageous since it opens up possibilities for solving the same problem in different ways; in another words, all one needs to do is to find the factorization that best suits the data at hand.

2.1.1 Expectation, Variance

It is well known that when an arbitrary RV X follows a distribution $p(x)$, the function $p(\cdot)$ is able to describe the probability of all possible values of X . In some cases, however, one might be interested in a single value that expresses the *expected*¹ value of X . As such, the expected value of X is computed by

$$E[X] = \int x \cdot p(x) dx. \quad (2.8)$$

For cases when it is necessary a measure of how much the values of the RV X vary from $E[X]$, it is used

$$\text{Var}[X] = E[X^2] - E[X]^2, \quad (2.9)$$

where $\text{Var}[X]$ is the variance around the mean, $E[X]$ is the first moment² of the distribution and $E[X^2]$ is the second moment. From Eq. (2.9) it is inferred that when $\text{Var}[X]$ is low, then the values of X are near $E[X]$.

2.1.2 Mode and Kurtosis

The mode is a measure of central tendency that corresponds to the most likely value occurring of the RV X . When the random variable is discrete, the mode is any value of x with the highest probability, which, may not be unique. On the other hand, for continuous random variables the mode is the value x for which the probability density function (measuring the concentration of x values) has a maximum (HOBSON, 2002). For a unimodal $p(x)$ we compute the mode by

$$\text{Mode}[X] = \arg \max p(x). \quad (2.10)$$

Yet another measure that characterizes shape of distributions is the Kurtosis. Conceptually, the kurtosis is used to compare the shape of a given probability density function against the normal distribution (WESTFALL, 2014); in other words, it is as a measure of departure from normality. Kurtosis is then defined as the standardized fourth population moment about the mean

$$\text{Kur}[X] = \frac{E[X - \mu]^4}{(E[X - \mu]^2)^2} = \frac{\mu^4}{\sigma^4}, \quad (2.11)$$

where $E[\cdot]$ is the expectation operator, μ is the mean, μ^4 is the fourth moment of the distribution, and σ is the standard deviation (DECARLO, 1997). Since the normal

¹ Often times also called the average or mean of X .

² In probability theory and statistics the moment-generating function of a RV X is yet another alternative to specify its probability distribution. The various moments of a distribution are an important class of expectations (CASELLA; BERGER, 2002).

distribution has $\text{Kur}[X] = 3$, it is possible to characterize shape of other distributions comparing their kurtosis to this value. The comparison presupposes the same variance on both distributions; otherwise, it will not simply reflect a difference in kurtosis (shape) but will also reflect the difference in variance, which might lead to false conclusions.

2.2 Some Probability Density Functions

Probability theory is applied in many areas of research. Amongst other reasons, this is possible due to probabilistic functions that can properly represent the data available. Some distributions that are particularly useful in this research are shown below.

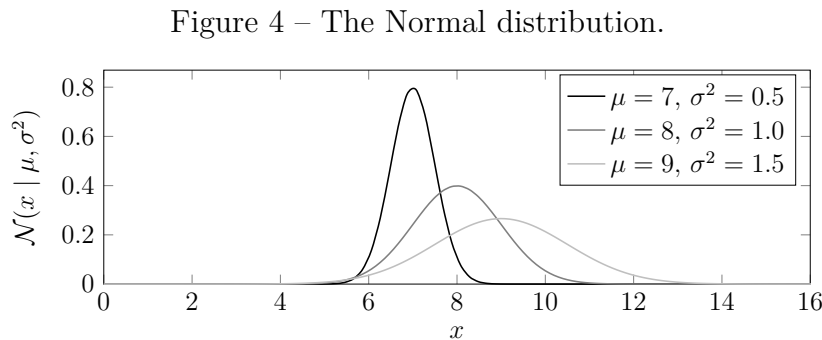
2.2.1 Normal Distribution

It was in the early nineteenth century that Laplace originally developed tools for large sample theory and proved the first general central limit theorem. In simple terms, the theorem shows how sums of random variables tend to behave like standard normal variables, as the sample size become large and it is standardized to have mean zero and unit variance (PATEL; READ, 1996).

According to Casella e Berger (2002), the *Normal* distribution or *Gaussian* distribution, is the most common distribution used over the real numbers, specially when the problem at hand gives no prior knowledge of its distribution. As such, it seems plausible to include this distribution in this research. Given that a RV X follows a normal distribution, its PDF is obtained by

$$\mathcal{N}(x | \mu, \sigma^2) = \sqrt{\frac{1}{2\pi\sigma^2}} \exp \left[-\frac{(x - \mu)^2}{2\sigma^2} \right], \quad (2.12)$$

with two parameters $\mu \in \mathbb{R}$ and $\sigma^2 \in \mathbb{R}^+$. In Figure 4 we show examples of densities when one varies these parameters.



Source: Author.

As it can be observed from Figure 4, the normal distribution is symmetric; as such, for a RV x that follows the normal distribution, we have that $\text{Mode}[X] = \text{E}[X] = \mu$. The variance is $\text{Var}[X] = \sigma^2$ and the shape is characterized by $\text{Kur}[X] = 3$.

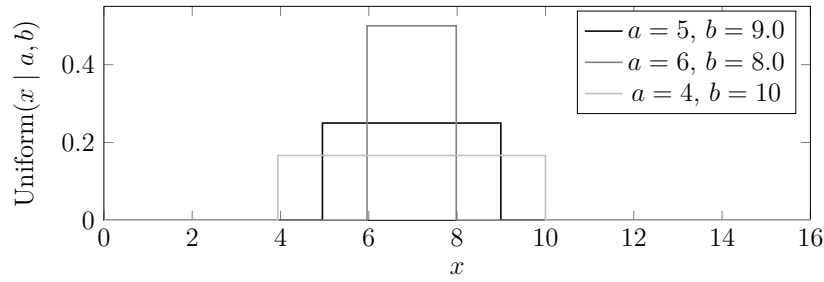
2.2.2 Uniform Distribution

Similarly, the *Uniform distribution* has two limiting parameters $-\infty < a < b < \infty$, which yields the support $x \in [a, b]$. The PDF of a uniformly distributed RV X is

$$\text{Uniform}(x \mid a, b) = \frac{1}{b - a}. \quad (2.13)$$

This probability function represent RVs that are symmetric and equally probable. Also, by choosing only values $0 \leq a < b < \infty$ we have parameters that make the uniform distribution to have a positive support. Figure 5 shows different parameters for this distribution.

Figure 5 – The Uniform distribution.



Source: Author.

Finally, summary values for this distribution are given by

$$\text{E}[X] = \frac{a + b}{2}, \quad (2.14)$$

$$\text{Var}[X] = \frac{(b - a)^2}{12} \text{ and} \quad (2.15)$$

$$\text{Kur}[X] = \frac{9}{5}. \quad (2.16)$$

Notice that any value along the line formed between a and b can be chosen as the mode of the distribution.

2.2.3 Gamma Distribution

The Gamma Distribution is a special case of the three parameter Generalized Gamma distribution, which can be used to generate many other distributions, such as Rayleigh and Maxwell-Boltzmann. The Generalized Gamma distribution has been used in

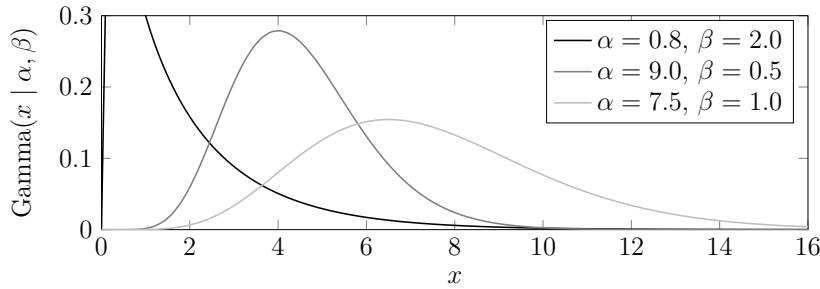
several research areas such as engineering, hydrology and survival analysis. (AYIENDA, 2014). In this research we focus on using the Gamma distribution, which is a continuous two-parameter function with shape parameter $\alpha > 0$, rate parameter $\beta > 0$ and $x \in (0, \infty)$. We represent a RV X as Gamma by

$$\text{Gamma}(x \mid \alpha, \beta) = \frac{x^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} \exp \left[-\frac{x}{\beta} \right]. \quad (2.17)$$

This distribution is used to model waiting times, lifespan testing and it is used in Poisson process³ in general (HUSAK; MICHAELSEN; FUNK, 2007), (HUSAK; MICHAELSEN; FUNK, 2007). As a result, this research finds it expedient to investigate how this distribution behaves given the proposed framework.

As showed in Figure 6 the two-parameter Gamma distribution can have many shapes depending on how those parameters are characterized (GUPTA; KUNDU, 2001). For instance, if $X \sim \text{Gamma}(x \mid \alpha = 1, \beta = 1/\lambda)$, then X follows the exponential distribution with rate parameter λ . On the other hand, if $X \sim \text{Gamma}(x \mid \alpha, \beta)$, then $1/X$ follows the Inverse-Gamma distribution. In summary, the Generalized Gamma, as well as the Gamma distribution can be used to represent many other functions, for example Maxwell-Boltzmann, Rayleigh, Nakagami, Erlang, etc. As a result, one can expect to represent a body of positive-supported distributions by simply using the Gamma distribution.

Figure 6 – The Gamma distribution.



Source: Author.

Finally, the summary values for the Gamma distribution are given by

$$E[X] = \alpha\beta, \quad (2.18)$$

$$\text{Var}[X] = \alpha\beta^2 \quad (2.19)$$

$$\text{Mode}[X] = (\alpha - 1)\beta, \text{ for } \alpha \geq 1 \text{ and} \quad (2.20)$$

$$\text{Kur}[X] = 3 + \frac{6}{\alpha}. \quad (2.21)$$

³ A Poisson process is a stochastic procedure with simple and wide applications for modeling times at which arrivals occur in a system (GARDINER, 2009).

2.3 Bayesian Reasoning

In inferential statistics, there are two ways of reasoning about the same target parameter θ . The first is to assume it is fixed and can be estimated after observations of many instances of the problem. The second way is to treat θ as a variable itself that depends on the data its coming from. The former is called the *frequentist statistics* and the latter is called *Bayesian statistics* (GELMAN et al., 2014), (EFRON, 2005). Concisely, repeatable problems might be solved using with a frequentist approach. On the other hand, for limited datasets where one has a prior knowledge of the problem, Bayesian statistics might be a good venue of investigation. Moreover, Bayesian statistics gives a more intuitive mechanism for updating beliefs upon arrival of new evidence. Naturally, there are more distinctions between these two approaches. The more interested reader should refer to Efron (2005) and Lehmann e Casella (2006) for a more exhaustive differentiation. This particular research investigates the Multilateration problem using the Bayesian machinery by computing a posterior distribution over the desired variable θ given the information $\mathcal{D}$.

In order to use the Bayesian reasoning, it is needed at first to assume that θ follows a given distribution $p(\theta)$, which is called a *prior probability distribution*. After that, the data $\mathcal{D}$ is observed and then the uncertainty about θ is updated via Bayes' rule

$$p(\theta | \mathcal{D}) = \frac{p(\mathcal{D} | \theta)p(\theta)}{p(\mathcal{D})}. \quad (2.22)$$

The probability $p(\mathcal{D} | \theta)$ is the *likelihood function* that conveys how probable $\mathcal{D}$ is for different situations of θ . Finally, $p(\mathcal{D})$ is called *evidence* or *normalization constant* and ensures that the posterior $p(\theta | \mathcal{D})$ is a valid PDF. For cases of point estimation the evidence can be fully neglected and the *unnormalized posterior* $\tilde{p}(\theta | \mathcal{D}) = p(\mathcal{D} | \theta)p(\theta)$ can be used; on the other hand, for the purposes of measuring uncertainty and for comparing models, the evidence can be computed by

$$p(\mathcal{D}) = \int p(\mathcal{D} | \theta)p(\theta)d\theta. \quad (2.23)$$

Naturally, Eq. (2.23) poses a hindrance for using Bayesian statistics because θ might be multidimensional. Nonetheless, advances in sampling methods and computational speed and memory capacities offer a elegant and satisfactory way for approximating the evidence.

Lastly, Bayesian statistics provides a posterior distribution over the desired θ . Nonetheless, in some cases (as in the Multilateration problem) the requested outcome is a point estimate. The two more common point estimators for this task are *Maximum a Posteriori Probability* (MAP) and Bayes estimates, where we use *Sampling Methods*.

2.4 Maximum A Posteriori

One possible way for obtaining this estimation is computing $p(\boldsymbol{\theta} \mid \mathcal{D})$, which selects the point of highest posterior probability by

$$\boldsymbol{\theta}_{MAP} = \arg \max_{\boldsymbol{\theta}} p(\boldsymbol{\theta} \mid \mathcal{D}) = \arg \max_{\boldsymbol{\theta}} p(\mathcal{D} \mid \boldsymbol{\theta})p(\boldsymbol{\theta}). \quad (2.24)$$

This, then, becomes an optimization problem that can be easily solved using a *Stochastic Gradient Descent* (SGD) algorithm. As for the posterior distribution mean one can use *Sampling Methods* as discussed in the next section.

2.5 Bayes Estimator via Sampling Methods

In most practical cases, computing the expected value in Eq. (2.8) and the evidence (2.23) might be intractable due to the continuous nature of variables or because of unknown aspects of the PDF at hand. As a consequence, algorithms for approximating those values are necessary. In this work, for computing the expected value of posterior distributions, also known as Bayes Estimates, we use two algorithms: *Importance Sampling* and *Metropolis-Hastings*. As for the evidence, we apply *Numeric Integration*.

2.5.1 Importance Sampling Method

Importance Sampling (IS) is particularly useful when draws of $\boldsymbol{\theta}$ from $p(\boldsymbol{\theta} \mid \mathcal{D})$ are not easily obtainable or cannot be drawn at all. The Importance Sampling method directly computes the posterior mean without generating samples from the distribution. More precisely, it is possible to compute

$$\mathbb{E}[\boldsymbol{\theta} \mid \mathcal{D}] = \int \boldsymbol{\theta} p(\boldsymbol{\theta} \mid \mathcal{D}) d\boldsymbol{\theta} \quad (2.25)$$

without generating samples from $p(\boldsymbol{\theta} \mid \mathcal{D})$. For that one needs a proposal probability density $h(\boldsymbol{\theta})$ from which drawing random samples is possible. As a result, Eq. (2.25) can be rewritten as

$$\begin{aligned} \mathbb{E}[\boldsymbol{\theta} \mid \mathcal{D}] &= \int \boldsymbol{\theta} p(\boldsymbol{\theta} \mid \mathcal{D}) d\boldsymbol{\theta} \\ &= \int \frac{\boldsymbol{\theta} p(\boldsymbol{\theta} \mid \mathcal{D}) h(\boldsymbol{\theta})}{h(\boldsymbol{\theta})} d\boldsymbol{\theta} \\ &= E_h \left[\frac{\boldsymbol{\theta} p(\boldsymbol{\theta} \mid \mathcal{D})}{h(\boldsymbol{\theta})} \right], \end{aligned} \quad (2.26)$$

where $E_h[\cdot]$ is the expectation for a probability measure for which the density is $h(\boldsymbol{\theta})$. The density $p(\boldsymbol{\theta} \mid \mathcal{D})$ is called the *nominal* or *target distribution*, $h(\boldsymbol{\theta})$ is the *importance* or *proposal distribution* and $p(\boldsymbol{\theta} \mid \mathcal{D})/h(\boldsymbol{\theta})$ is the *likelihood ratio*.

The Importance Sampling algorithm is then as follows. Generate samples $\theta_1, \dots, \theta_S$ according to the proposal distribution $h(\theta)$. Then the estimator for $E[\theta | \mathcal{D}]_{\text{IS}}$ becomes

$$E[\theta | \mathcal{D}]_{\text{IS}} = \frac{1}{S} \sum_{s=1}^S \frac{\theta_s p(\theta_s | \mathcal{D})}{h(\theta_s)}. \quad (2.27)$$

When it is necessary to use unnormalized posterior $\tilde{p}(\theta | \mathcal{D})$, the computation of $E[\cdot]_{\text{IS}}$ can be done via *biased Importance Sampling*. Then, the expected value can be approximated by

$$E[\theta | \mathcal{D}]_{\text{BIS}} = \frac{\sum_{s=1}^S \theta_s w(\theta_s)}{\sum_{s=1}^S w(\theta_s)}, \quad (2.28)$$

where $w(\theta_s) = \tilde{p}(\theta_s | \mathcal{D})/h(\theta_s)$. When it is possible to choose $h(\theta)$ such that the ratio $p(\theta_s | \mathcal{D})/h(\theta_s)$ is roughly constant, we can obtain reasonably precise Bayes estimations.

2.5.2 Metropolis-Hastings Method

It is also possible to compute the expected value using the Metropolis-Hastings algorithm. Differently from Importance Sampling, this method is useful when there is interest in sampling from the posterior distribution $p(\theta | \mathcal{D})$, but we cannot draw samples:

- directly from a mathematical software,
- computing from inverse CDF method,
- using the accept-reject algorithm.

Nonetheless, when it is possible to obtain some *proportionality* in the target function $p(\theta | \mathcal{D})$, one can use Metropolis-Hastings sampling algorithm (GELMAN et al., 2014) as shown in algorithm 1.

Algorithm 1 Metropolis-Hastings

- 1: Pick an initial state $\mathbf{x}^{(1)}$.
- 2: **for** $s = 1$ to S **do**
- 3: Sample $u \sim \mathcal{U}(0, 1)$.
- 4: Sample $\mathbf{x}^* \sim h(\mathbf{x} | \mathbf{x}^{(s)})$.
- 5: Compute the acceptance probability:

$$\rho(\mathbf{x}^{(s)}, \mathbf{x}^*) = \min \left[1, \frac{p(\mathbf{x}^* | \mathcal{D})}{p(\mathbf{x}^{(s)} | \mathcal{D})} \frac{h(\mathbf{x}^{(s)} | \mathbf{x}^*)}{h(\mathbf{x}^* | \mathbf{x}^{(s)})} \right] \quad (2.29)$$

- 6: If $u < \rho(\mathbf{x}^{(s)}, \mathbf{x}^*)$, set $\mathbf{x}^{(s+1)} = \mathbf{x}^*$.
 Otherwise, set $\mathbf{x}^{(s+1)} = \mathbf{x}^{(s)}$.
 - 7: **end for**
-

From algorithm 1 we observe that

- step 4 characterizes this algorithm as being a Markov Chain, since it draws a new state $\mathbf{x}^*$ from the previous one, using the proposal distribution $h(\cdot)$.

- the sequence $\mathbf{x}^{(s)}$ converges to $\boldsymbol{\theta}$, where $\boldsymbol{\theta} \sim p(\boldsymbol{\theta} \mid \mathcal{D})$.
- the sequence $\mathbf{x}^{(s)}$ is not independent.
- the sample mean of the distribution is

$$\mathbb{E}[\boldsymbol{\theta} \mid \mathcal{D}]_{\text{MH}} = \frac{1}{S} \sum_{s=1}^S \mathbf{x}^{(s)}. \quad (2.30)$$

- the term $\frac{p(\mathbf{x}^* \mid \mathcal{D})}{p(\mathbf{x}^{(s)} \mid \mathcal{D})}$ is the ratio of the target function for the new proposed state $\mathbf{x}^*$. Note that this term can be the unnormalized posterior $\tilde{p}(\boldsymbol{\theta} \mid \mathcal{D})$, since the evidences cancels out.

In this work the proposal distribution $h(\mathbf{x})$ is given by a Multivariate Normal distribution.

2.5.3 Numeric Integration for Evidence Computation

When dealing with unidimensional named distributions (Normal, Gamma, Beta, Poisson, and so forth), the posterior distribution can be computed analytically in closed form. Usually this is done by simulations using combinations of preprogrammed routines for such distributions (GELMAN et al., 2014).

In general, the Bayesian approach deals with high-dimensional integration problems that naturally demand lot of computing power. This is actually why Bayesian estimation was first thought to be unsuitable for practical problems (PESONEN; PICHÉ,).

In order to have the normalized posterior, we need to compute Eq. (2.23), which normalizes the posterior. We make the approximation to $p(\mathcal{D})$ via basic Monte Carlo sampling, that is

$$\begin{aligned} p(\mathcal{D}) &= \int_{-\infty}^{\infty} \tilde{p}(\boldsymbol{\theta} \mid \mathcal{D}) d\boldsymbol{\theta} \\ &= \int_{-\infty}^{\infty} p(\mathcal{D} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \\ &\approx \frac{1}{S} \sum_{s=1}^S p(\mathcal{D} \mid \boldsymbol{\theta}_s), \end{aligned} \quad (2.31)$$

where $\boldsymbol{\theta}_s$ are samples drawn from a known probability density function, known as proposal distribution. In this research the proposal distribution is given by a Multivariate Normal.

2.6 Closing Remarks

In this chapter we introduced concepts upon which we built our probabilistic Multilateration method. From the basic definition of a random variable to the basic rules that govern probabilistic reasoning we built a solid foundation for the upcoming method. In addition, we showed important statistical values that will play a fundamental role in defining parameters for the distributions of interest. Moving onward we concluded this

chapter by presenting the classic methods for combining probabilistic assertions to perform principled estimations about the desired variable. In the following chapter we bring all these concepts together and present the proposed probabilistic Multilateration method.

3 Probabilistic Multilateration

As we presented in Section 1.2, the main components of this research are the target point $\mathbf{q} = [q_1, q_2, \dots, q_D]^T \in \mathbb{R}^D$, the reference points $\{\mathbf{r}^{(k)}\}_{k=1}^K$, and the set of distance observations $\mathcal{D} = \{\delta^{(k)}\}_{k=1}^K$ between the target $\mathbf{q}$ and each reference point $\mathbf{r}^{(k)}$. The reference points $\mathbf{r}^{(k)}$ are known and fixed positions of devices that exchange signals with $\mathbf{q}$. Distances $\delta^{(k)}$ come from the original signals transformed in some metric of measurement. In other words, signals of RSSI, decibel, etc must undergo a conversion to some representation of distance in order to be used in this research.

The probabilistic model for the Multilateration problem consists of the joint probability density of all the involved random variables: that is, the latent variable $\mathbf{q}$ (the target) and the set of K observed variables¹ $\{\delta^{(k)}\}_{k=1}^K$ (the distances between target and reference points). The reference points $\{\mathbf{r}^{(k)}\}$ will be assumed to be fixed and known; thus, they shall be treated as deterministic variables².

The joint probability density function of all variables is $p(\mathbf{q}, \{\delta^{(k)}\}_{k=1}^K)$ and, by using the product rule, it factors into products of conditional and marginal densities,

$$p(\mathbf{q}|\{\delta^{(k)}\})p(\{\delta^{(k)}\}) = p(\mathbf{q})p(\{\delta^{(k)}\}|\mathbf{q}). \quad (3.1)$$

Inference consists of computing the density $p(\mathbf{q}|\{\delta_{k=1}^{(K)}\})$ of the latent variable given the observations; that is,

$$p(\mathbf{q}|\{\delta^{(k)}\}) = \frac{p(\mathbf{q})p(\{\delta^{(k)}\}|\mathbf{q})}{p(\{\delta^{(k)}\})}, \quad (3.2)$$

in which the probability density function $p(\{\delta^{(k)}\}|\mathbf{q})$ of the observations given the latent variables is combined, in the form of likelihood function, with a prior density $p(\mathbf{q})$ and then inverted to give the posterior density. As the marginal probability density of the data, or marginal likelihood or evidence, $p(\{\delta^{(k)}\}) = \int_{\mathbf{q}} p(\mathbf{q})p(\{\delta^{(k)}\}|\mathbf{q})$, does not depend on $\mathbf{q}$, it may be treated as a normalization constant.

By assuming that the distances $\delta^{(k)}$ are conditionally independent given $\mathbf{q}$, the resulting likelihood can be written as the product of the individual densities $p(\delta^{(k)}|\mathbf{q})$, and

¹ From this point on we simplify the notation by representing the assignment and the RV interchangeably. For instance, the RV of the distance observation $\Delta^{(k)} = \delta^{(k)}$ are used interchangeably.

² Clearly, this is a working assumption that could be easily relaxed by endowing also $\{\mathbf{r}^{(k)}\}$ with probability density functions. However, this extended formulation will not be pursued in this work.

the posterior is

$$\begin{aligned} p(\mathbf{q}|\{\delta^{(k)}\}) &= \frac{p(\mathbf{q}) \prod_{k=1}^K p(\delta^{(k)}|\mathbf{q})}{p(\{\delta^{(k)}\})} \\ &= \frac{p(\mathbf{q}) \prod_{k=1}^K p_{\boldsymbol{\theta}^{(k)}}(\delta^{(k)}|\mathbf{q})}{p(\{\delta^{(k)}\})}, \end{aligned} \quad (3.3)$$

in which we also assumed that each distance $p(\delta^{(k)}|\mathbf{q})$ has parameters $\boldsymbol{\theta}^{(k)}$ that can be tied to $\mathbf{q}$ and $\mathbf{r}^{(k)}$.

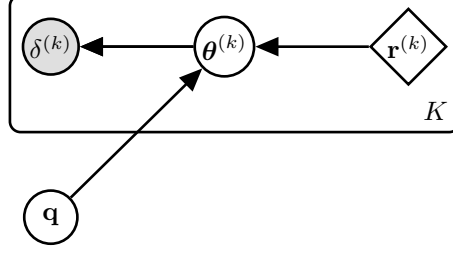


Figure 7 – Multilateration: Probabilistic graphical model.

The probabilistic graphical model of Multilateration is depicted in Figure 7 using a belief network in which each random variable is denoted by a node and links represent probabilistic relationships between joined variables. Observed variables $\delta^{(k)}$ are depicted by shaded nodes, while latent variables $\mathbf{q}$ are unshaded. For clarity, we included deterministic variables $\mathbf{r}^{(k)}$ and denoted them by a diamond. A plate labeled with K indicates that there are K triplets of variables of those kinds.

As shown in Eq. (3.3) good choices for prior and likelihood functions are paramount to model the Multilateration problem appropriately. Additionally, the proposed method is not concerned in finding parameters $\boldsymbol{\theta}^{(k)}$; the method is rather parameterized by values taken from the Multilateration problem itself to create the desired posterior distribution. In the following sections it is presented choices for distribution and the strategies adopted directly obtain $\boldsymbol{\theta}^{(k)}$.

3.1 Prior Distribution

In many applications the Normal distribution is usually one of the first choices for modeling a RV X when there are no prior knowledge of the problem. In addition, the prior distribution refers to an estimation of $\mathbf{q}$ itself. As such, it is necessary a distribution with support $\mathbb{R}^D$ since the point of interest can be anywhere in that interval.

As a result, for the prior distribution, we assume a Normal distribution centered in the given set of reference points; that is, $\mathbf{q} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with mean

$$\boldsymbol{\mu} = \frac{1}{K} \sum_k \mathbf{r}^{(k)} \quad (3.4)$$

and covariance matrix

$$\Sigma = \frac{1}{K-1} \sum_k (\mathbf{r}^{(k)} - \boldsymbol{\mu})(\mathbf{r}^{(k)} - \boldsymbol{\mu})^T. \quad (3.5)$$

Although the Normal distribution seems to be the most reasonable choice when there exists no prior knowledge of the problem (GOODFELLOW; BENGIO; COURVILLE, 2016), any continuous distribution with infinite support can be used to model the prior belief over $\mathbf{q}$. In addition, we use the Normal distribution as prior since its non-informative nature results in a posterior distribution that is mainly characterized by the data model using our proposed framework. Following, we discuss how distributions are used to model the likelihood function $p_{\boldsymbol{\theta}^{(k)}}(\delta^{(k)}|\mathbf{q})$.

3.2 Likelihood Function

In the context of Multilateration systems, distance measurements taken between two points can only be positive; as a consequence, distributions with positive support (as shown in Section 2.2) are chosen to model each $\delta^{(k)}$ that comprises the likelihood function of the proposed method. Additionally, we assume that the distance observations can be located at the mode or even the mean of this given distribution.

In this work we use the Gamma distribution to model $p(\delta^{(k)} | \mathbf{q})$. However, any positive supported distribution can be used as a candidate to model the set of distance observations $\{\delta^{(k)}\}$. In the following section we show guidelines for computing values³ $\boldsymbol{\theta}^{(k)}$ for this two-parameter distribution and we briefly discuss some special cases of the Generalized Gamma distribution to use other distributions to model distance observations.

To model distance observations as $\delta^{(k)} \sim \text{Gamma}(\boldsymbol{\theta}^{(k)}(\mathbf{q}))$, with $\boldsymbol{\theta}^{(k)} = \{\alpha_k, \beta_k\}$ in the Multilateration problem, we propose to tie the latent variable (the target point $\mathbf{q}$) to the individual parameters of the observation models $p(\delta^{(k)}|\mathbf{q})$. This process can be done by assuming that the mean of the distribution $p(\delta^{(k)}|\mathbf{q})$, or conversely its mode, is equal to the distance measurement $\|\mathbf{q} - \mathbf{r}^{(k)}\| = \delta^{(k)}$. More specifically, we need parameters that satisfy

$$\mathbb{E}[\delta^{(k)} | \mathbf{q}, \mathbf{r}^{(k)}] = \|\mathbf{q} - \mathbf{r}^{(k)}\| \quad (3.6)$$

or

$$\text{Mode}[\delta^{(k)} | \mathbf{q}, \mathbf{r}^{(k)}] = \|\mathbf{q} - \mathbf{r}^{(k)}\|. \quad (3.7)$$

Since we use the two parameter Gamma distribution, centering the distance measurement at one of these values will raise ambiguity on the parameters. For instance,

³ The number of parameters of the distribution defines the size of $\boldsymbol{\theta}^{(k)}$.

taking again parameters α and β from the Gamma distribution and assuming that the distance $\|\mathbf{q} - \mathbf{r}\|$ is at the mean of the Gamma distribution we have

$$\mathbb{E}[\delta|\mathbf{q}, \mathbf{r}] = \alpha\beta \quad \implies \quad \|\mathbf{q} - \mathbf{r}\| = \alpha\beta \quad \implies \quad \beta = \frac{\|\mathbf{q} - \mathbf{r}\|}{\alpha}.$$

In this case, even though β has been defined, α could take any positive value and the equation would still hold. This is clearly a problem since we want to uniquely define both parameters by the distance observation δ . As a result, this approach alone is not indicated for modeling distance observations when the distribution has more than one parameter.

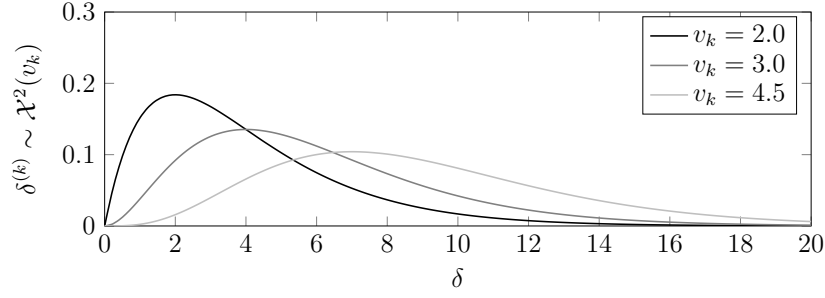
In the following we present three alternatives for solving this ambiguity when centering the distance observation at the mean of the distribution. As explained earlier, however, it is also possible to use the mode. The procedure to obtain those equations are the same as explained here. We opted not to include them here in order to avoid redundancy. Table 1 shows parameters for assuming distance observations at the mean of the Gamma distribution. The full description of parameters for mean and mode can be found at Appendix A.

Chi-Square $\mathcal{X}^2$ Distribution. The first possible way to avoid this ambiguity is by using another special case of the Gamma distribution, where we set the parameters to obtain a one-parameter distribution, for instance the $\mathcal{X}^2$ distribution, as described in (AYIENDA, 2014). Making $\alpha_k = v_k/2$ and $\beta_k = 2$, we have the $\mathcal{X}^2$ distribution with parameter v_k . Since this distribution has only the parameter v_k and the mean is the parameter itself, we can use the Eq. (3.6) to properly define the parameter by

$$v_k = \|\mathbf{q} - \mathbf{r}^{(k)}\|. \quad (3.8)$$

The downside of this approach is that the density function of the $\mathcal{X}^2$ distribution always starts at the origin⁴, which means that the shape of the distribution flattens as the distance observation is further from the origin. This behavior is depicted in Figure 8.

⁴ In fact, all one-parameter distributions that we analyzed: Rayleigh, Nakagami and Maxwell-Boltzmann presented the same behavior.

Figure 8 – The χ^2 distribution behavior.

Source: Author.

Fixed Variance ν . Another possible way for removing ambiguity from the Gamma distribution is to keep track of how these distance observations are dispersed around mean or mode; that is, we set the variance of the distribution to be exactly ν , with $\nu > 0$. In this case we make

$$\text{Var}[\delta^{(k)} | \mathbf{q}, \mathbf{r}^{(k)}] = \nu.$$

Then, for distances modeled as Gamma variables we make $\delta^{(k)} \sim \text{Gamma}(\boldsymbol{\theta}^{(k)}(\mathbf{q}))$, where parameters $\boldsymbol{\theta}^{(k)} = \{\alpha_k, \beta_k\}$ are computed using the distribution mean as

$$\alpha_k \beta_k = \|\mathbf{q} - \mathbf{r}^{(k)}\| \quad \implies \quad \beta_k = \frac{\|\mathbf{q} - \mathbf{r}^{(k)}\|}{\alpha_k}; \quad (3.9)$$

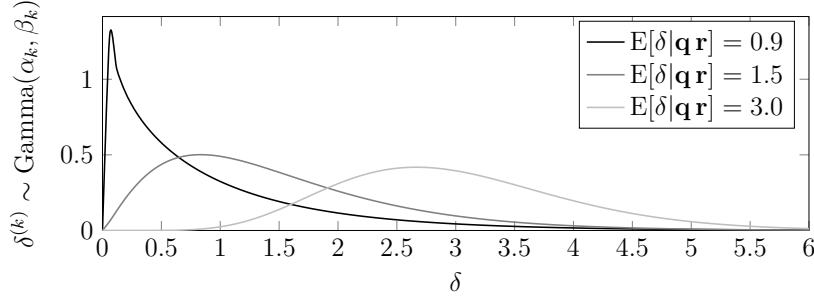
we then solve the ambiguity on α_k by

$$\text{Var}[\delta^{(k)} | \mathbf{q}, \mathbf{r}^{(k)}] = \alpha_k \beta_k^2 \quad \implies \quad \alpha_k \beta_k^2 = \nu \quad \implies \quad \alpha_k = \frac{\nu}{\beta_k^2}. \quad (3.10)$$

Solving the system for $\alpha_k, \beta_k > 0$ we get

$$\alpha_k = \frac{\|\mathbf{q} - \mathbf{r}^{(k)}\|^2}{\nu} \quad \text{and} \quad \beta_k = \frac{\nu}{\|\mathbf{q} - \mathbf{r}^{(k)}\|}. \quad (3.11)$$

As a result, we associate parameters of the Gamma distribution to distance observations and a confidence on them represented by ν . In addition, note that both parameters are uniquely defined given that they both must be greater than 0. The resulting behavior of the distribution is shown in Figure 9.

Figure 9 – Gamma distribution for $\nu = 1$.

Source: Author.

The obvious pitfall of this approach is that the shape of the distribution does not remain the same for near-zero distance observations. This might be problematic because in such cases the distribution does not present the bell-shape curve. Nonetheless, in comparison to the previous technique we are able to remove ambiguity from both parameters without having the density always starting from 0.

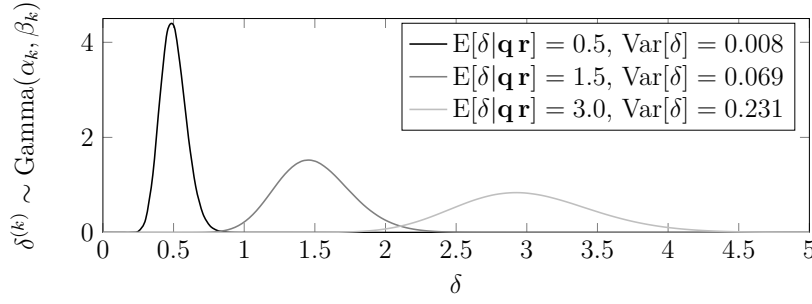
Approximation to the Normal Distribution. The last way for removing ambiguity presented in this work is by approximating the Gamma distribution to the Normal distribution. As a result, the closer the kurtosis of the Gamma distribution gets to 3, closer the mean gets to the mode of the distribution. For that, in addition to having $\text{Var}[\cdot] \leq \nu$ we need to make sure that the kurtosis t of the Gamma distribution is approximately equal to that of the Normal distribution. That is,

$$\text{Kur}[\delta^{(k)}|\mathbf{q}, \mathbf{r}^{(k)}] = 3 + \frac{6}{\alpha_k} \implies t = 3 + \frac{6}{\alpha_k} \implies \alpha_k = \frac{6}{t-3}. \quad (3.12)$$

Now the final parameters are given by

$$\alpha_k = \frac{\|\mathbf{q} - \mathbf{r}^{(k)}\|^2}{\nu} + \frac{6}{t-3} \quad \text{and} \quad \beta_k = \frac{\|\mathbf{q} - \mathbf{r}^{(k)}\|}{\alpha_k}. \quad (3.13)$$

In this way we always uniquely define parameters, guarantee variance at most ν and make sure that the shape is close to the Normal distribution even for values of $\delta^{(k)}$ that are near the origin, as shown in Figure 10.

Figure 10 – Gamma distribution for $\nu = 1$ and $t = 3.2$.

Source: Author.

We point out that using this schema when $\delta^{(k)} \rightarrow 0$ the Gamma distribution behaves in the same manner as the Dirac Delta Distribution⁵. If all distance observations had this particular distribution, we would interpret the Multilateration problem as deterministic. That is, all distance observations would have no variance and if the intersection between them existed, this intersection would represent the target point.

Table 1 – Parameters for the Gamma distribution and their variations

Assumption	Technique	Parameter	Constraint
$E[\delta^{(k)} \mid \mathbf{q}, \mathbf{r}^{(k)}]$	Chi-Square	$\alpha_k = 2\ \mathbf{q} - \mathbf{r}^{(k)}\ $ $\beta_k = 2$	None
	Fixed variance ν	$\alpha_k = \frac{\ \mathbf{q} - \mathbf{r}^{(k)}\ ^2}{\nu}$ $\beta_k = \frac{\nu}{\ \mathbf{q} - \mathbf{r}^{(k)}\ }$	$\nu > 0$
	Normal approximation	$\alpha_k = \frac{\ \mathbf{q} - \mathbf{r}^{(k)}\ ^2}{\nu} + \frac{6}{t-3}$ $\beta_k = \frac{\ \mathbf{q} - \mathbf{r}^{(k)}\ }{\alpha_k}$	$\nu > 0, t > 3$

Source: Author.

Remarks on other distributions. Alternatively to the Gamma distribution, it is possible to define the Uniform distribution in such way that it has a positive support and it is suitable to be used for modeling distance observations. The interpretation of using the Uniform Distribution is to determine a equally probable region of interest around the distance observation. However, applying gradient optimization in this case is not profitable due to the lack of inflection on the Uniform distribution. A possible and reasonable way

⁵ The Dirac Delta distribution models the density of an ideal point mass, where the function equals to 0 everywhere except for zero and its integral is equal to one (ARATYN; RASINARIU, 2005).

of estimating the latent variable in this setup is to compute the intersecting area that the set of distance observations make with the prior distribution. Finally, distributions presented in this research are by no means exhaustive and a myriad of positive supported distributions can be easily found in the literature to model distance observations following the procedures described in this work.

3.3 Inference

Up to this point we have offered a fully developed method for finding the distribution of $\mathbf{q}$ given distance observations and reference points. In practical applications, however, one might be interested in a single estimation for the position of $\mathbf{q}$. In this research the point of interest $\mathbf{q}$ is estimated via two main venues: computing the mode and mean of the posterior distribution. For estimating the we use MAP inference and for the mean we use Bayes estimate.

3.3.1 Computation of MAP Estimates

In order to use MAP inference the Multilateration problem can be cast as the problem of computing the most likely assignment to the variable $\mathbf{q}$ given the values of $\{\delta^{(k)}\}$, that is,

$$\hat{\mathbf{q}}_{\text{MAP}} = \arg \max_{\mathbf{q}} p(\mathbf{q} | \{\delta^{(k)}\}) = \arg \max_{\mathbf{q}} g(\mathbf{q}), \quad (3.14)$$

where $g(\mathbf{q}) = p(\mathbf{q})p(\{\delta^{(k)}\}|\mathbf{q})$. Notice that for the MAP inference task the evidence term can be fully neglected since it does not depend on $\mathbf{q}$. The MAP inference now can be written as

$$\hat{\mathbf{q}}_{\text{MAP}} = \arg \max_{\mathbf{q}} p(\mathbf{q}) \prod_i p(\delta^{(k)}|\mathbf{q}). \quad (3.15)$$

For computing $\arg \max_{\mathbf{q}} g(\mathbf{q})$ in Eq. (3.15) we take $\arg \max_{\mathbf{q}} \log g(\mathbf{q})$ due to the monotone nature of the log function. For $\delta^{(k)} \sim \text{Gamma}(\cdot)$, we have

$$\log g(\mathbf{q}) = \log(C) - \frac{1}{2}(\mathbf{q} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{q} - \boldsymbol{\mu}) + \sum_k^K \left[\log \left(\frac{\delta^{(k)\alpha_k-1}}{\Gamma(\alpha_k) \beta_k^{\alpha_k}} \right) - \frac{\delta^{(k)}}{\beta_k} \right], \quad (3.16)$$

with parameters given in Sections 3.1 and 3.2.

In order to make point estimates using MAP inference, it is necessary to optimize Eq. (3.16) via *Gradient Ascent* (RUDER, 2016). Thus, after parameters for each distribution are put in place, it is necessary computation of gradients for each optimization function formed from the various possibilities of parametrization. Annex A shows partial derivatives for the Gamma distribution with respective parameters presented in this work.

3.3.2 Computation of Bayes estimate

The mean of the posterior distribution is referred to as Bayes estimate and is computed by

$$\hat{\mathbf{q}}_{Bayes} = \int \mathbf{q} p(\mathbf{q} | \{\delta^{(k)}\}) d\mathbf{q}. \quad (3.17)$$

This result can be attained by approximation methods such as Importance Sampling and Metropolis-Hastings. For Importance Sampling we compute

$$\hat{\mathbf{q}}_{BIS} = \frac{\sum_{s=1}^S \frac{\mathbf{q}_s g(\mathbf{q}_s)}{h(\mathbf{q}_s)}}{\sum_{s=1}^S \frac{g(\mathbf{q}_s)}{h(\mathbf{q}_s)}}, \quad (3.18)$$

where $g(\mathbf{q}_s)$ is the unnormalized posterior computed for the $\mathbf{q}_s$ sample drawn from the proposal distribution $h(\mathbf{q})$. In this research we choose the multivariate normal $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ as the proposal distribution, with $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ given by Eq. (3.4) and (3.5), respectively.

Alternatively, one can obtain the Bayes estimate via Metropolis-Hastings. For that, we first generate samples $\mathbf{q}_s$ following algorithm 1, then we compute the sample mean by

$$\hat{\mathbf{q}}_{MH} = \frac{1}{S} \sum_{s=1}^S \mathbf{q}_s. \quad (3.19)$$

As means of simplification the proposed method is described in Algorithm 2.

Algorithm 2 Proposed Method

INPUT: Sets $\{\delta^{(k)}\}$, $\{\mathbf{r}^{(k)}\}$

OUTPUT: Estimations $\hat{\mathbf{q}}_{MAP}$, $\hat{\mathbf{q}}_{BIS}$, $\hat{\mathbf{q}}_{MH}$

- 1: Compute parameters for the prior distribution. Eq. (3.4) and (3.5).
 - 2: Choose parameter and distribution for distance observations. Table 1.
 - 3: Compute $\hat{\mathbf{q}}_{MAP}$ via gradient optimization. Eq. (3.16)
 - 4: Compute $\hat{\mathbf{q}}_{BIS}$ via Importance Sampling . Eq. (3.18).
 - 5: Compute $\hat{\mathbf{q}}_{MH}$ via Metropolis-Hastings . Algorithm 1 and Eq. (3.19)
-

3.4 Closing Remarks

In this chapter we presented the probabilistic framework to solve the Multilateration problem. From determining the prior distribution to nuances that parameters for the likelihood function might have, we showed a well-reasoned and elegant solution for the D -dimensional posterior distribution Multilateration problem. Nonetheless, having the posterior distribution alone might not suffice in cases where a point estimation is necessary. For such cases, we also offered the means for estimating the coordinates of the target point via MAP inference and Bayes estimate. In the following chapter we validate the proposed method and compare its accuracy to other solutions that solve the Multilateration problem.

4 Experiments

In this chapter we present some results for the proposed method using one synthetic dataset and another dataset that comes from a real-world application. The synthetic dataset will serve as proof-of-concept piece to validate the reasoning behind the proposed method. The second portion of experiments will compare and validate the proposed method against related techniques for solving the Multilateration problem.

4.1 Validating the Proposed Method

For the synthetic dataset we generated different D -dimensional Multilateration systems. For each D -dimensional system we generated the target point $\mathbf{q}$ and the set of reference points $\{\mathbf{r}^{(k)}\}$. We tested sets of reference points with $K \in \{3, 16\}$ for $D = 2$, with $K \in \{4, 16\}$ for $D = 3$ and for $D = 5$ we tested with $K \in \{6, 32\}$.

The final element to complete the dataset is the collection of distance observations $\{\delta^{(k)}\}$. We used the Gamma distribution with parameters from Section 3.2 to randomly generate $\{\delta^{(k)}\}$ for each Multilateration system. For example, using target and reference points with $D = 2$ and $K = 2$ we randomly generated three collections of distance observations $\{\delta^{(k)}\}$, and therefore, three datasets. The first collection has each distance observation modeled as $\delta^{(k)} \sim \text{Gamma}\left(\alpha_k = \frac{\|\mathbf{q} - \mathbf{r}^{(k)}\|}{2}, \beta_k = 2\right)$, the χ^2 distribution, to have $\|\mathbf{q} - \mathbf{r}^{(k)}\|$ centered at the distribution mean. The second collection has $\delta^{(k)} \sim \text{Gamma}\left(\alpha_k = \|\mathbf{q} - \mathbf{r}^{(k)}\|^2, \beta_k = \frac{1}{\|\mathbf{q} - \mathbf{r}^{(k)}\|}\right)$ to have $\|\mathbf{q} - \mathbf{r}^{(k)}\|$ centered at the distribution mean and constant variance $\nu = 1$. Lastly we generated the collection of distance observations $\delta^{(k)} \sim \text{Gamma}\left(\|\mathbf{q} - \mathbf{r}^{(k)}\|^2 + 30, \beta_k = \frac{\|\mathbf{q} - \mathbf{r}^{(k)}\|}{\alpha_k}\right)$ to have $\|\mathbf{q} - \mathbf{r}^{(k)}\|$ centered at the distribution mean with variance $\nu \leq 1$ and kurtosis $t = 3.2$. We systematically repeated this process for each combination of target and reference points aforementioned.

The relevance of these datasets is to assess

- the ability of the method to converge to an estimation;
- whether the proposed method is able to make *reasonable* inferences on datasets that came from one of the Gamma distributions presented in this research;
- how the inference is affected by the dimension of the Multilateration system;
- the affect that the number of reference points have on making *good* approximations of the target point.

The reason behind this validation is to justify using this method to solve Multilateration problems. If the proposed method does not meet all the above requisites we would have no foundational reason to compare it with other methods. Having this primary

Table 2 – Probabilistic Multilateration Method Results.

System	$\delta^{(k)}$ Distribution	Metrics		
		$\ \mathbf{q} - \hat{\mathbf{q}}_{\text{MAP}}\ $	$\ \mathbf{q} - \hat{\mathbf{q}}_{\text{BIS}}\ $	$\ \mathbf{q} - \hat{\mathbf{q}}_{\text{MH}}\ $
$D = 2, K = 3$	$\mathcal{X}^2$	$4.37 \pm 2.23\text{e-}9$	$3.87 \pm 5.0\text{e-}2$	$3.58 \pm 3.3\text{e-}2$
	$\nu = 1$	1.80 ± 2.16	$3.16 \pm 2.4\text{e-}2$	$3.23 \pm 1.7\text{e-}2$
	$\nu \leq 1, t = 3.2$	$2.57 \pm 4.0\text{e-}7$	$5.40 \pm 1.2\text{e-}1$	$5.43 \pm 2.0\text{e-}2$
$D = 2, K = 16$	$\mathcal{X}^2$	$0.39 \pm 7.85\text{e-}7$	$0.36 \pm 4.15\text{e-}2$	$0.35 \pm 3.5\text{e-}2$
	$\nu = 1$	$0.41 \pm 4.16\text{e-}10$	$0.40 \pm 2.2\text{e-}2$	$0.41 \pm 2.2\text{e-}2$
	$\nu \leq 1, t = 3.2$	33.85 ± 15.84	$0.10 \pm 2.2\text{e-}2$	$0.12 \pm 4.2\text{e-}2$
$D = 3, K = 4$	$\mathcal{X}^2$	$0.78 \pm 3.21\text{e-}6$	$1.54 \pm 4.8\text{e-}2$	$1.65 \pm 3.4\text{e-}2$
	$\nu = 1$	$1.16 \pm 4.43\text{e-}7$	$1.19 \pm 2.6\text{e-}2$	$1.22 \pm 1.6\text{e-}3$
	$\nu \leq 1, t = 3.2$	1.70 ± 0.69	1.40 ± 0.29	1.42 ± 0.96
$D = 3, K = 16$	$\mathcal{X}^2$	$1.01 \pm 1.04\text{e-}8$	$0.97 \pm 3.4\text{e-}2$	$0.99 \pm 1.4\text{e-}2$
	$\nu = 1$	$0.63 \pm 7.21\text{e-}8$	$0.68 \pm 3.7\text{e-}2$	$0.69 \pm 2.7\text{e-}2$
	$\nu \leq 1, t = 3.2$	2.67 ± 1.61	$0.32 \pm 7.8\text{e-}2$	$0.31 \pm 5.8\text{e-}2$
$D = 5, K = 6$	$\mathcal{X}^2$	$1.61 \pm 2.27\text{e-}6$	3.81 ± 0.83	3.71 ± 0.53
	$\nu = 1$	$4.05 \pm 2.71\text{e-}6$	$4.26 \pm 1.168\text{e-}2$	$4.5 \pm 1.2\text{e-}2$
	$\nu \leq 1, t = 3.2$	$2.11 \pm 3.94\text{e-}7$	$2.85 \pm 2.54\text{e-}2$	$2.58 \pm 2.12\text{e-}2$
$D = 5, K = 32$	$\mathcal{X}^2$	$1.83 \pm 5.24\text{e-}8$	$1.40 \pm 5.24\text{e-}8$	$1.41 \pm 3.24\text{e-}7$
	$\nu = 1$	$0.33 \pm 2.55\text{e-}9$	$0.66 \pm 1.65\text{e-}1$	$0.65 \pm 6.5\text{e-}2$
	$\nu \leq 1, t = 3.2$	$2.55 \pm 0.23\text{e-}9$	$0.73 \pm 3.68\text{e-}1$	$0.75 \pm 4.67\text{e-}1$

validation enables us to compare the proposed method with other solutions where we do not know the distribution (if any) distance observations came from.

To realize tests on these synthetic datasets we used the approximation methods discussed in Section 2.5 and in Chapter 3. For computing $\hat{\mathbf{q}}_{\text{BIS}}$ the proposal distribution $h(\mathbf{q})$ is given by a Multivariate Normal $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $S = 500$. In addition, for generating samples $\mathbf{q}_s$ to compute $\hat{\mathbf{q}}_{\text{MH}}$ using Metropolis-Hastings, the proposal distribution $h(\mathbf{q})$ is given by a Multivariate Normal $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ given by Eq. (3.5) and 3.4, respectively and $S = 1000$.

For the Gradient Optimization to obtain $\hat{\mathbf{q}}_{\text{MAP}}$ we run the proposed method 10 times on both systems with constant step-size $\eta = 0.00125$, maximum of 5000 iterations and $\|\nabla \log [g(\mathbf{q})]\| \leq 10^{-6}$ as the termination criteria for the Gradient optimization. The complete summary of results for this validation portion is shown in Table 2.

As shown in Table 2, the metric $\|\mathbf{q} - \hat{\mathbf{q}}_{\text{MAP}}\|$ measures the distance from the real target point $\mathbf{q}$ to $\hat{\mathbf{q}}_{\text{MAP}}$, the average estimation over the 10 runs, with respective standard deviation shown below. When compared to the analogous metric for Biased Importance Sampling $\|\mathbf{q} - \hat{\mathbf{q}}_{\text{BIS}}\|$, the MAP inference usually outperforms the the Biased Importance Sampling technique in our experiments. As for the Metropolis-Hasting method, $\|\mathbf{q} - \hat{\mathbf{q}}_{\text{MH}}\|$, we see that it is close to the estimation given the Biased Importance Sampling technique; which is expected since both approximate the mean of the posterior distribution.

Regarding computational time, both Biased Importance Sampling and Metropolis-Hastings demand higher computational time than the MAP estimation. In that regard MAP estimation is, generally, 5 times faster than Biased Importance Sampling estimator and about 7 times faster than the Metropolis-Hastings. We point out that choices for the proposal distribution $h(\cdot)$ will directly affect the convergence time for the Biased Importance Sampling and Metropolis-Hastings. It is worth noting that for both estimators we use log densities to avoid computational problems regarding small numbers.

Lastly, Figure 11 shows results for probabilistic Multilateration systems with points of interest. We noticed that for fewer reference points, prior distributions tends to have a greater affect on the posterior compared to $K = 32$. In addition, the presence of flat regions require many runs of the MAP inference with different starting points each time, so that the estimation with higher density is chosen to be the final estimation for $\mathbf{q}$. Areas near reference points tends to $-\infty$; as a result, the target point $\mathbf{q}$ cannot be located at any reference point when using the Gamma distribution to model distance measurements $\{\delta^{(k)}\}$.

4.2 Validation via real-world application Dataset

The dataset provided at <https://www.kaggle.com/liwste/indoor-positioning> comes from three Bluetooth beacons (A, B, C) installed in a 2.74 meters wide by 4.38 meters long room. Readings¹ (in meters) were made using a Sony Xperia E3 smartphone in five different locations of the room. In Figure 12 we have more insight about the problem and regarding where the dataset comes from.

¹ This work used the converted data from RSSI to meters provided by the author of the dataset.

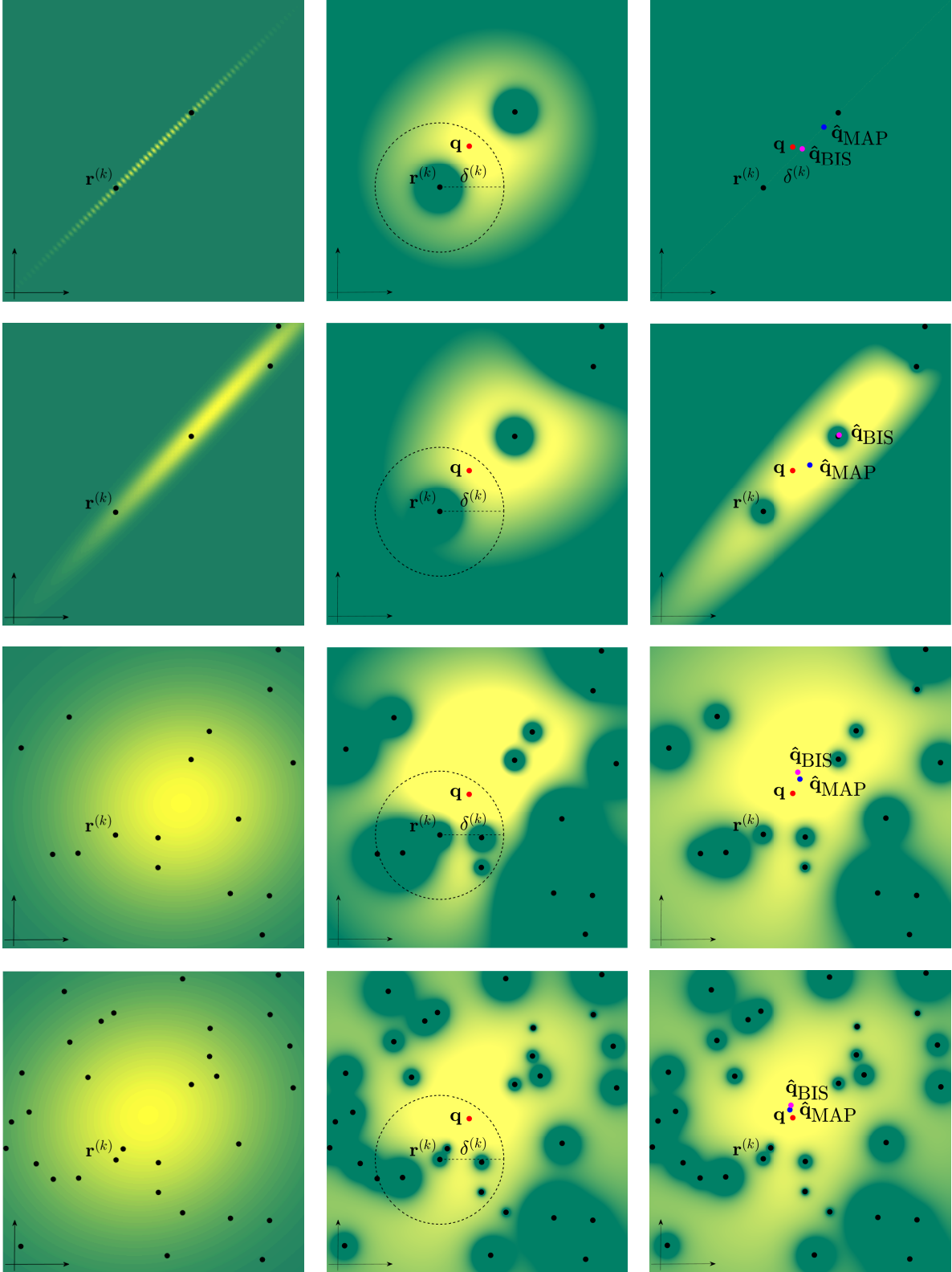


Figure 11 – Probabilistic Multilateration: Left column, $\log p(\mathbf{q})$; Middle column, $\log \sum_{k=1}^K p(\delta^{(k)}|\mathbf{q})$, and; Right column, $\log [p(\mathbf{q}) \prod_{k=1}^K p(\delta^{(k)}|\mathbf{q})]$, for $D = 2$ and $K \in \{2, 4, 16, 32\}$.

Figure 12 – Room where dataset was generated.



Source: EVANS; SWARTZ.

As it can be seen in Figure 12 all beacons are positioned in a room with static barriers (walls) and possibly dynamic ones (people). The position where the person is located, is the target point $\mathbf{q}$ that has to be estimated. Moreover, as stated by the author of the dataset, the smartphone was located in five different positions; and, in each position, the system records several readings with interval of 30 to 60 seconds between them. Table 3 shows an excerpt of the dataset.

As shown in Table 3 coordinates X, Y represent all possible positions for the query point $\mathbf{q}^{(i)}$. That is, $\mathbf{q}^{(1)} = \{1.22, 1.80\}$, $\mathbf{q}^{(2)} = \{0.79, 1.80\}$ and so forth. Columns Distance A, Distance B and Distance C show measurements from each respective beacon (reference point) to $\mathbf{q}^{(i)}$. Each position $\mathbf{q}^{(i)}$ has about 50 measurements. That is, there are about 50 readings from the same set of three beacons to the same $\mathbf{q}^{(i)}$. We proceeded the tests by taking the average of this measurements. This results in $K = 3$ reference points and distance observations derived from the beacons. Then, we compute the estimators: $\mathbf{q}_{MAP}^{(i)}$, $\mathbf{q}_{BIS}^{(i)}$ and $\mathbf{q}_{MH}^{(i)}$.

With regard to values for the approximation methods, discussed in Section 2.5, the proposal distribution $h(\mathbf{q})$ for Importance Sampling is, still, given by a Multivariate Normal $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. However, the sample size becomes $S = 500$. In addition, for the Metropolis-Hastings algorithm, the proposal distribution $h(\mathbf{q})$ continues to be a Multivariate Normal $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ given by Eq. (3.5) and 3.4, respectively. In Table 4 we show results for using the proposed to solve this Multilateration problem. These results come from 10 runs of each procedure described formerly.

We also compared results from methods described in Section 1.1. That is, Related

Table 3 – Sample table of measurements from beacons to each location

Distance A	Distance B	Distance C	Position X	Position Y
0.8774	0.7686	1.4572	1.22	1.80
1.4011	1.2588	2.4600		
1.3913	1.2726	2.4287		
1.1250	1.3023	2.1765	0.79	1.80
1.1327	1.3295	2.2154		
1.1374	1.3387	2.1804		
1.1250	1.2867	2.1951	1.65	1.80
1.1327	1.2962	2.0740		
1.2016	1.0529	2.1125		
1.0383	1.3098	1.3864	1.22	1.37
1.0342	1.3387	1.4161		
1.0456	1.3098	1.3864		
0.9545	1.0793	1.9280	1.22	2.23
0.9044	1.0870	1.9014		
0.9044	0.7686	1.9093		

Source: Author.

Table 4 – Table of errors from estimates to the target point $\mathbf{q}^{(i)}$.

Position	$\delta^{(k)}$ Distribution	Metrics		
		$\ \mathbf{q} - \hat{\mathbf{q}}_{\text{MAP}}\ $	$\ \mathbf{q} - \hat{\mathbf{q}}_{\text{BIS}}\ $	$\ \mathbf{q} - \hat{\mathbf{q}}_{\text{MH}}\ $
$\mathbf{q}^{(1)}$	χ^2	$0.82 \pm 1.36\text{e-}6$	3.78 ± 0.46	3.76 ± 0.32
	$\nu = 1$	$0.43 \pm 1.39\text{e-}7$	1.42 ± 0.07	1.39 ± 0.15
	$\nu \leq 1, t = 3.2$	$1.06 \pm 10.39\text{e-}7$	0.63 ± 0.06	0.60 ± 0.05
$\mathbf{q}^{(2)}$	χ^2	$0.99 \pm 0.24\text{e-}7$	3.69 ± 0.24	3.53 ± 0.18
	$\nu = 1$	$0.88 \pm 3.06\text{e-}7$	1.41 ± 0.02	1.46 ± 0.02
	$\nu \leq 1, t = 3.2$	1.51 ± 0.19	0.89 ± 0.04	0.89 ± 0.04
$\mathbf{q}^{(3)}$	χ^2	$1.12 \pm 9.45\text{e-}7$	3.40 ± 0.23	3.53 ± 0.24
	$\nu = 1$	$1.08 \pm 1.58\text{e-}7$	1.59 ± 0.06	1.58 ± 0.10
	$\nu \leq 1, t = 3.2$	1.17 ± 0.05	1.11 ± 0.05	1.08 ± 0.06
$\mathbf{q}^{(4)}$	χ^2	$1.46 \pm 0.20\text{e-}7$	3.83 ± 0.20	3.78 ± 0.18
	$\nu = 1$	$1.44 \pm 1.76\text{e-}7$	1.89 ± 0.06	1.76 ± 0.14
	$\nu \leq 1, t = 3.2$	1.81 ± 0.12	1.45 ± 0.07	1.46 ± 0.08
$\mathbf{q}^{(5)}$	χ^2	$0.34 \pm 8.03\text{e-}7$	2.79 ± 0.35	2.81 ± 0.31
	$\nu = 1$	$0.16 \pm 2.02\text{e-}7$	0.68 ± 0.05	0.63 ± 0.16
	$\nu \leq 1, t = 3.2$	0.41 ± 0.39	0.18 ± 0.05	0.20 ± 0.05

Source: Author.

Table 5 – Comparison between results from related works and the proposed method

Method	Metrics				
	$\ \mathbf{q}^{(1)} - \hat{\mathbf{q}}^{(j)}\ $	$\ \mathbf{q}^{(2)} - \hat{\mathbf{q}}^{(j)}\ $	$\ \mathbf{q}^{(3)} - \hat{\mathbf{q}}^{(j)}\ $	$\ \mathbf{q}^{(4)} - \hat{\mathbf{q}}^{(j)}\ $	$\ \mathbf{q}^{(5)} - \hat{\mathbf{q}}^{(j)}\ $
PALEM	0.83	0.84	1.19	1.03	0.72
LIWSN	0.51	0.83	1.01	1.34	0.10
DPEMM	1.15	1.22	1.24	1.58	0.72
$\hat{\mathbf{q}}_{\text{MAP}}$	0.43 (1.39e-7)	0.88 (3.06e-7)	1.08 (1.58e-7)	1.44 (1.76e-7)	0.16 (1.76e-7)

Source: Author.

Work 1 refers to the method PALEM by PINEDA-BRISE et al. (2015); Related Work 2 refers to the solution LIWSN by Ramadurai e Sichitiu (2003); and Related Work 3, DPEMM, is the method by Naganawa, Miyazaki e Tajima (2017). Briefly, they all work for 2D and 3D cases; moreover, they use the probabilistic framework to compute an estimation for the target point using Maximum Likelihood. They differ from each other in the way that distance measurements are treated. In PALEM, the authors use range-free radio frequency signals to come up with the distance observation. On the other hand, in LIWSN the authors use range-based signals to assert the certainty of distance observations. Lastly, in DPEMM the authors use satellites as receivers from aircraft a position monitoring system to compute the Time of Arrival (ToA) of radio frequency signals exchanged between satellites and the aircraft. These three solutions come from many different applications and they solve the problem at hand well in their given environment: at most 3D cases and minimally controlled scenarios.

Table 5 shows results for using this dataset with the proposed method and related solutions. Each column refers to the error estimation given by each target position $\mathbf{q}^{(i)}$ and estimation from the related works, respectively described by their abbreviations. The estimation $\hat{\mathbf{q}}_{\text{MAP}}$ is the best result given by the proposed method. In this case we assumed distance observations coming from the Gamma distribution with constant variance $\nu = 1$.

In summary, we see that the proposed method is not only on par with known methods but surpasses them in some cases. We also point out that related methods do not give support for systems beyond 3 dimension as does the proposed method.

5 Conclusion

In this work we propose a novel method to solve the classic Multilateration problem via a probabilistic structure. Even though there exist in the literature working methods that take this probabilistic stance, we perceived the lack of the specific approach taken here. Whilst some methods take the Normal distribution to model distance measurements errors, we take the position of modeling distance observations as following a specific positive supported distribution, namely the Gamma distribution.

Through an extensive study on profitable parametrizations and methods for point estimation, we offer a method that is fully capable of representing the posterior distribution of the target point; in addition to make point estimations via optimization and approximation venues.

This research accomplishes its main and specific objectives by providing not only the probabilistic method for the Multilateration problem but also the mathematical framework that supports it. Results on artificial and real datasets give enough information to support the claim that this method is not only compatible to existing Multilateration methods but it also surpasses them in some cases. In addition, we show plots of the log posterior density of the target position. They help the reader gaining a better insight on the method's output and what to expect from the method when using it.

Finally, even though there are many solutions for the Multilateration problem, this method offers an elegant and coherent solution for providing the location of interest with a notion of uncertainty. This uncertainty naturally arises, as it can be perceived, in the real dataset utilized, where the author made measurements for his position using Bluetooth beacons. He compared these distance observations with the real distance taken from a measuring tape with and confirmed the uncertainty nature of distance observations using such methodology.

The repository with codes and more information about this work can be found at Vasconcelos (2018).

Bibliography

ABEL, J.; SMITH, J. The spherical interpolation method for closed-form passive source localization using range difference measurements. In: *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'87*. IEEE, 1987. v. 12, p. 471–474. Disponível em: <<https://doi.org/10.1109/ICASSP.1987.1169674>>. Cited in page 28.

ARATYN, H.; RASINARIU, C. *A short course in mathematical methods with Maple*. River Edge, NJ, USA: World Scientific Publishing Company, 2005. Disponível em: <<https://doi.org/10.1142/5918>>. Cited in page 53.

AYIENDA, K. C. *Gamma and Related Distributions*. Nairobi: BoD–Books on Demand, 2014. Cited 2 times on pages 40 and 50.

BAHL, P.; PADMANABHAN, V. N.; BALACHANDRAN, A. Enhancements to the radar user location and tracking system. *Microsoft Research*, Franz Steiner Verlag, London, v. 2, n. MSR-TR-2000-12, p. 775–784, 2000. Cited in page 29.

BAJAJ, R.; RANAWEERA, S. L.; AGRAWAL, D. P. Gps: location-tracking technology. *Computer*, IEEE, v. 35, n. 4, p. 92–94, 2002. Disponível em: <<https://ieeexplore.ieee.org/document/993780>>. Cited in page 28.

BECK, A.; STOICA, P.; LI, J. Exact and approximate solutions of source localization problems. *IEEE Transactions on signal processing*, IEEE, v. 56, n. 5, p. 1770–1778, 2008. Cited in page 28.

CASELLA, G.; BERGER, R. L. *Statistical inference*. Canada: Duxbury Pacific Grove, 2002. v. 2. Cited 2 times on pages 37 and 38.

CHAN, Y.-T.; HO, K. A simple and efficient estimator for hyperbolic location. *IEEE Transactions on signal processing*, IEEE, v. 42, n. 8, p. 1905–1915, 1994. Cited in page 28.

DECARLO, L. T. On the meaning and use of kurtosis. *Psychological methods*, American Psychological Association, v. 2, n. 3, p. 292, 1997. Cited in page 37.

EFRON, B. Bayesians, frequentists, and scientists. *Journal of the American Statistical Association*, Taylor & Francis, v. 100, n. 469, p. 1–5, 2005. Cited in page 41.

EVANS, M.; SWARTZ, T. Methods for approximating integrals in statistics with special emphasis on bayesian integration problems. *Statistical science*, JSTOR, p. 254–272, 1995. Disponível em: <<https://doi.org/10.1214/ss/1177009938>>. Cited in page 61.

FANG, B. T. Simple solutions for hyperbolic and related position fixes. *IEEE transactions on aerospace and electronic systems*, IEEE, v. 26, n. 5, p. 748–753, 1990. Cited in page 28.

GALATI, G. et al. Wide area surveillance using ssr mode s multilateration: advantages and limitations. In: *Radar Conference, 2005. EURAD 2005. European*. Spain: IEEE, 2005. p. 225–229. Cited in page 29.

GARDINER, C. *Stochastic methods*. Germany: springer Berlin, 2009. v. 4. Cited in page 40.

GAVIRIA, I. A. M. *New strategies to improve multilateration systems in the air traffic control*. Tese (Doutorado) — Editorial Universitat Politècnica de València, Spain, 2013. Cited 2 times on pages 29 and 30.

GELMAN, A. et al. *Bayesian data analysis*. CRC press Boca Raton, FL, 2014. v. 2. Disponível em: <<https://doi.org/10.1201/b16018>>. Cited 3 times on pages 41, 43, and 44.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. MIT Press, 2016. Disponível em: <<http://www.deeplearningbook.org>>. Cited 2 times on pages 35 and 49.

GUPTA, R. D.; KUNDU, D. Exponentiated exponential family: an alternative to gamma and weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, Wiley Online Library, v. 43, n. 1, p. 117–130, 2001. Cited in page 40.

HE, X. *A Probabilistic Framework for Multi-Sensor Fusion Based Indoor Localization on Mobile Platform*. Tese (Doutorado), Oakland University, 2017. Cited in page 28.

HELEN, M. et al. Using calibration in rssi-based location tracking system. In: *Proc. of the 5th World Multiconference on Circuits, Systems, Communications & Computers (CSCC20001)*. London: Franz Steiner Verlag, 2001. Cited in page 29.

HOBSON, A. Just the maths. *Coventry University, Engineering and Computing*. <http://www.coventry.ac.uk/jtm/5/Lec5p1.pdf> [31 January 2008], 2002. Cited in page 37.

HUSAK, G. J.; MICHAELSEN, J.; FUNK, C. Use of the gamma distribution to represent monthly rainfall in africa for drought monitoring applications. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, Wiley Online Library, v. 27, n. 7, p. 935–944, 2007. Cited in page 40.

LEHMANN, E. L.; CASELLA, G. *Theory of point estimation*. Canada: Springer Science & Business Media, 2006. Cited 2 times on pages 35 and 41.

MANTILLA-GAVIRIA, I. A. et al. Localization algorithms for multilateration (mlat) systems in airport surface surveillance. *Signal, Image and Video Processing*, Springer, v. 9, n. 7, p. 1549–1558, 2015. Cited 2 times on pages 27 and 29.

NAGANAWA, J.; MIYAZAKI, H.; TAJIMA, H. Detection probability estimation model for wide area multilateration. In: *IEEE. Integrated Communications, Navigation and Surveillance Conference (ICNS), 2017*. Germany: Springer, 2017. p. 2B1–1. Cited 3 times on pages 27, 29, and 63.

NORRDINE, A. An algebraic solution to the multilateration problem. In: *Proceedings of the 15th International Conference on Indoor Positioning and Indoor Navigation, Sydney, Australia*. Germany: Springer, 2012. v. 1315. Cited in page 28.

PASCHALIDIS, I. C.; LI, K.; GUO, D. Model-free probabilistic localization of wireless sensor network nodes in indoor environments. In: *Mobile Entity Localization and Tracking in GPS-less Environments*. Berlin, Heidelberg: Springer, 2009. p. 66–78. Cited in page 29.

PATEL, J. K.; READ, C. B. *Handbook of the normal distribution*. New York: CRC Press, 1996. v. 150. Cited in page 38.

- PESONEN, H.; PICHÉ, R. Numerical integration in bayesian positioning. In: *Progress in Industrial Mathematics at ECMI 2006*. [S.l.: s.n.]. p. 908–912. Cited in page 44.
- PINEDA-BRISE, A. et al. A probabilistic approach to location estimation in manets. *Ad Hoc & Sensor Wireless Networks*, Old City Publishing, France, v. 28, p. 97–114, 2015. Cited 4 times on pages 27, 28, 29, and 63.
- RAMADURAI, V.; SICHITIU, M. L. Localization in wireless sensor networks: A probabilistic approach. In: *International conference on wireless networks*. New York, NY, USA: ACM, 2003. p. 275–281. Disponível em: <<https://doi.org/10.1145%2F1236360.1236368>>. Cited 2 times on pages 29 and 63.
- RAMSEY, F. P. Truth and probability. In: *Readings in Formal Epistemology*. [S.l.]: Springer, 2016. p. 21–45. Cited in page 35.
- RUDER, S. An overview of gradient descent optimization algorithms. *CoRR*, abs/1609.04747, 2016. Disponível em: <<http://arxiv.org/abs/1609.04747>>. Cited in page 54.
- SAVVIDES, A.; PARK, H.; SRIVASTAVA, M. B. The n-hop multilateration primitive for node localization problems. *Mobile Networks and Applications*, Springer, Germany, v. 8, n. 4, p. 443–451, 2003. Cited in page 27.
- SMITH, A.; HULSTROM, R. *Report on EATMP TRS 131/04*. Euro-control, 2005. Disponível em: <<https://pdfs.semanticscholar.org/6b06/9be56e155b32cec318a30f49a83150f5e372.pdf>>. Cited 2 times on pages 28 and 30.
- STAFF, C. C. *Triangulation, Trilateration, or Multilateration?(EE Tip# 125)*. France: Old City Publishing, 2014. Cited in page 27.
- STEFANŃSKI, J. Hyperbolic position location estimation in the multipath propagation environment. In: *Wireless and Mobile Networking*. Berlin, Heidelberg: Springer, 2009. p. 232–239. Cited in page 28.
- VASCONCELOS, D. *Probabilistic Multilateration*. Brazil: GitHub, 2018. <<https://github.com/DanielJean007/ProbabilisticMultilateration>>. Cited in page 65.
- VOGEL, K. A surveying problem travels from china to paris. *From China to Paris: 2000 Years Transmission of Mathematical Ideas*, Franz Steiner Verlag, London, v. 46, p. 1, 2002. Cited in page 27.
- WESTFALL, P. H. Kurtosis as peakedness. *The American Statistician*, Taylor and Francis, v. 68, n. 3, p. 191–195, 2014. Disponível em: <<https://doi.org/10.1080/00031305.2014.917055>>. Cited in page 37.
- YAO, K. et al. Blind beamforming on a randomly distributed sensor array system. *IEEE Journal on Selected Areas in Communications*, IEEE, v. 16, n. 8, p. 1555–1567, 1998. Cited in page 28.
- ZEKAVAT, R.; BUEHRER, R. M. *Handbook of position location: Theory, practice and advances*. IEEE, 2011. v. 27. Disponível em: <<https://ieeexplore.ieee.org/servlet/opac?bknumber=8633728>>. Cited in page 28.

Appendix

APPENDIX A – Complete list for Gamma Distribution Parameters

The complete list of parameters for the Gamma distribution and their variations is described below.

Table 6 – Parameters for the Gamma distribution and their variations

Assumption	Technique	Parameter	Constraint
$E[\delta^{(k)} \mid \mathbf{q}, \mathbf{r}^{(k)}]$	Chi-Square	$\alpha_k = 2\ \mathbf{q} - \mathbf{r}^{(k)}\ $ $\beta_k = 2$	None
	Fixed variance ν	$\alpha_k = \frac{\ \mathbf{q} - \mathbf{r}^{(k)}\ ^2}{\nu}$ $\beta_k = \frac{\nu}{\ \mathbf{q} - \mathbf{r}^{(k)}\ }$	$\nu > 0$
	Normal approximation	$\alpha_k = \frac{\ \mathbf{q} - \mathbf{r}^{(k)}\ ^2}{\alpha_k} + \frac{6}{t-3}$ $\beta_k = \frac{\nu}{\alpha_k}$	$\nu > 0, t > 3$
$\text{Mode}[\delta^{(k)} \mid \mathbf{q}, \mathbf{r}^{(k)}]$	Chi-Square	$\alpha_k = 2\ \mathbf{q} - \mathbf{r}^{(k)}\ - 4$ $\beta_k = 2$	$\delta^{(k)} > 2$
	Fixed variance ν	$\alpha_k = \frac{1}{2\nu} \sqrt{4\nu\ \mathbf{q} - \mathbf{r}^{(k)}\ ^2 + \ \mathbf{q} - \mathbf{r}^{(k)}\ ^4} + 2\nu + \ \mathbf{q} - \mathbf{r}^{(k)}\ ^2$ $\beta_k = \frac{\ \mathbf{q} - \mathbf{r}^{(k)}\ }{\alpha_k - 1}$	$\nu > 0$
	Normal approximation	$\alpha_k = \frac{6}{t-3} + \frac{1}{2\nu} \sqrt{4\nu\ \mathbf{q} - \mathbf{r}^{(k)}\ ^2 + \ \mathbf{q} - \mathbf{r}^{(k)}\ ^4} + 2\nu + \ \mathbf{q} - \mathbf{r}^{(k)}\ ^2$ $\beta_k = \frac{\ \mathbf{q} - \mathbf{r}^{(k)}\ }{\alpha_k - 1}$	$\nu > 0, t > 3$

Source: Author.

Annex

ANNEX A – Partial Derivatives for the Gamma Distribution

For MAP inference computation, partial derivatives with respect to $\mathbf{q}$ are necessary. Following we show derivations for the Gamma distributions for the three types of parametrization discussed in Section 3.2 and with distance observations $\delta^{(k)}$ centered at the mean of each case.

A.1 Gradient for measurements centered at the mean

Table 7 – Gradient for optimization functions with $\delta^{(k)}$ centered at the mean.

Likelihood function	
$\frac{\partial}{\partial \mathbf{q}} \log g(\mathbf{q}) = \frac{\partial}{\partial \mathbf{q}} \left(\log [p(\mathbf{q})] + \log [\prod_k p(\delta^{(k)} \mathbf{q})] \right)$	
Gamma, with $\alpha_k = \ \mathbf{q} - \mathbf{r}^{(k)}\ /2$ and $\beta_k = 2$	
$\frac{\partial}{\partial \mathbf{q}} \log g(\mathbf{q}) = \sum_k \left(\frac{(\mathbf{q} - \mathbf{r}^{(k)})^T}{2\ \mathbf{q} - \mathbf{r}^{(k)}\ } \left[\log \left(\frac{\delta^{(k)}}{2} \right) - \psi \left(\frac{\ \mathbf{q} - \mathbf{r}^{(k)}\ }{2} \right) \right] \right) - \Sigma^{-1}(\mathbf{q} - \boldsymbol{\mu})^T$	
Gamma, with $\text{Var}[\cdot] = \nu$	
$\begin{aligned} \frac{\partial}{\partial \mathbf{q}} \log g(\mathbf{q}) = \sum_k \left(\frac{(\mathbf{q} - \mathbf{r}^{(k)})^T}{\nu \ \mathbf{q} - \mathbf{r}^{(k)}\ } \left[2\ \mathbf{q} - \mathbf{r}^{(k)}\ \left[\log \left(\frac{\delta^{(k)} \ \mathbf{q} - \mathbf{r}^{(k)}\ }{\nu} \right) - \right. \right. \right. \\ \left. \left. \left. \psi \left(\frac{\ \mathbf{q} - \mathbf{r}^{(k)}\ ^2}{\nu} \right) \right] - \delta^{(k)} + \ \mathbf{q} - \mathbf{r}^{(k)}\ \right] \right) - \Sigma^{-1}(\mathbf{q} - \boldsymbol{\mu})^T \end{aligned}$	
Gamma, with $\text{Var}[\cdot] \leq \nu$ and $\text{Kur}[\cdot] = 3.2$	
$\begin{aligned} \frac{\partial}{\partial \mathbf{q}} \log g(\mathbf{q}) = \sum_k \left(\frac{(\mathbf{q} - \mathbf{r}^{(k)})^T}{\nu \ \mathbf{q} - \mathbf{r}^{(k)}\ ^3} \left[2\ \mathbf{q} - \mathbf{r}^{(k)}\ ^3 \left[\log \left(\frac{\delta^{(k)} (\ \mathbf{q} - \mathbf{r}^{(k)}\ ^2 + 30\nu)}{\nu \ \mathbf{q} - \mathbf{r}^{(k)}\ } \right) - \right. \right. \right. \\ \left. \left. \left. \psi \left(\frac{\ \mathbf{q} - \mathbf{r}^{(k)}\ ^2}{\nu} + 30 \right) \right] - (\delta^{(k)} - \ \mathbf{q} - \mathbf{r}^{(k)}\) (\ \mathbf{q} - \mathbf{r}^{(k)}\ ^2 - 30\nu) \right] \right) - \Sigma^{-1}(\mathbf{q} - \boldsymbol{\mu})^T \end{aligned}$	
Source: Author.	

An important remark is that $\psi(\cdot)$ is the same as $\psi^{(0)}(\cdot)$. It is important to point out that this function only supports positive values. Moreover, $\nu \in \mathbb{N}^*$ represents the maximum variability around the mean.