MASTER

Automatic commissioning of wireless sensor networks

Gong, L.

Award date: 2016

Link to publication
Automatic Commissioning of Wireless Sensor Networks

Master Thesis

Li Gong
l.gong@student.tue.nl

Supervisors:
prof. dr. J.J. Lukkien <j.j.lukkien@tue.nl>
ir. X. Wang <xiangyu.wang@philips.com>
dr. J. Schmaltz <j.schmaltz@tue.nl>

Eindhoven, August 2016
Abstract

Commissioning of luminaires in lighting systems is an essential step to make lighting systems work as desired. However, commissioning is often considered complex and error-prone. The ultimate is a system that can commission itself automatically without or with limited human intervention.

In this Masters thesis, we exploit the distance measurement capability provided by Atmel IEEE802.15.4 radios as an enabler to achieve automatic localization as part of commissioning. The overall flow of the localization process involves two major steps. First, there is a distance measurement step where the distances between luminaires are collected. Second, with the help of a known installation map and the distance results obtained in the first step, every luminaire is then localized.

There are three major contributions of the Masters thesis. First, certain methods in advanced filtering and averaging of distance measurement data are investigated and found to be beneficial in achieving improved accuracy. Second, three localization methods based on graph matching are studied. All the three methods give satisfactory results when dual antennas are used per luminaire. Third, we put this all together into a robust localization scheme that can be included in the commissioning procedure.
Preface

First of all, I would like to thank my supervisor prof.dr. J.J. Lukkien from the Eindhoven University of Technology for his continuous guidance and scientific support, especially for providing theoretical inspection and correction. Beside my supervisor, I would like to thank my tutor ir. X. Wang from Philips Lighting Research. During the master project, Mr. Wang gave me a lot of insights, as well as the knowledge of related topics. He was always ready for help and guiding me in a clear direction. My sincere thanks also go to my colleague P. Lin, who worked in another similar project and provided me many ideas. Last but not least, I would like to thank my family especially my Mum for supporting me spiritually during the thesis writing and graduation.
# Contents

## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Figures</td>
<td>ix</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xi</td>
</tr>
<tr>
<td>Acronym</td>
<td>xiii</td>
</tr>
</tbody>
</table>

## 1 Introduction

1.1 Background ........................................... 1
1.2 Current Luminaire Mapping Methods .................... 2
1.3 Motivation of Thesis .................................. 3
1.4 Contributions ......................................... 3
1.5 Organization .......................................... 4

## 2 Problem Statement

2.1 Application Scenario ................................. 5
2.2 Problem Statement: .................................. 6
2.3 Related Work ........................................ 6

## 3 Method Design

3.1 Measure Distance to Generate Adjacency Matrix .. 9
3.2 Localization by Graph Matching Methods ............ 11

## 4 Distance Estimation by Phase Measurement

4.1 AT86RF233 and Basic Ranging Principle .............. 13
4.2 Performance Exploration and Evaluation .............. 15
   4.2.1 Experiments on Ranging Parameters .............. 15
   4.2.2 Experiments in Different Environments .......... 20
   4.2.3 Experiments Summary and Overview .............. 21
4.3 Filtering and Averaging ................................ 25

## 5 Localization by Graph Matching

5.1 Localization by Point Matching ...................... 33
   5.1.1 Estimated Layout Generation ...................... 34
   5.1.2 Point Matching .................................. 36
5.2 Localization by Iterative Distance Matching ....... 37
   5.2.1 Distance Matrix Generation ...................... 37
   5.2.2 Iterative Distance Matching ...................... 38
5.3 Localization by Heuristic Edge Matching ............ 40
   5.3.1 Heuristic Edge Matching ......................... 40

Automatic Commissioning of Wireless Sensor Networks  vii
List of Figures

1.1 Mapping physical IDs to logical IDs. ........................................... 1
1.2 Luminaire mapping by using barcodes. ........................................ 3
3.1 Overall workflow. ........................................................................... 9
3.2 Adjacency matrix generation by thresholding. ............................... 10
3.3 Process of generating the adjacency matrix. ................................. 11
3.4 Process of localization. ................................................................. 12
3.5 An example of a symmetric floor plan. ......................................... 12
4.1 Hardware platform of AT86RF233 ranging evaluation kit. .......... 13
4.2 Phase measurement principle. ....................................................... 14
4.3 Distance measurement results. ..................................................... 15
4.4 Testbed 1 in HTC34.1.044. .......................................................... 16
4.5 Distance measurement results with different frequency step size. .... 17
4.6 Distance measurement results with different sweeping bandwidth. .. 18
4.7 Antenna rotation experiment setup. ............................................. 19
4.8 Antenna radiating patterns. ........................................................... 19
4.9 Distance measurement results with different antenna rotation. ...... 20
4.10 Distance measurement results with different multipath environments. 21
4.11 Overview of ranging performance. ............................................. 23
4.12 Distance measurement results in histograms. .............................. 24
4.13 Filtering and averaging methods. ............................................... 25
4.14 A two-level adaptive DQF filtering. ........................................... 26
4.15 Distribution of distance measurement ($D_{E,I}$). ......................... 26
4.16 Distribution of distance measurement ($D_{C,D}$). ......................... 27
4.17 An example of BottomSelection filtering ($D_{E,I}$) ...................... 28
4.18 An example of the Avg-Std. ....................................................... 29
4.19 Overview of different filtering and averaging methods in Testbed 1. 30
4.20 Testbed 2 with the unit distance of 170 cm. ............................. 31
4.21 Overview of different filtering and averaging methods in Testbed 2. 32
5.1 Workflow of localization by points matching. ............................. 33
5.2 Work flow of estimated layout generation. .................................. 35
5.3 An example of layout generation without orientation transformation. 35
5.4 An example of orientation transformation. ................................ 36
5.5 Work flow of iterative distance matching. .................................. 37
5.6 Iterative distance matching with three anchors. .......................... 38
5.7 Iterative distance matching of Location 6. ................................... 39
5.8 Localization result of the first iteration. .................................... 39
5.9 Localization of the second iteration. .......................................... 39
5.10 (a) Mapping results of third iteration and (b) final results. ........ 40
5.11 Work flow of localization by heuristic edge matching. .............. 41
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Percentage of correctness of adjacency matrix.</td>
<td>44</td>
</tr>
<tr>
<td>6.2</td>
<td>Localization results of Method 1.</td>
<td>45</td>
</tr>
<tr>
<td>6.3</td>
<td>Estimated layout of Adjacency matrix 7.</td>
<td>46</td>
</tr>
<tr>
<td>6.4</td>
<td>Localization results by different localization methods.</td>
<td>47</td>
</tr>
<tr>
<td>6.5</td>
<td>Localization results' cost of Method 2.</td>
<td>48</td>
</tr>
<tr>
<td>6.6</td>
<td>Estimated layout of Adjacency matrix 7.</td>
<td>48</td>
</tr>
<tr>
<td>6.7</td>
<td>Cost calculation based on Method 3.</td>
<td>49</td>
</tr>
<tr>
<td>6.8</td>
<td>Distance generation example.</td>
<td>49</td>
</tr>
<tr>
<td>6.9</td>
<td>Mapping results by different objective functions in Method 3.</td>
<td>50</td>
</tr>
<tr>
<td>6.10</td>
<td>Mapping results by different localization methods.</td>
<td>51</td>
</tr>
<tr>
<td>6.11</td>
<td>Estimated layout of Adjacency matrix 7 and 11.</td>
<td>52</td>
</tr>
<tr>
<td>6.12</td>
<td>Estimated layout of Adjacency matrix 9 and 12.</td>
<td>52</td>
</tr>
<tr>
<td>A.1</td>
<td>Phase difference measurement.</td>
<td>58</td>
</tr>
<tr>
<td>B.1</td>
<td>Pseudo-code of Fruchterman and Reingold force-directed graph drawing algorithm.</td>
<td>60</td>
</tr>
</tbody>
</table>
List of Tables

1.1 Mapping table between physical IDs and logical IDs. .................................. 2
4.1 Some default ranging setups. .......................................................... 14
Acronym

AOA Angle of Arrival
DQF Distance Quality Factor
LOS Line of Sight
NLOS Non Line of Sight
RF Radio Frequency
RIPS Radio Interferometric Positioning System
RSSI Radio Signal Strength Indicator
SRIPS Statistical Radio Interferometric Positioning System
TDOA Time Difference of Arrival
WSNs Wireless Sensor Networks
WLAN Wireless Local Area Network
Chapter 1

Introduction

In general, commissioning is to assure that the whole system and its components of an industrial plant or something newly produced are designed, installed, tested, operated, and maintained according to the operational requirements and reach a right working condition. As for the automatic commissioning of Wireless Sensor Networks (WSNs), it means the commissioning process is realized automatically without or with limited manual work, and it is based on wireless communication. Chapter 1 firstly introduces the background of this topic and its specific scenario in wireless lighting systems. Then, it lists several current methods for commissioning and introduces the motivation of the thesis. In the end, it summarizes the contributions and the organization of the thesis.

1.1 Background

Lighting systems in a modern building typically comprise lamps, sensors, control devices, and power drivers. Such systems are no longer a single wall switch and a bulb. They consist of hundreds of luminaries, sensors, and control devices. These discrete components need to be configured correctly during installation and commissioning. More precisely, the components are installed on the ceilings and mapped to a floor plan (e.g., an AutoCAD drawing). The floor plan specifies each device type, its logical ID (location) and its physical connections to the network. When the devices are installed and connected to the network, we need to figure out physical IDs (MAC address) of the devices. The physical IDs of the devices have to be mapped to the logical IDs (see Fig. 1.1 and Tab. 1.1). Then, appropriate control connections can be established between the devices properly.

![Mapping physical IDs to logical IDs](image)

Figure 1.1: Mapping physical IDs to logical IDs.

In conventional lighting systems, the commissioning work is still performed manually. During the commissioning procedure, an electrician first installs the physical luminaries and then manually
records the locations of the luminaries on the floor plan one by one. In this way, the commissioning work is quite tedious and time-consuming. Furthermore, the repetitive commissioning work often leads to human caused errors such as data input mistakes. The errors will result in wrong control connections, i.e., sending commands to incorrect luminaries. In this case, the commissioning errors will lead to additional losses such as more energy consumption, influence on occupant comfort and safety. In the end, errors need to be inspected and fixed by a highly skilled engineer thereby increasing the commissioning costs.

### 1.2 Current Luminaire Mapping Methods

The majority of luminaire commissioning method is performed manually. The commissioning method is a luminaire mapping method. The (manual) method requires electricians to turn on the luminaire one by one and figure out the identification one by one. In a complex lighting system, the workload will be so large that it is easy to lead to human-caused errors. Using barcodes is an alternative mapping method. Every luminaire has a unique barcode. Electricians first affix the barcodes on the floor plans during the installation. Then they scan the barcodes manually to map the luminaires’ physical IDs to logical IDs (Fig. 1.2). To some extent, this method is easier, but it still requires electricians to walk through the whole building. When the process involves tens or hundreds of luminaires, it is prone to errors and is still a time-consuming task.

In the past ten years, there have been attempts to auto-commissioning by indoor positioning methods. The idea is to map the luminaries’ physical IDs to the logical IDs by positioning results. Most of the methods are using radio frequency-based localization methods including Received Signal Strength Indicator (RSSI), Time-Different-of-Arrival (TDOA), and Angle-of-Arrival (AOA) [22] [14]. Investigations for such indoor localization methods are still ongoing. For RSSI, the principle is to translate the received signal strength to distance based on theoretical or empirical models. In the indoor environment, there exist problems such as multipath fading, background interference, and irregular signal propagation characteristics. Due to these problems, RSSI-based localization method can not guarantee the accuracy of the positioning results. For TDOA, these systems are based on measurement of time difference of arrival of a signal transmission to more than two receivers. The positioning result is accurate, but such systems often require extensive hardware which is both costly and energy-consuming. For AOA, the localization principle is based on measurement of the angle of incidence where signals arrive at the receivers. Similar to TDOA, AOA systems need complex hardware. Because of the above constraints and drawbacks, these indoor positioning methods are still under exploration and are rarely employed in real implementation.

In the thesis, we use another indoor positioning method Phase-of-Arrival (POA) [21] based on a radio transceiver AT86RF233 [2]. The basic software architecture and Application Program-

<table>
<thead>
<tr>
<th>Logical ID</th>
<th>Physical ID</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>$(x_1, y_1)$</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>$(x_2, y_2)$</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>$(x_3, y_3)$</td>
</tr>
<tr>
<td>4</td>
<td>D</td>
<td>$(x_4, y_4)$</td>
</tr>
<tr>
<td>5</td>
<td>E</td>
<td>$(x_5, y_5)$</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>$(x_6, y_6)$</td>
</tr>
<tr>
<td>7</td>
<td>G</td>
<td>$(x_7, y_7)$</td>
</tr>
<tr>
<td>8</td>
<td>H</td>
<td>$(x_8, y_8)$</td>
</tr>
<tr>
<td>9</td>
<td>I</td>
<td>$(x_9, y_9)$</td>
</tr>
<tr>
<td>10</td>
<td>J</td>
<td>$(x_{10}, y_{10})$</td>
</tr>
<tr>
<td>11</td>
<td>K</td>
<td>$(x_{11}, y_{11})$</td>
</tr>
<tr>
<td>12</td>
<td>L</td>
<td>$(x_{12}, y_{12})$</td>
</tr>
</tbody>
</table>

Table 1.1: Mapping table between physical IDs and logical IDs.
The transceiver uses phase difference of radio signals to estimate the distance. The hardware is not complex. The positioning accuracy is acceptable. Principles and performance analysis are described in Chapter 4.

Apart from using above methods for auto-commissioning, there are attempts to use photo sensors to achieve auto-localization. For instance in [5], Jonathan M. Barrilleaux proposes a methodology to determine the location map based on the direction of the light signal via directional sensors and photo sensors. Besides, they use a triangulation method for localization. However, the light signals will be blocked by the walls and ceilings, thereby this method is very limited in the indoor environment. Localization via triangulation algorithms is inaccurate and usually requires a significant number of reference nodes to get an acceptable result [10].

1.3 Motivation of Thesis

Commissioning in the lighting system is quite fundamental and significant since it is nearly the last step to present clients a fully functional-realized product. Also, a proper commissioning process is an essential step to decrease the manufacturing costs. For one thing, a proper commissioning helps avoid installation errors and reduce maintenance costs. For another thing, when the commissioning is automatically processed without or with limited manual work, plenty of labor costs can be reduced. There have been some attempts to extend the commissioning work to an automatic process by radio frequency (RF) technologies or photo sensors. However, the results of such attempts are not satisfactory due to various reasons such as accuracy, costs, or complexity. Therefore, to avoid installation errors and reduce commissioning costs, a new indoor positioning approach and an automatic commissioning method need to be proposed.

1.4 Contributions

In this thesis, we exploit the distance measurement (by phase difference) capability provided by Atmel IEEE 802.15.4 radios [1] as an enabler to achieve automatic commissioning. There are two major contributions in the thesis. First, certain methods in advanced filtering and averaging of distance measurement data are investigated and found to be beneficial in achieving improved accuracy. Second, three localization methods based on graph matching are studied. All the three
methods give satisfactory results when dual antennas are used per luminaire. To sum up, a robust localization scheme that can be included in the commissioning procedure are proposed and validated.

1.5 Organization

The thesis is organized as follows. Chapter 2 states the problem and formulates the commissioning problem as an assignment of nodes identifiers to locations (i.e., a localization problem). Based on the problem statement, Chapter 3 describes the overall workflow and introduces the designed approaches from a top level. Then, Chapter 4 illustrates the details of knowledge as well as solutions for distance estimation by phase measurement. Afterward, Chapter 5 specifies the localization methods in steps. In Chapter 6, it compares and evaluates the performance of the localization methods by different criteria. In the end, Chapter 7 ends the thesis with conclusions as well as future work.
Chapter 2

Problem Statement

Chapter 1 introduces the background of this thesis that is to solve the commissioning problem in the wireless lighting systems. The intended solution is based on indoor positioning by Phase-of-Arrival (POA). In this chapter, a more detailed application scenario will be described. Then, a brief summary and an explicit statement of this commissioning problem are illustrated. Afterward, related research literatures are listed.

2.1 Application Scenario

Modern buildings are integrated with multi-layered services infrastructures. The infrastructures often consist of various systems such as ventilation systems, heating systems, or lighting systems. In general, these systems need to be centrally integrated and configured during the installation for maintenance and controlling. In this thesis, the particular application scenario is the wireless lighting systems in Philips Lighting B.V. We want to simplify the maintenance and maximize the energy efficiency of the buildings. To be more specific, the initial concern falls in the automatic process of commissioning of lighting systems, which needs a proper connection and configuration.

With the enormous development of wireless communication, an increasing number of infrastructure systems including lighting systems are integrated with radio frequency devices. Consequently, the lighting systems are becoming wireless which brings lots of conveniences. The wires are eliminated to a large extent, which is extremely helpful during the installation and upgrade. However, the wireless connection also introduces difficulty of configuration. After the luminaires are installed and connected the network. In the next step, we need to identify every luminaire for controlling and maintenance. More precisely, the positions of the luminaires need to be figured out, and they need to be identified correctly in the installation layouts or floor plans.

Although the wireless connections result in the problem of difficult commissioning to some extent, this also introduces the opportunity to realize a kind of auto-commissioning based on wireless communication. It is the inspiration of this thesis research direction. Useful information can be directly derived from radio signals such as signal strength. For instance, RSSI-based ranging uses the signal strength for positioning. Alternatively, position information by distance estimation can be obtained through hidden information from radio signals(e.g., phase difference of radio signals). With the position information, the commissioning problem then can be solved easily.

In summary, the scenario of this thesis is to solve the localization as part of the commissioning problem. Precisely, it is to figure out which of the luminaries is in which location in the wireless lighting systems.
CHAPTER 2. PROBLEM STATEMENT

2.2 Problem Statement:

To avoid installation errors and reduce commissioning costs, an automatic commissioning method needs to be proposed. It is based on an indoor positioning approach by RF ranging. The commissioning problem is to work out which of the luminaries is in which location, so that appropriate control connections can be established between them.

In the application of a lighting system’s commissioning, the luminaires’ locations are fixed, and their positions can be obtained from floor plans. Therefore, the localization that computes and finds out the locations corresponding to particular luminaries can be replaced by finding the mapping between a set of luminaire IDs and a set of locations. As the luminaries are embedded with wireless devices, we also name luminaire IDs as node IDs.

To sum up, the problem specification is that we have a set of node IDs and a set of locations, and we would like to figure out the location of each ID. The formulation is shown as follows.

- \( L \) represents the physical floor plan, i.e., \( L \) is the set of physical locations. \( i \in L \) is represented by a pair of coordinates \((x_i, y_i)\). We refer also to elements of \( L \) as numbers: location \( i = (x_i, y_i) \). To simplify the problem analysis, the physical locations in the floor plan are in a grid layout.

- \( N \) is the set of names or IDs of the physical nodes. Also \( N \) is numbered.

- \( f \) is a localization function \( f : N \rightarrow L \), i.e., an assignment of nodes to locations.

- The actual localization, \( f_a \), maps nodes in \( N \) to their physical location in \( L \). This actual localization represents the state as it is: which physical node is found in which physical location. It is the localization we want to compute and it is the one we use as reference of quality. Without loss of generality we let numbering in \( L \) and \( N \) coincide through \( f_a \).

  Hence, \( f_a(i) = i \).

- \( \text{AvgDist} \) and \( \text{Accuracy} \) are two criteria to evaluate a localization. \( \text{AvgDist} \) is the average distance of a localization, i.e., \( \text{AvgDist}(f) = \frac{1}{|N|} \sum_{i \in N} ||f(i) - f_a(i)|| \). It is also the cost for the localization. \( \text{Accuracy} \) is another institute criterion, i.e., the percentage of correctly mapped nodes, \( \text{Accuracy}(f) = \frac{1}{|N|} \sum_{i \in N} \delta(f(i) - f_a(i)) \), where \( \delta(x) = \begin{cases} 1, & \text{if } x = (0, 0) \\ 0, & \text{otherwise} \end{cases} \).

The problem is to find the localization \( f \) (i.e., an assignment of nodes to locations) with minimal \( \text{AvgDist} \) and maximal \( \text{Accuracy} \); actually, the problem is to find \( f_a \).

2.3 Related Work

As stated in previous sections, the commissioning problem is a localization problem. Specifically, it is an indoor positioning problem based on sensor networks localization. Also, it is not a pure location computation but a location mapping problem. Related literature studies are listed as follows.

N. Patwari [20], H. Liu [15], G. Mao [16], et al. provide comprehensive overviews as well as general performance comparisons of the most sensor networks localization techniques including RSSI, TDOA, AOA. The localization schemes and evaluation are application specific. The criterion can be accuracy, precision, complexity, scalability, robustness, and cost. For RSSI, the principle is to translate the received signal strength to distance based on theoretical or empirical radio propagation models. RSSI benefits from its simplicity and low-cost. It does not need extra hardware. The tradeoff is an inaccurate measurement result on the order of several meters [4]. A general radio propagation model (i.e., Log-normal Shadowing Model [9]) is widely developed in this scheme. However, due to the complex indoor environments, the path-loss model does not always hold which makes the RSSI scheme unreliable. For TDOA, it computes the difference between arrival times to obtain distance by using two different signals with different propagation...
speeds. TDOA can address the time synchronization problem of Time-of-Arrival (TOA) and achieve a decent accuracy, especially under Line-of-Sight (LOS) conditions [3]. The downside of TDOA is that this technique requires calibration to ensure performance and needs expensive as well as energy-extensive hardware. For AOA, this method can achieve the accuracy to within a few degrees [23]. Unfortunately, AOA has a similar problem like TDOA. It requires even more expensive hardware in comparison with TDOA, and it cannot perform well without enough space for spatial diversity [20] [3].

Nikitin et al. use RF phase information to determine position and velocity of RFID tags in [19]. However, the authors show that the user case of this particular method and the RFID is limited to the indoor environment with Light-of-Sight (LOS) condition and fewer reflections such as a warehouse. Another attempt is that Bram et al. propose a statistical radio interferometric positioning system (SRIPS) which specified in [6]. This SRIPS approach addresses the limitation of the original Radio Interferometric Positioning System (RIPS), which are the shortage of available radio platforms, long measurement and calibration times. The authors enable the implementation of RIPS on a common platform CC2430 and achieve an accuracy of ±0.3 meters within 0.1 seconds in a 20 x 20m² outdoor environment. However, the authors present their results in an outdoor environment. They also show that their approach using the scalar propagation model ignores the polarization effects which will be encountered with obstacles in the indoor environments. Furthermore, their propagation model does not take multiple reflections into account. The SRIPS or RIPS cannot achieve a reliable localization performance in the indoor environments.

N. Jovanovic et al. [11] propose an iterative node localization approach for intelligent street lighting. Their scenario is similar to our topic that the authors translate the location computation problem to a location mapping problem. In the authors’ approach, they develop an iterative algorithm of graduated assignment based on graph matching. The mapping is established through a relation between RSSI properties and physical distance. From their test results, it proves the feasibility of this approach by a small error rate at around 10% within 200 nodes. Although the results from this scenario are acceptable, the authors indicate the requirement of some reference nodes. Another thing is that in the street lighting scenario, the average distance between nodes is far larger than that in the lighting systems. Furthermore, this scenario is fit in the outdoor environments, which has minimal or less multi-path problems.
Chapter 3

Method Design

Previous chapters introduce and describe the automatic commissioning problem in lighting systems. So as to improve the commissioning efficiency and avoid man-made errors, a method is proposed and designed in this chapter. This method is designed for automatic localization, which consists of two phases. First, there is a distance measurement phase where the distances between luminaires are collected. Second, with the help of a known floor plan and the distance results obtained in the first step, every luminaire is then localized. The overall workflow of the method is illustrated in this chapter, and the details are described in the following chapters.

The overall workflow of is shown in Fig. 3.1.

![Figure 3.1: Overall workflow.](image)

### 3.1 Measure Distance to Generate Adjacency Matrix

The problem of commissioning is to find the localization of node IDs to locations. It is a localization problem. The information we can retrieve from distance measurement is not where nodes are located but which nodes are neighbors. This neighborhood information represents a graph which needs to be mapped onto the (known) graph represented by the floor plan. Since the floor plan is given, we would like to use methods that map the adjacency graph to the floor plan graph to solve the localization problem instead of computing the nodes locations. The first phase of our
CHAPTER 3. METHOD DESIGN

approach is to generate an adjacency matrix, which represents the neighborhood information. The adjacency matrix is the input for graph matching methods in next subsection. Assume there are \( N \) luminaires. Then, the adjacency matrix \( \text{adj} : \mathbb{N} \times \mathbb{N} \to \{0, 1\} \) will be an \( N \times N \) matrix, where the element \( \text{adj}_{i,j} = 1 \) in the matrix represents luminaire \( i \) and luminaire \( j \) are within a certain distance. An example of adjacency matrix generation is shown in Fig 3.2 and it is illustrated in next paragraph.

![Adjacency matrix generation](image)

Figure 3.2: Adjacency matrix generation by thresholding.

Fig.3.3 shows the process of generating an adjacency matrix. In Step (a), the distance measurement is obtained based on phase measurement by the transceiver AT86RF233 [2]. Reasons to use this distance measurement technology are mainly for the sake of cost and accuracy. Fundamental principles and performance are discussed in Chapter 4. Step (b) is to filter and average raw measurement data. The filtering and averaging steps are based on distance measurement patterns. Details are presented in Chapter 4. Step (c) compares the measurement data with a distance threshold to generate the adjacency matrix (see Fig 3.2). When the distance between two luminaries (e.g., luminaire \( i,j \)) is less than the threshold, the two luminaries are neighbors (i.e., \( \text{adj}_{i,j} = 1 \)). Otherwise, they are not neighbors (i.e., \( \text{adj}_{i,j} = 0 \)). The distance threshold is decided by empirical analysis. The goal is to distinguish the measurements which are not from neighbors as much as possible.
3.2 Localization by Graph Matching Methods

The second phase applies graph matching methods between the graph represented by the computed adjacency matrix and the graph represented by the floor plan. The input is the computed adjacency matrix from Phase 1. The output is the localization function $f$. Fig 3.4 shows the overall workflow of Phase 2.

Step (a) pegs specific locations manually. A pegging is a mapping between node IDs and locations. More precisely, we map (localize) at least three nodes (i.e., three nodes can determine a plane) to their locations in the floor plan. This step is to avoid a symmetry problem in graph matching step. The localization methods are based on graph matching with the floor plan (see next step and Chapter 5). When the floor plan is rectangular symmetry (see Fig 3.5), the graph information such as adjacency information between Location 1, 2, 3, 4 and Location 9, 10, 11, 12 are identical. Therefore, a manual pegging is required to resolve this kind of symmetry problem. Also, the localization complexity is reduced after node pegging. Details are described in Chapter 5. The manual pegging is performed by existing luminaire mapping methods such as visually check or barcode scanning.

Step (b) applies graph matching methods. In this thesis, we introduce three different methods. The first one transforms an adjacency matrix into estimated locations. Then, a Hungarian algorithm for point matching is applied [18]. The second one uses an iterative matching based on distance properties. The last method employs a fast ant system algorithm [30] to heuristic search all the possible mappings based on the adjacency matrix and the distance properties. They are illustrated in Chapter 5. The localization (mapping) is evaluated by the two criteria $AvgDist$ and $Accuracy$ in Chapter 6.
Figure 3.4: Process of localization.

Figure 3.5: An example of a symmetric floor plan.
Chapter 4

Distance Estimation by Phase Measurement

This chapter describes Phase 1 which is distance estimation by phase measurement. First, it introduces the hardware (i.e., chipset AT86RF233) as well as its ranging principles. Then, a set of experiments is presented to evaluate the ranging performance (i.e., measuring accuracy) of the chipset. The last section illustrates several filtering and averaging methods.

4.1 AT86RF233 and Basic Ranging Principle

Fig 4.1 shows the chipset AT86RF233 for distance measurement [2]. The chipset consists of an extended radio board with two swivel antennas and a base controller board with batteries. The potential applications for this chipset are 2.4 GHz IEEE 802.15.4, ZigBee, ranging and localization applications.

![Figure 4.1: Hardware platform of AT86RF233 ranging evaluation kit.](image)

Fig 4.2 presents a basic scenario for the phase measurement. The principle of AT86RF233 distance measurement is based on repeating individual frequency measurements\(^1\). See Fig 4.2a, the initiator sends a signal at a frequency \(f_1\) while the reflector reflects another signal at frequency \(f_2\). Then, the chipset can measure a phase difference \(\Delta \varphi_1\). This procedure is repeated \(N\) times. The mathematical formula is shown as follows. Appendix A presents a basic deviation.

\(^{1}\)The principle is under a non-disclosure agreement (NDA). Only the basic principle is known.
CHAPTER 4. DISTANCE ESTIMATION BY PHASE MEASUREMENT

(a) Phase measurement devices

Figure 4.2: Phase measurement principle.

\[ d = \frac{c}{4 \cdot \pi} \cdot \frac{-1 \cdot \sum_{n=1}^{N-1} \Delta \phi_n}{(N-1) \cdot \Delta f} \] (4.1)

- \( d \) represents the distance between the initiator and the reflector,
- \( c \) represents the transmission speed of the radio signal,
- \( N - 1 \) represents the times of repeating individual frequency measurement,
- \( \Delta f \) represents the frequency offset,
- \( \Delta \phi_n \) represents the measured phase difference.

The software architecture and the Application Programming Interface (API) of the chipset are available. Based on the user manual [1], we can program a python script to invoke a peer-to-peer ranging measurement. For each measurement, users can select different ranging parameters. The major ranging parameters include frequency step size (i.e., the frequency offset \( \Delta f \)), sweeping bandwidth (i.e., the used frequency band after entire frequency measurements), and antenna diversity\(^1\). Tab 4.1 specifies the details of ranging parameters.

<table>
<thead>
<tr>
<th>Ranging setups</th>
<th>Default</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency Step Size</td>
<td>2</td>
<td>Ranging measurement frequency step in MHz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.5, 1, 2, 3]</td>
</tr>
<tr>
<td>Sweeping Bandwidth</td>
<td>40</td>
<td>Ranging measurement sweeping bandwidth in MHz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[5, 10, 20, 40, 80]</td>
</tr>
<tr>
<td>Antenna Diversity</td>
<td>1</td>
<td>Utilization of Antenna Diversity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0, 1]</td>
</tr>
</tbody>
</table>

Table 4.1: Some default ranging setups.

Fig 4.3 shows a measurement log indicating the distance measurement results. For each peer-to-peer ranging measurement, the chipset provides four distance measurements results corresponding
CHAPTER 4. DISTANCE ESTIMATION BY PHASE MEASUREMENT

to different antenna pairs. Based on the four pairs results, a weighted distance with pair number -1 is provided by the chipset. It is generated by a confidential algorithm provided by Atmel. Apart from the distance results, there is a DQF value\(^2\). It is a value to indicate the reliability of each distance measurement. A larger value indicates a more reliable distance measurement.

Figure 4.3: Distance measurement results.

4.2 Performance Exploration and Evaluation

The chipset is relatively new. There is no official evaluation about its ranging performance. Therefore, we need to conduct some experiments to explore and evaluate the chipset’s ranging performance.

A default testbed is shown in Fig 4.4. The location is Building 34 of High Tech Campus in Eindhoven. The testbed is a 4×3 network. The unit distance between the direct neighbors (i.e., nearest neighbors) is 100 cm.

4.2.1 Experiments on Ranging Parameters

First, we conduct three experiments to analyze the influence of different ranging parameters. The parameters are specified in Tab. 4.1. Antenna diversity analysis is replaced to antenna rotation analysis. The rotation is that the antenna of the chipset can be rotated to different directions. Default rotation is shown in Fig 4.4. Therefore, the three experiments are about the setup of frequency step size, sweeping bandwidth, and antenna rotation respectively. When there is no further indication, the experiments use default setups.

Note that the first two experiments are based on two central nodes of the testbed (i.e., Node \(F\) and Node \(G\)). The distance between them is 100 cm. Experiment on antenna rotation is based on four nodes (to include different orientation’s measurements). To reinforce the rotation impacts, the distance between direct neighbors is 170 cm. The four nodes are placed in the center of the room to mitigate reflections. For each peer-to-peer measurement, the sample size is 100.

\(^1\)Antenna Diversity is a feature to support the control of two antennas to select the most reliable RF signal path provide by the AT86RF233.

\(^2\)DQF is distance quality factor. It ranges from 0 to 100.
1) Setup of Frequency Step Size

The principle of the phase-based localization is by repeating individual frequency measurement. Therefore, the sweeping frequency step size plays a significant role in the ranging performance. For one thing, a smaller frequency step size indicates more repeating individual frequency measurements. For another thing, the frequency step size is in inverse proportion to distance (see $\Delta f$ in equation 4.1). The step size cannot be too large. Otherwise, it is not small enough to handle a large distance.

We set the first experiment based on four different frequency step size from 0.5 MHz to 1, 2, and 4 MHz. As mentioned, the experiment is conducted between two nodes with the distance of 100 cm. Explicit results contain 100 iterations distance measurements by different pairs (see Fig 4.3). Fig 4.5 shows the results in average and standard deviation by pairs.

According to the results, the frequency step size does influence the ranging performance. We evaluate the ranging performance based on the error to the actual distance. When the measurements are close to the actual distance 100 cm, it indicates the ranging performance is decent. For this experiment, when the step size is set larger than 1 MHz, the performance is becoming better and more stable in average results. However, the standard deviations increase slightly after step size 1 MHz. Hence, it is suggested to set the step size to 1 MHz in our testbed.
CHAPTER 4. DISTANCE ESTIMATION BY PHASE MEASUREMENT

(a) Average results
(b) Standard deviation results

Figure 4.5: Distance measurement results with different frequency step size.

2) Setup of Sweeping Bandwidth

The second experiment is about sweeping bandwidth. For every repeating individual frequency measurement, the initiator sweeps through a frequency step unit. When the initiator sweeps to the stop frequency, a complete measurement is finished. The sweeping bandwidth is the entire frequency range during the repeating individual measurement. Apparently, a larger sweeping bandwidth provides more frequency measurements. The final result will be more accurate due to the arithmetic average. Also, as the chipset uses the 2.4-GHz ISM band, co-existence and interference with other wireless devices need to be taken into account. A larger sweeping bandwidth indicates relatively fewer interference.

We set the second experiment based on five different sweeping bandwidths from 5 MHz to 10, 20, 40, and 80 MHz. Similarly, the experiment is conducted between two nodes with the distance of 100 cm. Fig 4.6 shows the results in average and standard deviation by pairs.

According to the results, the average distances are closer to actual distance 100 cm when the sweeping bandwidth is larger. The standard deviations become smaller with a larger sweeping bandwidth. Clearly, a larger sweeping bandwidth has a better ranging performance. Precisely, it is suggested to set to 40 MHz or 80 MHz. The results are as expected since a larger sweeping bandwidth grants more repeating measurements and less interference.
CHAPTER 4. DISTANCE ESTIMATION BY PHASE MEASUREMENT

(a) Average results

(b) Standard deviation results

Figure 4.6: Distance measurement results with different sweeping bandwidth.

3) Setup of Antenna Rotation

The third experiment is about antenna rotation. Different antenna rotations result in various multi-path patterns. The indoor Non-Line-of-Sight (NLOS) environment is typically small space and is easier to generate reflections. According to the user manual[2], it is suggested to rotate the two antennas to be perpendicular to each other. It is reasonable according to the principles of antenna theories [29]. For one thing, a perpendicular rotation can ensure the quality of RF links. In antenna theory, there is no signal in the direction of antenna pointing to which is called a null area (see Fig 4.8). The perpendicular rotation reduces the null area impacts by compensation. For another thing, the perpendicular rotation makes the multi-path patterns flatter. Similar to the benefit over RF links, the perpendicular rotation avoids uneven multi-path patterns through the compensation for each other.

We set the third experiment based on four antenna rotations. Fig 4.7 shows the four potential antenna rotations. The experiment is conducted among four nodes with the distance of 170 cm. Fig 4.9 shows the results in average and standard deviation by pairs. The average and standard deviation are based on all measurements of direct neighbors.

Likewise, when the average is close to actual distance and the standard deviation is small, the ranging performance is decent. From the results, Rotation 1 and Rotation 3 perform better than the other two rotations. It is as expected since the radio radiation patterns, as well as the multi-path patterns for these two rotations, are even and flat. Although the other two rotations’ radiation patterns are even, the multi-path patterns are not as even as Rotation 1 and Rotation...
CHAPTER 4. DISTANCE ESTIMATION BY PHASE MEASUREMENT

Figure 4.7: Antenna rotation experiment setup.

Figure 4.8: Antenna radiating patterns.
CHAPTER 4. DISTANCE ESTIMATION BY PHASE MEASUREMENT

(a) Average results

(b) Standard deviation results

Figure 4.9: Distance measurement results with different antenna rotation.

3 due to the reflections of walls. Between Rotation 1 and Rotation 3, Rotation 1 is slightly better from the results.

4.2.2 Experiments in Different Environments

The distance measurement is conducted in the indoor environment. Multi-path patterns are various in different environments. Moreover, the ranging performance will be influenced by different multi-path patterns. From the last subsection, several suggested ranging setups are found. Precisely, the frequency step size is 1 MHz. The sweeping bandwidth is 40 MHz. The antenna rotation is Rotation 1. Based on these setups, this subsection describes experiments in different environments.

We set the experiments based on three different environments. The first one is based on a static multi-path environment (i.e., no people or obstacles are moving around). The second one is based on a dynamic multi-path environment that one person is walking around in the room. The third one is based on a dynamic multi-path environment that there is an additional and mobile node under Node \( G \) (see Fig 4.4). The complete testbed 4.4 with 12 RF nodes is used. Likewise, Fig 4.10 shows the results in average and standard deviation by pairs. The average and standard deviation are based on all measurements of direct neighbors. The actual distance between them is 100 cm.

Considering the measurements by antenna pair 1, 2, 3, 4 are typically much larger than actual
distance, we focus on measurements by pair -1. The average results of Environment 1 and Environments 3 are similar. However, Environment 1 has a smaller standard deviation. Therefore, Environment 1 (i.e., the static multi-path environment) has better and more stable ranging performance. The reason is that either the walking people or the mobile node in dynamic multi-path environments will affect the multi-path patterns. As a result, the distance measurement contains more measurements of non-direct paths. If considering the wireless link quality, ranging performance under the environment with an additional and mobile node might have interference problems due to collisions. Therefore, to obtain better ranging performance, it is suggested to conduct the distance measurement under a static multi-path environment.

![Figure 4.10: Distance measurement results with different multipath environments.](image)

4.2.3 Experiments Summary and Overview

In the last two subsections, several ranging setups are explored, and some useful setups are identified. To achieve optimum ranging results, it is suggested to:

- apply ranging frequency management, i.e., select appropriate frequency step size and sweeping bandwidth.
- optimize antenna diversity to mitigate multi-path fading, i.e., use specific space diversity (antennas 90°).
- conduct distance measurement under a relatively static multi-path environment.

Automatic Commissioning of Wireless Sensor Networks 21
Apart from the three findings, there are other helpful setups which were found in the experiments. First, it is suggested to avoid placement of antennas near electromagnetic reflective surfaces (e.g., wall mounted nodes near the wall). The placement of node is preferred to be sufficiently high (e.g., wall mounted nodes near the ceiling). The reason is that such placements decrease the multi-path influence. Also, similar to the ranging frequency management, adaptive multiplexing management is necessary (e.g., to set an enough sweeping band different from the occupied bands). Detecting frequency bands occupied by WLAN, Bluetooth in advance helps avoid the collision problem.

An Overview of Ranging Performance

To evaluate measurement performance, we define several terminologies in this thesis. We translate the concept of distance into a step concept. One unit distance is one step, and the unit distance is the actual distance between direct neighbors. In Testbed 1, it is 100 cm. Therefore,

• 1 step connected nodes pairs: the distance between the two nodes is 100 cm. The two nodes are direct neighbors (see Fig 4.4, e.g., Node A and Node B is a pair of 1 step connected nodes, similar to Node A and Node E).

• 2 steps connected nodes pairs: the distance between the two nodes is 200 cm (see Fig 4.4, e.g., Node A and C is a pair of 2 steps connected nodes, similar to Node A and Node I).

• Cross connected nodes pairs: the distance between the two nodes is 141 cm. The two nodes are also 1.4 steps connected pairs. (see Fig 4.4, e.g., Node A and F is a pair of cross connected nodes, similar to Node B and Node E).

Similarly, there will be 3 steps, 4 steps, etc. connected nodes pairs. Due to the degrading performance of the larger steps measurements and for simplification, only the 1 step and the 2 steps measurement results are described to evaluate the ranging performance. As for the cross connected nodes pairs (e.g., Node A and Node F, Node A and Node G), for similar reasons, only the shortest cross connected nodes pairs (e.g., Node A and Node F) are described here.

The results overview is based on measurements of the Testbed 1 (see Fig 4.4) in the static multi-path environment. The ranging setups are 1 MHz frequency step size, 40 MHz sweeping bandwidth, and Rotation 1. The unit distance is 100 cm. Thus, 1 step is 100 cm, 2 steps is 200 cm, and 1.4 steps (cross connected) is 141 cm. The measurement sample size is 200 per pair (i.e., each direction 100 times per pair). The results are illustrated in Fig 4.11 by average and standard deviation of the three measurement types: 1 step connected pairs, 2 steps connected pairs, and cross connected pairs.

From the results, we find:

1. Expected averages of the three types data should be 100 cm, 200 cm and 141 cm respectively.
   The average error of 1 step connected pairs is less than 1 meter. The average error of 2 steps connected pairs is greater than 1 meter. The average error of cross connected pairs is greater than 2 meters. Obviously, the distance measurement results are noisy.

2. In terms of antenna pairs measurements, distance measurement results of pair 1,2,3,4 are worse than that of estimated pair -1.

3. In terms of pair types, distance measurement results of 1 step connected pairs are better than that of larger steps connected pairs; distance measurement results of cross connected pairs are worse than the others.

After the illustration by average and standard deviation, the detailed results of estimated pair -1 of the three measurement types are shown in Fig 4.12. As mentioned, we can get a DQF for each measurement. The legend \( \text{Distance data.dqf} > 90 \) indicates that the corresponding data is the measurement with a DQF larger than 90.
At a first glance, the measurement results of 1 step connected pairs are decent as many measurement results are close to the actual distance 100 cm. Also, the overall distributions tend to be narrow. The results are a bit worse concerning 2 steps and cross connected pairs. Besides, 1 step connected pairs have more measurements with good DQFs (i.e., $\text{Distancedata.dqf} > 90$), which means their measurements are more reliable.

To be more specific, the following observations are found:

1. Distance measurement results of 2 steps or cross connected pairs tend to be worse from the view of distribution, DQF, or peaks. Precisely, the distribution is wider. Fewer distance measurements have good DQFs. Peaks are not sharp.

2. Most distance measurement results tend to be larger than the actual distance, but there exist measurements shorter than the actual distance (e.g., $D_{G,H}$).

3. DQF indicator can tell the reliability information of the distance measurement results, but it is not a guarantee (e.g., $D_{I,J}$).

4. Horizontal distance measurement results (e.g., $D_{A,B}$) tend to be narrow and accurate, while vertical ones (e.g., $D_{A,E}$) tend to be wider and larger.

The results are observed as expected. As described in previous subsections, we explore the influence of antenna rotation and find that parallel links of specific rotations exert better performance. The parallel links pattern is the reason for Observation 1 and Observation 4 as well. As for Observation 2, the reason is that the distance measurements contain a lot of non-direct paths by multi-path influence. As for the shorter distance measurements, they are because of hardware errors.
CHAPTER 4. DISTANCE ESTIMATION BY PHASE MEASUREMENT

(a) Distance measurement results of 1 step connected pairs

(b) Distance measurement results of 2 steps connected pairs

(c) Distance measurement results of cross connected pairs (left diagonal)

Figure 4.12: Distance measurement results in histograms.
In summary, the overall ranging performance is considerably fair under a complex multi-path indoor environment, especially among 1 step connected pairs. Although the measurement results often contain non-direct paths measurements, a significant amount of the measurements is corresponding to or appropriate to the direct paths. The average accuracy is on the order of meters, while it is around 1 meter among 1 step connected pairs. The detailed accuracies will be presented and compared in the next section.

4.3 Filtering and Averaging

The previous sections describe the details of the distance estimation by phase measurement from its principles to its performance. From the performance evaluation, the accuracy is on the order of meters. The average error of direct neighbors (1 step connected pairs) is around 1 meter. Fig 4.12 shows more details. In general, the measurement data is noisy. On the one hand, the distributions are wide. On the other hand, most measurements are larger than the actual distance. Therefore, some filtering and averaging methods are required.

Fig 4.13 shows the main steps of filtering and averaging. These steps are described as follows.

![Diagram of Filtering and Averaging Methods]

- **DQF filtering**
  As mentioned (see Fig 4.3), the chipset provides users a distance quality factor (DQF) for each distance measurement. A higher DQF means the corresponding measurement is more reliable. From performance evaluation, this number is an easy and direct reliability indicator. Therefore, the first filtering method is to select out the distance measurements with good DQFs.

  Note that the DQF is not a guarantee (see Observation 3). Sometimes after this filtering, the sample size of the filtered data would be very small or even zero. To keep enough sample size and avoid misjudgment, this filtering is designed as an adaptive filtering method. Fig 4.14 shows a two-level adaptive DQF filtering. The first level sets the DQF threshold as 90 while the second level set it as 80. When the conditions are met, the measurement data is saved. Otherwise, the data is discarded.
CHAPTER 4. DISTANCE ESTIMATION BY PHASE MEASUREMENT

Figure 4.14: A two-level adaptive DQF filtering.

MinMax filtering

The distance measurement is in the indoor environments. It suffers from multi-path influence. In other words, a lot of measurements are based on reflection paths. Fig 4.15 shows the distribution of distance measurement between Node E and Node I. The actual distance between the two nodes is 100 cm. For one thing, we can find the distribution of measurement data is wide with a long tail. For another thing, most measurements are larger than the actual distance.

Figure 4.15: Distribution of distance measurement ($D_{E,I}$).

The second filtering method applies a minimal and maximal cut-off. The initial concern of MinMax filtering is to cut out those measurements which are too large or too small (see Fig 4.16). Therefore, a maximal threshold can be set as the maximal distance among the nodes in the network. Moreover, Phase 1 of distance measurement is to generate an adjacency matrix. It is required to figure out which node pairs are the direct neighbors. According to this purpose, the maximal threshold can be shortened. A possible maximal threshold is the actual distance of 2 steps. The maximal threshold setting needs to consider the sample size of measurements.

26 Automatic Commissioning of Wireless Sensor Networks
CHAPTER 4. DISTANCE ESTIMATION BY PHASE MEASUREMENT

It cannot be too small when there are not enough samples to be processed. As for the minimal threshold, it is typically set as 50% of the actual distance between the direct neighbors. The number is determined by experimental observation. We prefer not to filter out too many shorter measurements. The reason is that the distance measurement less than the actual distance mostly exists among the 1 step connected pairs. We do not want to filter out such measurements, but we cannot neglect the existing hardware errors. These errors lead to measurements equal to dozens of centimeters. Therefore, we set the minimal threshold between zero and the unit distance.

![Figure 4.16: Distribution of distance measurement (D_{C,D}).](image)

BottomSelection filtering

Similar to MinMax filtering, BottomSelection is another method to select out the distance measurements which are more appropriate to the direct paths. The principle is that the selection only picks the lowest measurement data. For example, if we bottom select the lowest 40% measurement data of $D_{E,I}$ (see Fig 4.17), the bars in red are the selected data after BottomSelection filtering. As for the threshold of this filtering, it also depends on the sample size of data. When the sample size is too small, the threshold can be relaxed. The reason is that we want to ensure enough data for the other filtering steps.
Averaging

One complete distance measurement often contains more than 100 samples. After previous filtering methods, a small number of distance measurements are kept. The last step applies an averaging strategy to average the retained distance measurements. Regarding the averaging methods, the simplest averaging is an arithmetical mean. We name it Avg. Besides Avg, we introduce another two methods which are WeightedAvg and Avg-Std:

- WeightedAvg: In every distance measurement the chipset provides a distance quality factor (DQF) which indicates the reliability of the measurement. In WeightedAvg, we translate this distance quality factor into weight information. For those measurements with higher DQFs more weight are placed and vice versa. More precisely, we first translate every DQF into a weighted DQF according to equation 4.2, and then we sum up the averaging distance by weighted DQF according to equation 4.3.

\[
DQF_{weighted,i} = \frac{DQF_i}{\sum_{i=1}^{n} DQF_i}
\]  

\[
d = \sum_{i=1}^{n} d_i \times DQF_{weighted,i}
\]

- Avg-Std: Avg-Std uses the subtraction between arithmetical average and standard deviation. This method is another strategy to bottom select the measurements and appropriate the measurements closer to actual distance. As mentioned, most measurements are much larger than the actual distance. After previous filtering steps, the arithmetical average of measurements is still often larger than the actual distance. Therefore, in the averaging step, we can apply this Avg-Std strategy to average the distance data closer to the actual distance. Assume the distribution of the distance data fits a normal distribution. Fig 4.18 shows an example of the Avg-Std. The principle of Avg-Std based on \(\mu - \sigma\) (see the following equations).

\[
\mu = \frac{1}{n} \sum_{i=1}^{n} d_i
\]  

\[
\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (d_i - \mu)^2}
\]
\[ d = \mu - \sigma \]  

(4.6)

**Comparison**

After a brief description of the filtering and averaging methods, two detailed comparisons of their performance are illustrated in Fig 4.19. The comparison criteria are based on mean and standard deviation among measurements of 1 step connected pairs in Fig 4.12a. The sample size is 200 per pair. There are 34 pairs in total. The actual distance of these 1 step connected pairs is 100 cm. Thus, in Fig 4.19a, when the mean is closer to 100 cm it indicates the better performance. As for standard deviation in Fig 4.19b, the lower standard deviation is better.

According to the overview figures, we observe:

1. In averaging methods (i.e., Avg, WeightedAvg, Avg-Std), Avg-Std has to be applied together with other filtering methods. Otherwise, the fluctuation due to the outliers would degrade its performance. A large fluctuation introduces a large standard deviation, which would lead to a small averaging number. This phenomenon is reflected in the methods involved Avg-Std (e.g., Avg-Std, DQF+Avg-Std, MinMax+Avg-Std, DQF+MinMax+Avg-Std).

2. MinMax filtering performs well in eliminating the outliers (e.g., WeightedAvg, Min-Max+WeightedAvg).

3. BottomSelection filtering performs well in appropriating the measurements to actual distance (e.g., WeightedAvg, Bottom+WeightedAvg).

4. Several good combinations of filtering and averaging methods are found (e.g., Min-Max+Bottom+WeightedAvg, DQF+MinMax+Bottom+WeightedAvg). It is preferred to apply all the filtering methods (i.e., DQF+MinMax+Bottom+WeightedAvg) if possible. The reason is that we want to avoid any potential outliers as much as possible.
CHAPTER 4. DISTANCE ESTIMATION BY PHASE MEASUREMENT

(a) Overview of different filtering and averaging methods by mean

(b) Overview of different filtering and averaging methods by standard deviation

Figure 4.19: Overview of different filtering and averaging methods in Testbed 1.

For the sake of validation, these filtering and averaging methods are applied to another testbed with the same layout but different unit distance of 170 cm (see Fig 4.20). Likewise, two results overview figures based on criteria of mean and standard deviation are shown in Fig 4.21. From the results, it validates the positive performance of these filtering and averaging methods.

In details,

1. For 1 step distance measurement in Testbed 1, baseline Avg of estimated pair -1 is 152.99 cm. As for individual pair 1,2,3,4, they are 140.88 cm, 219.47 cm, 160.98 cm.
cm, and 209.25 cm. When applying all the filtering and averaging methods (i.e., DQF+MinMax+Bottom+WeightedAvg), the number of estimated pair -1 becomes 98.02 cm. As for individual pair 1,2,3,4, they are 83.07 cm, 98.61 cm, 106.41 cm, and 96.96 cm.

2. For 1 step distance measurement in Testbed 2, baseline Avg of estimated pair -1 is 266.15 cm. As for individual pair 1,2,3,4, they are 340.94 cm, 424.71 cm, 313.78 cm, and 421.50 cm. When applying all the filtering and averaging methods (i.e., DQF+MinMax+Bottom+WeightedAvg), the number of estimated pair -1 becomes 179.78 cm. As for individual pair 1,2,3,4, they are 163.80 cm, 191.11 cm, 175.93 cm, and 182.96 cm.

The baseline is Avg. For 1 step distance measurement in Testbed 1, the average accuracy of the baseline is ±53 cm for estimated pair -1 and ±83 cm for individual pairs. When applying all the filtering and averaging methods (i.e., DQF+MinMax+Bottom+WeightedAvg), the average accuracy is ±2 cm for estimated pair -1 and ±7 cm for individual pairs. For 1 step distance measurement in Testbed 2, the average accuracy of the baseline is ±66 cm for estimated pair -1 and ±175 cm for individual pairs. When applying all the filtering and averaging methods (i.e., DQF+MinMax+Bottom+WeightedAvg), the average accuracy is ±10 cm for estimated pair -1 and ±12 cm for individual pairs.

To sum up, the above filtering and averaging methods improve the measuring accuracy significantly. The distance measurement based on estimated pair -1 is typically better than the individual pairs.
(a) Overview of different filtering and averaging methods by mean

(b) Overview of different filtering and averaging methods by standard deviation

Figure 4.21: Overview of different filtering and averaging methods in Testbed 2.
Chapter 5

Localization by Graph Matching

After filtering & averaging and thresholding (see Fig 3.2 in Chapter 3), an adjacency matrix is generated. This chapter describes methods of localization by graph matching with the adjacency matrix as input. In this thesis, we propose three different methods based on different properties of the graph. Details are described as follows.

5.1 Localization by Point Matching

Fig 5.1 shows the workflow of localization by point matching. Since the floor plan with actual locations is available, a possible matching relation can be established if we can generate an estimated layout with estimated locations. If we have the estimated location for each node $est : N \rightarrow \mathbb{R}^2$, we can find the localization $f : N \rightarrow L$ such that the objective function as follows is minimized.

![Workflow of localization by point matching](image)
CHAPTER 5. LOCALIZATION BY GRAPH MATCHING

\[
\sum_{i \in N} d_{\text{est}(i), f(i)}
\]  

(5.1)

where \(d\) is a distance function specified by Euclidean distance metric, i.e., \(d_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}\) represents the Euclidean distance between location \(i\) and location \(j\). Notice that this objective function means that we are looking for the localization that is the best “roundoff” of the estimated coordinates to the floor plan.

First and foremost, we have to generate the estimated layout with estimated locations. The details are explained in the following subsections.

5.1.1 Estimated Layout Generation

The purpose of estimated layout generation is to create an estimated layout and optimize the layout as much as possible to the actual layout. Given an adjacency matrix \(\text{adj} : N \times N \rightarrow \{0, 1\}\), we want to generate the estimated layout with estimated locations. Then, we can find the localization \(f\) by minimizing the objective function 5.1. In this thesis, we apply graph drawing algorithms (i.e., spectral graph drawing and force-directed algorithms [13] [7]) to generate the estimated layout.

**Graph Drawing Algorithms**

- Spectral graph drawing algorithm makes use of the eigenvalues and eigenvectors of matrices naturally associated with the graphs. It enables a rapid graph visualization by using the first two eigenvectors of the Laplacian matrix of the graph [13]. The first two eigenvectors correspond to the largest two eigenvalues. For the spectral graph drawing algorithm, it defines an objective function which indicates the energy of the graph, and then it optimizes the graph by minimizing the objective function [12]. In a 2D graph layout, spectral graph drawing algorithm solves this minimization problem by assigning the largest two eigenvalues to coordinates [13].

- Force-directed drawing algorithm [7] [12] generates graphs based on the topology relation between vertices and edges with assigned force. Assume every vertex represents a node. Every two of them are repulsive except for neighbors that are spring-like attractive (with a equilibrium length of the spring representing the distance). Based on these spring-like forces, joint forces can be computed. By repeatedly reducing the joint forces, the network becomes stable, and the graph is generated. In appendix B, it shows the pseudo-code figure of the classical Fruchterman and Reingold force-directed graph drawing algorithm.

According to [7] [12] [13] [27], spectral graph drawing algorithm has the advantage of computing fast while force-directed drawing algorithm typically optimizes the layout more elegant. Thus, in our approach, we first generate an initial graph based on adjacency matrix by using spectral graph drawing algorithm. Then, Fruchterman and Reingold force-directed drawing algorithm [7] is applied for aesthetically tuning. The reason is to reduce the complexity of the force-directed drawing algorithm by providing an initial layout. The flow chart is shown in Fig 5.2.

**Orientation Transformation**

As mentioned in Chapter 3 as well as the workflow (see Fig 5.1) of this section, we perform Step (a) to peg specific nodes. A pegging is a manual mapping (localization) between node IDs and locations. It is a method to avoid the ambiguity problem when encountering some symmetric layouts. An example is shown in Fig 5.3, this layout probably leads to a wrong localization \(<A, 9>, <B, 10>, <C, 11>, <D, 12>\) instead of a correct localization \(<A, 1>, <B, 2>, <C, 3>, <D, 4>\). Therefore, it is required to perform an orientation transformation based on the pegging nodes during layout generation. The orientation transformation is achieved by a matrix transformation with at least three reference nodes (i.e., anchors). Assume they are available after
Figure 5.2: Work flow of estimated layout generation.

Step (a), we have manually localized three corner nodes. It means we fit the three estimated locations to the actual locations.

Figure 5.3: An example of layout generation without orientation transformation.

The orientation transformation is performed through a rotation as well as a translation. The translation is intuitive that is a geometric transformation, i.e., moving every point by the same amount in a given direction. The rotation is based on a rotation matrix in Euclidean space. The rotation matrix $R$ is an equation 5.2 that rotates coordinates counter-clockwise through an angle $\theta$ about the origin of the coordinate system. The angle $\theta$ is derived from the anti-trigonometric function based on solving the rotation matrix. The rotation matrix is computed by solving the equation 5.3. $(x, y)$ are the coordinates of a reference node while $(x', y')$ are the coordinates after rotation. An example of the orientation transformation is shown in Fig 5.4.

$$R(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (5.2)$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (5.3)$$
5.1.2 Point Matching

After estimated layout generation, the estimated locations are available. The actual locations are also available from the floor plan. Then, a point matching method is applied by a Hungarian method [18]. The localization problem in the thesis is an assignment problem that assigns the node IDs to locations. The Hungarian method is a combinatorial optimization algorithm that solves the assignment problem in polynomial time. The objective function is equation 5.1. We want to find an assignment to minimize this objective function. Let \(d_{i,j}\) be the cost of assignment from the estimated location \(i\) to the actual location \(j\), we define an \(n \times n\) cost matrix.

\[
D = \begin{bmatrix}
d_{1,1} & d_{1,2} & \cdots & d_{1,n} \\
d_{2,1} & d_{2,2} & \cdots & d_{2,n} \\
\vdots & \vdots & \ddots & \vdots \\
d_{n,1} & d_{n,2} & \cdots & d_{n,n}
\end{bmatrix}
\]

An assignment is a set of \(n\) entry elements in the cost matrix, no two of which lie in the same row or column. The sum of the \(n\) entries of an assignment is the cost of the corresponding assignment cost. The optimal localization is the assignment with the lowest cost. Through Hungarian method, it will find the lowest-cost assignment given the cost matrix. Given theorem 1, the procedure of the Hungarian method is shown as follows.

**Theorem 1.** “If a number is added to or subtracted from all of the entries of any one row or column of a cost matrix, then an optimal assignment for the resulting cost matrix is also an optimal assignment for the original cost matrix.”[26]

1. “Subtract the smallest entry in each row from all the entries of its row.
2. Subtract the smallest entry in each column from all the entries of its column.
3. Draw lines through appropriate rows and columns so that all the zero entries of the cost matrix are covered and the minimum number of such lines is used.
4. Test for Optimality:
   a. If the minimum number of covering lines is \(n\), an optimal assignment of zeros is possible and we are finished.
   b. If the minimum number of covering lines is less than \(n\), an optimal assignment of zeros is not yet possible. In that case, proceed to 5.
5. Determine the smallest entry not covered by any line. Subtract this entry from each uncovered row, and then add it to each covered column. Return to 3.”[24]
5.2 Localization by Iterative Distance Matching

In this section, we propose another method which makes use of the distance properties. Assume the estimated layout with estimated locations is given based on graph drawing algorithms in the last section. The objective function is equation 5.4. We want to find the localization $f : N \rightarrow L$ such that the objective function is minimized.

\[ \sum_{i,j \in N} (d_{\text{est}(i), \text{est}(j)} - d_{f(i), f(j)})^2 \]  

(5.4)

Similarly, the localization $f$ is the best “roundoff” of the estimated coordinates to the floor plan. The objective function involves in quadratic permutations. In other words, minimizing the objective function is NP-Hard due to too many permutations. Therefore, this section introduces an iterative matching method. A flow chart of the localization by this iterative distance matching is shown in Fig 5.5. Details are described as follows.

![Figure 5.5: Work flow of iterative distance matching.](image)

5.2.1 Distance Matrix Generation

The localization in this section requires distance properties between the estimated layout and the actual layout. Then, distance matrices (i.e., estimated distance matrix and actual distance matrix) are introduced. A distance matrix is a matrix whose elements $D(i, j)$ represent the distance between location $i$ and location $j$. Assume the estimated layout is given based on graph drawing (see the last section). The estimated locations are available from the generated layout. Therefore, the distance matrix can be computed easily based on the Euclidean metric (i.e., $d_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ specifies the Euclidean distance between location $i$ and location $j$). To some extent, this distance matrix generation relies heavily on the generation of estimated layout. When the generated layout is inaccurate, the distance matrix is unreliable. As a result, it is difficult to generate a correct localization based on the distance-driven objective function.
CHAPTER 5. LOCALIZATION BY GRAPH MATCHING

For this reason, we implement another distance matrix generation method by shortest paths algorithms [8]. The principle is to transform the adjacency matrix \( \text{adj} : N \times N \rightarrow \{0, 1\} \) into the shortest distance between the nodes. The distance between the direct neighbors is 1 specified by \( \text{adj}(i, j) = 1 \). The distance between those who are not direct neighbors is computed by the Dijkstra’s algorithm [8]. Also, there exist isolated nodes. There are no paths to such nodes. In this case, the shortest distance to such nodes is defined as the average distance of all other node pairs. Normally, the isolated node is probably due to extremely large distance. The distance is better to be the maximal distance of all other node pairs. However, the isolated nodes in our application are typically due to measurement variation. Moreover, the localization based on the average distance outperforms the localization based on the maximal distance by further result analysis. Based on the shortest paths principle, we can compute a full distance matrix. Broadly speaking, the distance matrix is in hops (i.e., the distance is \#hops). Note that only the estimated distance matrix is generated through this method. The actual distance matrix of the floor plan (i.e., actual layout) is still computed through Euclidean metric.

5.2.2 Iterative Distance Matching

Given an estimated distance matrix and an actual distance matrix, we want to figure out the localization \( f \) to minimize the objective function 5.4. It is a matching problem. This kind of matching problem in graph theory is proved to be NP-Hard [17], because there are too many permutations which can not be searched in polynomial time. Therefore, we propose the iterative matching approach to simplify the problem. An example is shown in this subsection to illustrate how this approach works.

The localization method is performed iteratively. The subsequent localization must depend on previous localizations. Only when the previous localizations are correct can the following localizations be correct. So, the iterative method often starts with manually localized anchors. Assume after Step \( (a) \), we manually peg three anchors located on the left margin of the floor plan (see Fig 5.6). The three nodes with name \( A, E, I \) in the estimated layout are mapped to the anchors 1, 5, 9 in the floor plan.

![Figure 5.6: Iterative distance matching with three anchors.](image)

The first iteration is to find out which node in the estimated layout is mapped to Location 6 in the floor plan. The location is determined by the sum of distances from itself to anchors in the floor plan. In other words, the location is the nearest location to the three anchors. It has the minimal distance sum to the anchors. In this example, it is Location 6. Based on the floor plan information, we can compute its distances to the three anchors (see Fig 5.7). In the estimated layout, we obtain the distance properties by distance matrix generation (see the last subsection). Ideally, only the correct localization \( < F, 6 > \) can lead to a minimum according to the objective function 5.4. Then, the localization \( < F, 6 > \) is this step’s result (see Fig 5.8).
When a localization is found, its corresponding location is added as an anchor (see Fig 5.8). The following localization follows the same criterion by distance sum. When several locations have the same number of connections, we pick a minimal location ID. Fig 5.9 shows a location of the second iteration. Location 2 is the location waiting to be mapped, and the nodes in light yellow color in the estimated layout are all possible mappings.
Based on the same principle, the localization $< B, 2 >$ is found. By iterative matching, the complete localizations will be found. It is shown in Fig 5.10.

Figure 5.10: (a) Mapping results of third iteration and (b) final results.

### 5.3 Localization by Heuristic Edge Matching

In the last two sections, the two localization methods apply different approaches to simplify the location computation by graph matching. These two methods rely on transformation of adjacency matrix (i.e., estimated layout generation and distance matrix generation). The performance of such transformations has a direct impact on the localization. Given an adjacency matrix $\text{adj} : N \times N \rightarrow \{0, 1\}$ and a distance matrix from the floor plan, in this section, we introduce another method by direct matching between the adjacency matrix and the distance matrix. Precisely, we want to find the localization $f : N \rightarrow L$ such that the following objective function is minimized.

$$
\sum_{i,j \in N} \text{adj}_{i,j} \cdot d_{f(i), f(j)}
$$

Additionally, we apply a fast ant system (FANT) algorithm [30] to heuristically minimize the objective function. Fig 5.11 shows the workflow for this method.

### 5.3.1 Heuristic Edge Matching

Similar to the minimization of the objective function in the last section, the minimization problem in this section is also NP-Hard [17]. In the last section, we apply an iterative matching which is a kind of greedy best-first search algorithm [25]. It is quite easy to produce a local optimum since it only explores the matching by expanding the most promising node. To reduce the possibility of producing a local optimum, we implement a more sophisticated search algorithm, i.e., the fast ant system (FANT) by Taillard [30].

The fast ant system is a variation of the traditional ant system algorithms. This method simulates the foraging of the ant systems. It optimizes the matching process by implementing an
intensification strategy. More precisely, the ant system contains two processes which are an ant process and a queen process.

- **Ant process**: the ant process constructs a localization by a random localization initially. Each localization corresponds to a cost based on the objective function 5.5. Afterward, the following random localizations are generated probabilistically. The probability is proportional to the goodness of the localization cost. Moreover, in each random localization, there is a local search. This local search then swaps maximal two nodes randomly. If there is a localization with a lower cost, the better localization will be sent to the Queen process.

- **Queen process**: the queen process initializes a $n \times n$ memory matrix $M$ (also a weight matrix). The memory matrix element $m_{n,l}$ indicates the probability of the possible localization $< N, L >$. In a complete search, to “memorize” the found localization the corresponding memory elements increase a variable $v$ one step in every iteration. When a better localization with a lower cost is found, the corresponding memory elements $m_{n,l}$ increases with a constant $R$. In the meanwhile, the variable $v$ is re-initialized. By this method, the queen process maintains the best localization by reinforcing the attractive memories (i.e., larger weights specified by larger $m_{n,l}$) and by giving “less” weight to the unattractive memories (i.e., fewer weights specified by smaller $m_{n,l}$).

The details are shown as follows.

1. “Activate an Ant process,”

2. Wait for a localization $f : N \rightarrow L$ from an Ant process

3. Assume $f^*$ is the minimal cost localization. If $f = f^*$ then set $v \leftarrow v + 1$ and $m_{n,l} \leftarrow v, \forall n,l$,

4. If $f$ is better than $f^*$ then set $f^* \leftarrow f, v \leftarrow 1$ and $m_{n,l} \leftarrow v, \forall n,l$,

5. Set $m_{i,f_i} \leftarrow m_{i,f_i} + v, \forall i$,

6. Set $m_{i,f_i^*} \leftarrow m_{i,f_i^*} + R, \forall i.”$ [31]
Chapter 6

Results and Discussion

In this chapter, the results based on the three localization methods are shown and discussed.

6.1 Adjacency Matrix Evaluation

The adjacency matrix in this section is the input of the three localization methods in the following section. As mentioned in Chapter 3, the adjacency matrix is derived after thresholding. The results in this chapter are based on Testbed 1 when there is no further indication (see Fig 4.4). The distance threshold is set as 135 cm by empirical observation. When distance measurement between two luminaries (i.e., locations) is less than the threshold, the two luminaries are neighbors (i.e., \( \text{adj}_{i,j} = 1 \)). Otherwise, they are not neighbors (i.e., \( \text{adj}_{i,j} = 0 \)). Before thresholding, we apply different filtering and averaging methods to the distance measurement. The filtering and averaging methods are the twelve combination methods mentioned in Chapter 4. Each combination method corresponds to five adjacency matrices, as there are five distance measurements. Each five adjacency matrices by one combination method form an adjacency matrix set. For the sake of simplification and avoiding ambiguity, we name the adjacency matrix sets as numbers.

To evaluate the correctness of the adjacency matrix, we define a percentage of correctness (PoC) as follows.

\[
\text{PoC} = \frac{1}{|N| \cdot (|N| - 1)} \sum_{i,j \in N, i \neq j} \delta(\text{adj}(i,j) - \text{ref}(i,j))
\]

where \( \delta(x) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{otherwise} \end{cases} \). \( \text{ref} : N \times N \rightarrow \{0, 1\} \) is a reference adjacency matrix (i.e., representing the actual layout). A higher PoC indicates the percentage of correctness (of adjacency matrix elements) is larger. In other words, the adjacency matrix is better.

The evaluation of the adjacency matrix is shown by a scatter plot (see Fig 6.1). For better visualization, the scatter plot is displayed with smooth lines. Numbers in x axis represent different adjacent matrix sets. Each adjacency matrix set consists of five adjacency matrices corresponding to five distance measurements (i.e., individual pair 1,2,3,4 and estimated pair -1). For instance, the red point in column 11 represents the PoC of Adjacency matrix 11 based on individual pair 2. Precisely, it is based on “Adjacency matrix 11_pair 2”. PoC in y axis is the evaluation criteria. An ascending order is applied based on the PoC of estimated pair -1. A higher PoC indicates a better adjacency matrix.

From the plot, we can observe that the PoC of estimated pair -1 is typically better than that of individual pairs 1,2,3,4. The PoCs of individual pairs are not linear with the PoCs of estimated pair. For example, PoCs of Adjacency matrix (set) 5 and Adjacency matrix (set) 6 are very different between individual pairs and estimated pairs. The two matrix sets are based on MinMax+Bottom+WeightedAvg and MinMax+Avg-Std. As mentioned in Chapter 4, distance...
measurements based on individual pairs typically involved in more errors. The standard deviations are usually larger. Therefore, applying MinMax filtering to individual distance measurements would lead to a worst result sometimes. However, this phenomenon can be avoided or alleviated when using measurements of estimated pair. The measurement of estimated pair is a weighted combination based on the four individual measurements. Then, the fluctuations of individual measurements can be largely neutralized. As a result, the adjacency matrix of estimated pair is better.

6.2 Localization Results

6.2.1 Criteria

As mentioned in the problem statement, we define the actual localization as $f_a$, which maps nodes in $N$ to their actual location in $L$. $f_a$ is the localization we want to compute, and it is the one that we use as the reference of quality. Then, we define $AvgDist$ as follows.

$$AvgDist(f) = \frac{1}{|N|} \sum_{i \in N} ||f(i) - f_a(i)||$$  \hspace{1cm} (6.2)

The $AvgDist(f)$ represents the average distance between the estimated localization and actual localization. Ideally, when the localization is 100% correct (i.e., $f_a$), the $AvgDist$ is zero. A lower $AvgDist$ indicates a better localization. Apart from $AvgDist$, we define another intuitive criterion which is a localization mapping accuracy.

$$Accuracy(f) = \frac{1}{|N|} \sum_{i \in N} \delta(f(i) - f_a(i))$$  \hspace{1cm} (6.3)

where $\delta(\vec{x}) = \begin{cases} 1, & \text{if } \vec{x} = (0, 0) \\ 0, & \text{otherwise.} \end{cases}$ Ideally, when the localization is 100% correct, the $Accuracy$ is 1. A higher $Accuracy$ indicates a better localization.

6.2.2 Results

Results of Method 1

For simplification, we name the localization by point matching as Method 1; we name the localization by iterative distance matching as Method 2; we name the localization by heuristic edge

![Evaluation of adjacency matrix](image.png)
matching as Method 3. First of all, we use the $12 \times 5$ adjacency matrices (see Fig 6.1) as inputs and apply Method 1. The localization results are shown by $AvgDist$ as well as $Accuracy$ (see Fig 6.2).

![AvgDist of Method 1](image1)

![Accuracy of Method 1](image2)

(a) Figure 6.2: Localization results of Method 1.

From the results of Method 1 by point matching, we observe:

- $AvgDist$ and $Accuracy$ behave oppositely, but they show a nearly identical trend. A lower $AvgDist$ indicates a better localization. A higher $Accuracy$ means a better localization. $Accuracy$ is more intuitive while $AvgDist$ is for inspection. For instance, the accuracies of Adjacency matrix 3 & 4 are identical of individual pair 2 and estimated pair -1. From results of $AvgDist$, localization by estimated pair -1 is better than that by individual pair 2. Individual pair 2 has larger $AvgDist$s. It means they mistaken node pairs with longer distance.

- The results of estimated pair -1 are usually better than that of individual pairs 1,2,3,4. The reason is that the adjacency matrices of estimated pair -1 are typically better than that of individual pairs 1,2,3,4 (see Fig 6.1).

- For the relation between result and $PoC$ (see Fig 6.1), a decent result usually has a decent $PoC$. However, a good $PoC$ does not guarantee a good result. Take Adjacency matrix 7 & 8 of individual pair 2 for example. Their $PoCs$ are both 0.92. However, the results of them either according to $AvgDist$ or $Accuracy$ are quite poor. The reason is that neither $AvgDist$ nor $Accuracy$ is in a linear correlation with $PoC$. Precisely, $PoC$ takes every element in the
adjacency matrix into account. However, adjacency elements (i.e., $ref(i, j) = 1$) are more error-prone. These errors have more impacts on layout generation. An example is shown in Fig 6.3. There are not many wrong edges in the graph. Nevertheless, the graph misses many edges which should be connected such as Edge(A, E), Edge(C, D). As a result, the graph is very different from the actual layout. Consequently, point matching based on this graph is unreliable.

![Figure 6.3: Estimated layout of Adjacency matrix 7.](image)

**Results of Three Methods**

Based on results from Method 1, we notice the results from the individual pairs are quite poor. Finding any helpful or useful correlation is hard. For this reason, in the comparison of the three localization methods, the results are based on adjacency matrices of estimated pair -1. When there is no further indication, the following adjacency matrix refers to the one of the pair -1. Fig 6.4 shows the overall comparison.

At a first glance, results of Method 3 by heuristic edge matching are always better than that of Method 1 by point matching and Method 2 by iterative distance matching. Method 2 by iterative distance matching is the worst, as only three adjacency matrices correspond to 100% correct.
Method 2 For Method 2, if we check the adjacency matrices of the three correct localizations, they all have good PoCs (see Fig 6.1 Adjacency matrix 11 or Adjacency matrix 12). From analysis of Method 1, a decent PoC does not guarantee a decent localization. In Method 2, Adjacency matrix 10 is an example. The PoC is around 0.95, but Adjacency matrix 10 still does not have 100% correct results.

Method 2 applies distance matching by minimizing the objective function 5.4. Different localizations correspond to different “minimums”. Take Adjacency matrix 10 for example. Based on the objective function, we compute the minimum by estimated localization. It is 48.30. Based on the same function, the minimum by actual (reference) localization is 42.67. Obviously, the minimum by estimated localization is not a global minimum. The reason is that the iterative matching method only expands the most promising node. It is quite easy to miss the correct matchings. Similar phenomena are found in Adjacency matrix 5,6,7,8. Fig 6.5 shows the two minimums scatter plot. Minimum_{est} is the minimum derived from the estimated localization by given adjacency matrix, whereas Minimum_{ref} is the minimum derived from the actual localization by reference adjacency matrix.
CHAPTER 6. RESULTS AND DISCUSSION

Figure 6.5: Localization results’ cost of Method 2.

Method 1 & 3  As for Method 1 and Method 3, in most cases, Method 3 outperforms Method 1. The initial concern of introducing Method 3 is to avoid the negative influence of estimated layout generation. We adapt the objective function to direct multiplication between adjacency matrix and distance matrix.

Take the Adjacency matrix 7 for example. The result of Method 1 is slightly worse than that of Method 3 (see Fig 6.4). If we check the exact localization, the mistakes are \( f(K) = (x_{10}, y_{10}) \), \( f(J) = (x_{11}, y_{11}) \). It means it confuses Node J and Node k. The estimated layout is shown in Fig 6.6. Apparently, the incorrectness is caused by edges involved Node I, J, K, L. More precisely, there is no edge between J and F, which results in the marginalization of J. Due to the marginalization, it leads to the incorrect localization of \( < K, 10 > \) and \( < J, 11 > \) by point matching. As for the edge matching, we continue the analysis based on the similar edges (see Fig 6.7). Method 3 applies edge matching by minimizing the objective function 5.5. The edges are specified by the adjacency matrix from estimated graph and the distance matrix from actual graph (i.e., floor plan). Apparently, the correct localization has a lower “minimum”. The missing edge(J, F) does not bring extra values based on the objective function 5.5. As long as the other edges are still correct, it ensures the lower “minimum”. Therefore, in comparison with Method 1 by point matching, Method 3 by edge matching is more robust.

Figure 6.6: Estimated layout of Adjacency matrix 7.
CHAPTER 6. RESULTS AND DISCUSSION

Method 2 & 3 In Method 3, we also adapt the objective function from the distance matching to the edge matching. Similar to layout generation, Method 2 involves in distance generation through shortest path algorithms. The distance generation has a direct influence on the distance matching.

In other words, the distance generation is defective. For one thing, the generated distance between two nodes is a naive estimation. The distance is in hops (i.e., the distance is #hops). See Fig 6.8, the estimated distance between A and F is 200 cm rather than 141 cm. Because it is the sum of distance between (A, B) and (B, F) (or (A, E) and (E, F)). For another thing, errors are introduced when there is an isolated node (e.g., see Node L in Fig 6.8). There is no any edge to the isolated node. As mentioned in Chapter 5, the distance to such a node is defined as the average distance of all other node pairs. Due to these two reasons, the distance generation by the shortest path algorithms may introduce unexpected errors. Thus, in Method 3 we employ the objective function based on edge matching. The edge matching directly makes use of adjacency properties rather than distance properties by generation. Fig 6.9 shows a heuristic matching comparison by different objective functions (i.e., distance matching and edge matching). In general, the results based on edge matching objective function are better.
Figure 6.9: Mapping results by different objective functions in Method 3.
To sum up, Method 3 by heuristic edge matching is relatively better among the three localization methods. For one thing, it avoids the negative impacts introduced either by graph drawing or distance generation. For another thing, it optimizes the matching algorithm from a simple best-first search to a sophisticated heuristic search by FANT (fast ant system).

6.3 Validation

For validation, we use another 12 adjacency matrices based on Testbed 2 (see Fig 4.20 in Chapter 4) to apply the three localization methods. The 12 adjacency matrices are based on distance measurement of estimated pair -1. Localization results by AvgDist and Accuracy are shown in Fig 6.10. Similarly, the results of Method 3 by heuristic edge matching remains better. Method 2 by iterative distance matching is the worst.

![Mapping Results by Different Localization Methods](image)

Figure 6.10: Mapping results by different localization methods.

In comparison with the results from the Testbed 1. Results of Method 1 and Method 2 are a bit worse. The reason is due to the defect of experiment environment. Particularly, there is a metallic bar in the middle of Testbed 2 (see the curtain bar in Fig 4.4 or Fig 4.20). As a result, the distance measurements in-between nodes are typically far larger than the average measurement. Because of the particular reason, the adjacency correctness of in-between nodes is mostly mistaken. It is also reflected on the localization results. For example, the results of Adjacency matrix 7 and 11 are terrible. Their estimated layouts are shown in Fig 6.11. Clearly, the missing edges in between lead to the incorrectness of Method 1 or Method 2. Similar phenomena happen in other adjacency
matrices (see Fig 6.12). Fortunately, the Edge(F, G) is correct of Adjacency 9 and 12. Therefore, the results of them are better. However, they still cannot achieve 100% correct due to the other two missing edges in between.

Although the results from Testbed 2 are a bit different, it validates the robustness of Method 3. As mentioned, it avoids adverse impacts from graph drawing and distance generation. Also, the heuristic search algorithm FANT is more capable of finding out the best localization.

![Figure 6.11: Estimated layout of Adjacency matrix 7 and 11.](image1)

![Figure 6.12: Estimated layout of Adjacency matrix 9 and 12.](image2)
Chapter 7

Conclusions

The goal of this thesis is to solve the localization problem for commissioning in the wireless lighting systems. To put it simply, we have a set of luminaire IDs and a set of locations, and we would like to know the location of each ID. In the thesis, we use an RF chipset to sense the distance, and then transform the distance properties into adjacency properties. Afterward, we apply graph matching methods to achieve localization. In details, the conclusions are summarized as follows.

A complete localization scheme is proposed and validated. It involves two major steps. First, there is a distance measurement phase where the distances between luminaries are collected. Second, with the help of a known floor plan and the distance results obtained in the first phase, every luminaire is then localized by graph matching methods.

The capabilities of distance measurement by AT86RF233 is explored and evaluated. For accurate distance measurement, the RF chipset requires some proper setups including frequency step size, sweeping bandwidth, antenna rotations. In our testbeds, the frequency step size is set to 1 MHz. The sweeping bandwidth is recommended to be configured as large as possible. The antennas are proposed to be rotated perpendicularly.

The average accuracy of 1 step distance measurement in the two testbeds can achieve \( \pm 60 \) cm. Advanced filtering and averaging methods such as DQF filtering, Minmax filtering, BottomSelection filtering, and WeightedAvg improve the average accuracy significantly (i.e., \( \pm 10 \) cm).

Three localization methods based on graph matching are studied. The first one transforms the adjacency properties into estimated locations. Then, a point matching algorithm by Hungarian assignment is applied. The second method employs an iterative search algorithm to match the distances. The last method implements a fast ant system algorithm to heuristic search all the possible mappings based on the adjacency properties and the distance properties. In comparison, Method 3 by heuristic edge matching is the most robust.

According to the results, all the methods give satisfactory results when dual antennas are used per luminaire. To sum up, the proposed method by distance measurement and graph matching is feasible to solve the localization problem for commissioning in our testbeds.

7.1 Future Work

More complex layouts should be tested such as L-shape or U-shape layouts. Also, the localization methods are only validated in two 12 nodes testbeds. An extension to a larger network is required. It is to verify the scalability and the robustness. In a real implementation, the lighting systems typically contain hundreds of luminaries while RF measuring has a limited range. A grouping strategy needs to be developed.

Current localization methods do not use the full topology information of the network. The methods only take the adjacency properties (i.e., one hop topology information) into account. In theory, full topology information should provide more useful information and have better localization performance. However, the problem is that the full topology information needs to consider the
poor multi-step distance measurements. A sophisticated method to exploit such poor multi-step measurements should be developed.
Bibliography


[2] Atmel Corporation. AT86RF233: Low Power, 2.4GHz Transceiver for ZigBee, RF4CE, IEEE 802.15.4, LoWPAN, and ISM Applications; Datasheet, 7 2014. Rev. 8315E. 2, 10, 13, 18


Appendix A

Mathematical Model

A phase difference measurement is illustrated in Fig A.1. Assume we have two devices A and B, which are the initiator and reflector respectively. For the sake of simplification, we only describe one iteration phase measurement with two frequencies. First, the two devices set their carrier frequency to a $f_1$. Then, Device A transmits a signal, and Device B will measure the transmitted phase $\varphi_{B1}$ from Device A. Afterward, these two devices exchange their roles: Device A becomes a receiver, and Device B is a transmitter. Device A will measure the transmitted phase $\varphi_{A1}$ from Device B. Next, the two devices switch to another carrier frequency $f_2 = f_1 + \Delta f$, and similar measurements will be repeated. In the last step, the two devices will exchange their phase measurement results to remove a common phase drift caused by the frequency switch. Then, the phase difference can be derived as equation A.1.

$$\Delta \varphi = (\varphi_{A2} - \varphi_{B2}) - (\varphi_{A1} - \varphi_{B1}) \quad (A.1)$$

With measured phase difference, derivation of distance estimation can be described as follows. Above all, assume a radio signal is in a far field radiation. The mathematical description of a sinusoidal signal is given by [28]

$$y_1(d, t, f_c) = A \cdot \sin[2\pi f_c \left(\frac{d}{c} - t\right) + \varphi_1] \quad (A.2)$$

where $A$ is the amplitude of the signal, $c$ is the transmission speed, $d$ is the distance from the transmitter to the receiver, $t$ is the transmission time, $\varphi_1$ is an unknown phase offset, and $f_c$ is the dedicated carrier frequency.

Then, the signal is down mixed to an intermediate frequency signal:

$$y_2(t, f_l) = B \cdot \sin(2\pi f_l t + \varphi_2)$$

with an amplitude of $B$, an intermediate frequency of $f_l$, and another unknown phase offset $\varphi_2$. Afterward, we obtain $y_m(d, t) = y_1(d, t, f_c) \cdot y_2(t, f_l)$:

$$y_m(d, t) = \frac{AB}{2} \cdot \left[ \cos[(2\pi f_c \left(\frac{d}{c} - t\right) + \varphi_1) - (2\pi f_l t + \varphi_2)] - \cos[(2\pi f_c \left(\frac{d}{c} - t\right) + \varphi_1) + (2\pi f_l t + \varphi_2)] \right] \quad (A.3)$$

Since the relation between phase difference and distance is what we are interested, we select the higher frequency side and rewrite equation A.3 as:

$$y_m(d, t) = \cos[2\pi t(f_c - f_l) - (2\pi f_c \left(\frac{d}{c} + \varphi_1 + \varphi_2)) \quad (A.4)$$

with $\varphi(d, f_c) = 2\pi f_c \left(\frac{d}{c} + \varphi_1 + \varphi_2)$. In general, frequency $f_c$ can be set as arbitrary values. For the sake of simplification, we assume $f_c - f_l$ as a constant, and redefine this frequency $f_c = f_{\text{base}} + i \Delta f$ which allows us to control the frequency offset.
APPENDIX A. MATHEMATICAL MODEL

Figure A.1: Phase difference measurement.

\[
\varphi(d, i) = 2\pi(f_{\text{base}} + i\Delta f)d/c + \varphi_1 + \varphi_2 \tag{A.5}
\]

where \(f_{\text{base}}\) is the base frequency, \(\Delta f\) is the frequency offset, and \(i \in N\). Then, based on this equation with \(i\) and \(i + 1\) we perform a subtraction, and we obtain

\[
\Delta \varphi(d) = \varphi(d, i + 1) - \varphi(d, i) = 2\pi \Delta f \frac{d}{c} \tag{A.6}
\]

After this, the distance estimation equation through a phase difference between two frequencies is derived. In real hardware, we need to take the influence of drifts of local oscillators [21] into account. Frequencies are typically derived from a fixed local oscillator and phase lock loops (PLL). In that case, the phase change due to frequency switch will occur, equation A.5 becomes:

\[
\varphi(d, i) = 2\pi(f_{\text{base}} + i\Delta f)d/c + \varphi_1 + \varphi_2 + \varphi_{\text{PLL}}(i) \tag{A.7}
\]

Likewise, we want to perform the subtraction and eliminate the unwanted term \(\varphi_{\text{PLL}}(i)\). But now we cannot achieve it by the same subtraction due to the different \(i\). As mentioned in measuring principles, the devices will switch roles and then exchange phase measurements. Since the two devices have to use the same frequency per time, the additional phase term will be canceled out. As for the retained terms, the coefficients are doubled due to the different vector direction. We have

\[
\Delta \varphi(d, i) = 4\pi(f_{\text{base}} + i\Delta f)d/c + 2(\varphi_1 + \varphi_2) \tag{A.8}
\]

Now we perform similar subtraction between \(i\) and \(i + 1\). In the end, we derive the distance equation by phase measurement of two frequencies:
APPENDIX A. MATHEMATICAL MODEL

\[ d = \frac{c}{4\pi \Delta f} \cdot \Delta \varphi \]  \hspace{1cm} (A.9)
Appendix B

Force-directed Drawing Algorithm

Algorithm 2: Fruchterman-Reingold

\begin{verbatim}
area ← W * L ; /* frame: width W and length L */
initialize G = (V, E) ; /* place vertices at random */
k ← \sqrt{area/|V|} ; /* compute optimal pairwise distance */
function f_r(x) = k^2/x ; /* compute repulsive force */
for i = 1 to iterations do
    foreach v ∈ V do
        v.disp := 0 ; /* initialize displacement vector */
        for u ∈ V do
            if (u ≠ v) then
                Δ ← v.pos - u.pos ; /* distance between u and v */
                v.disp ← v.disp + (Δ/|Δ|) * f_r(|Δ|) ; /* displacement */

function f_a(x) = x^2/k ; /* compute attractive force */
for each e ∈ E do
    Δ ← e.v.pos - e.u.pos ; /* e is ordered vertex pair v and u */
    e.v.disp ← e.v.disp - (Δ/|Δ|) * f_a(|Δ|);
    e.u.disp ← e.u.disp + (Δ/|Δ|) * f_a(|Δ|);

foreach v ∈ V do
    /* limit max displacement to frame; use temp. t to scale */
    v.pos ← v.pos + (v.disp/|v.disp|) * min(v.disp, t);
    v.pos.x ← min(W/2, max(-W/2, v.pos.x));
    v.pos.y ← min(L/2, max(-L/2, v.pos.y));
    t ← cool(t) ; /* reduce temperature for next iteration */
\end{verbatim}

Figure B.1: Pseudo-code of Fruchterman and Reingold force-directed graph drawing algorithm.

Automatic Commissioning of Wireless Sensor Networks