LEARNING BASED HYPERBOLIC POSITION BOUNDING IN WIRELESS NETWORKS

by
Eldai El Sayr

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for the degree of
MASTER OF SCIENCE

School of Computer Science
at
CARLETON UNIVERSITY
Ottawa, Ontario
April 2009
© Copyright by Eldai El Sayr, 2009
NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.
Abstract

One of the recent schemes that has been put forward to localize a node is Hyperbolic Position Bounding (HPB) [4]. This scheme uses the relative Received Signal Strength (RSS) obtained by a number of trusted nodes to determine the position of a transmitter. Utilizing the RSS values, HPB uses a probabilistic model to estimate the range of distance differences between receiver pairs and a transmitter. Then, these distance differences are used to construct hyperbolas that bound the transmitter with some degree of confidence. Due to shadowing effect and other environmental factors, not all of the receiver pairs in HPB perform at the same level. Therefore, we proposed a modification of HPB that improves its performance. Our proposed system is called Hyperbolic Position Bounding with Learning Ability (HPB-LA). It consists of two phases, a learning phase and an operational phase. In the learning phase the system uses real data to measure the accuracy of the receiver pairs. Then, it ranks the receiver pairs according to their accuracy. After that, it groups the receiver pairs in subsets, evaluates the performance of these subsets, and finally selects the subset that provides the highest accuracy. In the operation phase, the subset of receiver pairs that is identified in the learning phase is used to localize nodes.

We also developed a new heuristic for the transmitter Effective Isotropic Radiated Power (EIRP) range estimation and evaluated it with experimental data and simulation data. This proposed heuristic adopts the receiver pair-based concept. For each receiver pair that includes the Maximal RSS Receiver (MRR), our heuristic assumes that the transmitter is located between the MRR and the other receiver, and then uses the RSS values to determine the EIRP range used by each receiver.

We conducted an experiment to evaluate the performance of HPB-LA. Our analysis results from the experimental data show that HPB-LA can achieve a higher success rate compared to HPB. The evaluation of our proposed transmitter EIRP range estimation process using the experimental data shows that HPB and HPB-LA have experienced a large reduction in the candidate area size, where the candidate area is the area bounded by the hyperbolas that are constructed by the receiver pairs. The analysis from the simulation results show that for any number of receivers, our proposed heuristic always gives a smaller candidate area in comparison to the heuristic used in the original HPB.
To my parents and my daughter Aasya
Acknowledgements

I would like to thank my wife Fatima for her love, understanding, and support. Next, I would like to express my gratitude to my supervisor, Prof. Michel Barbeau and thank him for his support and guidance. I also owe a great deal of thanks to my colleagues who helped me carry out the experiment: Paul Boone and Christine Laurendeau whose suggestions were valuable to my research.
Table of Contents

Abstract ii
Dedication iii
Acknowledgements iv
Table of Contents v
List of Figures viii
List of Tables xiv
Abbreviations xv

1 Introduction 1
  1.1 Background ............................................. 1
  1.2 Definition of the Problems ................................ 8
  1.3 Highlight of the Results ................................ 10
1.3.1 Enhancing the Success Rate of HPB ......................... 10
1.3.2 Improving the Candidate Area of HPB ....................... 11
1.3.3 The Simulation Results .................................. 12
1.4 Outline of the Thesis ...................................... 12

2 Literature Review .............................................. 14
2.1 Radio Propagation ........................................... 16
2.2 Range-based Localization Techniques ......................... 19
    2.2.1 Angle of Arrival (AoA) ............................. 19
    2.2.2 Time of Arrival (TOA) ............................. 21
    2.2.3 Time Difference Of Arrival (TDOA) ................. 22
    2.2.4 Received Signal Strength (RSS) ....................... 22

3 Hyperbolic Position Bounding (HPB) .......................... 27
3.1 Factors that Affect the Hyperbola Pair Positions ............... 29
3.2 Factors that Affect the HPB Performance ....................... 39

4 HPB with Learning Ability (HPB-LA) .......................... 41

5 Evaluation of HPB with Learning Ability Scheme ................. 57
5.1 The Experiment Setup .................................... 58
    5.1.1 Hardware ........................................ 58
    5.1.2 Software ........................................ 60
List of Figures

1.1 Example of Minimum and Maximum Hyperbolas. Minimum and maximum hyperbolas constructed by a pair of receivers in an area of 800 m × 800 m. $H^-(R_1, R_2)$ and $H^-(R_2, R_1)$ are the minimum hyperbolas, and $H^+(R_1, R_2)$ and $H^+(R_2, R_1)$ are the maximum hyperbolas of $(R_1, R_2)$ receiver pair. ........................................... 4

1.2 Example of Candidate area. Candidate area (shaded) bounded by minimum hyperbolas constructed by two pair of receivers in an area of 800 m × 800 m. $H^-(R_1, R_3)$ and $H^-(R_3, R_1)$ are the minimum hyperbolas constructed by receiver pair $(R_1, R_3)$, and $H^-(R_2, R_4)$ and $H^-(R_4, R_2)$ are the minimum hyperbolas constructed by receiver pair $(R_2, R_4)$. The figure also demonstrates a successful localization. ...................... 5

1.3 Example a Failure in Localization. A transmitter is out side the candidate area. .......................................................... 7
2.1 Angle of Arrival. If the orientation is known, angles $\theta$ and $\beta$ represent the relative AoA measured at $N_u$ (location unaware node) from signals sent by $N_A$ and $N_B$ (nodes with known locations). Then triangulation is used to determine the position of $N_u$. 20

2.2 Example of TDOA. Hyperbola $H_{(R_1,R_2)}$ is constructed from the distance difference calculated from TOA of the transmitter $T$ signal at receiver $R_1$ and receiver $R_2$. Similarly, Hyperbola $H_{(R_1,R_3)}$ is constructed from the distance difference calculated from TOA of the transmitter $T$ signal at receiver $R_1$ and receiver $R_3$. 23

3.1 Example of Minimum Hyperbolas. A transmitter is bounded by minimum hyperbolas constructed by a pair of receivers in an area of 800 m x 800 m. $H^{-}(R_1,R_3)$ and $H^{-}(R_3,R_1)$ represent the minimum hyperbolas for receiver pair $(R_1,R_3)$. $H^{-}(R_2,R_4)$ and $H^{-}(R_4,R_2)$ represent the minimum hyperbolas for receiver pair $(R_2,R_4)$. 35

3.2 Example of Minimum Hyperbolas for all possible receiver pairs. The figure also shows the candidate area (shaded). 36
4.1 Example of Successful Bounding. A transmitter is successfully bounded by minimum and maximum hyperbolas constructed by a pair of receivers in an area of 800 m × 800 m. \( H^-(R_1, R_2) \) and \( H^-(R_2, R_1) \) are the minimum hyperbolas, and \( H^+(R_1, R_2) \) and \( H^+(R_2, R_1) \) are the maximum hyperbolas of \((R_1, R_2)\) receiver pair. The direction of the open part of the hyperbola depends on the distance difference sign (+/-). If the minimum distance difference and maximum distance difference have opposite signs then the open parts of the minimum hyperbola and maximum hyperbola have opposite directions.

4.2 Transmitter \( T \) is between receiver \( R_k \) (MRR) and receiver \( R_i \).

5.1 The experiment area setup.

5.2 The TRENDnet Wireless PCI adapter used as a receiver device in all the experiment.

5.3 The HWUG1 Wireless-G USB adapter used as a transmitter device in all the experiment.

5.4 Eldai El Sayr placing the laptop antenna on the top of a tripod while he is preparing to send some packets.

5.5 The success rate and candidate area size of receiver pairs.

5.6 Receiver pairs ordered according to their success rate. The receiver pairs on the left have a higher success rates than the receiver pairs on the right. For example, receiver pair \( P_1 \) has a success rate of 100%, while receiver pair \( P_6 \) has a success rate of 91%.
5.7 Receiver pairs ordered according to their candidate area size. The receiver pairs on the left have a smaller candidate area size than the receiver pairs on the right of the chart. For example, receiver pair P₂ has a candidate area size of 74%, while receiver pair P₁ has a candidate area size of 90%.

5.8 Receiver pair subsets for success rate optimization. Subsets with fewer receiver pairs that have high success rates give the highest success rate. For example, S₄ which consists of the two receiver pairs that have the highest success rate produces the highest success rate among all the other subsets.

5.9 Receiver pair subsets for candidate area optimization. Subsets with large number of receiver pairs are more likely to give a candidate area with a size less than the size of a candidate area that is given by a subset that has fewer receiver pairs, as explained in Chapter 4. We notice that, the candidate area size increases when we move from the left to the right of the chart, at the same time the number of receiver pairs in the subsets decreases from left to right.

5.10 Receiver Pair subsets PRS. Subset S₁ has the highest RPS (5.53), while S₅ has the lowest RPS (zero).

5.11 A Comparison between HPB and HPB-LA.
5.12 Comparison between the performance of HPB when the original transmitter EIRP range estimation heuristic was used (HPB-A) and performance of HPB when the Receiver Pair-Based EIRP Range Estimation heuristic was used (HPB-B). ................................. 75

5.13 Success rate and the candidate area size of receiver pairs. The success rate and candidate area size of each pair is shown. .................. 77

5.14 Receiver pairs ordered according to their success rate. ............... 78

5.15 Receiver pairs ordered according to their candidate area size. ...... 79

5.16 Receiver pair subsets for success rate optimization. ................. 80

5.17 Receiver pair subsets for candidate area size optimization. ........ 81

5.18 Receiver pair subsets PRS. Subset G5 has the highest RPS (3.26), while G8 has the lowest RPS (0.52). ........................................... 83

5.19 A Comparison between HPB and HPB-LA when the Receiver Pair-Based EIRP Range Estimation heuristic is used. .................... 84

5.20 A Comparison between HPB and HPB-LA when two different EIRP range estimation heuristics are used. HPB(A) is HPB with the HPB original transmitter EIRP range estimation heuristic, HPB-LA(1) is HPB-LA with the HPB original transmitter EIRP range estimation heuristic, and HPB-LA(2) is HPB-LA with the Receiver Pair-Based EIRP Range Estimation heuristic. ........................................... 85

5.21 A snap shot from the simulation. ........................................... 88
5.22 Simulation results. HA-SR and HA-CAS are the success rate and candidate area size produced when HPB used its original transmitter EIRP range estimation heuristic, while HB-SR and HB-CAS are the success rate and candidate area size produced when HPB used the Receiver Pair-Based EIRP Range Estimation heuristic.
List of Tables

5.1 Receiver pairs and their symbols. ................................. 65
5.2 Receiver pairs subsets for success rate optimization. .......... 69
5.3 Receiver pair subsets for candidate area size optimization. .... 69
5.4 Receiver pairs and their symbols. ................................. 76
5.5 Receiver pairs subsets for success rate optimization. .......... 82
5.6 Receiver pairs subsets for candidate area size optimization. .... 82
Abbreviations

AoA Angle of Arrival
EIRP Effective Isotropic Radiated Power
HPB Hyperbolic Position Bounding
HPB-LA Hyperbolic Position Bounding with Learning Ability
MAC Media Access Control
MRR Maximal RSS Receiver
PSVs Public Safety Vehicles
RF Radio Frequencies
RPCAS Receiver Pairs Candidate Area Size
RPS Receiver Pairs Score
RPSR Receiver Pairs Success Rate
RSS Received Signal Strength
RSUs Road Side Units
TDOA Time Difference Of Arrival
TOA Time Of Arrival
US Ultrasound
Chapter 1

Introduction

1.1 Background

After an attack is detected in a wireless network, reattributing the attacking message to its originator is very challenging. This problem becomes even harder when the rogue node is an insider who owns a valid identity like a MAC address or a digital certificate and has the ability to acquire forged identities. Moreover, in some wireless applications like vehicle safety where dynamic MAC addresses are supported [18], an insider may disguise its identity to broadcast fraudulent messages. In such situations finding the physical location of a rogue node is the first step in attributing an attacking message to the attacker. The process of estimating the physical position of a node in a wireless network is called localization. One of the recent schemes [4] that has been put forward to localize a node is the HPB. This scheme uses the RSSs obtained by a number of trusted nodes to determine the position of the transmitter. Trusted nodes are the
nodes that have been authenticated by some authentication mechanism. Since the EIRP used by the transmitter is unknown and the RSS values may fluctuate, the HPB uses a probabilistic model to estimate the range of distance differences between receiver pairs and the transmitter. The maximum and the minimum distance differences are used to construct a maximum and a minimum hyperbola respectively between a pair of receivers. Then, with a certain degree of confidence, the transmitter can be located between these two hyperbolas. By considering different pairs of receivers, multiple pairs of these hyperbolas may be constructed. The area formed by the intersection of these hyperbolas contains the transmitter with a compound degree of confidence.

Due to the shadowing effect and other environmental factors, not all the receiver pairs in HPB perform at the same level. Therefore, it is important to have a mechanism to rank the receiver pairs in HPB according to their accuracy rates, and to investigate the possibility of choosing a subset of these receiver pairs that would give a better performance than using all the receiver pairs as in the original HPB. Also, one of the biggest factors that affect the performance of HPB is the estimation of the EIRP range that is used by the transmitter. Finding a better way of estimating the EIRP range might boost the performance of HPB.

The performance of the receiver pairs in HPB depends on the accuracy by which the hyperbola pairs that are constructed by these receiver pairs locate the transmitter. This accuracy has two components. The first one is the success rate with which the minimum and maximum hyperbolas constructed by the receiver pairs can bound the actual location of the transmitter. The second component is the size of the candidate
area on which the transmitter might reside. According to these parameters, we expect that some of the hyperbola pairs to perform better than others. Therefore, we should define the candidate area as the area located by the subset of receiver pairs which has the highest success rate or/and the smallest candidate area.

**Maximum and Minimum Hyperbolas**

In HPB each receiver pairs construct a pair of minimum hyperbolas and a pair of maximum hyperbolas. These hyperbolas have the two receivers as foci and the distance difference is calculated from RSS values measured from a transmitter message. Figure 1.1 shows an example of minimum and maximum hyperbolas.

**Candidate Area Size**

Candidate area is an area bounded by hyperbolas constructed by a pair of receiver or a group of pair of receivers. The candidate area size is a percentage of the candidate area size to the size of the total area covered by the network. Figure 1.2 shows an example of a candidate area.

**Success in Localization**

A success in localizing a transmitter means that a transmitter is bounded by the hyperbolas constructed by a pair of receivers or a group of pair of receivers. Figure 1.2
Figure 1.1: Example of Minimum and Maximum Hyperbolas. Minimum and maximum hyperbolas constructed by a pair of receivers in an area of 800 m x 800 m. $H^-(R_1, R_2)$ and $H^-(R_2, R_1)$ are the minimum hyperbolas, and $H^+(R_1, R_2)$ and $H^+(R_2, R_1)$ are the maximum hyperbolas of $(R_1, R_2)$ receiver pair.
Figure 1.2: Example of Candidate area. Candidate area (shaded) bounded by minimum hyperbolas constructed by two pair of receivers in an area of 800 m × 800 m. \( H^-(R_1, R_3) \) and \( H^-(R_3, R_1) \) are the minimum hyperbolas constructed by receiver pair \((R_1, R_3)\), and \( H^-(R_2, R_4) \) and \( H^-(R_4, R_2) \) are the minimum hyperbolas constructed by receiver pair \((R_2, R_4)\). The figure also demonstrates a successful localization.
shows a transmitter is successfully localized by two pairs of receivers.

**Failure in Localization**

A failure in localizing a transmitter means that a transmitter is not bounded by the hyperbolas constructed by a pair of receivers or a group of pair of receivers. Figure 1.3 shows a failure in localizing a transmitter.

**Factors that affect the size of the candidate area constructed by HPB**

Trusted receivers are the nodes that are known to the security system and which can never become rogues. Also they are very difficult to be controlled by a rogue node. In applications such as vehicle safety, these nodes could be road-side-units (RSUs) or public safety vehicles (PSVs). RSS received by each pair of these nodes are used to construct a minimum and a maximum hyperbola. A trusted receiver with good performance is the receiver that gives consistent performance in constructing minimum and maximum hyperbolas when it is paired up with any other receiver. There are many factors that can affect the performance of a receiver. As the HPB scheme is used to locate nodes in an open environment, the environment surrounding the receiver has a great effect on it. Environmental influences, like the obstacles between the transmitter and receiver, can cause a ray of waves from the transmitter to the receiver to reflect or diffract from its path. Therefore, the localization accuracy of each receiver pair has to be evaluated individually and then compared to the others.
Figure 1.3: *Example a Failure in Localization. A transmitter is outside the candidate area.*
Accuracy of Localization

Deterministic schemes that use RSS to calculate the location of a transmitter sometimes fail badly in pinpointing its exact location. This is due mainly to the fact that the shadowing effect cannot be accounted for by a fixed amount. On the other hand probabilistic schemes are very good in estimating the area of the region in which the transmitter might fall (the candidate area). The HPB scheme measures the accuracy by which minimum and maximum hyperbola pairs constructed by the receiver pairs bound the actual location of the transmitter. However, the problem of finding and selecting a subset of these receiver pairs that give a better success rate and/or a smaller candidate area size compared to all receiver pairs ones have not been studied so far.

1.2 Definition of the Problems

Probabilistic RSS-based geometric schemes that localize nodes by estimating the area in which the localized node resides, such as Barbeau and Robert [7] and Laurendeau and Barbeau [4], usually do that by giving a candidate area with some degree of confidence. In Barbeau and Robert [7] for example, the candidate area is an area bounded by an intersection of annuli. In HPB, the candidate area is the area bounded by a number of hyperbolas pairs.

Due to the shadowing effect, not all of the receiver pairs in HPB are expected to give the same degree of accuracy when constructing the hyperbolic area. Since HPB uses all the receiver pairs to construct the hyperbolic area, it is expected that the accuracy
given by the localization process is not optimum. In this thesis, we propose a modification of HPB that we believe will improve its performance. Our proposed system is called HPB-LA. It consists of two phases, a learning phase and an operational phase. In the learning phase the system uses real data to measure the accuracy of the receiver pairs, then it ranks the receiver pairs according to their accuracy, and finally it selects a subset of these receiver pairs to construct the candidate area. In the operational phase, the subset of receiver pairs that is identified in the learning phase is used to localize nodes.

Since the estimation of the EIRP range used by the transmitter is a big factor in determining the size of the candidate area and the success rate in HPB, we proposed a heuristic for estimating the EIRP range to improve the performance of HPB and HPB-LA in terms of the candidate area size.

To validate this proposed modification in HPB, an experiment has to be conducted to measure the RSS signal received by a number of trusted receivers. Then different pairs of hyperbolas have to be constructed from the RSS obtained from these receivers. After that, each pair of receivers is evaluated for the degree of accuracy it provides in terms of the success rate and size of the candidate area. According to this evaluation, a subset of receiver pairs is selected to construct the minimum and maximum hyperbolas between which the actual location of the transmitter is expected to be pinpointed with a compound degree of confidence.
1.3 Highlight of the Results

In this thesis, a study of the factors that affect the performance of HPB scheme is conducted, and ways for improving the HPB scheme are investigated. According to our analysis, a modified model of HPB is proposed which can give better success rate than the original HPB scheme. We also introduce a new heuristic for calculating the EIRP range that is assumed by the HPB that the transmitter is using. This heuristic has proved that it can improve the performance of the HPB in terms of the size of the candidate area. The new heuristic is able to improve the performance of HPB because it incorporates a more dynamic structure in formulating the equations that lead to a better estimation for the transmitter’s EIRP range. To evaluate our proposed transmitter’s EIRP range estimation heuristic further, we simulate a vehicular network scenario.

1.3.1 Enhancing the Success Rate of HPB

We propose a modification in the HPB scheme that improves its performance. The HPB scheme uses all possible receiver pairs to construct minimum and maximum hyperbola pairs for each one of the receiver pairs. Then the accuracy results from the combination of these receiver pairs represent the final accuracy that the system could provide. From our analysis in Section 5.2, we notice that some of these receiver pairs perform better than other pairs. We then investigate the possibility that a subset of these pairs could produce a better accuracy than using the whole pairs together. According to the same analysis results, choosing five receiver pairs out of six receiver pairs (the total number
of receiver pairs in this study is six) reduces the error from 8% to only 1% (i.e., this step increases the accuracy by 7%) while it increases the size of the candidate area by only 2%.

1.3.2 Improving the Candidate Area of HPB

The accuracy of localizing a transmitter could be improved if the EIRP range with which the transmitter is sending its messages is known to the trusted receivers. In the HPB, the heuristic used to calculate the range of EIRP that the transmitter might have used in broadcasting its messages relies on the assumption that the transmitter is close to one of the receivers. This receiver is the one that has received the highest RSS value among all of the receivers in the range of the transmitter. We call this receiver as MRR. Then, the shadowing effect on the RSS values is considered in estimating the EIRP range. This heuristic might give a good estimation for the lower and higher values of the EIRP range of the transmitter. However, this estimation process is quite static and does not use all the information provided by the group of receivers in the range of the transmitter. We propose a new heuristic to estimate the EIRP range. For each receiver pair that includes the MRR our heuristic assumes that the transmitter is located between MRR and the other receiver. By using this heuristic to estimate the EIRP possible ranges, the size of the candidate area is relatively decreased by 41% while the success rate decreases relatively by only 16%.
1.3.3 The Simulation Results

To evaluate our proposed transmitter EIRP range estimation heuristic further, we augment our experiment results with a vehicular scenario simulation. In this simulation two heuristics for estimating the EIRP range are used. The first heuristic is the heuristic that was used in the original HPB. The second heuristic is the heuristic we introduce in this thesis and is described in Chapter 4. The parameters used in this simulation are adopted from the large scale parameter values given by Liechty et al. [20], [21]. The values of path loss exponent and standard deviation used are 2.76 and 5.62 respectively. We augment the simulated RSS values independently by a randomly generated shadowing. Our simulation results agree with our experimental results in the fact that our proposed heuristic gives a smaller candidate area size than the heuristic used in the original HPB. Moreover, the same simulation results which have been presented in Section 5.3 show that, for any number of receivers, our proposed heuristic always gives a smaller candidate area in comparison to the heuristic used in the original HPB.

1.4 Outline of the Thesis

The rest of the thesis is organized as follows. Chapter 2 focuses on some of the work done on localization in wireless networks. It discusses the location finding schemes, including the HPB scheme.

Chapter 3 gives the theoretical details of the hyperbolas accuracy estimation using
HPB in Section 3.1. The factors that affect the hyperbola pair positions in HPB are discussed in Section 3.2.

Chapter 4 discusses our system of identifying and ranking the accuracies of the HPB hyperbola pairs. Also in this chapter a new heuristic for calculating an estimate of the EIRP range used by the transmitter is introduced.

Chapter 5 describes the experimental study that was carried out for measuring and validating the accuracy of the hyperbolic pair positions. The experiment was conducted in an open location. Received signal values from each transmission location were recorded for analysis. The hardware and software used to implement the experiment are described in Section 5.1. In Section 5.2, we give the analysis of the experimental results. The analysis of the simulation results is provided in Section 5.3.

Chapter 6 summarizes the conclusions and projects future work.
Chapter 2

Literature Review

Determining the physical location of a rogue node sometimes is the most important step in attributing an attack to its originator. RSS-based localization schemes in wireless networks have to model signal fading effects and random shadowing that come with the characteristics of signal propagation. Therefore, these localization techniques bear in their structure a randomized factor. As a result, the estimated location of the transmitter and its actual location may differ. Sometimes the level of confidence in the localization results in these schemes is given up front. However, ways of improving the accuracy level of some of these schemes have not been studied so far.

The main feature of the HPB scheme is the estimation of the transmitter position by constructing different pairs of minimum and maximum hyperbolas. The area between these different pairs of hyperbolas can be determined according to a given degree of confidence. Therefore, the size of the area in which the transmitter is expected to be found depends on the given level of confidence. As the results have shown, the HPB
scheme gives good accuracy in localizing a node. But ways of increasing the accuracy of this scheme have not been investigated.

We focus on improving the accuracy of the HPB scheme by making use of real data collected from a number of trusted receivers. We also examine the other factors that affect the performance of the HPB scheme, such as the heuristic used to calculate the range of EIRP that a transmitter might use for sending its messages. This chapter includes related work and discuss the different localization techniques and attempts that are carried on to improve the accuracy of some of these techniques.

Existing localization techniques in wireless networks can be categorized broadly into range-free [26], [27], [37], and range-based localization [14], [17]. This classification depends on the nature of the process that is used by them. The location of a node in range-based schemes is estimated by first estimating distances by utilizing direction, timing, or radio signal strength information. On the other hand, range-free localization schemes do not require this data for estimating a position of a node. Instead nodes connectivity and proximity information is used to estimate the positions of the nodes. Range-free schemes are used mainly in sensor networks. Section 2.1 discusses radio propagation and radio propagation mechanisms, while Section 2.2 examines the range-based localization techniques.
2.1 Radio Propagation

To understand signal fading we need to look at radio transmission. A radio transmission path between a transmitter and a receiver depends largely on the environment. This path ranges from a simple line-of-sight to ones that are obstructed by some objects like mountains, buildings, and/or foliage. Radio propagation is affected mainly by three factors: reflection, diffraction, and scattering. These three factors are defined by Rappaport [11] as follows:

(1) Reflection: Electromagnetic waves reflect when they encounter an object that has a very large dimensions compared to the travelling wavelength. Hence, objects like the earth, mountains, and buildings cause electromagnetic waves to reflect. Reflection causes electromagnetic waves to travel along different paths of varying lengths. Sometimes multiple reflections occur from different objects between the transmitter and receiver. This leads to what has been called multipath fading. For this reason, the signal strength decreases as the distance between the transmitter and receiver increases.

(2) Diffraction: Diffraction occurs when a surface with sharp edges obstructs the path between the transmitter and receiver. When electromagnetic waves encounter an object they bend around it.

(3) Scattering: Scattering occurs when the propagating electromagnetic waves have to travel through a medium different from the current one. This causes the electro-
magnetic waves to be partially reflected and partially transmitted. A typical example of scattering is radio waves move through rain or fogs.

To predict the average signal strength at a specific location, different models for electromagnetic waves propagation have been developed. There are two main propagation groups:

I. Large-scale propagation models:
Large-scale propagation models characterize signal strength for large distances (several hundreds or thousands of meters) between the transmitter and receiver.

II. Small-Scale propagation models:
Small-Scale propagation models characterize the rapid fluctuations of the received signal strength for short durations of time or for short travel distances (a few wavelengths).

Radio Propagation Models

One of the well known propagation models is the model that is used to predict received signal strength at any receiver in the range of the transmitter when there is a clear line of sight between the transmitter and that receiver. This model is known as free space propagation model [11]. The free space propagation model can be applicable for example to microwave line-of-sight links and satellite communication systems.

In the free space model, Friis equation calculates the power received by a receiver an-
tenna at a distance $d$ from the transmitter antenna.

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L}$$  \hspace{1cm} (2.1)

Where:

$P_t$ is the transmitted power.

$P_r(d)$ is the received power.

$G_t$ is the transmitter antenna gain.

$G_r$ is the receiver antenna gain.

$d$ is the distance between the transmitter antenna and the receiver antenna.

$L$ is the system loss factor which is not related to propagation, and it is always greater or equal to one.

$\lambda$ is the wavelength in meters.

We also have the following equation,

$$G = \frac{4\pi A_e}{\lambda^2}$$  \hspace{1cm} (2.2)

Where $A_e$ is the effective aperture, which is related to the physical size of the antenna.

$\lambda$ is related to carrier frequency by the following equation:

$$\lambda = \frac{c}{f}$$  \hspace{1cm} (2.3)

Where:

$f$ is the carrier frequency in Hertz.

$c$ is the speed of light.
2.2 Range-based Localization Techniques

In this section, we discuss range-based localization techniques. In these schemes, nodes are localized by first estimating their distances by utilizing direction, timing, or radio signal strength information.

2.2.1 Angle of Arrival (AoA)

A location of a wireless node can be calculated by estimating the AoA. The AoA is measured with respect to other nodes in the range of the node that needs to be localized. Figure 2.1 shows an example of AoA.

In practice, the AoA measurement is strongly affected by the fact that magnetic waves follow multiple paths. As the AoA measurement estimate is obtained from the strongest path, its measurement can have a substantial error. The AoA is measured by a couple of ways. One way of estimating the angle is by phase differences in signal received by two or more antennas as described by Eren [3]. The other way of estimating AoA is by using RSS ratio between multiple directional antennas [5], which is called array of antennas. In the following explanation, we assume that only two directional antennas are used. The AoA estimation is done by having the two directional antennas point in different directions in the way that their major rays overlap. Then the ratio of RSS
Figure 2.1: Angle of Arrival. If the orientation is known, angles $\theta$ and $\beta$ represent the relative AoA measured at $N_u$ (location unaware node) from signals sent by $N_A$ and $N_B$ (nodes with known locations). Then triangulation is used to determine the position of $N_u$.

Values obtained from these two major rays is used to estimate the AoA. The accuracy of the AoA depends mainly on the radio propagation environments.

The main drawback of the AoA method of localization is that all the channels have to have the same effect on the received signal. Moreover, array sizes affect very largely the cost of the receivers. The accuracy of the AoA localization system and the number of antennas are directly related, i.e. decreasing the number of antennas decreases the accuracy.
2.2.2 Time of Arrival (TOA)

The time taken for the signal to travel from one node to another is called the TOA. This time can be used as a parameter for localization. The result of multiplying this time by the propagation speed of the signal gives the distance between transmitter and receiver. Delay at the processing queues of the transmitter node or the receiver node can affect the measurement of TOA. Localization using the method of TOA can be carried out by using ultrasound (US) frequencies or radio frequencies (RF). The frequency at which RF is working is much higher than the one at which US is working. Therefore, RF has much higher time resolution requirement for the TOA schemes. Hence, TOA localization schemes that use RF have a higher accuracy than the ones that use US.

Another way of calculating the time of arrival is by measuring the time of flight using two different signals with two different propagation speeds such as RF and US. The RF signal is used to synchronize the transmitter and receiver while the US signal is used for measuring the time of flight. TOA information collected by a number of trusted nodes can be used to estimate the location of a non-collaborative node as outlined in Waters and Felten [30], Brands and Chaum [28], and Sastry et al. [32]. However, TOA has a serious problem in installing and maintaining hardware for time synchronization as outlined by Park et al. [38].
2.2.3 Time Difference Of Arrival (TDOA)

The TDOA technique is also known as the hyperbolic position location technique. The TDOA utilizes the cross-correlation process to calculate the time of arrival (TOA) of a transmitter signal at two or more base stations or access points as outlined by Ralph Bucher and D. Misra in [34], and Cosma et al. in [36]. The receivers represent the foci and the distance differences between the receivers are used to construct the hyperbolas. The intersection of all these hyperbolas provides the estimated location of the transmitter. Ho [29] proposed an algebraic approach to estimate a location of a transmitter by solving non-linear hyperbolic equations. TDOA, however, suffers the same problem of time synchronization as TOA as it has been described in Section 2.2.2.

2.2.4 Received Signal Strength (RSS)

Wireless nodes communicate with each other using radio signals. The power of a radio signal measured at a receiver is called Received Signal Strength. RSS can easily be measured and most of the wireless nodes don’t require additional hardware to measure RSS. Therefore, RSS became a very popular method for estimating locations of nodes. The relationship between the transmission power and RSS is proportional to the distance between the receiver and the transmitter. In theory, this relationship is formalized by Frii’s transmission equation (1) as shown by Liu et al. [8].

Environmental conditions such as obstacles, reflection, diffraction, scattering and in-
Figure 2.2: Example of TDOA. Hyperbola $H_{(R_1,R_2)}$ is constructed from the distance difference calculated from TOA of the transmitter $T$ signal at receiver $R_1$ and receiver $R_2$. Similarly, Hyperbola $H_{(R_1,R_3)}$ is constructed from the distance difference calculated from TOA of the transmitter $T$ signal at receiver $R_1$ and receiver $R_3$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
terferences make the ideal transmission difficult or impossible to achieve. Therefore, different propagation models have been constructed. These models are usually known as path loss models.

Attenuation of the signal strength level from a transmitter to a receiver, as it has been defined by a path loss model for an environment, is quantified either through theoretical assumptions or experimental analysis. These path loss models make it possible to calculate the distance between a transmitter and a receiver from the RSS. Localization techniques that are based on RSS can be divided broadly into two categories: signature dependent schemes and geometric base schemes.

Localization using Signalprints schemes

Localization in signature dependent schemes relies on a training set that is available for the specific environment. The signalprints of the training set are developed or constructed during an offline phase. During the inline phase the data in the training phase is consulted to estimate a location of a node. In other words, the signalprints method is a pattern matching method. An example of this scheme is the scheme proposed by Faria et al. [19] to localize a rogue mobile node in an indoor environment using WiFi/802.11 standards.

The node that needs to find its location measures the RSSs from each of the trusted nodes and stores them in a tuple, then this tuple is compared with the signal prints from the database of the signalprints for matching. Conversely, a number of trusted
nodes can localize a node by comparing the signalprint produced by the RSS received by them to the signalprints found in the database, as outlined by Faria et al. [19], Ross et al. [22], Ladd et al. [25], and Bahl and Padmanabhan [1].

The main disadvantage of this technique is the amount of the initial effort required to collect the signalprints data from experiments. This technique is used mainly in indoor environments. Also this scheme assumes that the node that needs to be localized collaborates [39] with the trusted nodes and reliably measures the signal strength from the beacon messages it receives from trusted nodes.

**Localization using Geometric RSS-based schemes**

In this scheme, a node position is estimated by utilizing the RSS information received by a number of trusted receivers as outlined in Bo-Chieh Liu et al. [31] and Chong Liu et al. [23]. Barbeau and Robert [7] introduced a probabilistic RSS-based geometric model to localize a rogue base station. This scheme is used for outdoor environments such as WiMAX/802.16. HPB [4] on which our proposed scheme is based is another example of a probabilistic RSS-based geometric scheme. In the HPB scheme, the trusted receivers form pairs. Each pair calculates the minimum and maximum distance differences from a transmitter to the two receivers. Then these distances are used to construct the minimum and maximum hyperbolas. The hyperbolas constructed by all the receivers contain the transmitter with a combined degree of confidence.

Many attempts have been made to improve the accuracy of some of the geometric RSS-based schemes. For example: Bo-Chieh Liu et al. [33] tried to correct the error
resulting from the shadowing, multiple-access interference, and nonline-of-sight effects in the scheme proposed by them in [31].

In this thesis we are aiming to improve the accuracy of HPB by employing a learning process. Our approach depends on real data collected from a number of trusted receivers. We call this approach HPB-LA. In HPB-LA, each receiver pair in the network is evaluated for its performance. Then, the receiver pairs are ranked according to their measured accuracy, mainly the success rate and candidate area size. A number of subsets of receiver pairs are formed and their performance is evaluated. Finally, the subset with the highest performance is selected for constructing the final hyperbolas. We believe this approach will give a better accuracy than HPB.
Chapter 3

Hyperbolic Position Bounding (HPB)

Our proposed scheme is based on HPB, which was put forward by Laurendeau and Barbeau in [4]. In this chapter, we give the theoretical details of HPB, and we also discuss the main factors that affect the HPB performance.

How does HPB work?

HPB estimates the probabilistic minimum and maximum transmitter-receiver distances. This estimation is done utilizing the RSS values reported by a number of trusted receivers. Then, the minimum and maximum transmitter-receiver distances are employed to compute the minimum and maximum bounds on the distance difference from a transmitter to a pair of receivers. Each pair of receivers calculates the
minimum and maximum distance differences between the two receivers. Then, the corresponding minimum and maximum hyperbolas are constructed by each pair of receivers. The transmitter is bounded in the intersection of these hyperbolas with a compound degree of confidence.

**HPB Threat Model**

The HPB identifies the threat model as an attacker broadcasts a message as an RF signal. This radio signal would reach all the receivers in its range. In this threat model, the attacker is assumed to be a mobile node. It is also assumed that there are a number of trusted receivers within the range of the attacker signal. These trusted receivers have the ability to communicate securely between them. Part of the communication between the receivers is dedicated for aggregating the RSS values received from the attacker's messages. Moreover, the threat model assumes that the coordinates of the trusted receivers are known globally. This threat model is feasible in many of the wireless networks. In vehicular networks [18] as an example, trusted receivers can be Road-Side units or On-Board units. Also WiMAX/802.16 [40] is another example, where Base Stations could serve as trusted receivers.

The threat model does not have any assumption for the EIRP employed by the attacker. This is due to the fact that the rogue node may change its power level from time to time to disguise its actual location.
3.1 Factors that Affect the Hyperbola Pair Positions

HPB follows the large scale path loss model. In this section we are going to outline the large scale path loss model as described by Rappaport [11], and then discuss the factors that affect this model.

The Log-Normal Shadowing Model

The log-normal shadowing model given by Rappaport [11] is a statistical path loss model. In other words, it computes statistically the path loss value for a signal received at distance $d$ from a transmitter. To compute the path loss for a signal, first the model estimates the signal loss at a number of different distances from the transmitter. This estimation of signal loss is based on three parameters. The first parameter is a pre-defined reference distance $d_0$ close to the transmitter. The second parameter is the path loss exponent $\nu$. The value of $\nu$ is dependent on the propagation environment. The third parameter is the standard deviation of the path loss $\sigma$. Usually the values of $\nu$ and $\sigma$ are estimated by a regression analysis technique using the values of the path losses calculated from experimental measurements.

In the log-normal shadowing model, the path loss $L(d)$ of a signal at a distance $d$ is described by Rappaport, as a Normal distribution random variable. The mean of this random variable is $\mu = \bar{L}(d)$, and $\sigma$ as standard deviation. The mean of the path
loss is defined by the following equation:

\[ L(d) = L(d_0) + 10\nu \log(d/d_0) \]  
(3.1)

Where \( L(d_0) \) is the average path loss at the \( d_0 \) (the reference distance). Therefore, and according to Rappaport, the log-normal shadowing can be expressed by the equation below:

\[ L(d) = \bar{L}(d) + Y_\sigma \]  
(3.2)

Where \( Y_\sigma \) is a Normal distribution zero-mean random variable with standard deviation \( \sigma \).

The following equation is given by Rappaport for a free space propagation, which is actually summarizing Equation 3.1 and Equation 3.2:

\[ L(d) = \bar{L}(d_0) + 10\nu \log(d/d_0) + Y_\sigma \]  
(3.3)

For any selected level of confidence \( C \), the value of \( Y_\sigma \) is in the interval \([ -z\sigma, +z\sigma ]\).

The value of \( z \) can be calculated as follows:

\[ z = \Phi^{-1}((1 + C)/2) \]  
(3.4)

Then \( z \) can be obtained from a Normal distribution table. Therefore, and as indicated by Laurendeau and Barbeau [4], for a confidence level \( C \), the path loss \( L(d) \) of a signal at a distance \( d \) from the transmitter is defined as follows:
Transmitter-Receiver Distance Range with Log-Normal Shadowing

The distance range is the difference between the minimum distance and the maximum distance from a transmitter to a receiver. This range can be calculated from the path loss using Rappaport’s log-normal shadowing model as it has been demonstrated by Barbeau and Robert in [7].

For a confidence level $C$, the minimum distance and maximum distance between a transmitter and a receiver, as demonstrated by Laurendeau and Barbeau in [4], can be calculated as follows:

\[
    d^- = d_0 \times 10^{\frac{L(d) - L(d_0) - z\sigma}{10\sigma}} \\
    d^+ = d_0 \times 10^{\frac{L(d) - L(d_0) + z\sigma}{10\sigma}}
\]

Obtain Path Loss from RSS value:

The RSS measurements obtained by the receivers in HPB threat model can be used to calculate the path loss if the EIRP value is known, and that is as follows:

\[
    L(d) = \text{EIRP} - \text{RSS}_i
\]
Where $RSS_i$ is the signal strength measured at receiver $i$.

The value of EIRP is unknown, as the HPB threat model has indicated earlier. Therefore, the EIRP is replaced with the possible EIRP range which is bounded by a minimum and a maximum EIRP. If we denote the minimum EIRP as $P^-$ and the maximum EIRP as $P^+$, similarly if we denote the minimum distance as $d^-$ and the maximum distance as $d^+$ we can conclude that the transmitter is in the interval $[d^-, d^+]$ with a confidence level $C$. We can compute $d^-$ and $d^+$ as follows:

$$d^- = d_0 \times 10^{\frac{P^- - RSS_i - L(d_0) + z\sigma}{10z\sigma}}$$  \hspace{1cm} (3.8)

$$d^+ = d_0 \times 10^{\frac{P^+ - RSS_i - L(d_0) + z\sigma}{10z\sigma}}$$  \hspace{1cm} (3.9)

The fact that a transmitter $T$ can be located from a given trusted receiver $k$ in the interval $[d^-, d^+]$ with a confidence $C$ can be stated as follows:

$$Pr(d^- \leq d(T, R_k) \leq d^+) = C$$

From what we have shown above, it is clear that the minimum distance and maximum distance between a transmitter with a given EIRP range and a receiver are bounded on the shadowing range $[-z\sigma, +z\sigma]$, where $z$ is derived from the confidence level $C$. 

32
Distance Differences for a given signal shadowing range

For a shadowing range \([-z\sigma, +z\sigma]\) for a confidence level \(C\) the distance differences can be given by the following two equations as it has been proved by Laurendeau and Barbeau in [4]:

\[
\Delta d_{k,t}^- = d_0 \times 10^{\frac{P^- - \text{RSS}_k - L(d_0) - z\sigma}{10\sigma}} - d_0 \times 10^{\frac{P^- - \text{RSS}_k - L(d_0) + z\sigma}{10\sigma}} \tag{3.10}
\]

\[
\Delta d_{k,t}^+ = d_0 \times 10^{\frac{P^+ - \text{RSS}_k - L(d_0) + z\sigma}{10\sigma}} - d_0 \times 10^{\frac{P^+ - \text{RSS}_k - L(d_0) - z\sigma}{10\sigma}} \tag{3.11}
\]

Plotting the Hyperbolas:

The minimum and the maximum bounding hyperbolas can be plotted by the following two equations:

\[
\sqrt{(x - x_k)^2 + (y - y_k)^2} - \sqrt{(x - x_t)^2 + (y - y_t)^2} = \Delta d_{k,t}^- \tag{3.12}
\]

\[
\sqrt{(x - x_k)^2 + (y - y_k)^2} - \sqrt{(x - x_t)^2 + (y - y_t)^2} = \Delta d_{k,t}^+ \tag{3.13}
\]

Then the location of the transmitter is bounded by these two hyperbolas.

For a number of trusted receivers, minimum and maximum hyperbolas are constructed between each of two different receivers (pair) \(R_k\) and \(R_t\). The transmitter is bounded in the intersection of these hyperbolas with a compound degree of confidence.
Figure 3.1 shows an example of minimum hyperbolas for one receiver pair, while Figure 3.2 shows an example of minimum hyperbolas for all possible receiver pairs for four trusted receivers.

**EIRP Range Estimation**

HPB estimates the range of EIRP \([P^-, P^+]\) that the rogue might use for sending its messages heuristically. The heuristic for estimating the EIRP range uses the fact that the trusted receivers have the ability to communicate using a secure channel. This fact implies that the trusted receivers can aggregate the information collected from the attacker messages. This information for HPB is mainly the RSS measurements. Then the receiver which receives the highest power becomes known to the other receivers as the MRR. It is likely that the MRR is the nearest receiver to the attacker. With each receiver using a confidence level \(C\) that produces the shadowing range \([-z\sigma, +z\sigma]\) the \(P^-_k\) and \(P^+_k\) for any receiver \(k\) can be calculated using the following two equations which are derived from the log-normal shadowing model:

\[
P^-_k = \bar{L}(d0) + 10n \log\left(\frac{d_k}{d0}\right) + rss_k - z\sigma
\]

\[
P^+_k = \bar{L}(d0) + 10n \log\left(\frac{d_k}{d0}\right) + rss_k + z\sigma
\]

Where \(d_k\) is the distance between \(R_k\) and the MRR.
Figure 3.1: Example of Minimum Hyperbolas. A transmitter is bounded by minimum hyperbolas constructed by a pair of receivers in an area of 800 m x 800 m. $H^-(R_1, R_3)$ and $H^-(R_3, R_1)$ represent the minimum hyperbolas for receiver pair $(R_1, R_3)$. $H^-(R_2, R_4)$ and $H^-(R_4, R_2)$ represent the minimum hyperbolas for receiver pair $(R_2, R_4)$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 3.2: Example of Minimum Hyperbolas for all possible receiver pairs. The figure also shows the candidate area (shaded).
The Heurisitic

Let:

The set of all trusted receivers be \( R \), and the size of \( R \) be \( n \).

The set that contains the \( P_k^- \) for \( R_k \) belong to \( R\{MRR\} \) is \( \Psi^- \) sorted in ascending order.

The set that contains the \( P_k^+ \) for \( R_k \) belong to \( R\{MRR\} \) is \( \Psi^+ \) sorted in ascending order.

1: \( i \leftarrow n-1 \) \{i is the index of the current supremum of the EIRP values with minimal shadowing, it starts from n-1, where n-1 is the size of \( R\{MRR\} \}\}

2: \( j \leftarrow 1 \) \{j is the index of the current infimum of the EIRP values with maximal shadowing, it starts from 1}\}

3: \textbf{while} \( i > 0 \) and \( j < n \) \textbf{do} \{Iterate through the ordered sets \( \Psi^- \) and \( \Psi^+ \}\}

4: \textbf{if} \( \Psi_i^- < \Psi_j^+ \) \textbf{then} \{If the current supremum of the EIRP values with minimal shadowing is lower than the current infimum of the EIRP values with maximal shadowing\}

5: \( P^- = \Psi_i^- \) \{The transmitter minimum EIRP is the current supremum of the EIRP values with minimal shadowing\}

6: \( P^+ = \Psi_j^+ \) \{The transmitter maximum EIRP is the current infimum of the EIRP values with maximal shadowing\}

7: \textbf{Exit} \{Exit the while loop\}

8: \textbf{end if}

9: \textbf{if} \( i > 1 \) \textbf{then}
10: if $\Psi_{i-1}^- < \Psi_j^+$ then \{If the next supremum of the EIRP values with minimal shadowing is lower than the current infimum of the EIRP values with maximal shadowing\}

11: \( P^- = \Psi_{i-1}^- \) \{The transmitter minimum EIRP is the next supremum of the EIRP values with minimal shadowing\}

12: \( P^+ = \Psi_j^+ \) \{The transmitter maximum EIRP is the current infimum of the EIRP values with maximal shadowing\}

13: Exit \{Exit the while loop\}

14: end if

15: end if

16: \( i = i - 1 \) \{Update i\}

17: \( j = j + 1 \) \{Update j\}

18: end while
3.2 Factors that Affect the HPB Performance

There are mainly two factors that affect the HPB performance:

(A) Trusted Receivers Locations

The performance of HPB depends mainly on RSS values measured at the trusted receivers in the range of the transmitter. The fluctuation in RSS or the shadowing has been taken care of in the calculation of the distance differences for the receiver pairs. However, the error in RSS due to local shadowing conditions as described by Liu et al. [33] are not accounted for. Therefore, the receivers with high such RSS fluctuations are expected to give less accurate results when they are paired up with each other.

(B) The Transmitter EIRP range Estimation

The estimation of the minimum distance and maximum distance between a transmitter and receiver depends largely on the estimation of the minimum EIRP and maximum EIRP that may be used by the transmitter as indicated by the following two equations as were given in Section 3.1:

\[ d^- = d_0 \times 10^{\frac{P^-_{\text{RSS}} - L(d_0) - \phi}{10}} \]  \hspace{1cm} (3.16)

\[ d^+ = d_0 \times 10^{\frac{P^+_{\text{RSS}} - L(d_0) + \phi}{10}} \]  \hspace{1cm} (3.17)
Where \( \text{RSS}_i \) is RSS at receiver \( i \).

A smaller \([P^-, P^+]\) range is expected to give a smaller candidate area between receiver pairs and that is for the following reason:

We have \((P^- - \text{RSS}_i)\) and \((P^+ - \text{RSS}_i)\) represent the path losses when the transmitter uses the minimum power and maximum power respectively. If the value \( P^− \) is so close to the value of \( P^+ \) then we will have \( d^- \) as close as possible to \( d^+ \). This also gives the same effect in the calculated values of \( \Delta d_{i,j}^−, \Delta d_{j,i}^−, \Delta d_{i,j}^+, \) and \( \Delta d_{j,i}^+ \) for receiver pair \((R_i, R_j)\). Therefore, the area between the hyperbolas constructed by receiver pairs is minimum.
Chapter 4

HPB with Learning Ability

(HPB-LA)

In this chapter we discuss our system of modifying HPB to improve its accuracy. We call this system HPB-LA. Also in this chapter we introduce a new heuristic for calculating an estimate of the EIRP range used by the transmitter. We expect this heuristic to improve the candidate area size in both HPB and HPB-LA.

How does the HPB-LA work?

The HPB-LA scheme is based on HPB, details of which we presented in chapter 3. HPB-LA is exploring the possibility of having a subset of receiver pairs to construct the minimum and the maximum hyperbolas that may bound the location of the transmitter with a success rate greater than the success rate that could be provided by
the original HPB. At the same time the candidate area is to increase, but by a small fraction. HPB-LA consists of two phases, a learning phase and an operational phase. In the learning phase, real data is used to train the system to select a subset of receiver pairs that gives better accuracy in localizing a transmitter than the set of all possible receiver pairs. Then in the operation phase, the subset of receiver pairs that is identified in the learning phase is used to localize nodes.

The learning phase consists of four major steps. But before explaining these steps we want to provide the following definitions:

Receiver Pairs Success Rate (RPSR):
RPSR is the ratio of the number of times the transmitter is successfully located in the area bounded by the minimum and the maximum hyperbolas that are constructed by a receiver pair or a group of receiver pairs to the total number of localization attempts. Figure 4.1 shows a transmitter successfully bounded by minimum and maximum hyperbolas constructed by a pair of receivers.

Receiver Pairs Candidate Area Size (RPCAS):
RPCAS is the size of the candidate area bounded by the minimum and the maximum hyperbolas that are constructed by a receiver pair or a group of receiver pairs. The shaded area in Figure 4.1 represents the candidate area that bounded by minimum and maximum hyperbolas constructed by two receiver pairs.
Figure 4.1: Example of Successful Bounding. A transmitter is successfully bounded by minimum and maximum hyperbolas constructed by a pair of receivers in an area of 800 m × 800 m. $H^-(R_1, R_2)$ and $H^-(R_2, R_1)$ are the minimum hyperbolas, and $H^+(R_1, R_2)$ and $H^+(R_2, R_1)$ are the maximum hyperbolas of $(R_1, R_2)$ receiver pair. The direction of the open part of the hyperbola depends on the distance difference sign (+/-). If the minimum distance difference and maximum distance difference have opposite signs then the open parts of the minimum hyperbola and maximum hyperbola have opposite directions.
Receiver Pairs Score (RPS):

We define RPS for a set $G$ of receiver pairs as the ratio of the difference between $RPSR(G)$ and $RPSR(HPB)$, and the difference between $RPCAS(G)$ and $RPCAS(HPB)$, i.e.

$$RPS = \frac{RPSR(G) - RPSR(HPB)}{RPCAS(G) - RPCAS(HPB)}$$

Where $RPSR(G)$ is the $RPSR$ of the receiver pairs in $G$, $RPSR(HPB)$ is the average $RPSR$ of all possible receiver pairs, $RPCAS(G)$ is the $RPCAS$ of the receiver pairs in $G$, and $RPCAS(HPB)$ is the $RPCAS$ of all possible receiver pairs.

The Learning Phase

Step 1: Measuring Receiver Pairs Accuracies

For each receiver pair, we measure the $RPSR$ and $RPCAS$ using half of the data collected from an experiment.

Step 2: Ranking of Receiver Pairs

After each receiver pair is evaluated for the RPSR rate and RPCAS, then the pairs are ranked according to the following two parameters:

(A) Optimizing Receiver Pairs Success Rate

Here the receiver pairs are ranked according to their RPSR, with the receiver pair that
has the highest RPSR at the top of the list and the receiver pair with the lowest RPSR at the bottom of the list.

(B) Receiver Pairs Candidate Area Optimization
In this process the receiver pairs are ranked according to their RPCAS, with the receiver pair with the smallest RPCAS at the top of the list and the receiver pair with the largest RPCAS at the bottom of the list.

Step 3: Receiver Pairs Subsets Construction
After ranking the receiver pairs according to their performance, the receiver pairs are grouped into subsets. These subsets are formed to optimize either the success rate or the candidate area size.

Step 4: Selection of HPB-LA Receiver Pairs
The RPSs of the subsets formed in Step 3 are calculated. Then the subsets are compared according to their PRS. The subset with the highest RPS is selected to be the set that will be used to construct the hyperbolas of HPB-LA.

Finally, HPB-LA is evaluated using the other half of the data. The following example shows how HPB-LA is implemented.

Example:
Suppose that there are four receivers labelled $R_a$, $R_b$, $R_c$, and $R_d$ in the range of a
transmitter. Then assume that the RSS values from the transmitter's messages are recorded by the receivers. Furthermore, assume that the first half of the collected RSS values are used to calculate the RPSR and the RPCAS for all possible receiver pairs of the four receivers individually. Then these RPSR and RPCAS values are measured and recorded in the following table:

Step 1:

<table>
<thead>
<tr>
<th>Receiver Pair</th>
<th>Receiver Pairs Success Rate</th>
<th>Receiver Pairs Candidate Area Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>((R_a, R_b))</td>
<td>100</td>
<td>84</td>
</tr>
<tr>
<td>((R_a, R_c))</td>
<td>99</td>
<td>82</td>
</tr>
<tr>
<td>((R_a, R_d))</td>
<td>100</td>
<td>86</td>
</tr>
<tr>
<td>((R_b, R_c))</td>
<td>81</td>
<td>70</td>
</tr>
<tr>
<td>((R_b, R_d))</td>
<td>94</td>
<td>89</td>
</tr>
<tr>
<td>((R_c, R_d))</td>
<td>100</td>
<td>88</td>
</tr>
</tbody>
</table>

Table 4.1 The success rate and the candidate area size of receiver pairs.

Step 2a: Ranking the receiver pairs according to RPSR (Receiver Pairs Success Rate Optimization):
Step 2b: Ranking the receiver pairs according to Candidate Area Size (Receiver Pairs Candidate Area Size Optimization):

<table>
<thead>
<tr>
<th>Receiver Pair</th>
<th>Receiver Pairs Success Rate</th>
<th>Receiver Pairs Candidate Area Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R₀, R₅)</td>
<td>81</td>
<td>70</td>
</tr>
<tr>
<td>(R₆, R₇)</td>
<td>99</td>
<td>82</td>
</tr>
<tr>
<td>(R₆, R₇)</td>
<td>100</td>
<td>84</td>
</tr>
<tr>
<td>(R₆, R₉)</td>
<td>100</td>
<td>86</td>
</tr>
<tr>
<td>(R₆, R₃)</td>
<td>100</td>
<td>88</td>
</tr>
<tr>
<td>(R₆, R₆)</td>
<td>94</td>
<td>89</td>
</tr>
</tbody>
</table>

Table 4.3 Receiver pairs ordered according to their candidate area size.
Step 3:

For optimizing the RPSR, we form the following subsets (total of $N-2$ subsets, where $N$ is number of the receiver pairs):

1. $S_1 = \{ (R_a, R_b), (R_a, R_d) \}$
2. $S_2 = \{ (R_a, R_b), (R_a, R_d), (R_c, R_d) \}$
3. $S_3 = \{ (R_a, R_b), (R_a, R_d), (R_c, R_d), (R_a, R_c) \}$
4. $S_4 = \{ (R_a, R_b), (R_a, R_d), (R_c, R_d), (R_a, R_c), (R_b, R_d) \}$

Similarly, for optimizing the RPCAS, we form the following subsets:

1. $S_5 = \{ (R_b, R_c), (R_a, R_c) \}$
2. $S_6 = \{ (R_b, R_c), (R_a, R_c), (R_a, R_b) \}$
3. $S_7 = \{ (R_b, R_c), (R_a, R_c), (R_a, R_b), (R_a, R_d) \}$
4. $S_8 = \{ (R_b, R_c), (R_a, R_c), (R_a, R_b), (R_a, R_d), (R_c, R_d) \}$

Step 4:

The RPS for the subsets $S_1, S_2, S_3, \ldots, S_8$ are calculated and compared to each other. Then, the subset with the highest RPS is selected to be the set that will construct the hyperbolas of HPB-LA.

Finally: Using the second half of the collected data, the accuracy of HPB-LA is evaluated.
The relationship between the number of receiver pairs and the success rate

For a subset of receiver pairs, the location of the transmitter is bounded by the minimum and maximum hyperbolas constructed by each one of the receiver pairs. To have the transmitter successfully localized by this subset of receiver pairs, we have to have the transmitter successfully localized by each one of the receiver pairs. Therefore, adding more receiver pairs is more likely to decrease the success rate of a subset and vice versa. Also, a subset of receiver pairs that is composed of receiver pairs with high success rates is more likely to give a high success rate.

The relationship between the number of receiver pairs and the candidate area size

For a subset of receiver pairs, the candidate area is the area bounded by the minimum and maximum hyperbolas constructed by each one of the receiver pairs. This implies that the candidate area that is constructed by a subset of receiver pairs is the intersection of the areas that are constructed by each one of the receiver pairs. Therefore, the size of the candidate area is more likely to become smaller if more receiver pairs are added to the subset, and vice versa. Also, a subset of receiver pairs that is composed of receiver pairs with smaller candidate area is more likely to give a smaller size candidate area.
The Complexity of the Learning Phase

The complexity of HPB-LA learning phase is $O(n^2)$, where $n$ is the number of receivers.

The Scalability of the Learning Phase

HPB-LA is easily scalable because the number of measurements required to train the system is not big as for example in the signalprints schemes.

Update of the Learning Phase

Update of HPB-LA learning phase is required only when the success rate drop below a predefined threshold or the environment around the receivers changes.

Transmitter EIRP Level

As the transmitter might change its power level to disguise its location, no assumption is made about its power. The transmitter EIRP range is estimated heuristically.

Cases not handled by HPB-LA

When there are two transmitters that are transmitting at the same time, the system becomes busy.
Receiver Pair Based EIRP Range Estimation in HPB

The trusted receivers use a secure channel to aggregate the information they have collected from the attacker's messages. This information for HPB is mainly the RSS measurements. The receiver that receives the highest signal becomes known to the other receivers as the MRR. This implies that the \( (MRR) \) is the nearest receiver to the attacker. In this heuristic we assume that each receiver uses a confidence level \( C \) which gives a shadowing range \([-z\sigma, +z\sigma]\). The heuristic first considers the pairs that include the \( MRR \). For each one of these pairs the heuristic assumes that the transmitter is between the \( MRR \) and \( R_k \) (the receiver paired with the \( MRR \)), and that the transmitter is nearer to the \( MRR \) than to \( R_k \). The relationship between the two distances is captured by the following equation:

\[
d_{(k,MRR)} = d_{(MRR,T)} + d_{(k,T)}
\]

Where \( d_{(k,MRR)} \) is the distance between the \( MRR \) and \( R_k \), and \( d_{(MRR,T)} \) and \( d_{(k,T)} \) is the distances of the \( MRR \) and receiver \( R_k \) respectively from the estimated position of the transmitter.

The value of \( d_{(k,T)} \) can be found by solving the following two equations (which are derived from the log-normal shadowing model):

\[
P_k^* = \bar{L}(d0) + 10n \log(d_{(k,T)}/d0) + rss_k - z\sigma \tag{4.1}
\]

\[
P_m^* = \bar{L}(d0) + 10n \log(d_{(MRR,T)}/d0) + rss_m - z\sigma \tag{4.2}
\]
Where $P^+_k$ is the minimum power used by the transmitter to reach the receiver $k$ with a signal strength $r_{ss_k}$, and $P^-_m$ is the minimum power used by the transmitter to reach the MRR with a signal strength $r_{ss_m}$. Since $r_{ss_k}$ and $r_{ss_m}$ are measured for the same message then we have to have:

$$P^-_k = P^-_m$$

We can calculate the value of $P^-_k$ by first calculating the value of $d(k,T)$ and then substituting it back in equation 4.1.

Similarly, we have:

$$P^+_k = \bar{L}(d_0) + 10n \log(d(k,T)/d_0) + r_{ss_k} + z\sigma \quad (4.3)$$

$$P^+_m = \bar{L}(d_0) + 10n \log(d(MRR,T)/d_0) + r_{ss_m} + z\sigma \quad (4.4)$$

Where $P^+_k = P^+_m$, and $P^+_k$ is calculated by substituting the value of $d(k,T)$ in Equation 4.3.

**Solving Equation 4.1 and Equation 4.2 from above:**

Subtract Equation 4.2 from Equation 4.1:

$$0 = 10n \log(d(k,T)/d_0) + r_{ss_k} - 10n \log(d(MRR,T)/d_0) - r_{ss_m}$$

$$r_{ss_m} - r_{ss_k} = 10n \log(d(k,T)/d_0) - 10n \log(d(MRR,T)/d_0) = 10n \log(d(k,T)/d(MRR,T))$$

$$\log(d(k,T)/d(MRR,T)) = (r_{ss_m} - r_{ss_k})/10n$$

$$d(k,T) = d(MRR,T) \times 10^{(r_{ss_m} - r_{ss_k})/10n}$$

$$d(k,T) = (d(k,MRR) - d(k,T)) \times 10^{(r_{ss_m} - r_{ss_k})/10n}$$

$$d(k,T) + d(k,T) \times 10^{(r_{ss_m} - r_{ss_k})/10n} = d(k,MRR) \times 10^{(r_{ss_m} - r_{ss_k})/10n}$$
\[ d_{(k,T)} = \left( d_{(k,MRR)} \times 10^{(\text{rss}_{m}-\text{rss}_{k})/10n} \right) / \left( 1 + 10^{(\text{rss}_{m}-\text{rss}_{k})/10n} \right) \]

Receiver \( k \) uses \( P_k^- \) as its minimum power and \( P_k^+ \) as its maximum power for calculating the distance differences for the MRR and receiver \( k \) pair. The receivers in the pairs that do not include MRR use \( \max\{P_k^\pm\} \) as their minimum power and \( \min\{P_k^\pm\} \) as their maximum power.

The Heuristic

Let:

- \( R_k \) be the MRR.
- \( d_{IT} \) be the estimated distance between receiver \( R_i \) and transmitter \( T \).
- \( P_{ik}^- \) be the power used to calculate the distance \( d_{IT} \).
- \( P_{ik}^+ \) be the power used to calculate the distance \( d_{IT} \).
- \( P_{\text{lower}} = \phi \) be the set of lower powers used by the receivers.
- \( P_{\text{higher}} = \phi \) be the set of higher powers used by the receivers.
- \( d_{kl} \) be the distance between receiver \( R_k \) and receiver \( R_l \).
- \( n \) be the number of the receivers.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
1: for $i = 1$ to $n$ do
2: 
3: \textbf{if} $R_i \neq R_k$ \textbf{then}
4: 
5: $P_{ik}^{-} = \text{calculate } P_{ik}^{-}$
6: 
7: 
8: \textbf{end if}
9: 
10: \textbf{end for}
11: for $s = 1$ to $n$ do
12: 
13: \textbf{for} $t = s + 1$ to $n$ do
14: 
15: \textbf{if} $P_{st}^{-} \notin P_{lower}$ \textbf{then}
16: 
17: $P_{st}^{-} = \text{max of } P_{lower}$
18: 
19: \textbf{end if}
20: 
21: \textbf{if} $P_{st}^{+} \notin P_{higher}$ \textbf{then}
22: 
23: $P_{st}^{+} = \text{min of } P_{higher}$
24: 
25: \textbf{end if}
26: 
27: \textbf{end for}
28: \textbf{end for}
Calculate $P_{ik}^-$ by solving the following two equations:

$$P_{ik}^- = \bar{L}(d0) + 10\nu \log(d_{IT}/d0) + rssi - z\sigma$$ \hspace{1cm} (4.5)

$$P_{ki}^- = \bar{L}(d0) + 10\nu \log(d_{KT}/d0) + rss_k - z\sigma$$ \hspace{1cm} (4.6)

Where $P_k^- = P_i^-$ (because it is the same power used by the transmitter to reach $R_k$ and $R_i$)

Calculate $P_{ik}^+$ by solving the following two equations:

$$P_{ik}^+ = \bar{L}(d0) + 10\nu \log(d_{IT}/d0) + rssi + z\sigma$$ \hspace{1cm} (4.7)

$$P_{ki}^+ = \bar{L}(d0) + 10\nu \log(d_{KT}/d0) + rss_k + z\sigma$$ \hspace{1cm} (4.8)

Where $P_k^+ = P_i^+$ (because it is the same power used by the transmitter to reach $R_k$ and $R_i$)
Figure 4.2: Transmitter $T$ is between receiver $R_k$ ($MRR$) and receiver $R_i$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Chapter 5

Evaluation of HPB with Learning Ability Scheme

We conducted an experiment in an open environment to evaluate our proposed scheme of HPB-LA. The experiment setup is described in Section 5.1. The analysis of the data collected from the experiment is provided in Section 5.2: The analysis of the HPB-LA simulation is given in Section 5.2.1, while the analysis of the Receiver Pair-Based EIRP Estimation in HPB and HPB-LA using the experimental data is given in Section 5.2.2 and Section 5.2.3.

To validate further our heuristic of calculating the EIRP range of the transmitter which has been described in Chapter 4, we have augmented our experiment results with a vehicular scenario simulation. The description and analysis of this simulation are given in Section 5.3.
5.1 The Experiment Setup

The experiment was conducted in a parking lot at Carleton University. Four computers, each one equipped with a receiver, were placed as shown in Figure 5.1. These computers were used to collect the RSS from the transmitter messages from different transmitter locations. The four receivers were labelled $R_1$, $R_2$, $R_3$, and $R_4$. Three of these receivers were powered from cars while the fourth one was powered by an electrical generator. The distance between these receivers are shown in Figure 5.1. The transmitter was a laptop computer carried on hand.

5.1.1 Hardware

Four desktop computers were used as receivers. A TRENDnet Wireless PCI adapter, as shown in Figure 5.2, was used in each of these computers. These adapters are based on Atheros 5212 chipset. We used MADWiFi [16], an open source wireless adapter driver for Linux, to set up the wireless adapters in monitor mode. To extract the RSS value from each message, we configured the receivers to capture the radiotap header instead of the prism2 header.

The transmitter was a laptop. The wireless adapter that was used by the laptop was HWUG1 Wireless-G USB adapter from HAWKING (Figure 5.3) with removable antenna. The laptop wireless adapter was set up in ad hoc mode. The antennas that were used by the four receivers and the laptop were D-LINK 7 dBi Omni-Directional antennas. During the transmission, the antenna of the laptop was placed on the top
Figure 5.1: The experiment area setup.
of a tripod. This was done to avoid the obstruction of the transmission by the laptop operator as shown in Figure 5.4.

5.1.2 Software

The receivers and the laptop computer ran Fedora 9. We have chosen Fedora 9 as an operating system because the MADWiFi open source driver was working well with Fedora 9 and satisfies our needs. The network protocol analyzer we used for capturing and dumping the raw frames by the receivers was Wireshark [42]. Wireshark is an open source analyzer and supports multi-platform. The filtering feature in Wireshark was very helpful to us, as it allowed us to dump only the packets that were sent by the laptop.

The utility program that used to send packets from the laptop was written in C language. It used UDP to send a thousand packets for each execution.

5.1.3 Data Collection

The data collection process was started by aggregating the dump files from the four receivers. To prepare the data for analysis, a Java program was used to extract the relative information.
Figure 5.2: The TRENDnet Wireless PCI adapter used as a receiver device in all the experiment.
Figure 5.3: The HWUG1 Wireless-G USB adapter used as a transmitter device in all the experiment.
Figure 5.4: Eldai El Sayr placing the laptop antenna on the top of a tripod while he is preparing to send some packets.
5.2 Analysis of the Experimental Results

Using the data collected from the experiment and linear regression techniques, we measured the path loss exponent $\nu$ and standard deviation $\sigma$. The value of $\nu$ was 1.89 while the value of $\sigma$ is 4.57. The confidence level used for the analysis in this section is 95 percent.

5.2.1 HPB-LA Simulation

Using the first half of the data collected by the four receivers, we train the HPB-LA scheme. Then, using the second half of the data, we evaluate the accuracy of HPB-LA.

The Training of HPB-LA

In the learning phase, first we identified all the possible receiver pairs, examined each one of these receiver pairs individually and recorded their accuracy. Next, we ranked the receiver pairs in two different ways: (a) according to their success rate (RPSR), and (b) according to their candidate area size (RPCAS). After that, depending on the ranking parameter, we formed a number of subsets of the receiver pairs. Finally, these subsets were examined individually. Then, their accuracy is calculated and compared to each other to identify the subset that has the highest RPS, i.e. the subset that maximizes the success rate without a high increment in the candidate area size (RPCAS). We summarize the training process as the following four steps, and at the same time we give the analysis results.
Table 5.1: Receiver pairs and their symbols.

<table>
<thead>
<tr>
<th>Receiver Pair</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>((R_1, R_2))</td>
<td>(P_1)</td>
</tr>
<tr>
<td>((R_1, R_3))</td>
<td>(P_2)</td>
</tr>
<tr>
<td>((R_1, R_4))</td>
<td>(P_3)</td>
</tr>
<tr>
<td>((R_2, R_3))</td>
<td>(P_4)</td>
</tr>
<tr>
<td>((R_2, R_4))</td>
<td>(P_5)</td>
</tr>
<tr>
<td>((R_3, R_4))</td>
<td>(P_6)</td>
</tr>
</tbody>
</table>

Step 1:
We have four receivers which are labelled \(R_1\), \(R_2\), \(R_3\), and \(R_4\). These four receivers give us a total of six receiver pairs: \((R_1, R_2)\), \((R_1, R_3)\), \((R_1, R_4)\), \((R_2, R_3)\), \((R_2, R_4)\), and \((R_3, R_4)\), which are denoted as \(P_1\), \(P_2\), \(P_3\), \(P_4\), \(P_5\), and \(P_6\) in Table 5.1. The results obtained with these receiver pairs are shown in Figure 5.5.

Step 2a: The receiver pairs of Step 1 are ranked according to their success rate, as shown in Figure 5.6.

Step 2b: The receiver pairs of Step 1 are ranked according to their candidate area size, as shown in Figure 5.7.

Step 3: From the ranking of the receiver pairs in Step 2a and Step 2b the subsets \(S_1\), \(S_2\), \(S_3\), \(S_4\), \(S_5\), \(S_6\), \(S_7\), and \(S_8\) of receiver pairs are formed. Then these subsets are simulated individually and their accuracy is calculated (Figures 5.8 and 5.9). The subsets \(S_1\), \(S_2\), \(S_3\), and \(S_4\) (Table 5.2) give the highest success rates, while the subsets \(S_5\), \(S_6\), \(S_7\), \(S_8\) (Table 5.3) give the smallest candidate areas.

Receiver pairs with the highest success rates are grouped together to give a higher
Figure 5.5: The success rate and candidate area size of receiver pairs.
Figure 5.6: Receiver pairs ordered according to their success rate. The receiver pairs on the left have a higher success rates than the receiver pairs on the right. For example, receiver pair \( P_1 \) has a success rate of 100\%, while receiver pair \( P_6 \) has a success rate of 91\%.
Figure 5.7: Receiver pairs ordered according to their candidate area size. The receiver pairs on the left have a smaller candidate area size than the receiver pairs on the right of the chart. For example, receiver pair $P_2$ has a candidate area size of 74%, while receiver pair $P_1$ has a candidate area size of 90%.
<table>
<thead>
<tr>
<th>Subset</th>
<th>Receiver Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>((R_1, R_2), (R_1, R_3), (R_1, R_4), (R_2, R_4), (R_2, R_3))</td>
</tr>
<tr>
<td>S2</td>
<td>((R_1, R_2), (R_1, R_3), (R_1, R_4), (R_2, R_4))</td>
</tr>
<tr>
<td>S3</td>
<td>((R_1, R_3), (R_1, R_3), (R_1, R_4))</td>
</tr>
<tr>
<td>S4</td>
<td>((R_1, R_2), (R_1, R_3))</td>
</tr>
</tbody>
</table>

Table 5.2: Receiver pairs subsets for success rate optimization.

<table>
<thead>
<tr>
<th>Subset</th>
<th>Receiver Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>S5</td>
<td>((R_1, R_3), (R_2, R_4), (R_2, R_3), (R_3, R_4), (R_1, R_4))</td>
</tr>
<tr>
<td>S6</td>
<td>((R_1, R_3), (R_2, R_4), (R_2, R_3), (R_3, R_4))</td>
</tr>
<tr>
<td>S7</td>
<td>((R_1, R_3), (R_2, R_4), (R_2, R_3))</td>
</tr>
<tr>
<td>S8</td>
<td>((R_1, R_3), (R_2, R_4))</td>
</tr>
</tbody>
</table>

Table 5.3: Receiver pair subsets for candidate area size optimization.

success rate subset. For example, receiver pairs \((R_1, R_2), (R_1, R_3)\) have the highest success rates among all, so when these two receiver pairs are grouped together, they are likely to give the highest success rate among all of the receiver pairs subsets. This high probability is explained in Chapter 4.

**Step 4:** The RPS for the subsets formed in Step 3 are calculated (Figure 5.10). Then the subsets are compared to each other according to their RPS. The subset with the highest RPS is selected to build the hyperbolas in HPB-LA.

Based on the RPS results, as defined in Chapter 4, the subset with the highest RPS is selected to be the subset of receiver pairs that will form the hyperbolas in HPB-LA. Here, the subset \(S1\) which contains the receiver pairs: \((R_1, R_2), (R_1, R_3), (R_1, R_4), (R_2, R_4), (R_2, R_3)\) is the subset that is used to construct the hyperbolas of the
Figure 5.8: Receiver pair subsets for success rate optimization. Subsets with fewer receiver pairs that have high success rates give the highest success rate. For example, S4 which consists of the two receiver pairs that have the highest success rate produces the highest success rate among all the other subsets.
Figure 5.9: Receiver pair subsets for candidate area optimization. Subsets with large number of receiver pairs are more likely to give a candidate area with a size less than the size of a candidate area that is given by a subset that has fewer receiver pairs, as explained in Chapter 4. We notice that, the candidate area size increases when we move from the left to the right of the chart, at the same time the number of receiver pairs in the subsets decreases from left to right.
Figure 5.10: Receiver Pair subsets PRS. Subset $S1$ has the highest RPS (5.53), while $S5$ has the lowest RPS (zero).
HPB-LA during the operational phase.

Comparison

Using the second half of the data, we measure the accuracy of HPB-LA and accuracy of HPB and then compare the two schemes. Figure 5.11 shows the success rate and candidate area size for both the HPB-LA and HPB schemes. It is clear from Figure 5.11 that HPB-LA gives a substantial relative increase in the success rate (8%) with a relative increase in the candidate area size of only 4%.
5.2.2 Evaluation of the Receiver Pair-Based EIRP Range Estimation in HPB

To evaluate the heuristic we introduced in Chapter 4, we carried out the following steps:

(a) We used all the experimental data to compute the success rate and candidate area size of HPB when the original transmitter EIRP range estimation heuristic was used.

(b) With the same experimental data as in (a), we computed the success rate and candidate area size of HPB when the Receiver Pair-Based EIRP Range Estimation heuristic was used.

(c) The results from (a) and (b) were compared (Figure 5.12).

When the transmitter EIRP range was estimated by the Receiver Pair-Based EIRP Range Estimation heuristic in HPB, there was a large relative decrease in the candidate area size (41%) with a relative decrease in the success rate of only 16%.
Figure 5.12: Comparison between the performance of HPB when the original transmitter EIRP range estimation heuristic was used (HPB-A) and performance of HPB when the Receiver Pair-Based EIRP Range Estimation heuristic was used (HPB-B).
5.2.3 The Analysis of the Receiver Pair-Based EIRP Range Estimation in HPB-LA

To evaluate the effect of Receiver Pair-Based EIRP Range Estimation process on the performance of HPB-LA, we followed the HPB-LA training and comparison steps, as they have been outlined in Section 5.2.1.

The Training of HPB-LA

Step 1: We have four receivers which are labelled $R_1$, $R_2$, $R_3$, and $R_4$. These four receivers give a total of six receiver pairs: $(R_1, R_2)$, $(R_1, R_3)$, $(R_1, R_4)$, $(R_2, R_3)$, $(R_2, R_4)$, and $(R_3, R_4)$, which are denoted as $P_1$, $P_2$, $P_3$, $P_4$, $P_5$, and $P_6$, as shown in Table 5.4. The evaluation results of these receiver pairs are shown in Figure 5.13.

Step 2a: The receiver pairs in Step 1 are ranked according to their success rate, as shown in Figure 5.14.

Step 2b: The receiver pairs in Step 1 are ranked according to their candidate area size, as shown in Figure 5.15.

<table>
<thead>
<tr>
<th>Receiver Pair</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(R_1, R_2)$</td>
<td>$P_1$</td>
</tr>
<tr>
<td>$(R_1, R_3)$</td>
<td>$P_2$</td>
</tr>
<tr>
<td>$(R_1, R_4)$</td>
<td>$P_3$</td>
</tr>
<tr>
<td>$(R_2, R_3)$</td>
<td>$P_4$</td>
</tr>
<tr>
<td>$(R_2, R_4)$</td>
<td>$P_5$</td>
</tr>
<tr>
<td>$(R_3, R_4)$</td>
<td>$P_6$</td>
</tr>
</tbody>
</table>

Table 5.4: Receiver pairs and their symbols.
Figure 5.13: **Success rate and the candidate area size of receiver pairs.** The success rate and candidate area size of each pair is shown.
Figure 5.14: Receiver pairs ordered according to their success rate.
Figure 5.15: Receiver pairs ordered according to their candidate area size.

**Step 3:** From the ranking of the receiver pairs in Step 2a and Step 2b the subsets $G_1, G_2, G_3, G_4, G_5, G_6, G_7,$ and $G_8$ of receiver pairs are formed. Then these subsets are evaluated individually and their accuracy is recorded (Figures 5.16 and 5.17). The subsets $G_1, G_2,$ and $G_3$ give the highest success rates, while the subsets $G_5, G_6, G_7,$ and $G_8$ (Table 5.5) give the smallest candidate areas.

**Step 4:** The RPS for the subsets formed in Step 3 are calculated (Figure 5.18). Then the subsets are compared to each other according to their RPS. The subset with the
Figure 5.16: Receiver pair subsets for success rate optimization.
Figure 5.17: Receiver pair subsets for candidate area size optimization.
Subset | Receiver Pairs
--- | ---
G1 | \((R_1, R_2), (R_1, R_4), (R_1, R_3), (R_2, R_4), (R_3, R_4)\)
G2 | \((R_1, R_2), (R_1, R_4), (R_1, R_3), (R_2, R_4)\)
G3 | \((R_1, R_3), (R_1, R_4), (R_1, R_3)\)
G4 | \((R_1, R_2), (R_1, R_4)\)

Table 5.5: Receiver pairs subsets for success rate optimization.

Subset | Receiver Pairs
--- | ---
G5 | \((R_1, R_3), (R_2, R_4), (R_2, R_3), (R_3, R_4), (R_1, R_4)\)
G6 | \((R_1, R_3), (R_2, R_4), (R_2, R_3), (R_3, R_4)\)
G7 | \((R_1, R_3), (R_2, R_4), (R_2, R_3)\)
G8 | \((R_1, R_3), (R_2, R_4)\)

Table 5.6: Receiver pairs subsets for candidate area size optimization.

The highest RPS is selected to build the hyperbolas in HPB-LA. Here, the subset G5 which contains the receiver pairs: \((R_1, R_2), (R_1, R_3), (R_1, R_4), (R_2, R_4), (R_3, R_4)\) is the subset that will be used to construct the hyperbolas of the HPB-LA.

Comparison

Using the second half of the data we measured the accuracy of the HPB-LA and accuracy of the HPB and then compared the two accuracies. Figure 5.19 shows the success rate and candidate area size for both HPB-LA and HPB schemes when both of them used the Receiver Pair-Based EIRP Range Estimation heuristic.

We find that HPB-LA has achieved a relative increase in the success rate of 9% with
Figure 5.18: Receiver pair subsets PRS. Subset G5 has the highest RPS (3.26), while G8 has the lowest RPS (0.52).
Figure 5.19: A Comparison between HPB and HPB-LA when the Receiver Pair-Based EIRP Range Estimation heuristic is used.
Figure 5.20: A Comparison between HPB and HPB-LA when two different EIRP range estimation heuristics are used. HPB(A) is HPB with the HPB original transmitter EIRP range estimation heuristic, HPB-LA(1) is HPB-LA with the HPB original transmitter EIRP range estimation heuristic, and HPB-LA(2) is HPB-LA with the Receiver Pair-Based EIRP Range Estimation heuristic.
a relative increase in the candidate area size of 8% when both HPB and HPB-LA used the Receiver Pair-Based EIRP Range Estimation heuristic. However, if we compare the HPB with the original transmitter EIRP range estimation heuristic to HPB-LA with the Receiver Pair-Based EIRP Range Estimation heuristic we find that HPB-LA has achieved a relative decrease of 34% (from 50% to 33%) in the candidate area size, while the success rate decreased relatively by only 7% (from 92% to 86%), as shown by Figure 5.20.

Also if we compare the HPB-LA with the original transmitter EIRP range estimation heuristic to HPB-LA with the Receiver Pair-Based EIRP Range Estimation heuristic we find that HPB-LA has achieved a relative decrease of 37% (from 52% to 33%) in the candidate area size, while the success rate relatively decreased by only 13% (from 99% to 86%), as shown by Figure 5.20. Therefore, we believe that our proposed transmitter EIRP range estimation process can extend HPB-LA to deliver a wide range of accuracies, from a success rate of 99% and a candidate area size of 52%, to a success rate of 86% and a candidate area size of 33%.
5.3 The Simulation

The analysis results from Sections 5.2.2 and 5.2.3 show that Receiver Pair-Based EIRP Range Estimation used in HPB gives a smaller size candidate area. To validate further this idea, we simulated a vehicular network. The parameters used in this simulation were adopted from the large scale parameter values given by L. C. Liechty [20], [21]. The value of path loss exponent and signal shadowing standard deviation used were 2.76 and 5.62 respectively. The confidence level used in the simulation is 0.95. For each receiver, the RSS values were generated from the log-normal shadowing model. We augmented the simulated RSS values independently by a randomly generated shadowing. We executed HPB with the original EIRP range estimation heuristic and we also executed HPB with our proposed heuristic (described in Chapter 4).

Figure 5.21 shows a snap shot from the simulation. The red rectangles represent the Road-Side Units (RSUs). These are road-side infrastructure units that are equipped with onboard processing and wireless communication modules. The black rectangles represent the ON-Board Units (OBUs) which are vehicles that are equipped with onboard processing and wireless communication modules. The red hollow ellipses represent the transmitter previous positions while the solid ellipse represents the transmitter current position. The shaded area represents the candidate area.

Analysis

Our simulation results, presented in Figure 5.22, show that, for any number of receivers the Receiver Pair-Based EIRP Range heuristic always gives smaller candidate
Figure 5.21: A snapshot from the simulation.
Figure 5.22: Simulation results. HA-SR and HA-CAS are the success rate and candidate area size produced when HPB used its original transmitter EIRP range estimation heuristic, while HB-SR and HB-CAS are the success rate and candidate area size produced when HPB used the Receiver Pair-Based EIRP Range Estimation heuristic.

area than the heuristic used in the original HPB. This confirms our earlier results from the experimental analysis.
Chapter 6

Conclusions and Future Work

In this chapter, we present the conclusions and project the future work. We demonstrated using real data that HPB-LA can be employed to improve the success rate of HPB in localizing a rogue node. Also, we introduced a new heuristic for calculating an estimate of the EIRP range of a transmitter. This heuristic is a receiver pair based estimation process. The results show that this heuristic was able to reduce the size of the candidate area produced by both HPB and HPB-LA schemes considerably.

Although the localization performance of HPB-LA is constrained by a fundamental limitation caused by the nature of radio waves, enhancements that could improve the performance of HPB-LA are feasible. At the same time, further research is required in order to: test the performance of the localization model in real-time systems, obtain more accurate transmitter EIRP range estimation, and handle the assumption that the transmitter might employ directional antennas in broadcasting its messages.
6.1 Success Rate Optimization

HPB-LA, as a geometric RSS-based model, was proposed and tested. The model represents an enhancement to the HPB. To demonstrate the degree of improvement that HPB-LA has provided, the two models were evaluated in terms of success rate and candidate area size. The results showed that HPB-LA was very successful in optimizing the success rate of localizing a transmitter. The minimum relative increase in success rate that HPB-LA gave was 8%. In the analysis of HPB-LA with the experimental data and a confidence level of 95%, HPB gave a success rate of 92%. However, HPB-LA managed to increase this success rate to 99% with a relative increase in the candidate area size of 4% only (from 50% to 52%).

6.2 Candidate Area Size Optimization

When the original HPB transmitter EIRP range estimation heuristic, described in Section 3.1, was used by HPB-LA, the analysis results showed that HPB-LA managed to give a higher success rate than the one given by HPB. But HPB-LA could not decrease the size of the candidate area. In fact the size of the candidate area had gone up by a small fraction. When our proposed transmitter EIRP range estimation, described in Chapter 4, was used by HPB, the analysis result showed that the size of the candidate area has a relative decrease of 41% with a relative decrease in the success rate of 16%. Also, if we compare the results of Section 5.2.1 and of Section 5.2.3, we find that HPB-LA with our proposed transmitter EIRP range estimation had managed
to achieve a relative decrease in the candidate area size of 34% (from 50% to 33%) with only a 7% relative decrease in success rate (from 92% to 86%). Therefore, we believe that our proposed transmitter EIRP range estimation process can extend HPB-LA to deliver a wide range of accuracies, from a success rate of 99% and a candidate area size of 52%, to a success rate of 86% and a candidate area size of 33% as was shown in Section 5.2.3.

6.3 Transmitter EIRP Range Estimation

As the results have shown, the transmitter EIRP range estimation has a great effect in HPB and HPB-LA. Our proposed transmitter EIRP range estimation process has considerably boosted the performance of both HPB and HPB-LA in terms of the candidate area size, as it has been demonstrated in Section 6.2 above. The analysis results from Section 5.2.2 demonstrated that when our proposed transmitter EIRP range heuristic was used by HPB, the candidate area had a relative decrease of 41% with a relative decrease of 16% in the success rate. Moreover, our simulation results in Section 5.3 showed that for any number of receivers in a network, our proposed transmitter EIRP range estimation process always gives a smaller candidate area than the transmitter EIRP range estimation process that was first used by HPB.
6.4 Future Work

Future experiments are needed for testing the performance of HPB-LA in real-time systems. As our analysis results from Section 5.2 showed, the transmitter EIRP range estimation was a big factor in the performance of HPB-LA. Therefore, future transmitter EIRP range estimation algorithms are needed for the improvement of the HPB-LA accuracy. Another area that needs to be investigated in the future is the integration of the trusted nodes authentication process in the design of HPB-LA and use of the schemes that detect the presence of liars such as the one proposed by Ke Li [6]. Also studies should be initiated to evaluate the HPB-LA performance with the inclusion of the assumption that the transmitter might employ directional antennas.
Bibliography


[40] http://www.wimaxforum.org/
