LOCATION TRACKING MITIGATION FOR HONEST NODES
AND LOCATION ESTIMATION OF UNCOOPERATIVE DEVICES
IN WIRELESS MOBILE NETWORKS

by

Christine Laurendeau

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

DOCTOR OF PHILOSOPHY

School of Computer Science

at

CARLETON UNIVERSITY

Ottawa, Ontario
June, 2009

© Copyright by Christine Laurendeau, 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.
In loving memory of my grandmother

Blanche Dubé
Abstract

Wireless mobile networks are on the cusp of actualizing revolutionary advances in many areas, including public safety. Vehicular communications, for example, can convey real-time, life-preserving notifications of impending collisions and other hazards, and thus reduce the daily carnage on our roads. To bolster public faith in mobile networks and ensure that these don’t fall prey to unscrupulous attackers, their security vulnerabilities must be exposed in a methodical fashion and suitably mitigated. Yet many novel wireless technologies face a difficult quandary in their need to meet two seemingly orthogonal requirements: network security and user privacy. Security measures must guarantee that the originators of broadcast messages are authorized, and that their transmissions have integrity and are non-repudiable. Concurrently, honest cooperative mobile devices must be assured of a reasonable degree of privacy without the threat of location tracking over time and space.

We aim to harmonize wireless network security requirements with the preservation of location privacy by addressing three interdependent issues: the identification and assessment of serious risks to emerging wireless technologies; the mitigation of the location tracking threat in vehicular networks; and the localization of uncooperative wireless devices for the purpose of attack attribution or emergency assistance.

Our contributions in reconciling security requisites with location privacy in wireless mobile networks are thus three-fold:

1. We conduct a threat analysis into the risks to vehicular networks and compare the results with known threats to wireless access technology in order to assess the security of mobile networks in general. We identify location tracking and insider attacks as critical threats.

2. We examine the risk of location tracking in vehicular networks, and propose a scheme to secure vulnerable broadcast messages while preserving driver anonymity. We find that our approach achieves this goal, with smaller message sizes and more computationally efficient cryptographic operations.

3. (i) We explore the threat posed by uncooperative mobile network devices,
including malicious insiders. We put forth a probabilistic mechanism for the attribution of a given message to its originator by estimating the source location of its signal. Our Hyperbolic Position Bounding (HPB) algorithm combines the relative received signal strength between pairs of trusted receivers to construct a candidate area in Euclidian space. We localize an uncooperative transmitter within an area approximately $250 \times 250 \text{ m}^2$ in size, with a confidence level of 95%.

(ii) We propose a novel heuristic scheme to aggregate candidate area probability so that the combined degree of confidence is computed according to a paradigm of supportive rather than competitive evidence. Our aggregation model yields a probability distribution better fitted to HPB simulation results than a redistributive paradigm.

(iii) We describe an outdoor experiment with a WiFi/802.11 network to harvest measured Received Signal Strength (RSS) values. These are subsequently used by HPB to bound the location of a mobile transmitter with a degree of confidence. Experimental results are compared against prior simulation results and found to be consistent.

(iv) We describe alternate HPB algorithms to improve localization precision and computational efficiency, and provide a mobile attacker tracking capability in vehicular networks. We find that HPB can position a mobile transmitting device within stipulated guidelines for emergency services localization accuracy.

(v) We propose two new location estimation algorithms that extend the concept of centroid localization. Our mechanisms weigh each receiver’s coordinates based on a message’s relative RSS at that receiver, with respect to the span of RSS values over all receivers. Our simulation results demonstrate that our exponentially weighted variation meets the emergency services location accuracy regulations.
Acknowledgements

Doctoral studies are often likened to a journey of discovery, a period of progressive growth culminating in a state of relative academic enlightenment. My experience was also akin to a rebirth, a gentle blossoming from who I was to who I was meant to be, an irreversible passage between two worlds and two lifetimes. It was my good fortune to be sustained and abetted by those closest to me.

Principal thanks are extended to my immediate family: to Paul, Claire and Colin, for graciously enduring the long work days and study weekends, the endless chores, and the mountains of laundry; to my self-billed #1 cheerleader, my mother Jacqueline, and my irrepressible stepfather Donald, for their love and encouragement; and most importantly, to my sister Sonja, my constant, for always being there when I needed her, for her wisdom and her fearlessness, and for helping me to pick up the pieces.

None of the work contained herein would have been possible, but for the enthusiastic support and endless good humour of my faculty advisor Dr. Michel Barbeau. He has not only taught me the technical concepts fundamental to our field of research, but also life-long lessons on the nature and scope of elemental human qualities: generosity, kindness and loyalty; patience, perseverance and courage. With his unwavering faith in my abilities, he has instilled in me a firm and lasting belief in my own potential, and opened up countless new avenues for growth that I had hitherto deemed out of reach. For this gift that will benefit me for the length and breadth of all my future endeavors, I am forever grateful.

Sincere gratitude is also due to the professors and students whose contributions, both tangible and whimsical, fortified and encouraged me along the way: to Dr. Evangelos Kranakis, whose consistently constructive feedback and novel theoretical viewpoints were invaluable to my progress and formation as a researcher; to Dr. Pros- enjit Bose, for the laughs and Belgian chocolates; to my fellow doctoral students and lab-mates Paul Boone and Michel Paquette, for alternately cheering and commiserating on our respective fates.

Warmest thanks go to the School of Computer Science staff for their tireless help and cheerful demeanors: to Joni Campbell, whose comforting hugs were always
sincerely needed; to Claire Ryan, whose sense of class was a model to us all; and to fellow working mom Sharmi, who always understood.

Finally, thankful appreciation is due to the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Automobile of the 21st Century (AUTO-21) and Mathematics of Information Technology and Complex Systems (MITACS) Networks of Centers of Excellence (NCEs) for their financial contributions.
Table of Contents

Abstract iii

Acknowledgements v

List of Tables xi

List of Figures xii

List of Acronyms xv

I Preface 1

Chapter 1 Introduction 2

1.1 Motivation 2

1.2 Problems of Interest 4

1.2.1 Threats to Wireless Mobile Networks 4

1.2.2 Secure Anonymity in Vehicular Networks 5

1.2.3 Localization of Uncooperative Nodes 6

1.3 Summary of Contributions 7

1.4 Organization of Thesis 10

Chapter 2 Background 12

2.1 Threat Analysis Methodology 12

2.1.1 Types of Methodologies 12

2.1.2 Risk Assessment 13

2.2 Vehicular Network Model 17

2.2.1 DSRC/WAVE Architecture 17

2.2.2 WAVE Security 18

2.2.3 Vehicle Safety Application Requirements 19
List of Tables

Table 2.1 ETSI Risk Assessment ........................................... 14
Table 2.2 Adapted Risk Assessment ....................................... 16
Table 3.1 Localization Scheme Assumptions .......................... 31
Table 4.1 DSRC/WAVE Threat Analysis ................................. 44
Table 5.1 SAB vs. WAVE Beacon Message Size (in bytes) ........ 58
Table 5.2 SAB vs. WAVE Execution Time (in ms) .................... 59
Table 9.1 Average HPB Variations Execution Time (seconds) ....... 131
Table 10.1 Average Number of Receivers per Node Density ........ 148
Table 10.2 Average Location Error for Experimental Transmitter Locations 156
List of Figures

Figure 6.1 Example of Log-Normal Shadowing Model .............. 67
Figure 6.2 Example of TDOA ........................................ 70
Figure 6.3 Example of Minimum Bounds of Distance Difference Ranges . 72
Figure 6.4 Example of Maximum Bounds of Distance Difference Ranges . 72
Figure 6.5 Example of Distance Difference Areas .................. 72
Figure 6.6 Example of Minimum Hyperbolas for $R_1, R_2$ ............ 76
Figure 6.7 Example of Minimum Hyperbolas for $R_1, R_2$ and $R_3, R_4$ . 77
Figure 6.8 Example of Hyperbolic Areas for $R_1, R_2$ and $R_3, R_4$ .... 78
Figure 6.9 HPB Simulation Grid ...................................... 79
Figure 6.10 HPB Success Rate for $C = 0.95$ ...................... 81
Figure 6.11 HPB Success Rate ......................................... 81
Figure 6.12 HPB Success Rate for Aggregate Range Points ........... 82
Figure 6.13 HPB Candidate Area Size ............................... 83
Figure 6.14 HPB Candidate Area Size per Success Rate ............. 83
Figure 7.1 Example of Hyperbolic Area Partitions .................... 87
Figure 7.2 Compound Probability for Hyperbolic Area Partitions ..... 93
Figure 7.3 CP Simulation Grid Ranges ............................... 96
Figure 7.4 Hyperbolic Area Probability Distributions for $C = 0.95$ .... 97
Figure 7.5 Hyperbolic Area Probability Distributions for $C = 0.90$ .... 97
Figure 7.6 Kolmogorov-Smirnov Statistic for Distribution Differences . 98
Figure 7.7 Cumulative Probability Distributions for $C = 0.95$ ....... 100
Figure 7.8 Cumulative Probability Distributions for $C = 0.90$ ....... 100
Figure 7.9 Compound and Multiplicative Cumulative Probability ....... 101
Figure 7.10 Central Range Cumulative Probability .................. 101
Figure 7.11 Paired Hyperbolic Area Probability Distributions for $C = 0.95$ 102
Figure 7.12 Paired Hyperbolic Area Probability Distributions for $C = 0.90$ 102
Figure 8.1 Outdoor Experiment at Carleton University ............. 103
Figure 10.8 Algorithm Location Error CDF for Four Receivers (2.4 GHz) . 154
Figure 10.9 Algorithm Location Error CDF for Four Receivers (5.8 GHz) . 154
Figure 10.10REWL Location Error CDF by Node Density (2.4 GHz) . . . 155
Figure 10.11REWL Location Error CDF by Node Density (5.8 GHz) . . . 155
Figure 10.12Algorithm Location Error for Experimental Data .......... 156

Figure 11.1 Example of Directional Antenna Radiation Pattern .......... 164
List of Acronyms

**AES**  Advanced Encryption Standard

**AES-CCM**  AES in Counter with CBC-MAC (CCM) Mode

**AK**  Authorization Key

**AR**  Aggregate Range

**BS**  Base Station

**CA**  Certificate Authority

**CAMP**  Crash Avoidance Metrics Partnership

**CBC-MAC**  Cipher Block Chaining Message Authentication Code

**CDF**  Cumulative Distribution Function

**CL**  Centroid Localization

**CP**  Compound Probability

**CPU**  Central Processing Unit

**CR**  Central Range

**CRL**  Certificate Revocation List

**CRTC**  Canadian Radio-television Telecommunications Commission

**DES**  Data Encryption Standard

**DoS**  Denial of Service

**DSRC**  Dedicated Short Range Communications

**E911**  Enhanced 911
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECDSA</td>
<td>Elliptic Curve Digital Signature Algorithm</td>
</tr>
<tr>
<td>ECIES</td>
<td>Elliptic Curve Integrated Encryption Scheme</td>
</tr>
<tr>
<td>EIRP</td>
<td>Effective Isotropic Radiated Power</td>
</tr>
<tr>
<td>EMA</td>
<td>Exponentially Weighted Moving Average</td>
</tr>
<tr>
<td>ETSI</td>
<td>European Telecommunications Standards Institute</td>
</tr>
<tr>
<td>FCC</td>
<td>Federal Communications Commission</td>
</tr>
<tr>
<td>GA</td>
<td>Grid Candidate Area</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>HMAC</td>
<td>key hashed Message Authentication Code</td>
</tr>
<tr>
<td>HPB</td>
<td>Hyperbolic Position Bounding</td>
</tr>
<tr>
<td>IDS</td>
<td>Intrusion Detection System</td>
</tr>
<tr>
<td>IP</td>
<td>Internet Protocol</td>
</tr>
<tr>
<td>IPv6</td>
<td>Internet Protocol, version 6</td>
</tr>
<tr>
<td>IT</td>
<td>Information Technology</td>
</tr>
<tr>
<td>LBS</td>
<td>Location-Based Service</td>
</tr>
<tr>
<td>LOS</td>
<td>Line of Sight</td>
</tr>
<tr>
<td>MAC</td>
<td>Medium Access Control</td>
</tr>
<tr>
<td>MP</td>
<td>Multiplicative Probability</td>
</tr>
<tr>
<td>MS</td>
<td>Mobile Station</td>
</tr>
<tr>
<td>NIST</td>
<td>National Institute of Standards and Technology</td>
</tr>
</tbody>
</table>
OBU  On-Board Unit
PDF  Probability Distribution Function
PKI  Public Key Infrastructure
PPS  Precise Positioning System
PSOBU  Public Safety On-Board Unit
REWPL  Relative Span Exponential Weighted Localization
RF  Radio Frequency
RFID  Radio Frequency Identification
RSS  Received Signal Strength
RSU  Road-Side Unit
RWL  Relative Span Weighted Localization
SAASM  Selective Availability Anti-Spoofing Module
SAB  Secure Anonymous Broadcasting
SeVeCom  Secure Vehicular Communication
SHA  Secure Hash Algorithm
SPS  Standard Positioning System
TDOA  Time Difference of Arrival
TIPHON  Telecommunications and Internet Protocol Harmonization Over Networks
TOA  Time of Arrival
TPM  Trusted Platform Module
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-R</td>
<td>Transmitter-Receiver</td>
</tr>
<tr>
<td>USD</td>
<td>U.S. dollars</td>
</tr>
<tr>
<td>V2I</td>
<td>Vehicle-to-Infrastructure</td>
</tr>
<tr>
<td>V2V</td>
<td>Vehicle-to-Vehicle</td>
</tr>
<tr>
<td>VA</td>
<td>Vehicular Candidate Area</td>
</tr>
<tr>
<td>VANET</td>
<td>Vehicular Ad-hoc Network</td>
</tr>
<tr>
<td>VSCC</td>
<td>Vehicle Safety Communications Consortium</td>
</tr>
<tr>
<td>WAVE</td>
<td>Wireless Access in Vehicular Environments</td>
</tr>
<tr>
<td>WCL</td>
<td>Weighted Centroid Localization</td>
</tr>
<tr>
<td>WiMAX</td>
<td>Worldwide Interoperability for Microwave Access</td>
</tr>
<tr>
<td>WSMP</td>
<td>WAVE Short Message Protocol</td>
</tr>
</tbody>
</table>
Part I

Preface
Chapter 1

Introduction

As newer varieties of wireless technology continue to emerge and reshape our daily lives, it becomes more imperative that the architecture and infrastructure upon which they are based be reliable and secure. Yet in any wireless technology, the twin requirements of network security and user privacy constitute a fundamental clash of values. Security mechanisms are necessary to deter and detect attackers, and require the disclosure of participant identity for the purpose of authentication. Concurrently, individual users should retain a measure of control over their personal information, such as identity and location, and its dissemination. This requires the obfuscation of user information in order to achieve anonymity and thus safeguard privacy. The fine balance between the two seemingly orthogonal requirements of security and privacy is a concern in any type of technology and may shift accordingly.

We examine the provision of security mechanisms in emerging wireless mobile networks and their impact on location privacy. We separate two distinct classes of network participants: honest nodes, whose behaviour fully complies with network authority expectations; and uncooperative devices, which may comprise attackers, including malicious insiders, and malfunctioning or energy-drained nodes. We subsequently propose solutions to achieve an equilibrium between the security and location privacy requirements, through the obscurement of honest nodes' geographical coordinates from other devices, and the localization of uncooperative nodes for attack attribution or emergency assistance.

1.1 Motivation

Given the open nature of radio signals, wireless communications are particularly vulnerable to security exploits. Attacks upon wireless networks safeguarding public safety can have dire consequences, possibly resulting in loss of life. The remote
manipulation of tram system control signals by a young hacker in Poland nearly resulted in disastrous consequences for the passengers of several trains [12]. Yet in ensuring appropriate security measures, it is crucial that an adequate amount of user privacy be maintained. Assuming that we can distinguish the honest nodes from the uncooperative devices, we can strive to protect the anonymity of the former while exposing and exacting retribution on the latter.

Unproven technologies often feature security vulnerabilities hitherto unknown. Potential security exploits must be anticipated and effectively mitigated. We begin by tackling the analysis of security risks to novel wireless technologies. We conduct a methodical threat analysis into a proposed vehicular communications architecture and revisit an existing analysis into the risks to wireless access networks. We identify a number of critical risks, including a threat to the privacy of drivers in vehicular networks, in the form of location tracking. We propose an alternate security mechanism for vehicle safety broadcast messages that achieves both adequate security and driver anonymity. While honest cooperative devices must be ensured of sufficient privacy, malicious or uncooperative nodes cannot be afforded the same privilege. For example, an authenticated wireless device may abuse its valid credentials to launch disinformation attacks on unsuspecting nodes. Similarly, a malfunctioning or nearly energy-drained device may require assistance in a critical emergency. In either case, we must locate the source of a signal, assuming no cooperation from the originating device. In addition, if the signal source is a malicious node engaging in an insider attack, we may not assume knowledge of the radio equipment used. We address the attribution of an attack or distress message to its originator, through a location estimation mechanism that fulfills our minimalist assumptions.

We propose solutions for both the identity and location protection of honest nodes and the location exposure of uncooperative devices. We thus aim to provide a comprehensive approach to the secure management of device location information across wireless mobile technologies.
1.2 Problems of Interest

We focus on three problems of interest pertaining to the security of wireless mobile networks: the identification of serious threats to vehicular and wireless access networks, the coexistence of security and anonymity requirements in vehicular network broadcast messages, and the location estimation of uncooperative nodes in wireless networks.

1.2.1 Threats to Wireless Mobile Networks

The ubiquity of wireless networks in every aspect of our lives and the blind trust vested in them require the evolution of security threats to be continually monitored. In fact, a threat must first be evaluated in terms of importance, and only then does it need to be mitigated. Perspective can be achieved in a predictable, methodical manner by determining which threats are real and which are of little consequence. Since aiming at perfect security is a Utopian pursuit, and attempting to approach it can be very costly, the chimeras of perceived risk must be dispelled and the real threats unmasked so that research efforts can be effectively focused on the irrefutable menaces jeopardizing wireless networks.

The emerging field of vehicular communications is poised to become the technology of the next decade. With new vehicles being outfitted with on-board equipment which essentially renders each one into a mobile device capable of communicating with other vehicles and with fixed roadside stations, drivers are able to call upon a veritable panoply of new and exciting applications. In addition to gaining access to an unending stream of navigational and localized information available through Location-Based Services (LBSs), vehicles can benefit from a tremendous increase in traffic safety when they receive real-time notification of upcoming road hazards, imminent collisions and the movements of emergency vehicles. The Dedicated Short Range Communications (DSRC) Wireless Access in Vehicular Environments (WAVE) architecture [8] proposes an infrastructure for communications between vehicles and between vehicles and the associated infrastructure to support applications aimed at maximizing vehicle safety and driver convenience.

Emerging technologies such as DSRC/WAVE must be submitted to a rigorous
analytical process in order to identify the security vulnerabilities in their proposed architecture. The identification of important threats in DSRC/WAVE can only assist in strengthening its security measures and preventing attacks which may undermine the public’s faith in this potentially life-saving technology. While such an exercise has been conducted for wireless access networks [13] such as WiMAX/802.16 [77], the results have not been correlated with others from similar mobile technologies. Serious threats that are common to multiple wireless platforms are indicative of deeply rooted problems that must be addressed and swiftly mitigated.

1.2.2 Secure Anonymity in Vehicular Networks

Identity-based mechanisms are common for restricting network resource access to authorized parties, as well as for providing the means of detecting and revoking attackers. Public Key Infrastructure (PKI) systems stand out as a prime example, where in addition to ensuring message integrity and confidentiality, unique digital certificates identify their possessors to other parties. Balancing network security requirements with participant need for anonymity is a challenging endeavor, and vehicular networks are no exception.

In order to secure vehicular communications, the DSRC/WAVE architecture mandates the use of PKI mechanisms, where service application messages are encrypted and broadcast vehicle safety messages are digitally signed. Unfortunately in the latter case, the use of digital signatures entails the transmission of the sender’s unique certificate, thus providing the means for other devices, authorized or unauthorized, to potentially track a particular vehicle and its driver over a period of time and space. Location privacy is a highly prized commodity, and the lack of it may be exploited in unexpected fashion, as suggested by Schilit et al. [126]. Since the most common usage for broadcast messages involves vehicle safety applications, WAVE’s vulnerability to location tracking may hinder the wide-spread adoption of this innovative technology. An alternate scheme to broad-spectrum PKI is required to secure vehicle safety messages while preserving driver anonymity.
1.2.3 Localization of Uncooperative Nodes

While honest nodes should expect to retain control over the dissemination of their location information, there are cases where wireless nodes must forfeit their right to anonymity and location privacy. For example, insider attacks pose an often neglected threat scenario when devising security mechanisms. As the owner of a valid logical identity such as a Medium Access Control (MAC) address or digital certificate in an open medium, a malicious insider in a wireless network can broadcast with impunity verifiable messages containing falsified information. It may also evade retribution once an attack is detected, especially if its identity is fraudulently obtained, for example through theft, and its credentials are untraceable. Our only resort in attributing an exploit to a malicious insider may lie in locating the source of its attack message.

In other situations, for example in sensor networks, a malfunctioning device may need to be localized before retrieval. A cellphone user may require emergency assistance, yet be unaware of its position. A recent spate of such cases in Canada have highlighted the tragic consequences of failing to locate the source of an emergency 911 cellphone call. In one incident, a New Year’s Eve reveler lost in a snowstorm in the middle of the British Columbia woods called 911 for help, but the police were only able to find the teen over 12 hours later, after he had perished [41]. A more recent case had two children lost in snowy conditions who were lucky to survive when discovered several hours after their initial 911 call [123].

Such scenarios highlight the need for estimating the position of a device that is either unwilling or unable to assist in localization efforts. For example, an attacker may willfully provide erroneous information in order to foil attempts at localizing it. A cellphone with a nearly drained battery may be unable to compute and report its coordinates to other nodes. If we cannot rely on or trust the self-knowledge provided by a transmitting device, localization must be achieved using information that is obtained surreptitiously. Unlike the scenarios addressed by existing location estimation schemes, we cannot assume the cooperation of a transmitting node, nor make assumptions regarding its radio equipment.

It should be noted that any localization mechanism that can position a device without its knowledge or cooperation is susceptible of being abused by an attacker.
bent on locating an honest, unsuspecting node. So while our proposed location estimation approaches must be effective when employed by legitimate network entities, the requirements on their execution must offer a certain degree of protection for cooperative nodes, lest our solutions become victims of their own success.

In our threat model, we presume that the uncooperative transmitter to be localized is a mobile device in an outdoor environment and that a number of trusted receiving stations are within Line of Sight (LOS) radio range and can communicate with each other over a secure channel. The coordinates of the receivers are globally known. Such a scenario is feasible in a number of wireless technologies, for example with WiMAX/802.16 Mobile Stations (MSs) and Base Stations (BSs) or the On-Board Units (OBUs) and Road-Side Units (RSUs) in the DSRC/WAVE vehicular architecture. A single message may be transmitted by the target device. In order to evade retribution from network authorities, a malicious node may effect changes to its transmission power to modify the signal gain and obfuscate its true position. As a result, no assumptions can be made regarding the Effective Isotropic Radiated Power (EIRP) employed to transmit an attack message. However, an omnidirectional propagation pattern is presumed.

1.3 Summary of Contributions

Our contributions encompass three interrelated issues in promoting a balance between the security and privacy requirements of wireless mobile devices: the identification of threats to mobile technologies and their classification by degree of severity; a solution to the location tracking threat in vehicular networks, where PKI-equivalent security is maintained while driver anonymity is protected; and a mechanism to estimate the location of an uncooperative device assuming the message EIRP is unknown.

1. In order to expose the serious threats to wireless mobile networks, we assess the risks to the developing DSRC/WAVE vehicular network architecture using our adaptation of an existing threat analysis methodology. We thus provide a risk analysis of the more critical vulnerabilities inherent in vehicular networks. We also examine a previous analysis of threats to WiMAX/802.16 [13]. Since the
integration of both technologies is envisioned for multimedia applications over vehicular communications links [94], we correlate the results of both analyses and reach common conclusions for the generalized threats to mobile networks.

2. The delicate balance between security and privacy is highlighted in our threat analysis of the DSRC/WAVE architecture, where location tracking is uncovered as a critical threat. We reconcile both requirements and introduce the Secure Anonymous Broadcasting (SAB) protocol for the transmission of vehicle safety broadcast messages that are at once secure and preserve driver privacy. A hybrid key infrastructure is described, employing both symmetric and asymmetric mechanisms. Through the use of a shared network authorization key supplemented with an encrypted sender identifier, the privacy of devices requiring anonymity is protected and the means for authorities to identify a sending device in case of an attack is retained. Devices which can and should be identified, for example infrastructure units and public safety vehicles such as ambulances and police cars, retain the use of digital signatures to uphold their status as the more trusted devices. We assess SAB’s security by revisiting the previously identified threats to DSRC/WAVE and ensuring that the risk associated with location tracking has been reduced, while no other risk has increased as a result.

3. One of our principal contributions involves the development of a solution for the location estimation of a malicious, malfunctioning or otherwise uncooperative transmitter by a set of receiving devices within radio range. This contribution spans a number of related offerings, as follows:

(i) We put forth the Hyperbolic Position Bounding (HPB) mechanism for trusted receivers to probabilistically estimate the source location of a single transmitted message using the relative Received Signal Strength (RSS) measured at a number of trusted receivers. Since the EIRP of the transmitter is unknown and RSS values may fluctuate, a probabilistic model is used to estimate the range of distance differences between receiver pairs and the transmitter. These differences are used to construct minimum and maximum hyperbolas between two receivers where the transmitter may be
located, with a certain degree of confidence. We evaluate the performance of HPB using simulated RSS values. We find that since the success rate and localization granularity of the HPB approach is contingent upon the number of receiving devices, a lone attacker or a collusion of a small number of eavesdroppers meet with little success in exploiting HPB to localize an honest node.

(ii) A related issue emerges when multiple receiver pairs are considered by HPB. The confidence levels in each separate hyperbolic area must be aggregated for a higher degree of confidence associated with a smaller intersecting area as more receiver pairs are included. When combining evidence supporting a given candidate area for a transmitter, the intuitive expectation is that an area endorsed by a large number of receivers should be assigned a greater confidence than the areas advocated by few. We propose a novel heuristic scheme to compound the degree of confidence endorsed by multiple receiver pairs in such a way that additional evidence supports the probability assignment in the common hyperbolic area intersection rather than redistributing it outside the intersection.

(iii) We conduct an outdoor experiment using a WiFi/802.11 network to harvest RSS reports emitted by a mobile transmitter. We evaluate HPB with these experimental RSS values and compare the localization results against those obtained with simulated RSS reports.

(iv) Two additional variations of HPB are proposed to improve computational efficiency and localization granularity. An adaptation is also provided for tracking a mobile attacker in a vehicular network.

(v) Open-form location estimation approaches such as HPB approximate the position of a node within an area with a suitable degree of granularity for certain types of applications. However, other scenarios may require a more precise result. We thus propose an extension to an existing weighted centroid localization approach for this purpose. We outline two mechanisms to estimate the location of a transmitter by averaging the known receiver
coordinates, where each receiver's position is weighted based on a transmitted message's relative RSS at that receiver, with respect to the span of RSS values over all receivers. The weights may decrease from the maximal RSS receiver either linearly or exponentially.

In summary, our contributions can be expressed in terms of the following publications. We have three papers in refereed journals: a description of the original HPB algorithm evaluated with simulated RSS values [82], published in *Security and Communication Networks*; a mechanism to aggregate HPB probabilistic evidence in the intersection of multiple hyperbolic areas [83], published in the *Journal of Networks*; and a proposal for additional HPB algorithm variations with an extension for tracking a mobile device in a vehicular network [84], to appear in a special issue of the *EURASIP Journal on Wireless Communications and Networking* devoted to wireless network security. We also have three papers in refereed conferences and one submitted for consideration: a threat analysis for the DSRC/WAVE architecture [79], published in the proceedings of ADHOC-NOW (2006); a description of our Secure Anonymous Broadcasting protocol [80], published in the proceedings of ON-MOVE (2007); an evaluation of HPB with experimental RSS values measured in a WiFi/802.11 network [81], published in the proceedings of ON-MOVE (2008); and our weighted centroid localization scheme [85], submitted to WASA (2009). We have also published a book chapter revisiting a threat analysis for WiMAX/802.16 [16].

1.4 Organization of Thesis

Chapter 2 provides background material for our problem-solving approaches. Chapter 3 outlines the existing research in our areas of interest. Chapter 4 analyzes the threats common to wireless mobile networks. Chapter 5 describes our countermeasures for the location tracking problem in vehicular networks. Chapter 6 describes the HPB mechanism for bounding the position of an uncooperative device, based on the RSS of its transmitted message. Chapter 7 outlines a scheme for aggregating the HPB localization probability distributions ascribed to a transmitting device's location when RSS values from multiple receiver pairs are considered. Chapter 8 evaluates our HPB mechanism with experimental data obtained from a practical outdoor
experiment using a WiFi/802.11 network. Chapter 9 extends HPB with additional algorithm variations in a vehicular network and proposes a mechanism for tracking a mobile transmitter. Chapter 10 extends the use of averaged weighted receiver coordinates to locate a transmitter, where weights are RSS-based and may be computed in either a linear or an exponential fashion. Chapter 11 concludes the thesis and proposes directions for future work.
Chapter 2

Background

We introduce the background material implicit in the remainder of the thesis. We describe the threat analysis methodology employed to identify and rank the risks to wireless mobile networks. We outline the DSRC/WAVE vehicular network architecture examined in our analysis, as well as its provisions for security. For the location estimation problem, we examine a number of radio signal propagation models, and we justify the selection of an empirical large scale path loss model for use in the HPB mechanism.

2.1 Threat Analysis Methodology

A vulnerability in a given type of system constitutes a weakness in its design or implementation that allows unauthorized usage of the system resources. An attacker may subsequently exploit a vulnerability for personal gain to the detriment of authorized users. A threat represents a specific attack that may be perpetrated on a system due to a vulnerability. The risk of a particular threat indicates its estimated degree of severity.

The primary goal of a threat analysis is to determine the security vulnerabilities inherent to a given technology and ascertain the level of risk posed by each identified hazard. With this information, efforts to devise countermeasures can be efficiently focused on the more critical threats.

2.1.1 Types of Methodologies

There are two types of threat analysis methodologies, quantitative and qualitative in nature. Quantitative threat analysis methodologies [59, 110] offer an objective means for estimating the risk posed by a threat by using the statistical probability of its occurrence. However, a significant drawback of this type of methodology lies in its
reliance on historic data of past occurrences of a threat in order to predict future events. In the case of emerging technologies such as wireless access or vehicular networks, no such history of past incidents exists, therefore quantitative threat analysis methodologies cannot be applied. Qualitative threat analysis methodologies [2, 40, 50], on the other hand, allow for discrete, estimated values to be assigned to a variety of risk factors, including the likelihood of occurrence and impact on the victims. Although this type of methodology has no dependence on pre-existing historic data, it is more subjective than its quantitative counterpart. Risk factor values may vary according to the authors of the analysis and the information available. If, however, the analysts clearly define the conditions and boundaries within which each risk factor value is employed, then the consequent threat analysis can achieve reproducible results. The qualitative methodology thus provides a good tool for identifying security flaws and ranking them in order of importance.

2.1.2 Risk Assessment

An example of a qualitative threat analysis methodology is the one put forth by the European Telecommunications Standards Institute (ETSI), which forms the basis for the method used herein to assess the threats to vehicular and wireless access network security. The ETSI developed its methodology in 2003 for the purpose of analyzing security threats to its meta-protocol, Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON) [50]. The methodology allows for identified threats to be ranked in terms of risk, using estimated values for the likelihood of occurrence and impact upon the user or system.

In the ETSI methodology, a threat is ranked as critical under the following conditions: if it is likely and has high impact, if it is likely and has medium impact, and if it is possible and has high impact. A threat is only assessed as major if it is possible and has medium impact, otherwise it is deemed a minor risk. In general, a minor threat requires no countermeasures, a major one needs to be dealt with, and a critical threat must be addressed with the highest priority. We set clear guidelines for assigning qualitative values to the likelihood and impact risk factors, as described below.
Likelihood of Occurrence

The Likelihood risk factor denotes the possibility that attacks associated with a given threat are carried out. Three discrete levels of likelihood are defined: Likely, Possible and Unlikely. In order to evaluate the likelihood of a threat, two additional risk factors are taken into account: the motivation for an attacker to carry out the attack and the technical difficulties that must be resolved by the attacker in order to do so, as shown in Table 2.1. A threat is Unlikely if there is little motivation for perpetrating the specific attack or if significant technical difficulties must be overcome. A threat is Possible if the motivation for an attacker is sufficiently high and the technical difficulties are solvable because the required theoretical and practical knowledge for implementing the attack is available. A threat is Likely if a user or system is almost assured of being victimized, given a high attacker motivation and lack of technical hurdles.

Table 2.1: ETSI Risk Assessment

<table>
<thead>
<tr>
<th>Motivation</th>
<th>Difficulty</th>
<th>Likelihood</th>
<th>IMPACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>None</td>
<td>Likely</td>
<td>Critical</td>
</tr>
<tr>
<td>Moderate</td>
<td>None</td>
<td>Possible</td>
<td>Major</td>
</tr>
<tr>
<td>Low</td>
<td>Any</td>
<td>Unlikely</td>
<td>Minor</td>
</tr>
<tr>
<td>Any</td>
<td>Strong</td>
<td>Unlikely</td>
<td>Minor</td>
</tr>
</tbody>
</table>

Attacker Motivation. The topic of what motivates a computer hacker to conduct attacks has been addressed in the literature in order to devise better countermeasures. In his insightful study [21], anthropologist Roger Blake adapts basic social stratification theory to suggest that hackers are motivated primarily by wealth, power and prestige, although these may be sought in the acquisition of knowledge rather than money. The financial gains to be made through the sale of private information or the disruption of a business rival’s network, the power that private or secret information can afford an individual over others, and the prestige to be garnered before the hacker community are thus powerful motives for attacks. Schifreen [125] proposes five possible motives for hackers:
• Opportunity: the temptation of systems with poor security mechanisms
• Revenge: attacks perpetrated by a disgruntled employee, for example
• Greed: financial gain through the sale of information or blackmail
• Challenge: the prestige obtained through trumping a security system
• Boredom: the lack of better activities to occupy the hacker’s time

To these, Barber [18] adds:
• Vandalism: the defacing of corporate web sites
• Hacktivism: attacks conducted for making political statements
• Information Warfare: control of information by governments

In our application of the ETSI methodology, we assess the possibility of gain in terms of money and power to be somewhat more motivating than prestige. We therefore associate a *High* motivation with an attacker reaping significant financial or power-based gains, a *Moderate* motivation with limited gains or with creating mischief for the purpose of garnering prestige, and a *Low* motivation with little gain for the attacker.

**Technical Difficulty.** Technical difficulty refers to the technological hurdles encountered by an attacker in his attempts to implement a threat. It should be noted that such difficulties are dynamic in nature. What seems like an unsurmountable obstacle today may not be so in a few years’ time. In our threat analyses, we assign a Strong technical difficulty to threats to security mechanisms which currently may not be defeated because some theoretical elements for perpetrating an attack upon them are missing. A Solvable technical difficulty is associated with a security mechanism which may be countered or has been defeated in a related technology. A technical difficulty of None is assigned when a precedent for an attack already exists.

**Impact on Victim**

The *Impact* criterion evaluates the consequences for a user or system if a given threat is carried out. From the single user’s point of view, the impact of a threat is rated as *Low* if an attack results in only annoyance and the consequences, if there are any, are reversible and can be repaired. From the point of view of a system serving several
users, a threat is ranked with *Low* impact if the possible outages are very limited in scope, for example with few users affected for a short duration. For the user, the impact is *Medium* if a loss of service occurs for a short amount of time. For a system, the consequences of a *Medium* impact exploit consist of outages that are limited in both scope and possible financial losses. A threat carries a *High* impact for a user if an attack causes a loss of service for a considerable period of time. If targeted at a system, an attack associated with a *High* impact threat results in outages over a long period of time with a large number of users affected, possibly accompanied by law violations or substantial financial losses.

**Adapted Risk Assessment**

We found, through experience with the ETSI methodology, that many threats are over-classified as critical when they are better ranked as major. For example, our initial analysis of the threats to DSRC/WAVE, using the ETSI methodology as defined in Table 2.1, found eight critical risks and a single major one. Because of the restrictive conditions under which a threat is ranked as major, this classification remains under-utilized. Of the eight critical threats, we estimate that more than half represent a far less immediate risk than the others and are thus more logically ranked as major.

We revisit and update the ETSI methodology to prevent the over-classification of a large number of threats from obscuring and trivializing the serious nature of the more dangerous risks. By assigning the critical designation only to high impact threats that are likely to occur, our adaptation enables our threat analyses to place greater emphasis on the truly urgent risks, as shown in Table 2.2.

<table>
<thead>
<tr>
<th>Likelihood</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likely</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>Critical</td>
</tr>
<tr>
<td>Possible</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>Major</td>
</tr>
<tr>
<td>Unlikely</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Minor</td>
</tr>
</tbody>
</table>
2.2 Vehicular Network Model

We describe the emerging DSRC/WAVE architecture, along with its security measures. We review the requirements for vehicle safety applications, which constitute one of the two broad classes of applications envisioned for vehicular networks.

2.2.1 DSRC/WAVE Architecture

DSRC was conceived to provide an architecture for nodes within a vehicular network to communicate with each other and with the infrastructure in a secure and efficient manner [8]. In DSRC, subsequently specialized as WAVE [65, 66, 67, 68], vehicles are equipped with On-Board Units (OBUs) which can communicate with each other to propagate safety warnings through Vehicle-to-Vehicle (V2V) communications. As well, OBUs have access to infrastructure devices called Road-Side Units (RSUs) which are located intermittently along city streets and highways to provide access to a variety of services and applications, such as LBSs and wireline access, through Vehicle-to-Infrastructure (V2I) communications. By convention, since vehicles can only engage in V2V and V2I exchanges through their OBU device, we use the terms vehicle and OBU interchangeably. A special class of OBUs, called Public Safety On-Board Units (PSOBUs), are entrusted with special capabilities. PSOBU-equipped vehicles, which include police cars, fire trucks and ambulances, may operate as OBUs or RSUs, as circumstances dictate. All OBUs and PSOBUs are expected to be Global Positioning System (GPS) enabled, so that they are continuously aware of their location and have access to a synchronized source of time. DSRC/WAVE operates in the 5.9 GHz band and has 75 MHz of bandwidth allocated for vehicular communications, which are based on line of sight with a range of up to one kilometer and vehicle speeds of up to 140 km/h.

Two types of vehicular applications have been identified. Transaction applications allow OBUs (via RSUs) access to remote service providers hosting a variety of information services such as navigation information systems and location finders. Broadcast applications typically convey vehicle safety information, for example in collision avoidance and driver assistance systems. At the network layer [66], transaction applications employ the Internet Protocol, version 6 (IPv6), to carry point-to-point
transaction messages between OBU's and RSUs, and ultimately to the remote service providers. Broadcast applications, on the other hand, use the custom WAVE Short Message Protocol (WSMP) in order to minimize message size and delivery latency for critical safety broadcast messages.

2.2.2 WAVE Security

The WAVE security standard [67] outlines the PKI mechanisms entrusted with securing both transaction and broadcast messages. Given that transaction messages typically contain personal or financial information and that their delivery latency is not subject to the same stringent conditions as broadcast messages, they are secured through symmetric encryption using the AES in Counter with CBC-MAC (CCM) Mode (AES-CCM) [48] with a random key. This key is in turn encrypted with the receiver's public key using the asymmetric encryption algorithm Elliptic Curve Integrated Encryption Scheme (ECIES) [35]. Broadcast messages, on the other hand, are meant to convey information that is at once directed for public consumption and bounded by strict latency limits. As a result, they are not encrypted but rather sent in the clear and digitally signed by the sender with the Elliptic Curve Digital Signature Algorithm (ECDSA) [4] as a means of non-repudiation and message integrity insurance. In this manner, sending WAVE devices are held accountable for the messages they transmit. Every signed message includes the current timestamp, obtained from the OBU's internal clock which is synchronized with the GPS time. This timestamp is used by the receiver to verify the message against a cache of recently received messages as a defense against replay attacks.

In order for vehicles to sign messages, they must possess a digital certificate. The security standard recommends that procedures for PSOBU and RSU certificate enrollment include a manual component, but OBU certificate enrollment procedures have not yet been put forth. When signing a message, a vehicle may also include a certificate chain, i.e. the certificate that authorized the vehicle's certificate, as well as the certificate that authorized the certifying certificate, and so on, up to a root certificate which is issued by a trust anchor such as a governmental agency responsible for licensing vehicles. Upon receipt of a signed message, each node must verify that
it recognizes the root certificate used to authorize the sender’s certificate. The node must also ensure that none of the certificate chain’s members have been revoked by comparing each one against a Certificate Revocation List (CRL). These CRLs and the collection of valid root certificates are stored locally on each node and must be updated periodically.

2.2.3 Vehicle Safety Application Requirements

Several requirements for vehicle safety applications are documented and must be considered in our efforts to reconcile the security of their broadcast messages with driver privacy.

The Vehicle Safety Communications Consortium (VSCC) of automakers and its Crash Avoidance Metrics Partnership (CAMP) project have identified a set of essential vehicle safety applications [32], which may be grouped into five separate categories:

- Collision avoidance
- Driver assistance (e.g. merging, lane changes, traffic signals, platooning)
- Traffic information (e.g. traffic jams, road closures, work zones)
- Hazards (e.g. emergency vehicles, curve speed, low structures)
- Notifications (e.g. alerts, vehicle recalls)

While the information dissemination applications comprising the last three categories may be initiated by infrastructure devices in V2I, multihop V2V communications are necessary to ensure propagation to vehicles beyond the infrastructure range. The first two types of applications require periodic broadcast messages from each vehicle in order to gather its coordinates, so that collision risks may be assessed. Factoring in human reaction times, these applications place limits on message delivery latency, in the order of 100 ms, which in turn is affected by message size. The size of safety messages is expected not to exceed 100 bytes.
2.3 Radio Propagation Models

We examine potentially suitable radio signal propagation models to assist in our efforts to bound the physical location of an uncooperative device, based on the RSS values of its transmitted message.

Radio Frequency (RF) signals are subject to attenuation as they propagate through the air. Large scale fading occurs when a signal encounters sizable terrain-based obstacles such as buildings, trees and hills. Small scale signal fading is caused by the movement of a mobile device. The cumulative effect of fading over time and distance is termed the *path loss*. A number of theoretical models for estimating path loss have been put forth in the literature for the purpose of simulating propagation environments, and these models may be classified into two categories. Empirical propagation models use probabilistic methods to predict received signal characteristics such as path loss and strength. Deterministic models are specific to a particular area and take into account the various obstacles therein. The dynamic nature of outdoor environments inherent to wireless access networks such as WiMAX/802.16 and vehicular networks lends itself better to empirical models rather than deterministic ones. As a result, we focus on the former approach.

Several empirical propagation models have been proposed for forecasting large scale path loss as a function of the distance between a transmitter and a receiver. These are of particular interest, since they lend themselves to our purpose. If the distances between an uncooperative transmitter and trusted receivers can be approximated from the path loss, which is directly proportional to the RSS values, the transmitter’s location can be estimated. The Okumura model [104] predicts path loss based on transmitter and receiver antenna height, as well as the mean attenuation and the environment-based gain which can be obtained from experimental results. Miyashita et al. [97] observe that the Okumura model is unsuitable over complex terrain due to the difficulty in ascertaining the required correction factors. The validity of the Hata model [61], also known as the Okumura-Hata model, has been demonstrated for frequencies between 150-1500 MHz, but not at the higher frequencies commonly used in newer technologies. For example, WiMAX/802.16 uses the 2-11 or 10-66 GHz bands, while DSRC/WAVE vehicular networks operate in the 5.9 GHz
band. The two-parameter Nakagami model [99] has been suggested as best suited for modeling channel characteristics in vehicular communications [137]. This model is dependent upon two parameters, the mean received power and a fading parameter, that are both obtained through experimental studies for a given discrete value of the Transmitter-Receiver (T-R) distance. If the T-R distance changes, so do the values of both parameters. As a result, the Nakagami model is unusable for predicting a T-R distance from the measured path loss, since the parameters required to compute the path loss are dependent upon the T-R distance. The log-normal shadowing model outlined by Rappaport [115] provides a simple model to estimate large scale path loss from a T-R distance. Since our aim is to approximate a T-R distance based on a measured path loss, the log-normal shadowing model is best suited to our purpose.
Chapter 3

Related Work

We survey the existing literature pertaining to our three problems of interest.

3.1 Threats to Wireless Mobile Networks

In their 2005 report on the CAMP project [132], the VSCC outlines a number of threats to DSRC, including replay attacks, Denial of Service (DoS), and sensor and GPS spoofing. They also highlight the possibility of malicious insiders with access to cryptographic material. Blum and Eskandarian [22] describe a number of security threats to vehicular networks, including DoS, impersonation and fabrication, as they pertain to broadcast messages. The authors outline an architecture secured using PKI for maintaining a Vehicular Ad-hoc Network (VANET)'s integrity and functionality through the use of vehicular clusterheads managing authorized access to the network. Hubaux et al. [63] identify privacy as an issue. This topic is further expounded in [117], where Raya and Hubaux identify threats to the broadcast messages exchanged between vehicles in VANETs. The authors propose a security architecture featuring the use of digital signatures and multiple anonymous key pairs, or pseudonyms. Raya et al. [118] subsequently identify jamming, forgery, traffic tampering, impersonation and privacy violation as vulnerabilities. To our knowledge, no previous research into threats to vehicular networks employs a formal threat analysis methodology or provides a ranking of the risks associated with possible attacks.

3.2 Secure Anonymity in Vehicular Networks

Various means of anonymizing network traffic to preserve sender privacy have been proposed in the literature. These may be categorized as infrastructure-based, cryptography-based or protocol-based.
**Infrastructure-based Anonymity.** Specialized network servers known as *anonymousizers* are used to hide a sender's identity from potential eavesdroppers, for example in the *onion routing* protocol introduced by Syverson *et al.* [136] and in the *Crowds* web transaction anonymizer scheme proposed by Reiter and Rubin [119]. Beresford and Stajano [20] adapt the anonymizer concept to the domain of pervasive computing by introducing common *mix zones* in which users change pseudonyms, so that observers cannot match a user identity departing the zone to its identity upon entry. Spatio-temporal cloaking and its related concept, *k-anonymity*, are suggested by Gruteser and Grunwald [60], as well as Gedik and Liu [57] and McDiarmid and Irvine [95], to provide a means of protecting a user's identity from LBS providers. By aggregating the requests bound for a given LBS provider until one user is indistinguishable from *k* - 1 other users, either with respect to physical location or the timing of the request, an anonymity server can protect a user's identity and precise location. However, the inherent necessity for all messages to be routed through middleware makes infrastructure-based anonymity measures infeasible for vehicular networks, given the requirement for low latency and scalability, as well as the need for precise vehicle locations to be available for many applications.

**Cryptography-based Anonymity.** One mechanism for rendering PKI anonymous is outlined by Zhang *et al.* [151]. It allows a Certificate Authority (CA) to issue multiple anonymous certificates based on an original one, so that any message signed with an anonymous certificate can be traced back to the original certificate owner by the CA. The additional certificates act as pseudonyms for the owner and can be changed periodically to retain anonymity. However, as with any alias, they can become trackable over time unless frequently changed, and an algorithm for managing suitable switching intervals is lacking, as are the means for certificate revocation. Group signatures, on the other hand, were pioneered by Chaum and van Heyst [37] to allow a message to be signed by any member of a pre-defined group with the member's private key, while allowing a recipient to verify the signature with a shared group public key. In this manner, no recipient or eavesdropper can know which private key was used to generate the signature, except for a group manager who can "open" a signature if necessary to trace it to the original signing private key. Unfortunately,
some group member revocation schemes, such as the one advanced by Bresson and Stern [28], require generating signatures of proportional size to the number of group members. In a vehicular network encompassing millions of vehicles, this is unscalable. Others, such as the scheme suggested by Ateniese et al. [9], require a group manager to re-issue and disseminate new group certificates to the remaining members for every revocation. Boneh et al.’s scheme using a fixed signature size [24] necessitates every group member to calculate a new private key and group public key based on an exhaustive revocation list each time a member leaves the group. The means for such a list to be effectively disseminated to remaining group members, in a reliable and scalable fashion, is not explored and represents an important obstacle to the success of any list-based revocation scheme. Another mechanism put forth by Boneh and Shacham [25] requires revocation lists to be distributed to likely verifiers in a localized area. While this approach may be suitable for devices of limited mobility, the potential for vehicles to travel across thousands of kilometers entails the global dissemination of revocation lists. Overall, group signatures appear to be better suited to smaller, more manageable groups than vehicular networks.

**Protocol-based Anonymity.** A number of PKI-based protocols have been suggested for introducing anonymity in VANETs. In [117], Raya and Hubaux suggest that each vehicle be equipped with a set of pre-loaded certificates and that the vehicle periodically switch to a new one in order to maintain anonymity. However, it is unclear how rogue device revocation could be successfully implemented. In the event that a rogue is identified and directed to destroy its compromised certificate, mechanisms must be put in place to ensure compliance. Another pitfall in this approach concerns its scalability, given the number of vehicular devices and potential for mobility across the breadth of an entire continent. Sufficient numbers of certificates must be issued for each vehicle to maintain anonymity over a significant period of time. As a result, the size of the certificate database to be searched by a CA in order to match a compromised certificate to its owner’s identity is daunting. The pseudonym-based approach described by Armknecht et al. [7] relies on revocation lists for the eviction of a rogue device from the vehicular network. The distribution of such lists presents an important management issue and a serious obstacle to the feasibility of
security measures relying upon them. Papadimitratos et al. [107] advocate dynamic pseudonyms for privacy purposes. Revocation of pre-authorized sets of pseudonyms is proposed through either CRLs or through timestamped proofs of the pseudonym’s current validity. The additional overhead associated with every message represents a drawback to the timely broadcast of vehicle safety messages. Kamat et al. [73] address the revocation issue by having pseudonyms issued by BSs to vehicles whose certificates don’t appear in CRLs. An interesting feature of this scheme is the scalability of its revocation scheme, where revocation lists are only distributed to the BSs. Since this is a PKI approach where a digital signature must be generated and verified for every broadcast message, the cryptographic operation overhead for every message sent and received remains significant. Parno and Perrig [109] propose a dynamic key distribution system where vehicles periodically request new certificates from a CA, but this places an additional load on the CA in addition to the network. In [122], Sampigethaya et al. outline a scheme where vehicles are organized into groups, and only group leaders broadcast messages while the remaining members maintain silence in order to promote anonymity. Given the high volume of messages required for vehicle safety applications, silent periods are not suitable for broadcast messages. Beresford and Stajano’s concept of mix zones [20] is adapted to vehicular networks by Freudiger et al. [53]. Mix zones are created at intersections, and the messages sent within the zone are encrypted using a local symmetric key obtained from the nearest RSU. The processing delay introduced by the encryption and decryption of vehicle safety messages constitutes a liability in this approach.

### 3.3 Localization of Uncooperative Nodes

A number of wireless network scenarios require a location estimation solution with minimal assumptions. Malfunctioning sensors or energy-drained devices in emergency situations, such as a cellphone requiring 911 assistance, may need to rely on other network nodes to locate them if they are unable to do so. Another important scenario envisioned for our localization approach addresses the threat of a malicious insider. Such an authorized network member in possession of valid credentials may
transmit with impunity attack messages containing false information in order to mislead honest nodes. In a recent survey of IT security professionals [111], nearly half of the respondents reported that security breaches committed by malicious insiders engaging in corporate sabotage was a frequent occurrence. Another survey conducted by CSO Magazine, the U.S. Secret Service and CERT reports that over 40% of the attributable security breaches uncovered were committed by malicious insiders [42]. Because of these attackers’ privileged position in the network, the mitigation of their exploits poses a formidable challenge.

We expose the serious nature of the malicious insider threat. We outline existing techniques for their detection and subsequent attribution to a perpetrator. We approach the attribution issue by exploring the means to probabilistically locate the physical source of an attack radio signal, assuming neither the cooperation of a malicious transmitter, nor knowledge of the transmitting power (EIRP) employed. To this end, we review various device and network-based location estimation schemes, as well as existing mechanisms for the combination of probabilistic evidence.

3.3.1 Insider Attacks

The advent of public key cryptography, where a node is authenticated through the possession of a public/private key pair certified by a trust anchor, has addressed the primary threat posed by an outsider without valid credentials. But a wireless network safeguarded through a PKI is only as secure as the means implemented to protect its member nodes’ private keys. For example, the DSRC/WAVE security standard [67] advocates the use of digital signatures to secure vehicle safety broadcast messages, with tamper proof devices storing secret keys and cryptographic algorithms in each vehicle. Yet a convincing body of existing literature questions the resistance of such devices to a motivated attacker, especially in technologies that are relatively inexpensive and readily available [5, 6]. In the absence of strict distribution regulations, for example if tamper proof devices for vehicular nodes are available off the shelf from a neighborhood mechanic, a supply chain exists for experimentation with these devices for the express purpose of extracting private keys. The National Institute
of Standards and Technology (NIST) has established a certification process to evaluate the physical resistance of cryptographic processors to tampering, according to four security levels [102]. However, tamper resistance comes at a price. High end cryptographic processors certified at the highest level of tamper resistance are very expensive, for example an IBM 4764 coprocessor costs in excess of 8000 USD [64]. Conversely, lower end tamper evident cryptographic modules, such as smartcards, feature limited mechanisms to prevent cryptographic material disclosure or modification and only provide evidence of tampering after the fact [145]. The European consortium researching solutions in vehicular communications security, SeVeCom, has highlighted the existence of a gap in tamper resistant technology for use in vehicular networks [129]. While low end devices lack physical security measures and suffer from computational performance issues, the cost of high end modules is prohibitive. The gap between the two extremes implies that a custom hardware and software solution is required, otherwise low end devices may be adopted and prove to be a boon for miscreants.

**Detection**

Before a malicious insider can be stopped, an attack must be detected. Exploits originating from network outsiders can generally be thwarted with either authentication mechanisms, such passwords and digital certificates, or resource access control methods to restrict access to a minimal set of users. However in many domains, for example in DSRC/WAVE vehicle safety applications, an attack may only be detected if an invalid digital certificate is used to sign the message. Attack broadcasts by a rogue insider can thus go unchallenged.

Malicious activity on the part of insiders can be exposed using Intrusion Detection Systems (IDSs) [45]. These specialized systems may be statistics-based, where usage patterns that deviate probabilistically from expected user behaviour are identified, or rule-based, where proper system usage is defined and any activity outside these confines is flagged [112]. Another method to preemptively protect resources against inside attackers is to establish a *honeypot*, as proposed by Spitzner [134], in order to lure rogues away from critical systems.
We deem that the nodes best situated to detect an insider attack in a mobile network are the ones that possess the most contextual information in order to flag an anomalous message. They are necessarily located in close proximity to the perpetrator and are therefore within radio range. These one-hop neighbors also have first-hand knowledge of the attack message RSS values, which can be combined to bound the possible location of the malicious insider with a degree of confidence.

Attribution

Once an insider attack is detected, it must be attributed to its perpetrator. The field of network forensics, a term coined by Ranum [114], explores mechanisms to proactively capture network packets and events, and to analyze them in case of an attack. Payload attribution in network forensics, as introduced by Shanmugasundaram et al. [131], identifies the source and destination of a malicious packet in order to attribute it to an attacker. This approach assumes multiple hops between a sender and a receiver, meaning that packets must be logged throughout the breadth of the network, placing an additional storage burden on every network device.

Other means of attributing an exploit to an insider node are based on its logical identity. This approach can be fraught with problems if the identity is forgeable. Dynamic MAC addresses are supported in some domains, for example in vehicular networks [65], to promote privacy. A malicious insider may easily abuse this feature to assume a new identity at will. Password-based access control and digital certificates associated with public/private key pairs constitute additional means of identification. However, if the associated revocation scheme is based solely on the originator's logical identity or other falsifiable credentials, the network cannot attribute the source of an attack message directly to the malicious insider, nor can it prevent the attacker from using a new identity in subsequent exploits. Determining the physical location of an attack message's originator can be an important first step in apprehending the perpetrator, possibly linking its location to additional logical identities, and alerting neighboring devices to its presence in order to preemptively contain the impact of further attacks. An additional advantage to the use of a node's physical location as a unique unforgeable identity allows the network to detect and circumvent Sybil attacks,
first described by Douceur [46], where one attacker may assume multiple colluding identities in order to mislead honest nodes. Linking a node to a geographical position ensures that it is a unique physical device in the network and not a virtual entity under an attacker’s control.

3.3.2 Device-based Location Estimation

The problem of wireless node localization may be approached from one of two main directions: device-based (also known as handset-based) and network-based. Device-based self-localization involves a node seeking to learn its own position, occasionally with the help of other trusted devices within radio range. For example, the use of GPS can be seen as a device-based approach, since a node uses information supplied by a set of satellites in order to determine its coordinates. In techniques based on Time of Arrival (TOA), a device may situate its position with respect the known locations of other nodes by correlating arrival time of received messages and thus determining its distance to each node. A large proportion of the localization techniques proposed for sensor networks assume a device-based approach. For example, the three/two neighbor algorithm proposed by Barbeau et al. [15] allows for a sensor of unknown position to estimate its location from the coordinates of neighboring nodes, based on their respective TOA-approximated distances.

Self-localization achieved with device-based techniques can produce granular results. For example with GPS, a precision of ten meters may be achieved [103]. But self-localization is not feasible in all scenarios. An uncooperative node is one that cannot be relied upon to determine its coordinates, for example a defective sensor, a malicious device engaging in a security exploit or a low-battery handset in a critical situation. A malicious node broadcasting an attack message cannot be expected to cooperate with efforts to uncover its position. In other situations, a malfunctioning device or one whose battery is nearly drained may be unable to compute and report its coordinates to other nodes. Given that such devices may supply erroneous location information, either willfully or accidentally, they must be localized by other network nodes using measurements that cannot easily be forged.
Recent events involving localization failure of 911 callers have spurred the Canadian Radio-television Telecommunications Commission (CRTC) to regulate the same wireless Enhanced 911 (E911) provisions [33] as the Federal Communications Commission (FCC) in the U.S. [52] Under Phase II of the FCC and CRTC plans, localization efforts based on a handset (device-based) must yield a location accuracy of 50 meters in 67% of cases and 150 meters 95% of the time. Network-based localization, where other nodes estimate the position of a device, must accurately reveal a target location within 100 meters 67% of the time and within 300 meters in 95% of cases.

3.3.3 Network-based Location Estimation

The lion’s share of existing research into wireless device location determination presumes the collaboration of the node being localized. Although multiple types of network-based localization mechanisms have been proposed in the literature, our HPB algorithm fills a significant void by addressing a novel threat scenario, where few assumptions are made regarding the assistance and radio equipment of an uncooperative transmitter. Table 3.1 compares the assumptions upon which existing schemes rely against those of HPB. For example, triangulation is predicated upon the use of directional antennas to pinpoint the source of a signal. Time-based schemes such as TOA and Time Difference of Arrival (TDOA) require fine resolution clocks and accelerated processors to measure time of flight with minimal processing delays. Existing RSS-based schemes generally presuppose a known EIRP. In assuming none of these capabilities, the HPB mechanism strives for applicability in a broad set of technologies and security scenarios.

We further outline a number of existing network-based location estimation approaches.

Triangulation

The concept of triangulation was first introduced by Frisius [55] for map surveying and locating far-off geographical points. In more recent years, this approach has also served as a network-based technique to localize a transmitting device using two receivers of known coordinates and the transmission’s angle of arrival. A significant
Table 3.1: Localization Scheme Assumptions

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Special Hardware</th>
<th>Local Knowledge</th>
<th>Cooperation</th>
<th>EIRP Knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Triangulation [55]</td>
<td>directional</td>
<td>none</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>receiver antenna</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time-based [27, 124, 144]</td>
<td>accelerated</td>
<td>none</td>
<td>round-trip</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>CPU, synchronized clocks</td>
<td></td>
<td>response</td>
<td></td>
</tr>
<tr>
<td>Signature-based [11, 51, 121]</td>
<td>none</td>
<td>signal profiles</td>
<td>none</td>
<td>training phase</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EIRP</td>
</tr>
<tr>
<td>Relative RSS [62, 92]</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>same as other nodes</td>
</tr>
<tr>
<td>RSS ring-based [17, 91]</td>
<td>none</td>
<td>distance to reference point</td>
<td>none</td>
<td>must be known</td>
</tr>
<tr>
<td>HPB [82]</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
</tr>
</tbody>
</table>

drawback of the triangulation method is the necessity that receivers be equipped with directional antennas, so that the angle at which an incoming transmission originates may be measured. Such specialized hardware poses an additional burden on network deployment costs. We thus focus our research efforts on mechanisms that assume the more commonly available omnidirectional antennas.

**Time-based**

Time-based mechanisms, where T-R distances are estimated from TOA round-trip beacons, include the location verification methods put forth by Brands and Chaum [27], Sastry et al. [124] and Waters and Felten [144]. These schemes make a number of assumptions that are unsuitable for our scenarios. They require the availability of accelerated processors at both the transmitter and receiver to factor out the relatively large processing delays compared with a beacon’s time of flight. They assume that the clocks at each node feature nanosecond precision and are synchronized with each other. They also require the cooperation of the node being localized, since the round-trip beacon must be returned.

The TDOA approach translates the difference in arrival times of a given message at two receivers into a distance difference, and plots a hyperbola with the receiver
coordinates as foci. Multiple receiver pairs yield multiple hyperbolas, with a transmitter location determined at the common intersecting point. With this technique, the clocks at receiving devices must be synchronized with nanosecond precision, otherwise a common intersecting point may not exist. Even highly correlated GPS clocks may exhibit up to one microsecond of clock drift between receivers [135]. At the speed of light, a one microsecond drift translates into a distance difference of 300 meters, resulting in a margin of location error greater than the FCC regulations for E911 location accuracy.

**RSS-based**

RSS-based localization algorithms come in two flavours: signature dependent and geometric. Signature dependent techniques rely on an existing training set of RSS signalprints established during an offline training phase. This map of known RSS measurements is subsequently consulted during the localization phase to estimate a node's position based on the similarity of the signals received at trusted base stations or access points to the signalprints found in the training set, as outlined in Bahl and Padmanabhan [11], Faria and Cheriton [51], Roos *et al.* [121] and Ladd *et al.* [76]. Experiments of this nature have been conducted in indoor environments, where typically few environmental variations occur. Their applicability in outdoor scenarios, for example in WiMAX/802.16 wireless access networks or vehicular networks, remains an open question.

Geometric RSS-based localization techniques aim to estimate a node’s location in Euclidian space based on the signal strength of messages received from trusted nodes within range, for example in the schemes proposed by Chong Liu *et al.* [92] and Bo-Chieh Liu *et al.* [91]. Such approaches generally assume the complicity of the target node in reliably measuring the RSS of received beacon messages. Therefore, they cannot be used to localize an uncooperative node.

Relative RSS values are used in the algorithms put forth by He *et al.* [62] and Chong Liu *et al.* [92], where a sensor seeking to learn its position compares the RSS values it receives from trusted anchors. But these schemes assume that the target device transmits at the same EIRP as all the other nodes. While this assumption is
reasonable in the realm of sensor networks, our threat model can make no such supposition of a malicious device. Zhong et al. [152] demonstrate how RSS ratios between a minimum of four receivers can be used to passively localize a node. Their scheme assumes a uniform propagation environment and fails to consider signal strength variations between receivers.

Existing hyperbolic localization schemes use TDOA time differences to estimate receiver distance differences. An algebraic approach is then applied to estimate a device’s coordinates by solving a set of non-linear hyperbolic equations, for example in Chan and Ho [36]. Signal strength fluctuations taken into account are commensurate with small scale fading effects, and multiple signal measurements over time are required, so that variations can be averaged out. Even in a moderately shadowed environment, these approaches may fail to yield any solution. As well, some of these techniques such as the ones outlined by Bo-Chieh Liu et al. [90, 91] assume a known distance from the transmitter to the closest receiver as a reference point.

Barbeau and Robert [17] employ RSS measurements in outdoor wireless access networks, such as WiMAX/802.16, to allow a MS to detect whether a BS advertising its availability for a handoff may be a rogue. A probabilistic RSS-based geometric model is presented where the RSS measurements obtained from neighboring BSs by a non-localized MS may be used to construct annuli whose non-empty intersection likely contains the MS. An empty intersection may indicate a RSS measurement originating from a rogue BS.

Our HPB scheme extends the RSS-based geometric model proposed in [17] to localize a single transmitter from multiple receivers, taking into account the unknown EIRP potentially employed by a malicious insider. In addition, since the compounding effect of a large EIRP range and probabilistic fluctuations in RSS measurements may result in a potentially extensive area for the target device, we propose the use of hyperbola pairs rather than annuli for the localization. In this manner, relative rather than absolute RSS values are used to effectively reduce the area containing an uncooperative transmitter’s probable location.
A number of publications related to the location determination of vehicular devices focus on self-localization, where a node seeks to learn its own position [108, 26]. Although the measurements and information provided to these schemes are presumed to be trustworthy, this assumption does not hold for finding an attacker invested in avoiding detection and eviction from the network.

Some mechanisms for the localization of a vehicular device by other nodes are based on the principle of location verification, where a candidate position is proposed, and some measured radio signal characteristic, such as time of flight or RSS, is used to confirm the vehicle’s location. For example in [63] and [34], Hubaux, Capkun and Luo adapt Brands and Chaum’s distance bounding scheme [27] for this purpose. Yet a degree of cooperation is expected on the part of a target device for supplying a position. Additionally, specialized hardware is necessary to measure time of flight, including nanosecond-precision synchronized clocks and accelerated processors to factor out relatively significant processing delays at the sender and receiver. Xiao et al. [146] employ RSS values for location verification, but they assume that all devices, including malicious ones, use the same EIRP. An attacker with access to a variety of radio equipment is unlikely to be constrained in such a manner.

Location verification schemes for detecting false position reports may be beacon-based or sensor-based. Leinmüller et al. [87] filter beacon information through a number of plausibility rules. Because each beacon’s claimed position is corroborated by multiple nodes, consistent information is assumed to be correct, based on the assumption of an honest majority of network devices. This presumption leaves the scheme vulnerable to Sybil attacks [46]. If a rogue insider can generate a number of Sybil identities greater than the honest majority, then the attacker can dictate the information corroborated by a dishonest majority of virtual nodes. In ensuring a unique geographical location for a signal source, our HPB-based algorithms can detect a disproportionate number of co-located nodes.

Tang et al. [138] put forth a sensor-based location verification mechanism where video sensors, such as cameras and Radio Frequency Identification (RFID) readers, can identify license plates. However, cameras perform sub-optimally when visibility is
reduced, for example at night or in poor weather conditions. This scheme is supported by PKI-based beacon verification and correlation by an honest majority, which is also vulnerable to insider and Sybil attacks. Another sensor-based mechanism is suggested by Yan et al. [148], using radar technology for local security and the propagation of radar readings through beacons on a global scale. Again, an honest majority is assumed to be trustworthy for corroborating the beacons, both locally and globally.

Some existing literature deals explicitly with mobile device tracking, including the RSS-based mechanisms put forth by Mirmotahhary et al. [96] and by Zaidi and Mark [149]. These presume a known EIRP and require a large number of transmitted messages so that the signal strength variations can be filtered out.

**Centroid Localization**

In the realm of sensor networks, Centroid Localization (CL) has been suggested as an efficient method that never fails to produce a solution. The original incarnation of CL is described by Bulusu et al. [30], and localizes the transmitting source of a message to the \((x, y)\) coordinates obtained from averaging the coordinates all receiving devices within range. Weighted Centroid Localization (WCL), as proposed by Blumenthal et al. [23], assigns a weight to each of the receiver coordinates, as inversely proportional to either the known T-R distance or the link quality indicator available in ZigBee/IEEE 802.15.4 sensor networks [78]. Behnke and Timmermann [19] extend the WCL mechanism for use with normalized values of the link quality indicator. Schuhmann et al. [127] conduct an indoor experiment to determine a set of fixed parameters for an exponential inverse relation between T-R distances and the corresponding weights used with WCL. Orooji and Abolhassani [105] suggest a T-R distance-weighted averaged coordinates scheme, where each receiver’s coordinates is inversely weighted according to its distance from the transmitter. But this approach assumes that the receivers are closely co-located and that the T-R distance to at least one of the receivers is known *a priori.*
3.3.4 Combination of Probabilistic Evidence

A set of receivers can use the HPB mechanism to localize a transmitter within a hyperbolic area with a given probability distribution. The confidence levels vested in common hyperbolic area intersections must then be aggregated so that additional evidence supports rather than weakens the probability ascribed to the intersections.

The seminal work of Dempster and Shafer in establishing the Dempster-Shafer theory of evidence [44, 130] formulates the foundations for belief functions and rules for the combination of evidence provided by a set of independent observers. Previously, a multiplicative probability approach reconciled varying degrees of confidence in a given event by multiplying together the probabilities assigned to the event by the observers. In contrast, the Dempster-Shafer rule of combination introduces a normalization mechanism to compensate for the lack of knowledge of an observer vis-à-vis an event and to resolve the degree of conflict between observers. Our scenarios require the combination of probabilities in a hyperbolic area intersection common to a large number of receivers, each assigning the intersection a high degree of confidence. By definition, the common intersections are the subject of little conflict in the evidence supplied by the receivers. In circumstances of little conflict, the Dempster-Shafer rule of combination reduces to simple multiplicative probability.

Variants on the Dempster-Shafer theory aim to address situations with even greater sources of conflict, for example in Yager [147] and Inagaki [69]. As with the Dempster-Shafer approach, these reduce to multiplicative probability in the absence of significant conflict. Proportional assignment of probability according to the percentage of observations supporting an event is presented in Zhang [150] and Cholvy [39]. Because we are interested in a small intersection of relatively large hyperbolic areas, proportional assignment distributes the probability over the vast extent of the hyperbolic areas, rather than concentrating it in the intersection. The consensus operator introduced by Jøsang [70] yields similar results by combining both positive and negative observations for an event, but attributing weight to each observer according to a proportional assignment basis.

Voorbraak [140, 141] and Wakker [142] advocate the Dempster-Shafer theory for
scenarios dealing with ignorance and Bayesian conditioning for those involving uncertainty. The open world principle assumed by the Dempster-Shafer theory of evidence is best suited to knowledge acquisition systems seeking to expand a corpus of beliefs and thus to alleviate ignorance. Our scenarios assume a closed world, where a transmitter must necessarily be positioned within the radio range of receivers. Although the transmitter's precise location within Euclidian space is uncertain, each receiver pair assigns this entire space a fully defined probability distribution. Because our scenarios model uncertainty rather than ignorance, the Bayesian conditioning model emerges as more suitable for our purposes.
Part II

Threats and Location Tracking
Mitigation
Chapter 4

Threats to Wireless Mobile Networks

We assess and classify the risks to the DSRC/WAVE vehicular network architecture using an adapted threat analysis methodology. We re-examine an existing analysis of threats to the WiMAX/802.16 wireless access technology with the same method. We correlate the results of both analyses and highlight the common threats.

As DSRC/WAVE and WiMAX/802.16 constitute prime examples of emerging wireless technologies, in that they operate in the higher frequency bands (5.9 GHz and 2-66 GHz respectively) and support mobile communications, the serious risks jointly identified may provide an important indication of dangerous threats in other novel wireless platforms.

4.1 Threat Analysis for DSRC/WAVE

Given the potential for attackers to wreak havoc in vehicular networks, there is a compelling need to identify and address the most severe security threats specific to DSRC/WAVE. We conduct an analysis of those threats in [79], and we refine it here using our adaptation of the ETSI methodology found in Table 2.2. Identified threats are ranked as critical, major or minor, depending on their likelihood of occurrence and impact on the user or the network.

In our analysis, we focus on the most basic security attributes to be preserved in vehicular networks: availability, authenticity and confidentiality. The collated list of threats, organized by risk category, can be found in Table 4.1.

4.1.1 Threats to Availability

Threats to the availability and consistent behaviour of a vehicular network include DoS attacks, the introduction of malicious software (malware), the presence of black holes, and the potentially high volume of messages introduced through spamming.
**DoS.** DoS attacks render a network unavailable to its users, for example by flooding the nodes with messages or by jamming signals at the physical layer. These attacks can be carried out either by network insiders turned rogue or by outsiders to the network. One way to incapacitate a vehicular network is to artificially generate such a high volume of false messages that the network’s nodes, OBUs and RSUs alike, cannot process the superfluous data. This technique is known as *flooding* and results in important messages being lost. Consequences may include accidents if collision warnings or platoon directives are not delivered. With *jamming*, an attacker hampers message delivery by creating interference on the communications channel, thereby disrupting the applications which depend upon it. Given that DoS represents a disruption rather than an opportunity for gain, the motivation required on the part of an attacker is rated as moderate according to our analysis criteria, described in Section 2.1.2. The technical difficulty involved is solvable, given that it is theoretically possible. Since DoS would result in temporary outages, the impact on a network is ranked as medium, and according to Table 2.2, the threat is assessed as major.

**Malware.** The introduction of malware, such as viruses or worms, into a vehicular network has the potential to cause serious disruptions to its operation. Since OBUs and RSUs are expected to receive periodic software and firmware updates, this threat is more likely to be carried out by a malicious insider than by an outsider. The associated motivation is ranked as moderate because it consists of a disruption in service. Since the threat is theoretically possible, the technical difficulty is a solvable one if countermeasures are not in place. The impact on a user is considered high due to the resulting long-lasting outages. As a result, the malware threat is ranked as major.

**Black Hole.** A black hole is formed by nodes that fail to propagate messages. Such an attack can only be carried out by malicious insiders, since network outsiders are not expected to repeat messages. The consequences of having a black hole in a network include dropped traffic messages, service requests and replies. With sufficient numbers of malicious nodes colluding to form a black hole past which no messages are propagated, it may be possible for attackers to partition a vehicular network in such a way that legitimate nodes never receive messages. If this scenario succeeds,
nodes may be prevented from receiving critical updates to their root certificate lists and CRLs, leaving them vulnerable to insider attacks from nodes using revoked or falsified certificates. With significant gains to be made from this attack, its technical difficulty solvable and its impact on a limited number of nodes, the threat is ranked as major.

**Spamming.** There is a risk in increased transmission latency due to the presence of spamming messages. The motivation for marketers to acquire a RSU for this purpose is best rated as moderate. On one hand, it is likely to be very lucrative, but on the other hand, the business is ultimately accountable to its customers who typically resent such a waste of their time and bandwidth. With the technical difficulty rated as low since a marketer is an insider, and with the impact on a user also low because it represents little more than an annoyance, the threat is ranked as minor.

### 4.1.2 Threats to Authenticity

Ensuring the authenticity of a vehicular network includes protecting legitimate nodes from malicious insiders and outsiders infiltrating the network under an assumed identity, thwarting the introduction of misinformation into the network, exposing spoofed GPS signals, and identifying attacks replaying legitimate interactions.

**Masquerading by Insider.** There is much to be gained by a malicious insider masquerading as an OBU or a RSU. By assuming a false identity, an attacker can create mischief with impunity, such as injecting false messages into the network and deceiving authorities into believing that another node was responsible. With PSOBUs possessing special privileges within the network and RSUs providing wireline access and LBS information, spoofing such nodes can be the first step in accessing personal information and possibly compromising privacy. Given the potential for financial and power-based gain, the motivation for this attack is ranked as high. If a malicious insider obtains access to a victim’s secret keys, the technical difficulty is low. The direct impact on the victim is of medium scope, since a single device is targeted, but the possession of a false identity allows a perpetrator to engage in attacks with far reaching consequences affecting multiple nodes. This high impact, combined with the likelihood of success for a malicious insider, renders this threat critical.
Masquerading by Outsider. The masquerading attack attempted by a network outsider is bound to fail when faced with the lack of access to a targeted device's cryptographic material. Without a victim’s secret key, an attacker cannot generate impersonated digital signatures or decrypt messages. With the strong technical difficulty in successfully carrying out this attack, the threat is ranked as minor.

Broadcast Tampering by Insider. It is possible that a malicious insider may attempt to inject false traffic safety messages into the network for the purpose of creating mischief, for example causing accidents by suppressing traffic warnings or manipulating the flow of traffic to clear a chosen route. Broadcast messages are meant for general consumption, but they need to be signed in order to deter attackers from generating false messages. With access to valid credentials, and thus the ability to generate verifiable digital signatures, the associated technical difficulty is low. Given the high impact of this threat and its potential to affect a large number of vehicles, the risk is assessed as critical.

Broadcast Tampering by Outsider. Without the means to digitally sign verifiable broadcast messages, an outsider faces strong technical difficulties in conducting this attack. This threat is thus ranked as minor.

Transaction Tampering by Insider. Another possible threat to message integrity consists of an attacker modifying the messages exchanged in V2I in order to falsify transaction application requests or forge the associated replies. With a little imagination, a malicious agent could create an entire alternate life or lifestyle for an unsuspecting user, resulting in potential financial or power-based gain. With access to a victim’s secret keys, an insider faces few technical difficulties in carrying out this attack. Since the scope of the exploit is limited to the victim, its impact is medium, and the threat is classed as major.

Transaction Tampering by Outsider. The same attack by an outsider cannot succeed without access to a target device’s secret keys. Strong technical difficulties render this a minor threat.

GPS Spoofing. By using a GPS satellite simulator to generate radio signals stronger than those received from a genuine GPS satellite, an attacker can lead nodes to believe they are in a different location than they actually are [143], potentially
causing collisions. Also, if GPS time is used to timestamp messages, a spoofing of the GPS clock could result in nodes accepting expired messages as new ones and could thus lead to a successful replay attack. Given the potential gains for an attacker, the solvable technical difficulties involved in this type of attack and the limited scope of its impact to a few nodes at most, the threat is ranked as major.

**Replay Attack.** Vehicular networks operating in the WAVE framework are protected from replay attacks by having each node maintain a cache of recently received messages against which new messages are compared. Messages older than a configurable time are discarded. The others are compared against this cache to ensure that they have not previously been received. This scheme assumes that an accurate source of time is available. The case where the clock accuracy is compromised is considered in the analysis on GPS spoofing. Therefore, despite the potential gains to be made in terms of network manipulation, there are strong technical difficulties in carrying out a replay attack, rendering this a minor threat.

### 4.1.3 Threats to Confidentiality

With the messages exchanged between the nodes of a vehicular network being accessible over the air, the threats to confidentiality include the collection of location information available through the transmission of broadcast messages, and the illegitimate collection of transaction information through eavesdropping. Broadcast messages generally pertain to traffic safety information and are therefore uninteresting for the purpose of eavesdropping.

**Location Tracking.** With the potential for vehicle locations to be constantly tracked, it is not difficult to imagine the temptation for attackers to exploit this new opportunity. By collecting an unsuspecting individual’s location trace over time, a malicious agent can use such information in stalking or building a potentially damaging profile of its victim. Every time an OBU propagates a broadcast message to alert other vehicles to traffic safety updates, it digitally signs the repeated message with its own certificate which can identify the OBU and its current position to the receiving nodes. Given an attacker’s gain in personal location information, the solvable technical difficulties associated with this threat and its tremendous impact on
the victim, it is ranked as critical.

**Transaction Eavesdropping by Insider.** In accumulating a series of transaction messages, a malicious insider can construct a profile of a given user by observing which services are used regularly, when, from which location and how much is spent. With the potential for gain and an attacker's access to a victim's secret cryptographic material, this exploit emerges as likely. Given that its scope is limited to the target node, the risk is assessed as major.

**Transaction Eavesdropping by Outsider.** If an attacker is a network outsider, the associated technical difficulty is quite strong, given the level of encryption required to protect transaction messages. The threat is therefore ranked as minor.

<table>
<thead>
<tr>
<th>Threat</th>
<th>Motivation</th>
<th>Difficulty</th>
<th>Likelihood</th>
<th>Impact</th>
<th>Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masquerading (I)</td>
<td>High</td>
<td>None</td>
<td>Likely</td>
<td>High</td>
<td>Critical</td>
</tr>
<tr>
<td>Location Tracking</td>
<td>High</td>
<td>Solvable</td>
<td>Likely</td>
<td>High</td>
<td>Critical</td>
</tr>
<tr>
<td>Broadcast Tampering (I)</td>
<td>Moderate</td>
<td>None</td>
<td>Likely</td>
<td>High</td>
<td>Critical</td>
</tr>
<tr>
<td>Malware</td>
<td>Moderate</td>
<td>Solvable</td>
<td>Possible</td>
<td>High</td>
<td>Major</td>
</tr>
<tr>
<td>Black Hole, GPS Spoofing</td>
<td>High</td>
<td>Solvable</td>
<td>Likely</td>
<td>Medium</td>
<td>Major</td>
</tr>
<tr>
<td>Transaction Tampering (I), Transaction Eavesdropping (I)</td>
<td>High</td>
<td>None</td>
<td>Likely</td>
<td>Medium</td>
<td>Major</td>
</tr>
<tr>
<td>DoS</td>
<td>Moderate</td>
<td>Solvable</td>
<td>Possible</td>
<td>Medium</td>
<td>Major</td>
</tr>
<tr>
<td>Masquerading (O)</td>
<td>High</td>
<td>Strong</td>
<td>Unlikely</td>
<td>High</td>
<td>Minor</td>
</tr>
<tr>
<td>Replay, Transaction Tampering (O), Transaction Eavesdropping (O)</td>
<td>High</td>
<td>Strong</td>
<td>Unlikely</td>
<td>Medium</td>
<td>Minor</td>
</tr>
<tr>
<td>Broadcast Tampering (O)</td>
<td>Moderate</td>
<td>Strong</td>
<td>Unlikely</td>
<td>High</td>
<td>Minor</td>
</tr>
<tr>
<td>Spamming</td>
<td>Moderate</td>
<td>None</td>
<td>Likely</td>
<td>Low</td>
<td>Minor</td>
</tr>
</tbody>
</table>

### 4.1.4 Countermeasures

We uncovered three critical security threats inherent in the WAVE architecture: masquerading and broadcast tampering by an insider, and location tracking. Several major threats surfaced as well: malware, black hole, GPS spoofing, DoS and transaction
tampering and eavesdropping by an insider. These problems need to be addressed before a WAVE implementation can be considered secure.

**Insider Attacks.** DSRC/WAVE is vulnerable to several critical and major threats from malicious insiders: masquerading as another device, eavesdropping on private transaction messages, and tampering with broadcast or transaction messages. All of these necessitate the appropriation of a valid public/private key pair, either at a targeted device or from the government transportation authorities filling the role of CA. The WAVE security standard suggests that a device’s cryptographic algorithms and secret keys be stored in a tamper-proof module. While access to this material at the device level may be difficult, it is not impossible. Equally likely is the eventuality of either a social engineering or a database hacking attack on transportation authorities in order to gain access to existing credentials.

**Location Tracking.** An important conundrum is posed by the very nature of vehicular networks: the balance of driver privacy versus accountability. On one hand, drivers have the right to expect a certain level of confidentiality with regards to their identity, location and application usage. The vehicular network must ultimately allow them to anonymously broadcast messages. On the other hand, it is imperative that malicious nodes be identified and neutralized in order to prevent serious harm to other vehicles. This means that nodes must be uniquely and persistently identified so that attackers can be flagged over a significant distance and period of time. There is currently no mechanism in WAVE to support anonymous broadcast messages. Much of the existing research into location privacy involves protecting users from LBS providers by blurring the user’s exact location in space and time. Such schemes are not applicable in vehicle safety applications such as platooning and collision avoidance systems, where each vehicle’s precise location is of paramount importance.

**Malware.** According to the security considerations outlined in the informative portion of the WAVE security standard, received software and firmware upgrades should only be permitted if they are sent and digitally signed by nodes possessing the required permission. In [86], a new security standard proposed by a large consortium of smart phone vendors is discussed. Although the details are sketchy, it appears that this standard will be based on the Trusted Platform Module (TPM) chip used in PCs
and laptops [139]. This chip protects a device from malware by providing authentication and cryptography services, allowing only trusted agents to selectively access the functions available on the device. This is in essence the same approach recommended (although not mandated) in the WAVE security standard. This countermeasure is considered to be state of the art for smart phones and other mobile devices. It may be equally effective in vehicular networks, provided that it is implemented. Otherwise, entry points for malware will need to be closed through future upgrades as they are discovered.

**Black Hole.** Broadcast messages in WAVE are propagated by flooding. All OBUs have the same range, so we may assume that links are symmetrical. It may therefore be feasible to detect a black hole by having the sending nodes listen to their neighbors’ retransmissions to ensure that they repeat messages. This concept is known as passive acknowledgement [71]. According to the DSRC/WAVE MAC layer standard [65], vehicular devices periodically send Nearby Station requests to ensure that none of their neighbors have a duplicate MAC address. If a given node sends a broadcast message and knows who its neighbors are, it can flag a neighboring node that fails to repeat the message. More conventional black hole prevention techniques, including multipath routing and the use of backup routes, are outlined in [1]. However, WAVE has a measure of resistance to black holes by design, due to the fact that it propagates messages by flooding without any route optimization mechanisms. With messages sent to every node within range, each of which repeats to all the nodes within its range, redundancy provides resilience to black holes. Alternately, it may be feasible to have RSUs repeat traffic messages as well. Although these infrastructure nodes are vulnerable to physical tampering, the non-responsiveness of known RSUs can raise an alarm.

**GPS Spoofing.** The WAVE security standard recommends the implementation of plausibility rules regarding changes in vehicle location, as well as the use of special calibration measures on the OBU clock so that updates to the time are performed by accelerating or decelerating the clock in a continuous manner, rather than in an abrupt, discrete fashion. Such rules offer a good basis for a countermeasure to the GPS spoofing threat. Similar plausibility-based countermeasures are offered in [143],
such as profiling satellite availability, locations and signal strength in order to detect unaccountable changes to the configuration. If an OBU knows that at a given set of coordinates it can usually receive signals from certain satellites, each with a specific signal strength, and it suddenly receives much stronger signals, there is a chance that a GPS satellite simulator has been set up nearby. A problem arises with this countermeasure when environmental conditions, such as the weather or ionosphere, affect the quality of the received signal. Changes in signal strength may be due to the presence of a GPS simulator or to natural phenomena. However, a countermeasure to the GPS spoofing threat already exists. It consists of using the encrypted Precise Positioning System (PPS) military signal rather than the Standard Positioning System (SPS) civilian signal. The state-of-the-art PPS-based GPS receiver known as Selective Availability Anti-Spoofing Module (SAASM) [103] uses a combination of symmetric and asymmetric encryption, where an asymmetric scheme is used to distribute a symmetric key, as outlined in [56]. As a result, if transportation authorities are granted permission from the U.S. military to use the SAASM receiver, as some other governmental agencies have been, vehicular networks can be well protected from GPS spoofing attacks.

**DoS.** DoS attacks in Internet applications and e-commerce can be mitigated by requiring sending nodes to perform a task such as solving a puzzle before receiving services from a server [72]. This idea ensures that legitimate nodes only have a simple computational chore to perform while attackers find themselves bogged down with the same task due to the sheer volume of messages sent. Ideally, DoS attacks should be mitigated at the lowest possible layer in the protocol stack. By providing link layer authentication, for example through the use of cryptographically generated MAC addresses, outsider attacks may be deterred. The concept of cryptographically generated IPv6 addresses, described in [10], may be applicable to MAC addresses in vehicular networks, although malicious insider nodes remain the greater threat. Further experiments need to be conducted to determine whether similar or other DoS prevention schemes can be applied to vehicular networks without adversely affecting transmission latency. It should be noted, however, that none of these countermeasures can prevent a wideband jammer from disrupting the transmitted signals. Such an
attack may be countered with the use of directional antennas which could allow vehicles to circumvent the jammed area. In addition, a vehicular network bogged down with a significant load of legitimate messages becomes more vulnerable to DoS attacks. A further measure to increase WAVE’s resilience to this threat would be the introduction of a routing protocol minimizing the number of broadcast messages.

4.2 Threat Analysis for WiMAX/802.16

Threats to WiMAX/802.16 security were uncovered in previous work [13], but we reassess them in [16] according to our adaptation of the ETSI methodology depicted in Table 2.2. With the potential for DSRC/WAVE and WiMAX/802.16 to share common infrastructure nodes, as proposed by Mar et al. [94] in the context of multimedia applications, we examine the risks common to both types of technology. The threat analyses of both DSRC/WAVE and WiMAX/802.16 [16, 79] reach similar conclusions that may be extended to wireless mobile networks in general.

We find that the DoS threat is consistently ranked as major due to the ensuing loss of service. But because the loss is temporary and the motivation behind an attack is for mischief rather than concrete gain, it is not critical.

Eavesdropping is a pervasive problem in wireless networks due to the open nature of the medium and the ease with which passive attacks can be perpetrated. But if an attacker is an outsider, security measures such as encryption, for example with DES or AES for WiMAX/802.16 and with ECIES for DSRC/WAVE, are sufficient to mitigate such attacks and render the risk minor. Without security measures, or if these are circumvented by a malicious insider, the threat is major.

Masquerading and point-to-point message tampering can also be addressed through appropriate security measures. DSRC/WAVE network members in possession of valid credentials digitally sign and encrypt the transaction messages they transmit, which ensures confidentiality, integrity and authentication. WiMAX/802.16 MSs can achieve authentication through X.509 certificates and the integrity of their messages can be ensured using message authentication codes or encryption. If these security measures are used and an attacker is an outsider, then both the masquerading and the message tampering threats are ranked as minor. However, if authentication measures
are not in place or if an insider can counter them, then the masquerading threat is a critical risk in a mobile network, while the risk of point-to-point message tampering is major due to its likelihood and the limited scope of its impact.

4.3 Summary

Location tracking and insider attacks are deemed the most serious threats for which few solutions currently exist in vehicular networks. Location tracking in DSRC/WAVE is facilitated by the use of PKI mechanisms, namely digital signatures, for securing broadcast messages. These messages, exchanged between vehicular safety applications, are sent unencrypted to minimize processing delays at the receiver, yet they must be digitally signed to ensure accountability and message integrity. In Chapter 5, we present a new mechanism to provide security equivalent to the DSRC/WAVE measures while protecting a driver’s privacy.

We deem insider attacks to be a critical threat in wireless mobile networks. The risk can be minimized by adequate security measures protecting cryptographic material both at the device and at the network authority levels. Strict access protocols for the protection of valid credentials must be implemented, including resilient tamper-proof modules at the device and adequate training and supervision of personnel at the CA. Beyond these basic measures, the imperative remains for insider attacks in mobile networks to be detected and attributed to their perpetrator, and for the attacker to be expelled from the network. We address the attack attribution issue with our solution for estimating the location of an uncooperative device in Chapter 6.
Chapter 5

Secure Anonymity in Vehicular Networks

Broadcast messages in vehicular networks epitomize the conflict between a crucial need for security and a reasonable expectation of driver privacy. Without reliable security measures, these potentially life-saving messages become exposed to outsider threats such as masquerading and tampering, as discussed in Chapter 4. However, the PKI mechanism mandated to secure vehicle safety messages, namely digital signatures, necessitates that a transmitting OBU's digital certificate accompany each message, so that a receiving device may authenticate the signature. This certificate uniquely identifies a sending vehicle and renders it trackable by an attacker.

Anonymity has been billed as a viable means of enhancing and protecting personal and informational privacy [93]. In [80], we describe an approach to ensure adequate security for vehicle safety broadcast messages while promoting driver anonymity, when warranted. As long as a vehicular device conforms to the expected network behaviour, it can enjoy anonymity from other nodes. If it engages in malicious conduct, its identity can be revealed by a network authority for the protection of honest nodes.

5.1 Secure Anonymous Broadcasting

Our SAB protocol offers an approach for anonymizing vehicle safety broadcast messages in a secure fashion. It proposes a hybrid key infrastructure which differentiates between the privacy requirements of different types of vehicular network devices.

In the DSRC/WAVE architecture, the encryption of transaction messages, whether signed or not, effectively hides a sender's identity. Consequently, there are few, if any, privacy concerns with transaction messages. Broadcast messages, however, pose a unique problem in that the sender's certificate, which is transmitted with every message for the receiver's verification, leaves the sending vehicle vulnerable to location tracking, as highlighted in Chapter 4. In addition, signed broadcast messages
in WAVE suffer from a size problem. They are predicted to contain at least 250 bytes, of which 125 bytes represent a signer’s certificate [67], in the best case scenario. Since a certificate chain may be included, up to five certificates of 125 bytes each may appear in a signed message for a size of nearly 750 bytes. It is clear that broad-spectrum PKI, as applied to every type of network device, expands the message size beyond the expected vehicular application requirements, outlined in Section 2.2.3, to the detriment of bandwidth usage and message latency. In SAB, we propose to narrow the use of PKI to the devices that must remain identifiable and that carry the onus of trust. We propose an alternate scheme for devices that should remain anonymous.

**Notation.** The protocol description makes use of the BAN logic notation [31]. Encryption of message $M$ using symmetric key $K$ is represented as $\{M\}_K$. Asymmetric encryption public and private key pairs take the form $\{K, K^{-1}\}$, where $K$ is the public key and $K^{-1}$ the private one. Therefore, if device $A$’s public/private key pair is denoted as $\{A, A^{-1}\}$ and the device communicates message $M$ to device $B$, the encrypted message takes the form $\{M\}_B$ where $M$ is encrypted with $B$’s public key. A digitally signed message is represented as $\{M\}_A^{-1}$ where $M$ is signed with $A$’s private key, and a signed and encrypted message is depicted as $\{\{M\}_A^{-1}\}_B$. As well, we denote the output of cryptographic keyed hash functions, such as keyed-Hash Message Authentication Code (HMAC), over message $M$ using key $K$ as $H_K(M)$.

### 5.1.1 Hybrid Key Infrastructure

RSUs and PSOBUs, as trusted devices, occupy a unique position of trust in a vehicular network, and as such cannot remain anonymous. As a result, they retain the use of PKI digital signatures on the broadcast messages they originate.

On the other hand, OBUs such as vehicles owned by private citizens and commercial organizations have a right to preserve their location privacy. Since these vehicles do not enjoy a position of trust within the network, they can remain anonymous, as long as they can be held accountable for the authenticity of the messages they originate, and a device can be revoked in case of an attack. We propose that anonymous devices be provided with a shared network authorization key $AK$ known to all
authorized members of the network. The AK grants the privilege to send broadcast messages. Each broadcast message $M$ includes a HMAC digest, computed using $AK$ by the sender and denoted as $H_{AK}(M)$.

In order to enable the revocation of rogues, a sending device identifier is included with each broadcast message. To maintain anonymity, this identifier must remain illegible for all nodes other than a trusted authority responsible for revocation, such as a CA. Device $A$ includes the following OBU identifier with each broadcast message:

$$\{Id_A, H_{SKA}(Id_A|H_{AK}(M))\}_{CA}$$

A unique identifier, such as an Electronic License Plate [63] and denoted by $Id_A$, is associated with each device and can be traced back to its logical identity by the CA. This identifier and a copy of the message digest $H_{AK}(M)$ are concatenated and further hashed with a secret key $SK_A$ known only to the sending device $A$ and the CA. This measure prevents an attacker from masquerading as another device or replaying another device’s identifier from a previous message. The resulting hash output and $Id_A$ are encrypted with the CA’s public key so that the sending device’s identity is kept secret from all nodes but the CA, which is tasked with revoking rogue devices. Semantically secure encryption, which means that a plaintext encrypted multiple times with the same key will produce different ciphertexts [58], ensures that every identifier generated by one device, even for the same message, is different.

5.1.2 Key Management

Transportation authorities are typically established at the state or provincial level of government. In our model, the task of managing network authorization keys and transaction message digital certificates remains with the same agencies, using multiple server-based CAs, or their delegates, communicating with each other over an infrastructure network. The contiguous area $R$ encompassed by a vehicular network is partitioned into a set of $n$ distinct regions or jurisdictions, such that $R = \{R_1, R_2, ..., R_n\}$, where $R_i \cap R_j = \emptyset$ for all $1 \leq i, j \leq n$ and $i \neq j$. Jurisdictions may map to a state or province, or to a county or municipality. A local CA, denoted as $CA_\ell$ where $\ell \in R$, is assigned to each jurisdiction. Because the use of a global network authorization key
is unscalable, every $CA_\ell$ manages a local authorization key $AK_\ell$ which is valid only in jurisdiction $\ell$. Each device is assigned to a home CA, denoted as $CA_h$ with $h \in R$, which stores its certificate information.

**Enrollment.** In order to propagate vehicle safety broadcast messages, trusted devices require a digital certificate, and anonymous devices need a network authorization key. While certificates are obtained manually through transportation authorities, the authorization key is requested and renewed automatically. Each anonymous device $A$ is assigned a unique identifier $Id_A$ by its $CA_h$, as well as a secret key $SK_A$ known only to the device and its $CA_h$. The $Id_A$ serves to identify broadcast message senders to their $CA_h$, and the secret key $SK_A$ is used to prove this identity.

**Renewal and Jurisdiction Handoff.** Whenever a local authorization key $AK_\ell$ expires, device $A$ must petition the $CA_\ell$ for a new authorization key $AK'_\ell$ using a transaction message. Since transaction requests and replies are exchanged over service channels rather than broadcast control channels [68], they do not impact the bandwidth available for broadcast messages. Similarly, if $A$ enters a new jurisdiction $\ell'$, it must forward its request to the new local $CA_{\ell'}$ to obtain its authorization key $AK_{\ell'}$. In either case, the local $CA_\ell$ obtains authorization from $A$'s home $CA_h$ before complying, to ensure that $A$ is still authorized to broadcast. A set of concurrent network authorization keys is used to ensure that key renewal requests are staggered in order not to overwhelm the network upon authorization key expiry. It should be noted that authorization key renewals depend on the device being within RSU range. For this reason, there must be a built-in tolerance in remote areas for broadcast messages hashed with expired authorization keys. As well, an optimal network authorization key renewal interval must be devised to maximize protection against attacks while minimizing the associated overhead.

**Revocation.** Certificate revocation for trusted devices is implemented through CRLs, as currently mandated by DSRC/WAVE. While CRLs are of questionable scalability when relied upon for revocation of all nodes, the restriction of their use to trusted devices can greatly alleviate the size of these lists and the issues surrounding the scalability of their distribution. Since trusted devices are either fixed or typically have their mobility patterns restricted to one or a few adjoining jurisdictions,
the required frequency and scope of CRL distribution is naturally limited. Network authorization key revocation is achieved primarily through expiry and renewal, with only authorized devices in good standing allowed access to a renewed key. In the event that an attack is detected and device $A$ must be revoked, the $CA_\ell$ receives the attack message and decrypts the associated OBU identifier. It then forwards both the message and identifier to the device’s $CA_h$ for revocation. The $CA_h$ retrieves the $SK_A$ associated with $Id_A$ and computes the hash value of $Id_A$ concatenated with the message digest. If this value matches the one in the decrypted OBU identifier, then the rogue device is identified as $A$, flagged in $CA_h$’s database as a rogue and denied further access to network authorization keys in all jurisdictions. If the hash values don’t match, then an insider attack has been detected, where a rogue with a valid authorization key has appropriated an identifier from a different device or message. In this situation, alternate means of attributing the attack message to its source must be employed, as explored in Chapter 6.

5.1.3 Safety Message Broadcast

A trusted device $S$ transmitting a broadcast message $M$ signs it with its private key $S^{-1}$ and transmits:

$$\{M, S, \{M\}_{S^{-1}}\}$$

The recipient verifies the signature using the unsigned message, $S$’s public key and the signature. It must also verify the signature on the sender’s certificate, the signature on the certificate signer’s certificate, and so on, until every certificate in the chain is verified and compared against CRLs to ensure its validity. If all signatures verify and the certificates are deemed valid, then the message is processed. Otherwise, it is discarded, as outlined in [67].

An anonymous device $A$ proves its authorization to transmit a broadcast message $M$ by hashing it with the current network authorization key to produce a message digest $H_{AK}(M)$. It then transmits the plaintext message, the message digest, its $CA_h$ identifier and its OBU identifier as described in Section 5.1.1:

$$\{M, H_{AK}(M), CA_h, OBUid\}$$
where
\[ OBUid = \{Id_A, H_{SK_A}(Id_A|H_{AK}(M))\}_{CA_e} \]
Each recipient verifies that the message digest was generated with the currently valid AK. If it was, the recipient processes the message. Otherwise, the message is forwarded as an alert to the local CA_e.

5.1.4 SAB Analysis

We gauge the security of the SAB protocol in terms of its effect on the risks uncovered in the DSRC/WAVE threat analysis put forth in Chapter 4. We compare the performance of SAB with WAVE in terms of message size and cryptographic operation execution time required for sending and receiving messages.

**Threat Analysis**

Of the threats to DSRC/WAVE identified in Chapter 4, five are affected by the proposed SAB protocol: location tracking and masquerading and broadcast tampering by an insider, which are critical threats, and masquerading and broadcast tampering by an outsider, which are minor threats.

**Location Tracking.** In adopting a shared network authorization key to prove membership in a vehicular network, SAB provides devices with anonymity since no visible identifier is associated with each node. Anonymous non-repudiation is further ensured with an identifier encrypted with a CA’s public key so that no other device can identify a message’s sender. While the attacker motivation and the impact on a victim associated with this threat are still ranked as high, the strong technical difficulties introduced by SAB downgrade this threat to a minor risk.

**Masquerading by Insider.** A malicious insider has access to a victim’s secret key and may spoof its OBU identifier at will. The technical difficulties involved remain negligible, and the risk level is unchanged at critical.

**Broadcast Tampering by Insider.** With the SAB and WAVE protocols, a sender’s authorization to broadcast a message is verified by every recipient. It is at least as difficult for an attacker to fraudulently obtain a SAB network authorization key as it is to acquire a digital certificate. SAB authorization keys have short lifetimes.
which curtail an attacker's window of opportunity. Devices petitioning for authorization key renewal are verified dynamically at their home CA to ensure that they are still authorized members of the vehicular network. On the other hand, WAVE digital certificates are assumed to be valid until they appear on CRLs, which may be outdated and suffer from scalable distribution issues. As a result, since SAB dynamically verifies the authorization of a petitioner directly at its home CA, a malicious insider may only conduct a single attack before its digital certificate is flagged at its CA and it is no longer allowed access to network authorization keys. With WAVE, the delay between the detection of an attack and the distribution of the CRL invalidating the attacker's digital certificate may enable the perpetrator to carry out additional exploits until the updated CRL reaches all the devices within the attacker's range. This time lag may be exacerbated if the malicious node is using a digital certificate originating from a remote jurisdiction. While the threat of broadcast tampering by an insider remains a critical risk, the scope of its impact may be further contained by the security measures put in place with SAB.

**Masquerading by Outsider.** The technical difficulties encountered by an outsider attempting to impersonate another device are considerable. Given that device identifiers are transmitted in encrypted form and are only known to a sending device and its home CA, an attacker cannot knowingly target a given identifier but may generate random identifiers and happen upon a valid one by chance. In this case, however, the attacker also needs the secret key known only to the device associated with the identifier and its home CA in order to generate a valid hash value of the identifier and message digest. As a result of these strong technical difficulties, the threat of masquerading by an outsider remains a minor one.

**Broadcast Tampering by Outsider.** A vehicular device without valid credentials is unable to obtain a network authorization key, and a message hashed without this key can be detected by every receiving device. Given the technical hurdles, the risk of outsider attacks remains minor.
Message Size

Vehicle safety applications indicate a limit of 100 bytes for each broadcast safety message, as outlined in Section 2.2.3. However, it is clear that signed WAVE messages fall short of meeting this criterion with a size of 250 bytes in the best case scenario. In Table 5.1, the components of WAVE and SAB broadcast messages are broken down and compared.

The message depicted in Table 5.1 represents a 32-byte position report, or beacon message, that a device is periodically required to transmit to convey its coordinates. The beacon is encapsulated in a WAVE plaintext message data structure [67], for a total of 67 bytes. Table 5.1 depicts the size of the WAVE broadcast message, including the plaintext (67 bytes), the signer’s certificate (125 bytes, although a certificate chain may be included of up to five certificates) and the digital signature (64 bytes), for a total over 256 bytes.

The corresponding 67-byte plaintext message in SAB would also be transmitted, just as with WAVE, but with a SHA-1 HMAC digest of the message (20 bytes), the sender’s home CA identifier (2 bytes), and an OBU identifier consisting of the sender’s unique identifier (8 bytes) and the hashed value of the identifier and message digest (20 bytes), as outlined in Section 5.1.1. The ECIES encryption of the OBU identifier adds a significant amount of overhead, depending on the elliptic curve chosen. Tests were conducted using the Crypto++ cryptographic functions library [43] with two of the elliptic curves specified by NIST [101]. Regardless of the size of the plaintext, the P-224 elliptic curve adds 77 bytes of overhead, and the P-256 curve adds 85 bytes. While WAVE mandates the use of the P-256 curve for encrypting transaction messages, the resulting overhead is a strong deterrent to using the same curve for encrypting OBU identifiers in broadcast messages. For this reason, either the P-224 curve can be used or another secure asymmetric encryption mechanism can be investigated in order to minimize the message overhead. Table 5.1 assumes the use of the P-224 curve. As a result, the total message size for SAB is 194 bytes.

Although SAB broadcast messages also fall short of the required 100 bytes, the total overhead included in each broadcast message, i.e. the total message size minus the payload, is 127 bytes with SAB, compared to a minimum of 189 with WAVE.
SAB thus achieves an improvement of at least 33% over WAVE in terms of bandwidth consumption and associated latency. SAB has an additional advantage over WAVE in that without the use of a certificate chain, the message overhead size is fixed.

<table>
<thead>
<tr>
<th>Data Structure</th>
<th>WAVE</th>
<th>SAB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsigned Message</td>
<td>67</td>
<td>67</td>
</tr>
<tr>
<td>Certificate Chain</td>
<td>125 – 625</td>
<td>–</td>
</tr>
<tr>
<td>Signature</td>
<td>64</td>
<td>–</td>
</tr>
<tr>
<td>HMAC</td>
<td>–</td>
<td>20</td>
</tr>
<tr>
<td>CA identifier</td>
<td>–</td>
<td>2</td>
</tr>
<tr>
<td>OBU Identifier</td>
<td>–</td>
<td>28</td>
</tr>
<tr>
<td>Encryption Overhead</td>
<td>–</td>
<td>77</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>&gt; 256</td>
<td>194</td>
</tr>
</tbody>
</table>

Cryptographic Operations

We examine the computational cost of cryptographic operations required for sending and receiving SAB broadcast messages, compared to WAVE messages. WAVE mandates the use of ECDSA for signature generation and verification based on the NIST elliptic curves P-224 or P-256 and ECIES for encryption.

Benchmark tests were conducted using the Crypto++ library on a Pentium 1.60 GHz processor, with the parameter values for the P-224 and P-256 curves as specified by NIST [101]. A message base of one thousand 67-byte messages was constructed, and each message was signed and its signature subsequently verified using the ECDSA algorithm for both the P-224 and P-256 curves. Further, one thousand 28-byte identifiers were encrypted with the ECIES algorithm using the same elliptic curves.

In addition to the standard implementation of the cryptographic operations, Crypto++ allows for the use of a pre-computed table of powers of fixed bases in order to speed up exponentiation. Since WAVE devices are not expected to be highly constrained in terms of memory, the use of pre-computed information appears to be a favorable option given its positive effect on performance. However, it was noted during the benchmark tests that the use of pre-computation had a negative impact on the variance of the results. Signature generation results using pre-computation yielded a
standard deviation between 70-95% of the mean signature computation time. With the P-224 curve, for example, the mean time for signature generation was 8.20 ms with a standard deviation of 7.81 ms. Without the use of pre-computation, signature generation with the P-224 curve took on average 15.73 ms with a standard deviation of 2.11 ms, which represents only 13% of the mean. Similar observations were gathered with respect to signature verification and encryption, albeit to a lesser extent. It is unclear whether this considerable variability is due to the Crypto++ specific implementation or to the nature of elliptic curve cryptographic operations in general. The goal of our benchmark tests was to compare the cryptographic operation cost between two protocols and not to identify the conditions for optimal performance. We therefore illustrate the results obtained with the P-224 and P-256 curves without the use of pre-computation because of the significant reduction in variability. It should be noted, however, that measures for performance improvement may be achieved through the use of pre-computed tables, or through optimizations such as those outlined by Brown et al. [29]. Table 5.2 illustrates the comparison results.

Table 5.2: SAB vs. WAVE Execution Time (in ms)

<table>
<thead>
<tr>
<th></th>
<th>Sending Device</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>WAVE P-224</td>
<td>P-256</td>
</tr>
<tr>
<td><strong>Sign</strong></td>
<td></td>
<td>15.73</td>
<td>18.53</td>
</tr>
<tr>
<td><strong>Encrypt</strong></td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>HMAC</strong></td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>15.73</strong></td>
<td><strong>18.53</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Receiving Device</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WAVE P-224</td>
<td>P-256</td>
<td>SAB</td>
</tr>
<tr>
<td><strong>Verify</strong></td>
<td>31.08 × 2</td>
<td>42.08 × 2</td>
<td>–</td>
</tr>
<tr>
<td><strong>HMAC</strong></td>
<td>–</td>
<td>–</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>62.16</strong></td>
<td><strong>84.16</strong></td>
<td><strong>0</strong></td>
</tr>
</tbody>
</table>

For the sending device, WAVE only requires a signature to be generated. On average, this can be accomplished with the P-224 curve in 15.73 ms and with P-256 in 18.53 ms. In contrast, SAB requires the sending device to generate a HMAC message digest, as well as an encryption of the OBU identifier. Since our tests revealed that
HMACs can be generated on one kilobyte of data in 17 microseconds, we deemed the HMAC operation cost to be negligible in comparison to the other operations measured in milliseconds. The ECIES encryption of 28-byte OBU identifiers took 31.05 ms on average with the P-224 curve and 36.56 ms with P-256. Overall, a sending device takes roughly twice as long to process an outgoing message with SAB than with WAVE.

Upon receiving a message, WAVE devices must verify at least two signatures, one on the message and one on the sender’s certificate. If a certificate chain is included, then every certificate in the chain must be verified, for a total of up to six signature verifications. With the P-224 curve, two verifications averaged 62.16 ms, while with the P-256 curve, they took 84.16 ms. By comparison, SAB only requires a hash value to be generated and compared against the corresponding received message digest, although if concurrent network authorization keys are used to stagger key renewals, several hash operations may be required. Decryption of an OBU identifier by a CA is required only in the event that the sender is a suspected rogue and not as a matter of course. As a result, the computational cost in SAB receiving devices is negligible.

The most frequently transmitted broadcast messages are likely to be position beacons. In a given time period, a device receives as many beacons as it has neighbors \( n \), while it only sends one. The cost of sending and receiving messages with WAVE is represented by two constants, \( s \) and \( r \) respectively. The cost of sending a SAB message is twice the one in WAVE (\( 2s \)), while its receiving cost is null. SAB is the better solution if the computational performance in sending one message (\( s \) for WAVE and \( 2s \) for SAB) and receiving as many messages as there are neighbors (\( n \times r \) for WAVE and 0 for SAB) is more efficient in SAB than in WAVE:

\[
\begin{align*}
\text{SAB} & \quad \text{WAVE} \\
(n \times 0) + (1 \times 2s) & < (n \times r) + (1 \times s) \\
s & < n \times r \\
\frac{s}{r} & < n
\end{align*}
\]

Since we know from our tests that \( s < r \), we know that \( \frac{s}{r} < 1 \). As a result, SAB is more efficient if \( n \geq 1 \), where a device has at least one other device in range. So
in very sparse traffic, if a device has no neighbors, WAVE’s performance is better. If, however, the device has one or more neighbors necessitating signature verifications on several messages, which is the more likely scenario especially in urban settings, SAB is computationally superior. In general, given the open nature of the wireless medium, nodes inherently receive more messages than they send.

5.2 Summary

The SAB protocol addresses the location tracking threat in vehicular networks by reducing its risk from critical with the use of WAVE to minor, without negatively affecting the risk levels posed by other threats. We show that SAB achieves a 33% reduction in overhead for broadcast messages over the WAVE protocol. We also demonstrate that SAB employs fewer computationally expensive cryptographic operations overall, most notably in the case where a device has at least one other node within range.

Although both WAVE and SAB effectively protect a vehicular network against outsiders, few mechanisms exist in either protocol for the mitigation of masquerading or tampered broadcast messages by rogue devices using valid digital certificates. With such credentials, a malicious insider can generate verifiable digital signatures in WAVE and obtain network authorization keys in SAB. The only current countermeasure for this case is the finite lifetime of the digital certificate itself and the presumed difficulty for an attacker to continue to obtain valid certificates.

Further, insider attacks may be complicated in SAB if a perpetrator attempts to avoid revocation by improperly formatting the OBU identifier included in an attack message, or by including a false identifier. One method to counter this attempt at evasion with SAB is to augment the role of a network authority to intermittently monitor the OBU identifiers of broadcast messages. This practice may provide a rudimentary means of flagging the presence of a malicious insider by the invalid OBU identifier on the messages it transmits. The possibility of such detection may therefore prove to be an additional deterrent to the generation of tampered messages in SAB, given that WAVE provides no mechanism to detect malicious insiders with valid credentials.
Another method to counter a malicious agent seeking to evade retribution consists of localizing the source of an attack message and attempting to link its physical location in Euclidian space with a logical identity, such the digital certificate used to obtain a network authorization key. If an insider’s assumed identity can be flagged at the CA, further attacks can be mitigated, and neighboring devices can be alerted to the presence and position of the perpetrator. Given that other mobile technologies, for example WiMAX/802.16 wireless access networks, are also vulnerable to insider attacks, they can similarly benefit from mechanisms to locate a malicious insider. In Chapter 6, we address the problem of attack message attribution by localizing the signal’s originator.
Part III

Localization of Uncooperative Nodes
Chapter 6

Hyperbolic Position Bounding in Wireless Networks

We introduce the Hyperbolic Position Bounding (HPB) mechanism to compute a probabilistic candidate area in Euclidian space for the location of a target transmitter [82]. Because we assume an uncooperative transmitting device, our localization efforts must rely on information that is inadvertently revealed by the target node, in this case the Received Signal Strength (RSS) of a received message.

In our adversary model, the target transmitter is a malicious insider or otherwise uncooperative device broadcasting a single message. The receivers are trusted infrastructure nodes, such as DSRC/WAVE RSUs or WiMAX/802.16 BSs. They are necessarily situated within radio range of the transmitter, and have globally known coordinates. The receiving devices also possess the means to communicate with each other over a secure channel, in order to aggregate RSS measurements and bound the transmitter position in a locally centralized fashion. We presume that the transmitter emits its target message with an omnidirectional antenna for a uniform propagation pattern. Since an attacker may seek to obfuscate its position, we assume that the power at which the message is transmitted, the Effective Isotropic Radiated Power (EIRP), is unknown. Our localization efforts also make no presumption of assistance from the transmitting device.

Previous research has been conducted where an RSS value measured at a receiver is used to estimate a probable range of Transmitter-Receiver (T-R) distances, as outlined in Chapter 3. We extend this concept with our HPB method to compute a probabilistic range of distance differences between a target transmitter and a pair of receivers. This approach enables us to effectively bound the transmitter position in the case where the EIRP of a received message is unknown.

The HPB algorithm uses a probabilistic path loss model and the RSS values of a target message to predict a range of distance differences between a transmitter and a
receiver pair, with a degree of confidence. Hyperbolas between each receiver pair are computed at the minimum and maximum bounds of this distance difference range. The area situated between the corresponding minimum and maximum hyperbolas bounds the position of the transmitter, with the given confidence level.

6.1 Estimating Distance From Signal Strength

We outline a large scale path loss model and describe how a probable distance range, i.e. the minimum and maximum distances, between a transmitter and a receiver can be computed from a measured RSS value. We relate how existing work employs these minimum and maximum distances to construct an annulus around the receiver that probabilistically contains the location of the transmitter. We extend this research to compute a T-R distance range when an unknown EIRP is used.

6.1.1 The Log-Normal Shadowing Model

The log-normal shadowing model predicts a radio signal's large scale propagation attenuation, or path loss, as it travels over a known T-R distance [115]. The variations in signal strength experienced in a particular propagation environment, also known as the signal shadowing, behave as a Gaussian random variable with mean zero and a standard deviation obtained from experimental measurements.

Definition 6.1. The log-normal shadowing model defines the path loss $L(d)$ (in dB) of a signal at T-R distance $d$ from a transmitter as:

$$L(d) = \bar{L}(d_0) + 10\eta\log\left(\frac{d}{d_0}\right) + X_\sigma$$

where $d_0$ is a pre-defined reference distance close to the transmitter, $\bar{L}(d_0)$ is the average path loss at the reference distance, and $\eta$ is a path loss exponent dependent upon the propagation environment. The signal shadowing is represented by a Gaussian random variable $X_\sigma$ with zero mean and standard deviation $\sigma$.

Values for $\eta$ and $\sigma$ can be obtained from experimental measurements, for example in [89] and [47], where linear regression techniques are used to ascertain parameters $\eta$ and $\sigma$ from actual path loss readings.
HPB is a probabilistic mechanism with roots in the log-normal shadowing model. The localization granularity in estimating the position of a transmitter is thus contingent upon the amount of signal shadowing taken into account. Accounting for greater amounts of shadowing translates into a higher degree of confidence in the computed candidate area for the transmitter’s position.

**Definition 6.2.** We define the confidence level \( C \), with \( 0 \leq C < 1 \), as the probability that the random Gaussian signal shadowing \( X_\sigma \) put forth in Definition 6.1 lies in the interval \([-z \sigma dB, +z \sigma dB]\), where \( z = \Phi^{-1}(\frac{1+C}{2}) \), with \( \Phi(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{x^2}{2\sigma^2} \right] dx \) The constant \( z \) can be obtained from a Normal distribution table.

From Definitions 6.1 and 6.2, we can specify the probabilistic path loss more precisely.

**Lemma 6.1.** The path loss \( L(d) \) of a signal at distance \( d \) from a transmitter is defined with a confidence level \( C \) as:

\[
L(d) = \bar{L}(d_0) + 10\eta \log\left(\frac{d}{d_0}\right) \pm z \sigma
\]

where \( z = \Phi^{-1}(\frac{1+C}{2}) \).

**Proof.** This can be derived directly from Definitions 6.1 and 6.2.

**Example.** Figure 6.1 illustrates an example of the probability distribution of path loss at a distance of \( d = 100 \) m from a transmitter. In the 2.4 GHz frequency band, the average free space path loss measured at \( d_0 = 1 \) m equals 40 dB. For a path loss exponent of \( \eta = 2.76 \), we obtain an average loss of 95 dB at 100 meters. With a standard deviation \( \sigma = 5.62 \), the shaded area in Figure 6.1 depicts 95% of the probability distribution around the average path loss. The signal shadowing lies in the interval \([-1.96 \times 5.62 dB, +1.96 \times 5.62 dB] = [-11 dB, +11 dB]\) with probability 0.95, and so \( L(100 \text{m}) \) is contained in the interval \([84 dB, 106 dB]\) with the same probability.

### 6.1.2 Distance Range with Log-Normal Shadowing

In [17], Barbeau and Robert demonstrate that the minimum and maximum distances from a transmitter to a receiver can be calculated from the path loss using the log-normal shadowing model.
Lemma 6.2. The minimum and maximum distances \((d^- \text{ and } d^+ \text{ respectively})\) from a transmitter to a receiver can be computed as a function of path loss, with confidence level \(C\), as:

\[
d^- = d_0 \times 10^{\frac{L(d)-L(d_0)-2\sigma}{10\eta}}
\]

\[
d^+ = d_0 \times 10^{\frac{L(d)-L(d_0)+2\sigma}{10\eta}}
\]

Proof. The proof is based on Lemma 6.1 and can be found in [17]. \(\Box\)

Two circles may be constructed around a receiver, one at a radius equal to the minimum distance \(d^-\), and another at the maximum distance \(d^+\). These circles thus form an annulus, or ring, around the receiver, containing the position of the transmitter with probability \(C\) [17].

In our scenarios, each receiver \(R_k\) obtains a RSS measurement \(RSS_k\), but the path loss \(L(d)\) value required to compute the distance range using the equations in Lemma 6.2 is not readily available. We can thus replace the path loss with its equivalent, based on the transmitting power (EIRP) and RSS.

Fact 6.1. The path loss \(L(d)\) (in dB) at T-R distance \(d\) can be stated in terms of the EIRP and RSS (both in dBm) at receiver \(R_k\) as follows:

\[
L(d) = EIRP - RSS_k
\]
A rogue may transmit at various EIRP levels in order to mask its location. In our solution, we address this issue by assuming an unknown value for the EIRP. We thus update Barbeau and Robert's minimum and maximum distance equations for a range of EIRP values, bounded by a minimum and a maximum EIRP, denoted as $\mathcal{P}^-$ and $\mathcal{P}^+$ respectively.

**Lemma 6.3.** The minimum and maximum distances, $d^k_-$ and $d^k_+$, between transmitter $T$ and receiver $R_k$, sent within an estimated EIRP interval $[\mathcal{P}^-, \mathcal{P}^+]$ can be computed, with confidence level $C$, as:

\begin{align*}
d^k_- &= d_0 \times 10^{\frac{p^- - RSS_k - L(d_0) - z\sigma}{10\eta}} \quad (6.1) \\
d^k_+ &= d_0 \times 10^{\frac{p^+ - RSS_k - L(d_0) + z\sigma}{10\eta}} \quad (6.2)
\end{align*}

Alternately, the probability that the distance $d_k$ between $T$ and $R_k$ is situated in the interval $[d^k_-, d^k_+]$ is $C$:

$$Pr[d^k_- \leq d_k \leq d^k_+] = C$$

**Proof.**

1. For a single EIRP value $\mathcal{P}$, Lemma 6.2 and Fact 6.1 can be combined to show that $d^k_-$ and $d^k_+$ are the minimal and maximal distances respectively.

2. For a range of EIRP values $[\mathcal{P}^-, \mathcal{P}^+]$, let $D_k(\mathcal{P}, \nu)$ represent the distance between a transmitter $T$ and receiver $R_k$, if the EIRP is $\mathcal{P}$ and the amount of signal variation within the shadowing interval $[-z\sigma, +z\sigma]$ is $\nu$. Therefore, $D_k(\mathcal{P}, \nu) = d_0 \times 10^{\frac{p^- - RSS_k - L(d_0) - \nu}{10\eta}}$.

Four possible distance boundaries between $T$ and $R_k$ can be computed using combinations of EIRP and shadowing interval bounds:

- (i) $D_k(\mathcal{P}^-, -z\sigma)$
- (ii) $D_k(\mathcal{P}^-, +z\sigma)$
- (iii) $D_k(\mathcal{P}^+, -z\sigma)$
- (iv) $D_k(\mathcal{P}^+, +z\sigma)$
These points form a lattice with an infimum, and so a minimum distance $d_k^-$, of $D_k(\mathcal{P}^-, -z\sigma)$, since $\mathcal{P}^-$ is the minimum EIRP and $-z\sigma$ is the minimum shadowing bound. Its supremum, and thus the maximum distance $d_k^+$, is the value $D_k(\mathcal{P}^+, +z\sigma)$, because $\mathcal{P}^+$ is the maximum EIRP and $+z\sigma$ is the maximum shadowing bound.

An additional observation may be gleaned from the results uncovered in Lemma 6.3.

**Lemma 6.4.** For a given EIRP value $\mathcal{P}$, the minimum and maximum distances between a transmitter and a receiver are bounded solely by the signal shadowing range $[-z\sigma, +z\sigma]$, with confidence level $C$.

*Proof.* With a constant EIRP value $\mathcal{P} = \mathcal{P}^- = \mathcal{P}^+$, the proof can be directly inferred from Lemma 6.3, since $-z\sigma$ and $+z\sigma$ are the lower and upper bounds respectively of the signal shadowing range. □

### 6.2 Estimating Location From Distance

Minimum and maximum distances between a transmitter and a receiver have been used, for example in [17], [91] and [92], to construct a pair of circles forming an annulus around the receiver within which the transmitter may be located. Multiple annuli may be computed around several receivers, and the location of a transmitter can be estimated within the annuli intersection area. However, this approach is more successful when the difference between the minimum and maximum distances is not significant. If it is, the annuli may be so wide that their intersection is too large to effectively locate the transmitter, even if multiple receivers are considered. With our assumption of an unknown EIRP, and thus a range of EIRP values for computing the minimum and maximum T-R distances, we cannot use the annuli method to localize an uncooperative transmitter, as exemplified in Chapter 8.

Our HPB approach thus relies on the relative distance difference between a pair of receivers and a transmitter, similar to the TDOA technique. In TDOA, a hyperbola is constructed with two points of known coordinates at the foci. The properties of hyperbolas are such that every point on the hyperbola is at the same distance difference.
of the two foci. For example, if the difference in distances from a transmitter $T$ to two receivers $A$ and $B$ is known, the corresponding hyperbola $\mathcal{H}_{A,B}$ can be plotted, as shown in Figure 6.2. The transmitter must necessarily lie on the hyperbola between $A$ and $B$. If a second distance difference is known, for example between receivers $B$ and $C$, a second hyperbola can be plotted, and the location of the transmitter $T$ is discovered at the intersection of the two hyperbolas.

![Figure 6.2: Example of TDOA](image)

In our scenarios, neither the transmitter EIRP nor the exact amount of signal shadowing that receivers experience are known. Therefore, we cannot determine the precise distance difference between a pair of receivers. Instead, our HPB algorithm uses a TDOA-based technique combined with the estimated minimum and maximum distances between a transmitter and receivers. We can thus define a candidate area $\mathcal{A}_{i,j}$ bounded by two hyperbolas between a pair of receivers $R_i$ and $R_j$: one hyperbola at the minimum bound of the distance difference range and another at the maximum bound.

### 6.2.1 Computing a Distance Difference Range

We use the minimum and maximum distance equations, as defined in Lemma 6.3, to compute the minimum and maximum bounds of the range of distance differences between a pair of receivers $R_i$ and $R_j$. Intuitively, the minimum bound of this range from $R_i$'s perspective is the difference between the distances closest to $R_i$ yet farthest
from $R_j$, at the minimum transmitting power. In turn, the maximum bound of the range from $R_i$'s perspective is the difference between the distances farthest from $R_i$ yet closest to $R_j$, at the maximum EIRP.

**Example.** Let us illustrate how the minimum and maximum distances between a transmitter and two receivers, with a possible range of EIRP, can be combined to obtain the minimum and maximum bounds of a corresponding distance difference range. Let two receivers $R_1$ and $R_2$ receive a single message from transmitter $T$, measuring RSS values $RSS_1 = -54$ dBm and $RSS_2 = -53$ dBm. Assuming a frequency of 2.4 GHz, let an EIRP range $[P^-, P^+]$ equal $[15, 23]$ dBm, with $\eta = 1.67$, $\sigma = 3.49$, $d_0 = 1$ m, and $L(d_0) = 40$ dB. For a confidence level of $C = 0.95$, we use $z = 1.96$. Additionally, we adopt the notation of Lemma 6.3, where $D_k(P, V) = d_0 \times 10^{\frac{P-RSS_k-L(d_0)+V}{10\sigma}}$.

The minimum bound of the distance difference range from $R_1$ to $R_2$ is computed as follows. We denote the closest that the transmitter can be situated to $R_1$, and so with the minimal amount of signal shadowing, presuming the minimum EIRP, as $D_1(P^-, -\sigma)$. We further represent the farthest that the transmitter can be located from $R_2$, thus with the maximal amount of signal shadowing, with the minimum EIRP, as $D_2(P^-, +\sigma)$. The combination of both quantities corresponds to the minimum bound of the distance difference range from $R_1$ to $R_2$, and is depicted as $\Delta d_{1,2}^-$ in Figure 6.3(a). The converse is the minimum bound of the distance difference range from $R_2$ to $R_1$, denoted as $\Delta d_{2,1}^-$. This set of points represents the closest that the transmitter can be to $R_2$, while the farthest from $R_1$, at the minimal EIRP, as illustrated in Figure 6.3(b). The maximum bound of the distance difference ranges from $R_1$ to $R_2$ and from $R_2$ to $R_1$ presume the maximal EIRP, $P^+$. From the perspective of $R_1$ to $R_2$, the maximum bound $\Delta d_{1,2}^+$ indicates the farthest that the transmitter can be located from $R_1$ at the highest EIRP, using the maximal amount of signal shadowing, yielding $D_1(P^+, +\sigma)$, while being closest to $R_2$ and so with the minimal amount of shadowing, for $D_2(P^+, -\sigma)$. The maximum bound of the distance difference range $\Delta d_{1,2}^+$ from $R_1$ to $R_2$ is shown in Figure 6.4(a), and the corresponding maximum bound $\Delta d_{2,1}^+$ from $R_2$ to $R_1$ is illustrated in Figure 6.4(b). One area delimiting the location of the transmitter is located between the minimum and maximum bounds of the distance difference range from $R_1$ to $R_2$, and the other is situated
Figure 6.3: Example of Minimum Bounds of Distance Difference Ranges

Figure 6.4: Example of Maximum Bounds of Distance Difference Ranges

Figure 6.5: Example of Distance Difference Areas
between the corresponding bounds from $R_2$ to $R_1$, as depicted in Figure 6.5.

We define these concepts more formally.

**Theorem 6.1.** Let $d_i$ and $d_j$ be the unknown distances between a transmitter $T$ and receivers $R_i$ and $R_j$ respectively.

1. The minimum bound $\Delta d_{i,j}^-$ of the distance difference range between $d_i$ and $d_j$, with confidence level $C$, is the distance difference at the minimal EIRP ($P^-$) over the full signal shadowing range $[-z\sigma, +z\sigma]$.

\[
\Delta d_{i,j}^- = \left( d_0 \times 10^{\frac{P^- - RSS_i - L(d_0) - z\sigma}{10\eta}} \right) - \left( d_0 \times 10^{\frac{P^- - RSS_j - L(d_0) + z\sigma}{10\eta}} \right)
\] (6.3)

2. The maximum bound $\Delta d_{i,j}^+$ of the distance difference range between $d_i$ and $d_j$, with confidence level $C$, is the distance difference at the maximal EIRP ($P^+$) over the full signal shadowing range $[+z\sigma, -z\sigma]$.

\[
\Delta d_{i,j}^+ = \left( d_0 \times 10^{\frac{P^+ - RSS_i - L(d_0) + z\sigma}{10\eta}} \right) - \left( d_0 \times 10^{\frac{P^+ - RSS_j - L(d_0) - z\sigma}{10\eta}} \right)
\] (6.4)

**Proof.** Because both $R_i$ and $R_j$ receive the same target message sent with a single (albeit unknown) transmitting power, the EIRP value used to calculate the distance difference between $R_i$ and $R_j$ must be the same. Thus the minimum bound of the distance difference range uses $P^-$, and the maximum bound uses $P^+$. From Lemma 6.4, we know that for each EIRP value, the range of minimum and maximum distances must encompass the full signal shadowing range between $-z\sigma$ and $+z\sigma$, with confidence level $C$.

Let $D_k(P, \nu)$ represent the distance between $T$ and $R_k$ with parameters $P$ as the EIRP and $\nu$ as the amount of signal shadowing. Thus according to Lemma 6.3, $D_k(P, \nu) = d_0 \times 10^{\frac{P - RSS_k - L(d_0) + \nu}{10\eta}}$.

1. At the minimal EIRP, the possible distance difference boundaries between $R_i$ and $R_j$ can be computed, with confidence level $C$, using combinations of shadowing bounds:

   (i) $D_i(P^-, -z\sigma) - D_j(P^-, -z\sigma)$

   (ii) $D_i(P^-, -z\sigma) - D_j(P^-, +z\sigma)$
(iii) \( D_i(P^-, +z\sigma) - D_j(P^-, -z\sigma) \)

(iv) \( D_i(P^-, +z\sigma) - D_j(P^-, +z\sigma) \)

These points form a lattice with infimum \( \Delta d_{i,j}^- = D_i(P^-, -z\sigma) - D_j(P^-, +z\sigma) \), since \(-z\sigma\) and \(-(+z\sigma)\) represent the minimum shadowing bound. The supremum of this lattice consists of \( D_i(P^-, +z\sigma) - D_j(P^-, -z\sigma) \), given that \(+z\sigma\) and \(-(-z\sigma)\) represent the maximum shadowing bound. It should be noted that this supremum equals \(-\Delta d_{j,i}^-\), which is the minimum bound of the distance difference range between \( d_j \) and \( d_i \).

2. At the maximal EIRP, the possible distance difference boundaries between \( R_i \) and \( R_j \) can be computed, with confidence level \( C \), as follows:

(i) \( D_i(P^+, -z\sigma) - D_j(P^+, -z\sigma) \)

(ii) \( D_i(P^+, -z\sigma) - D_j(P^+, +z\sigma) \)

(iii) \( D_i(P^+, +z\sigma) - D_j(P^+, -z\sigma) \)

(iv) \( D_i(P^+, +z\sigma) - D_j(P^+, +z\sigma) \)

These form a lattice with supremum \( \Delta d_{i,j}^+ = D_i(P^+, +z\sigma) - D_j(P^+, -z\sigma) \), since \(+z\sigma\) and \(-(-z\sigma)\) represent the maximum shadowing bound. The infimum of this lattice consists of \( D_i(P^+, -z\sigma) - D_j(P^+, +z\sigma) \), given that \(-z\sigma\) and \(-(+z\sigma)\) represent the minimum shadowing bound. This infimum equals \(-\Delta d_{j,i}^+\), which is the maximum bound of the distance difference range between \( d_j \) and \( d_i \).

6.2.2 Plotting Minimum and Maximum Hyperbolas

Given the definitions for the range of distance differences between a pair of receivers, we may construct the corresponding hyperbolas bounding the location of a transmitter.

**Definition 6.3.** Let a transmitter \( T \) be located at unknown coordinates \((x, y)\) and a pair of receivers \( R_i, R_j \) at known coordinates \((x_i, y_i)\) and \((x_j, y_j)\) respectively. Let \( \Delta d_{i,j}^- \) and \( \Delta d_{i,j}^+ \) be defined as the minimum and maximum bounds of the distance difference range between \( R_i \) and \( R_j \), with confidence level \( C \).
We define the minimum hyperbola $H_{i,j}^-$ between $R_i$ and $R_j$ as the closest boundary to $R_i$ where $T$ can be situated, while still in radio range of $R_j$, assuming the minimal EIRP ($P^-$). $H_{i,j}^-$ is constructed at the minimum bound $\Delta d_{i,j}^-$ of the distance difference range from $R_i$ to $R_j$, as put forth in Theorem 6.1:

$$\sqrt{(x - x_i)^2 + (y - y_i)^2} - \sqrt{(x - x_j)^2 + (y - y_j)^2} = \Delta d_{i,j}^-$$

Conversely, we define the maximum hyperbola $H_{i,j}^+$ between $R_i$ and $R_j$ as the farthest boundary from $R_i$ where $T$ can be situated, while in closest range to $R_j$, assuming the maximal EIRP ($P^+$). $H_{i,j}^+$ is thus constructed at the maximum bound $\Delta d_{i,j}^+$ of the distance difference range from $R_i$ to $R_j$:

$$\sqrt{(x - x_i)^2 + (y - y_i)^2} - \sqrt{(x - x_j)^2 + (y - y_j)^2} = \Delta d_{i,j}^+$$

**Theorem 6.2.** Let $H_{i,j}^-$ and $H_{i,j}^+$ be the minimum and maximum hyperbolas from $R_i$ to $R_j$, as stated in Definition 6.3.

Then a transmitter $T$ is located in the area $A_{i,j}$ between the hyperbolas $H_{i,j}^-$ and $H_{i,j}^+$, with confidence level $C$. Alternately, we say that $Pr[T \in A_{i,j}] = C$ and $Pr[T \in \overline{A}_{i,j}] = (1 - C)$, where $\overline{A}_{i,j}$ is the complement of $A_{i,j}$.

**Proof.** We define the distance between $T$ and $R_i$ as $d_i = \sqrt{(x - x_i)^2 + (y - y_i)^2}$ and the distance between $T$ and $R_j$ as $d_j = \sqrt{(x - x_j)^2 + (y - y_j)^2}$. If $\Delta d_{i,j} = d_i - d_j$ is defined as the exact difference between the distances from $T$ to $R_i$ and from $T$ to $R_j$, we obtain the equation for the hyperbola between $R_i$ and $R_j$:

$$\sqrt{(x - x_i)^2 + (y - y_i)^2} - \sqrt{(x - x_j)^2 + (y - y_j)^2} = \Delta d_{i,j}$$

We know from Theorem 6.1 that $\Delta d_{i,j}^-$ is the minimum bound of the distance difference between $d_i$ and $d_j$ and that $\Delta d_{i,j}^+$ is the maximum bound of this difference, with probability $C$. We can therefore deduce that the probability of $T$ being located in the area between $H_{i,j}^-$ and $H_{i,j}^+$ is $C$, and the probability of $T$ being located outside this area is $(1 - C)$.

An additional pair of minimum and maximum hyperbolas can be constructed between receivers $R_i$ and $R_j$, namely the hyperbolas based on the inverted order of
the receivers, \( R_j \) and \( R_i \). Thus each pair of receivers yields four hyperbolas to help determine the location of a transmitter. We have also noted in simulations that the maximum bound of the distance difference range between receivers is often too large for the corresponding hyperbola to be plotted. However, it is still required to delimit the candidate hyperbolic area for the transmitter.

![Figure 6.6: Example of Minimum Hyperbolas for \( R_1, R_2 \)](image)

**Example.** Let us compute the candidate hyperbolic areas \( A_{i,j} \) and \( A_{j,i} \) for the location of a transmitter \( T \) with confidence level \( C = 0.95 \), which yields the normal distribution constant \( z = 1.96 \). We assume a transmitter frequency of 2.4 GHz. For a reference distance \( d_0 = 1 \) m, we use the parameter values obtained by Liechty *et al.* [88, 89] for LOS propagation and a seven meter high antenna, where the path loss exponent is \( \eta = 2.76 \) and the standard deviation is \( \sigma = 5.62 \). The average path loss at \( d_0 \) is calculated with Friis' transmission equation for free space propagation [54], assuming isotropic transmitting and receiving antennas:

\[
\overline{L}(d_0) = \left( \frac{4\pi f d_0}{c} \right)^2 = 40 \text{ dB}
\]

We assume an example layout as depicted Figure 6.6, where receivers \( R_1 \) and \( R_2 \) receive an attack message from a transmitter \( T \) with signal strength \( RSS_1 = -79.20 \) dBm and \( RSS_2 = -74.27 \) dBm respectively, corresponding to an actual transmitted EIRP of 30 dBm. Further, we model the EIRP range with \( \mathcal{P}^- = 15 \) dBm and \( \mathcal{P}^+ = 40 \) dBm. Equations (6.1) and (6.2) reveal that the transmitter \( T \) is located between 37
m and 1848 m from $R_1$ and between 24 m and 1225 m from $R_2$, with probability $C = 0.95$. Using Equations (6.3) and (6.4), we compute the minimum bound of the distance difference range between $d_1$ and $d_2$ as $\Delta d_{1,2}^- = -115$ m and the maximum bound as $\Delta d_{1,2}^+ = 1653$ m. Conversely, the minimum bound of the range between $d_2$ and $d_1$ is calculated as $\Delta d_{2,1}^- = -205$ m and the maximum bound is $\Delta d_{2,1}^+ = 930$ m. The minimum hyperbolas $\mathcal{H}_{1,2}^-$ and $\mathcal{H}_{2,1}^-$ associated with $\Delta d_{1,2}^-$ and $\Delta d_{2,1}^-$ respectively are depicted in Figure 6.6. The candidate areas for transmitter $T$ include the area between $\mathcal{H}_{1,2}^-$ and $\mathcal{H}_{1,2}^+$, known as $\mathcal{A}_{1,2}$ and shown with dotted arrows, and the area between $\mathcal{H}_{2,1}^-$ and $\mathcal{H}_{2,1}^+$, named $\mathcal{A}_{2,1}$ and featured with dash-dotted arrows. $T$ is located within $\mathcal{A}_{1,2}$ with probability 0.95, and within $\mathcal{A}_{2,1}$ with the same probability.

Minimum and maximum hyperbolas can be constructed between multiple pairs of receivers, forming a number of intersecting areas within which a transmitter location can be further bounded. For example, Figure 6.7 illustrates the minimum hyperbolas between receiver pairs $R_1, R_2$ and $R_3, R_4$, and Figure 6.8 illustrates the hyperbolic areas between the same receiver pairs.

Figure 6.7: Example of Minimum Hyperbolas for $R_1, R_2$ and $R_3, R_4$
Figure 6.8: Example of Hyperbolic Areas for $R_1, R_2$ and $R_3, R_4$

### 6.3 Performance Evaluation

We convey the results obtained by executing our HPB algorithm using various transmitter locations, with two scenarios involving fixed receivers and simulated RSS values. We analyze the success rate of the algorithm in estimating the transmitter position and examine the corresponding candidate area size.

Results are presented for two separate scenarios, one with two receivers and the other with four receivers, located on a $1000 \times 1000$ m$^2$ grid. The transmitter location is simulated at each 100 meter interval in the grid from zero to 1000 meters on both the X-axis and Y-axis, as shown in Figure 6.9. The HPB algorithm is executed 1000 times for each of the possible 121 transmitter locations, given four individual confidence levels, $C \in \{0.95, 0.90, 0.85, 0.80\}$. For each execution, minimum and maximum hyperbolas are constructed between every possible pair of receivers, according to Theorem 6.2.

The simulation assumes a frequency of 2.4 GHz, as well as the values for the reference distance, path loss exponent and shadowing standard deviation determined for this frequency in an outdoor environment by Liechty et al. [88, 89], where $d_0$ equals one meter, $\eta$ equals 2.76 and $\sigma$ equals 5.62. For each HPB execution, the setup phase assumes a transmitter EIRP of 30 dBm, and employs the log-normal shadowing model in order to compute a simulated RSS value at each receiver. A
different amount of random signal shadowing is added to each RSS value, along a Gaussian distribution, with mean zero and Liechty’s shadowing standard deviation. In the position bounding phase, a probable EIRP range is determined dynamically by taking the closest receiver to the transmitter location, i.e. the receiving device with the highest RSS, as a reference receiver. The EIRP range is set to the intersection of EIRP ranges required for each remaining receiver to reach the reference receiver. The dynamic EIRP range computation is explored in greater depth in Chapter 9 (see Heuristic 9.1).

The performance of the HPB algorithm is evaluated along two metrics: the success rate in correctly localizing the transmitter within a bounded area, and the minimization of the candidate area size. Consequently, both metrics are gathered for each HPB execution.

The success rate reflects the percentage of executions for which the maximal probability candidate area, i.e. the intersection of hyperbolic areas constructed from all possible receiver pairs, contains the actual transmitter location. For all receiver pairs $R_i, R_j$, if the actual difference in T-R distances $\Delta d_{i,j}$ between $R_i$ and $R_j$ is contained in the distance difference interval $[\Delta d_{i,j}^-, \Delta d_{i,j}^+]$, then an HPB execution is successful.

The candidate area size represents the percentage of the 1000 × 1000 m² grid covered by the candidate area computed with HPB. We sample the simulation grid at regular intervals, with a pre-determined granularity. More precise results can be achieved with a small sampling interval, where the size of the candidate area is slightly
reduced, but to the detriment of computational efficiency. A trade-off between the area size and performance is thus reached with a sampling interval of 20 meters. We determine the candidate area size as the ratio of sample points whose distances from the transmitter are situated within the distance difference intervals \([\Delta d_{i,j}^-; \Delta d_{i,j}^+]\) for all receiver pairs \(R_i, R_j\), over the total number of sample points. Optimal results are obtained when the success rate is maximized and the area size is minimized.

Figure 6.10 illustrates the success rate for the four-receiver scenario, given \(C = 0.95\). Since metrics were gathered solely for the grid points shown in Figure 6.9, values for the intermediate points were intrapolated linearly between the values for the computed grid points. The success rate is highest in the cross shape between the receivers, which we term the aggregate range. This range constitutes the zone in which the transmitter is located between at least one pair of receivers, enabling the receivers to aggregate and support each other’s findings. The lowest success rates are achieved in the corners of the grid, outside the aggregate range, since these zones are not situated between any pair of receivers. Low success rates are also obtained when the transmitter is located precisely at the receiver locations. These special cases are eliminated from our subsequent analysis, because a zero meter distance is less than the reference distance of \(d_0\) (one meter) used in the path loss model. Figure 6.10 displays the success rates averaged over the 1000 executions of the HPB algorithm for each transmitter location. The success rate associated with each grid point within the aggregate range lies in a 90% statistical confidence interval of ±3% of the mean for that point. The non-aggregate range points are situated in an interval of ±4% from the grid point mean. Consequently, not only is the HPB success rate lower in the non-aggregate range, the results are also less reliable due to their greater variance. This bears out the intuition that a greater receiver coverage poses a significant advantage in the localization of an uncooperative device.

Figure 6.11 depicts the success rate as a function of the distance from the grid’s midpoint, located at coordinates (500, 500), for the four-receiver scenario. This success rate is shown for each of the four confidence levels tested, within a 90% statistical confidence interval depicted with each point for \(C = 0.95\). Intuitively, the farther the transmitter from the midpoint, the lower the expected success rate, and Figure 6.11
confirms this hypothesis for all confidence levels. At the midpoint, where the distance is zero, the highest success rate is achieved. However, because these success rates occur between the receivers in a cross shape rather than concentric circles, Figure 6.11 does not show a completely linear decrease. For example, noticeable dips in the graph occur at distances 361 m, 424 m, 500 m and 566 m. These correspond to the instances where the transmitter is located in the non-aggregate range at the corners of the grid, and thus within range of only one receiver. Figure 6.12 captures the same data as Figure 6.11, but with the non-aggregate range points excluded. A more linear success rate is achieved as the transmitter location moves away from the midpoint.

While the highest success rates are associated with higher confidence levels, so
are the larger candidate area sizes. Figure 6.13 shows the candidate area size as a percentage of the total simulation area depicted in the 1000×1000 m² grid for the two-receiver and four-receiver scenarios. In each case, the candidate area size decreases with the confidence level. This reflects the reduced value of the normal distribution constant $z$ for lower confidence levels. The shadowing interval is therefore smaller, resulting in reduced hyperbolic areas whose intersections are correspondingly smaller. However, the candidate areas for the two-receiver scenario are significantly larger than those for the four-receiver scenario, reaching 62% of the total simulation area in some instances. Clearly, this type of result is of very little use in locating an uncooperative device. The four-receiver scenario yields more promising results, where even a 0.95 confidence level produces a candidate area on average below 25% of the total grid size. This finding is consistent with the expectation that a higher number of receivers yields finer grained results and thus bounds a transmitter location to a smaller candidate area.

Figure 6.14 illustrates the average candidate area size for the success rates achieved with each confidence level in the four-receiver scenario. In general, the size of the candidate area is larger for a given success rate as the confidence level increases. For example, a success rate of 80% yields a candidate area of 23% for $C = 0.95$, an area size of 18% for $C = 0.90$ and an area of 17% for $C = 0.85$. Thus the average area size clearly decreases with the confidence level, due to the reduced shadowing interval.
6.4 Summary

We presented a hyperbolic location estimation mechanism for attributing a target message to an originating uncooperative device by estimating its position without its assistance or knowledge of the EIRP. Performance evaluation through simulations reveals a success rate commensurate with the confidence level, although the size of a candidate area also increases with the success rate. Correspondingly, a confidence level of 95% yields an average candidate area size slightly below 25% of the simulated grid area.

The HPB mechanism may provide an important first step in attributing an attack
message to a malicious insider in various types of wireless networks. Technology-based means to effect retribution on the transmitting device and possibly evict it from the network remain an open issue.

In Chapter 7, we investigate the combination of confidence levels ascribed to the intersection of HPB hyperbolic areas by multiple receiver pairs. We evaluate the performance of HPB using RSS values harvested from an outdoor experiment in Chapter 8. We then propose further improvements to HPB performance by extending its application to vehicular networks and devising a mechanism to track mobile devices, as described in Chapter 9.
Chapter 7

Probabilistic Evidence Aggregation

Theorem 6.2 provides the means to delineate a hyperbolic area to bound the possible location of a transmitter, with a desired confidence level. Given a number of receivers, hyperbolic areas can be constructed between multiple receiver pairs and the confidence levels aggregated together. Intuitively, an area endorsed by a greater number of receivers should be given a higher probability than areas advocated by fewer receivers. However in multiplicative probabilities, if two receiver pairs agree with confidence 0.95 that a transmitter is located in a particular area, for example, then that area is assigned confidence $0.95^2$. If four receiver pairs agree, then the confidence drops to $0.95^4$. We require a scheme to assign a probability distribution to the possible transmitter location, such that agreement from multiple receiver pairs compounds the probability inside the intersecting hyperbolic areas without redistributing it outside the intersections.

In [83], we partition a target transmitter's radio range according to the HPB-computed hyperbolic areas and their intersections. We revisit existing mechanisms to combine the confidence levels in these partitions. We then introduce a heuristic method to compound the evidence contributed by multiple receiver pairs, and thus the probability, in each hyperbolic area partition. We evaluate the performance of our compound probability approach using the HPB simulation results obtained in Chapter 6.

7.1 Partitioning Hyperbolic Areas

We first define the partitioning of the Euclidian space encompassing a transmitter's signal range, based on the number of intersecting HPB hyperbolic areas in which each sub-area lies.
**Definition 7.1.** Let $\xi$ be the Euclidian space encompassing the radio range of a transmitter, and let $\mathbb{R}$ be the set of all receiving devices in this range. Let $\mathcal{W}$ be the set of HPB hyperbolic areas computed using Theorem 6.2, for all combinations of receiver pairs in $\mathbb{R}$:

$$\mathcal{W} = \{A_{i,j} : \exists \mathcal{H}_{i,j}, \mathcal{H}_{i,j}^+ \text{ for all } R_i, R_j \in \mathbb{R} \text{ and } A_{i,j} \text{ is the area situated between } \mathcal{H}_{i,j} \text{ and } \mathcal{H}_{i,j}^+\}$$

$\xi$ is thus comprised of the union of all hyperbolic areas in $\mathcal{W}$ and their complements.

Let $n$ be equal to $|\mathcal{W}|$ and represent the number of hyperbolic areas. We note that because $\mathcal{W}$ always contains the hyperbolic area pair $A_{i,j}, A_{j,i}$, the value of $n$ is necessarily even. We define disjoint hyperbolic area partitions $S^k$ of $\xi$, such that for all $0 \leq k \leq n$, the partition $S^k$ contains the sub-areas $s_m$ of $\xi$ situated at once within the intersection of exactly $k$ hyperbolic areas and within the intersection of the complements of the remaining $n - k$ hyperbolic areas in $\mathcal{W}$. More formally, let $\mathcal{P}(\mathcal{W})$ denote the powerset of $\mathcal{W}$, and $I_k \subset \mathcal{P}(\mathcal{W})$ be the set of all combinations of exactly $k$ hyperbolic areas, such that $I_k = \{\alpha_i : \alpha_i \in \mathcal{P}(\mathcal{W}) \text{ and } |\alpha_i| = k\}$. Then:

$$S^k = \left\{s_m : s_m = \left(\bigcap_{h=1}^{k} a_h \text{ for every } a_h \in \alpha_i\right) \bigcap \left(\bigcap_{\ell=k+1}^{n} \overline{a}_\ell \text{ for every } a_\ell \in \{\mathcal{W} \setminus \alpha_i\}\right) \text{ for every } \alpha_i \in I_k\right\}$$

where $\bigcap_{h=1}^{k} a_h$ represents the intersection of any $k$ hyperbolic areas, and $\bigcap_{\ell=k+1}^{n} \overline{a}_\ell$ is the intersection of the complements of the remaining $n - k$ hyperbolic areas.

**Example.** Figure 7.1(b) depicts the partitioning of the Euclidian space covered by the $n = 4$ hyperbolic areas $\{A_{1,2}, A_{2,1}, A_{3,4}, A_{4,3}\}$ shown in Figure 7.1(a). To define sub-area $G$, for example, we compute the intersection of $k = 3$ hyperbolic areas $\{A_{1,2}, A_{2,1}, A_{4,3}\}$, which yields a zone we term $\Gamma$, corresponding to $D \cup G$ in Figure 7.1(b). The complement of the remaining $n - k = 1$ hyperbolic area $A_{3,4}$, denoted as $\overline{A}_{3,4}$, corresponds to $J \cup G \cup M$ in the same Figure. The intersection of $\Gamma$
and $\overline{A}_{3,4}$ defines sub-area $G$, since $\Gamma \cap \overline{A}_{3,4} = G$. So we see that $G = A_{1,2} \cap A_{2,1} \cap A_{4,3} \cap \overline{A}_{3,4}$. In Figure 7.1(b), the hyperbolic area partitions, as stated in Definition 7.1, are identified as follows: $S^4 = \{D\}; S^3 = \{E, F, G, H\}; S^2 = \{J, K, M, N\}$; and $S^1 = S^0 = \emptyset$. If the Euclidian space could be displayed in its entirety, additional sub-areas for $S^2, S^1$ and $S^0$ would be visible. However, we restrict the scope of this example to a $1000 \times 1000$ m$^2$ grid.

(a) Hyperbolic Areas
(b) Partitions

Figure 7.1: Example of Hyperbolic Area Partitions

7.2 Existing Evidence Combination Schemes

Every hyperbolic area partition $S^k$ is comprised of sub-areas situated in the intersection of the same number of hyperbolic areas $(k)$, and are thus equally likely sub-areas for the location of a transmitter. As a result, we assign a single probability to each partition $S^k$ based on the value of $k$. The intuitive idea is that the more hyperbolic areas a partition belongs to, the more likely it is that the transmitter is located in that partition, i.e. $Pr[T \in S^j] < Pr[T \in S^k]$ for all $0 < j < k < n$.

One method to compute the probability in a given partition involves multiplying together the probabilities assigned to its sub-areas by each receiver pair.

Lemma 7.1. Let $S^k$ be a set of hyperbolic area partitions, as stated in Definition 7.1, and let $X$ represent a random variable corresponding to a transmitter position over a Euclidian space $\xi$. Then the probability that the transmitter is located in sub-area $s_m$
of partition $S^k$ is the probability that it is situated in the intersection of $k$ hyperbolic areas and the intersection of the complements of the remaining $(n - k)$ hyperbolic areas:

$$Pr[x \in s_m] = C^k \times (1 - C)^{n-k} \quad \text{where } s_m \in S^k$$

Proof.

$$Pr[x \in s_m] = Pr \left[ \left( x \in \bigcap_{h=1}^{k} a_h \right) \cap \left( x \in \bigcap_{\ell=k+1}^{n} \bar{a}_\ell \right) \right] \quad \text{by Definition 7.1}$$

$$= (Pr[x \in a_h])^k \times (Pr[x \in \bar{a}_\ell])^{n-k}$$

$$= (Pr[x \in A_{i,j}])^k \times (Pr[x \in \bar{A}_{i,j}])^{n-k}$$

$$= C^k \times (1 - C)^{n-k} \quad \text{by Theorem 6.2}$$

Of even greater interest is the probability assigned to the set of all sub-areas situated within a given partition $S^k$. Each hyperbolic area computed by a receiver pair can be thought of as an independent Bernoulli trial with probability of success $C$ and probability of failure $(1 - C)$. The findings of each receiver pair are independent of each other even if they share a common receiver, because the distance difference is unique to each pair. The probability distribution with a total of $n$ hyperbolic areas is akin to $n$ Bernoulli trials, and the resulting distribution is binomial.

**Definition 7.2.** Let $S^k$ be a set of hyperbolic area partitions, as stated in Definition 7.1. Then we define the Multiplicative Probability (MP) distribution over a transmitter's radio range $\xi$ as follows. The probability that a transmitter is located in hyperbolic area partition $S^k$ is the probability that it lies in any of the sub-areas $s_m$ of $S^k$.

**Lemma 7.2.** Let a set of hyperbolic area partitions $S^k$ be computed over $\xi$, a transmitter's radio range in Euclidian space, as put forth in Definition 7.1. If the transmitter location represented by $X$ is a random variable with a multiplicative probability distribution over $\xi$, as set out in Definition 7.2, then the probability distribution function
(p.d.f.) \( f(x, k) \) of \( X \) for the location of the transmitter within \( k \) hyperbolic areas, with a constant confidence level \( C \) and total number of hyperbolic areas \( n \), is the following:

\[
f(x, k) = \Pr[x \in S^k] = \binom{n}{k} \times C^k \times (1 - C)^{n-k}
\]

**Proof.** Since there are \( \binom{n}{k} \) possible ways in which a transmitter can be located in \( k \) hyperbolic areas and in the complements of the remaining \( n - k \) hyperbolic areas, there are \( \binom{n}{k} \) possible sub-areas in partition \( S^k \). Therefore:

\[
Pr[x \in S^k] = \sum_{m=1}^{\binom{n}{k}} Pr[x \in s_m] \quad \text{where } s_m \in S^k \quad \text{by Definition 7.1}
\]

\[
= \sum_{m=1}^{\binom{n}{k}} [C^k \times (1 - C)^{n-k}] \quad \text{by Lemma 7.1}
\]

\[
= \binom{n}{k} \times C^k \times (1 - C)^{n-k}
\]

\( \square \)

However, the multiplicative probability method fails to effectively compound the confidence levels. Combined evidence from an increasing number of receiver pairs supporting the confidence in a given intersection actually decreases the overall probability in that area, since the multiplicative probability scheme re-distributes the associated probability outside the intersection. In the example depicted in Figure 7.1, with \( C = 0.95 \), two hyperbolic areas inform us that the probability of the transmitter located in the intersection is \( 0.95^2 = 0.90 \). But the evidence provided by four hyperbolic areas locates the transmitter in a smaller area with probability \( 0.95^4 = 0.81 \). In effect, additional evidence supporting the original finding reduces the probability of the transmitter’s location in the intersection, leading to the counter-intuitive conclusion that the transmitter is more likely to be found in the intersection of two hyperbolic areas than in the intersection of four areas.

The Dempster-Shafer method computes the plausibility \( Pl \) in a sub-area as its multiplicative probability, normalized to exclude any sub-areas of conflict. For our scenario, the Dempster-Shafer plausibility applies as follows:

\[
Pl[x \in S^k] = \binom{n}{k} \times \frac{C^k \times (1 - C)^{n-k}}{1 - K}
\] (7.1)
where $K$ equals $(1 - C)^k \times C^{(n-k)}$ and represents the degree of conflict.

For our scenarios, we are interested in localizing an uncooperative transmitter with a high degree of confidence, so $C$ will generally be close to one and $(1 - C)$ will be close to zero. The more useful probabilities are vested in the partitions with $k$ close to $n$, since that is where the transmitter is most likely to be located. With a relatively large $k$ compared to $(n - k)$ and a low value of $(1 - C)$, the value of $K$ is nearly equal to zero, so the denominator of Equation (7.1) is close to one. The Dempster-Shafer plausibility $Pl$ thus reduces to the multiplicative probability set forth in Lemma 7.2.

7.3 Compounding Hyperbolic Area Confidence

We propose an alternate, heuristic method for aggregating evidence in such a manner that supporting evidence increases the likelihood of a transmitter being located in a given area. The intuitive idea behind our mechanism is based on the Bayesian conditional approach and reflects the democratic principle. Given $n$ hyperbolic areas, if all $n$ agree that the transmitter is located in their common intersection with probability $C$, then the partition corresponding to this intersection is assigned probability $C$. Of the remaining probability $(1 - C)$, the proportion assigned to the intersection of $n - 1$ hyperbolic areas is $C$, for $C \times (1 - C)$. Of the remaining $(1 - C) \times (1 - C)$, $C$ is ascribed to the intersection of $n - 2$ hyperbolic areas, and so on until the intersection of zero hyperbolic areas is assigned the remainder $(1 - C)^n$.

**Definition 7.3.** Let $S^k$ be a set of hyperbolic area partitions, as stated in Definition 7.1. Let $\xi$ be the Euclidian space covered by $n$ hyperbolic areas and their complements. Then we define the Compound Probability (CP) distribution over a transmitter’s radio range $\xi$ as follows. The probability that a transmitter is located in a partition $S^k$ of $\xi$ lying within $k$ hyperbolic areas, for $k > 0$, is the probability that it is situated within $k$ hyperbolic areas, given the supporting evidence that it is within $k - 1$ of these areas, combined with the probability that it is outside the remaining $n - k$ hyperbolic areas. In the case where the partition is outside of all hyperbolic areas, i.e. for $k = 0$, the compound probability reduces to the multiplicative probability presented in Lemma 7.2.
Theorem 7.1. Let a set of hyperbolic area partitions $S^k$ be computed over $\xi$, a transmitter’s radio range in Euclidian space, as put forth in Definition 7.1. If the transmitter location represented by $X$ is a random variable with a compound probability distribution over $\xi$, as set out in Definition 7.3, then the p.d.f. $f(x,k)$ of $X$ for the location of the transmitter within $k$ hyperbolic areas, with a constant confidence level $C$ and total number of hyperbolic areas $n$, is the following:

$$f(x,k) = Pr[x \in S^k] = \begin{cases} C \times (1 - C)^{n-k} & \text{for } k > 0 \\ (1 - C)^n & \text{for } k = 0 \end{cases}$$

Proof.

1. (i) For $k > 0$:

$$Pr[x \in S^k] = Pr\left[ (x \in \bigcap_{h=1}^{k} a_h) \mid (x \in \bigcap_{h=1}^{k-1} a_h) \right] \times Pr\left[ (x \in \bigcap_{\ell=k+1}^{n} \bar{a}_\ell) \right]$$

by Definitions 7.1 and 7.3

$$= \frac{Pr\left[ x \in \bigcap_{h=1}^{k} a_h \bigcap \bigcap_{h=1}^{k-1} a_h \right]}{Pr\left[ x \in \bigcap_{h=1}^{k-1} a_h \right]} \times Pr\left[ (x \in \bigcap_{\ell=k+1}^{n} \bar{a}_\ell) \right]$$

by the definition of conditional probabilities

$$= \frac{Pr\left[ x \in \bigcap_{h=1}^{k} a_h \right]}{(Pr[x \in a_h])^{k-1}} \times (Pr[x \in \bar{a}_\ell])^{n-k}$$
\[ \frac{(Pr[x \in A_{i,j}])^k}{(Pr[x \in A_{i,j}])^{k-1}} \times (Pr[x \in \overline{A}_{i,j}])^{n-k} \]
\[ = \frac{C^k}{C^{k-1}} \times (1 - C)^{n-k} \text{ by Theorem 6.2} \]
\[ = C \times (1 - C)^{n-k} \]

(ii) For \( k = 0 \):

\[ Pr[x \in S^0] = Pr \left[ (x \in \bigcap_{h=1}^{0} a_h) \right] \times Pr \left[ (x \in \bigcap_{\ell=1}^{n} \overline{a_\ell}) \right] \]
\[ = Pr[x \in \xi] \times (Pr[x \in \overline{a_\ell})]^n \]
\[ = 1 \times (Pr[x \in \overline{A}_{i,j}])^n \]
\[ = (1 - C)^n \text{ by Theorem 6.2} \]

2. \( Pr[x \in S^k] \) is a probability distribution, since \( \sum_{k=0}^{n} Pr[x \in S^k] = 1. \)

\[ \sum_{k=0}^{n} Pr[x \in S^k] = (1 - C)^n + \sum_{k=1}^{n} [C \times (1 - C)^{n-k}] \]
\[ = (1 - C)^n + C \times \sum_{k=0}^{n-1} (1 - C)^k \]
\[ = (1 - C)^n + C \times \left( \frac{1 - (1 - C)^n}{1 - (1 - C)} \right) \]
\[ = (1 - C)^n + 1 - (1 - C)^n \]
\[ = 1 \quad \square \]

In contrast to the multiplicative probability which follows a binomial distribution, the compound probability reflects a geometric distribution. Every hyperbolic area within which a transmitter is located counts for a successful Bernoulli trial for MP, but these are compounded as one combined success with probability \( C \) for CP, due to the democratic agreement. Consequently, only the number of failures \((n - k)\) with probability \((1 - C)\) are counted.

**Example.** The compound probability for the example illustrated in Figure 7.1 is shown in Figure 7.2. The probability associated with partition \( S^4 = \{D\} \) is \( C = 0.95 \), with partition \( S^3 = \{E, F, G, H\} \) is \( 0.05 \times 0.95 = 0.0475 \) and with partition
$S^2 = \{J, K, M, N\}$ is $0.05^2 \times 0.95 = 0.002375$. If any sub-areas were associated with partition $S^1$, their probability would be $0.05^3 \times 0.95$, and the remainder would be assigned to $S^0$ with $0.05^4$.

Figure 7.2: Compound Probability for Hyperbolic Area Partitions

A refinement of Theorem 7.1 can be achieved by considering the probability of a transmitter's location within paired hyperbolic areas. The reason for this is one of symmetry. If a transmitter is centrally located between a pair of receivers $R_i$ and $R_j$, it must necessarily lie within the symmetric hyperbolic area pair $A_{i,j}$ and $A_{j,i}$. If the transmitter is not centrally situated, little granularity is lost by twinning the asymmetric pairs of hyperbolic areas within which it lies. We thus require a mechanism for combining the probabilities of paired hyperbolic areas, whether they correspond to symmetric receiver pairs or not.

**Definition 7.4.** Let $r = \frac{n}{2}$ be the number of paired hyperbolic areas associated with a total $n$ such areas. We define disjoint partitions $V^\gamma$ of a transmitter range $\xi$, such that for all $0 \leq \gamma \leq r$, the partition $V^\gamma$ contains the sub-areas of $\xi$ situated in the intersection of at least $\gamma$ paired hyperbolic areas. More formally:

$$V^\gamma = \begin{cases} 
S^{2\gamma} & \text{for } \gamma = r \\
S^{2\gamma} \cup S^{2\gamma+1} & \text{for } 0 \leq \gamma \leq r - 1
\end{cases}$$

**Example.** In the example shown in Figure 7.1(b), partition $V^2 = \{D\}$ is comprised of the sub-area situated within four hyperbolic areas and thus two paired areas,
in this case $\mathcal{A}_{1,2}, \mathcal{A}_{2,1}$ and $\mathcal{A}_{3,4}, \mathcal{A}_{4,3}$. Partition $V^1 = \{E, F, G, H, J, K, M, N\}$ includes the sub-areas located in the intersection of two or three hyperbolic areas and thus in at least one paired area. Some of these paired areas are symmetric, as for sub-areas $\{E, F, G, H\}$ appearing in $\mathcal{A}_{1,2}, \mathcal{A}_{2,1}$ or in $\mathcal{A}_{3,4}, \mathcal{A}_{4,3}$. Other sub-areas are situated in asymmetric paired areas, for example $\{J\}$ is located in the intersection of $\mathcal{A}_{2,1}$ and $\mathcal{A}_{4,3}$. Partition $V^0$ equals $\emptyset$, since every sub-area in Figure 7.1(b) appears in at least one paired hyperbolic area.

We extend the compound probability method proposed in Theorem 7.1 to consider paired hyperbolic areas.

**Corollary 7.1.** Let the paired compound probability $CP$ that a transmitter is located in a partition $V^\gamma$ of $\xi$ lying within $\gamma$ paired hyperbolic areas, for $\gamma < r = \frac{n}{2}$, be the probability that it is situated within partition $S^{2\gamma}$ or $S^{2\gamma+1}$, as put forth in Definition 7.4. In the case where the transmitter is located in the intersection of all hyperbolic areas, i.e. for $\gamma = r$, the compound probability is equal to that of $S^n$. Thus:

$$Pr[x \in V^\gamma] = \begin{cases} C \\
C \times (2 - C) \times (1 - C)^{(n-2\gamma-1)} \\
(1 - C)^{n-1}
\end{cases} \begin{array}{l} 
\text{for } \gamma = r \\
\text{for } 1 \leq \gamma \leq r - 1 \\
\text{for } \gamma = 0
\end{array}$$

**Proof.**

1. For $\gamma = r$:

$$Pr[x \in V^r] = Pr[x \in S^n] \quad \text{by Definition 7.4}$$

$$= C \times (1 - C)^0 = C \quad \text{by Theorem 7.1}$$

2. For $1 \leq \gamma \leq r - 1$:

$$Pr[x \in V^\gamma] = Pr[x \in S^{2\gamma} \cup S^{2\gamma+1}] \quad \text{by Definition 7.4}$$

$$= Pr[x \in S^{2\gamma}] + Pr[x \in S^{2\gamma+1}]$$

$$= C \times (1 - C)^{n-2\gamma} + C \times (1 - C)^{(n-(2\gamma+1))} \quad \text{by Theorem 7.1}$$

$$= C \times (1 - C)^{(n-2\gamma-1)} \times (2 - C)$$
3. For $\gamma = 0$:

$$Pr[x \in V^0] = Pr[x \in S^0 \cup S^1]$$

by Definition 7.4

$$= Pr[x \in S^0] + Pr[x \in S^1]$$

$$= (1 - C)^n + C \times (1 - C)^{n-1}$$

by Theorem 7.1

$$= (1 - C)^{n-1}$$

\[ \square \]

7.4 Performance Evaluation

We compare the HPB simulation results obtained in Chapter 6 against both the compound and multiplicative probability distributions. Our simulation scenario is modeled on a $1000 \times 1000$ m$^2$ grid. The transmitter location is simulated along each grid point located at 100 meter intervals. Four receivers are simulated, and the RSS values computed at each receiver follow the log-normal shadowing model [115], with a different random amount of signal shadowing added to each receiver RSS value, along a Gaussian distribution. We assume the radio signal frequency to be 2.4 GHz. The reference distance, path loss exponent and shadowing standard deviation measured experimentally for this frequency by Liechty et al. [88, 89] are used. One thousand executions of the HPB algorithm are computed for each of four confidence levels, $C \in \{0.95, 0.90, 0.85, 0.80\}$, for each possible transmitter location on the grid. The number of hyperbolic areas in which the transmitter is located is accumulated over the executions. The success rate of our results is deemed accurate within a 90% statistical confidence interval of $\pm 3\%$ to $\pm 4\%$ of the grid point mean, depending on the distance between the transmitter location and the center of the grid.

We partition the simulation grid into three ranges of grid points within which the HPB algorithm exhibits different behaviors. As with the position bounding mechanism described in Chapter 6, the compound probability approach performs optimally when the transmitter is located between at least one pair of receivers. Figure 7.3 illustrates the grid points at which the transmitter locations are simulated, as well as the positions of the four receivers, each depicted by a small cross. The central range comprises the centrally located grid points, the aggregate range includes the
grid points located between pairs of receivers so that the confidence levels can be combined, and the outer range includes the points outside the scope of any pair of receivers. The aggregate range is deemed to encompass the central range points as well.

With four receivers in the simulation scenario, six possible receiver pairs compute a total of 12 hyperbolic areas. Figures 7.4 and 7.5 compare the probability distributions within these hyperbolic areas for \( C = 0.95 \) and \( C = 0.90 \). The compound probability (CP) and multiplicative probability (MP) distributions are computed according to Theorem 7.1 and Lemma 7.2 respectively. The simulation results depicted correspond to HPB executions with transmitter locations within the central (CR) and aggregate (AR) ranges. While both computed probability distributions decrease noticeably below the maximum 12 hyperbolic areas, the CP distribution remains more proportional to the simulation results obtained within both the central and aggregate ranges. For example, the relative probability decrease exhibited in the simulation results from 12 to 11 hyperbolic areas for \( C = 0.95 \) averages around 93\%, and the corresponding decrease in compound probability is 95\%. By contrast, the multiplicative probability decreases by only 37\%.

The goodness-of-fit of the CP and MP distributions to the simulation results is measured using the \( D_N \) statistic associated with the Kolmogorov-Smirnov test [74, 133]. The Kolmogorov-Smirnov statistic expresses the difference between a hypothesized probability distribution \( F_0(x) \) and an empirical one \( F_N(x) \) as the least upper
Figure 7.4: Hyperbolic Area Probability Distributions for $C = 0.95$

Figure 7.5: Hyperbolic Area Probability Distributions for $C = 0.90$
bound of the absolute difference between all the corresponding points of both distributions: $D_N = \sup_x |F_N(x) - F_0(x)|$. Figure 7.6 depicts the Kolmogorov-Smirnov statistics with the computed distributions, CP and MP, as the hypothesized distribution, and the simulation results in the central and aggregate ranges as the empirical distribution. For all confidence levels, the maximum point-wise differences between the CP distribution and simulation results in both ranges are minimal, compared to their differences with the MP distribution. For $C \in \{0.95, 0.90\}$, the Kolmogorov-Smirnov statistics for CP and the simulation data remain below 15%, while they reach 60% when MP is considered. The performance of the CP algorithm in the central range alone is even better, with the maximum difference not exceeding 5% for $C \in \{0.95, 0.90\}$. The compound probability distribution is clearly the better model for the HPB simulation results than multiplicative probability, although compound probability performs even better for the central range points.

Figure 7.6: Kolmogorov-Smirnov Statistic for Distribution Differences

Figures 7.7 and 7.8 illustrate the cumulative probability distribution results for $C = 0.95$ and $C = 0.90$. Again, it can be seen that the CP distribution more closely models both the central and aggregate range simulation results than does the MP distribution. Figure 7.9 plots the cumulative probability distribution for both CP and MP, given each of four confidence levels $C \in \{0.95, 0.90, 0.85, 0.80\}$. It should be noted that as a binomial distribution, the multiplicative cumulative probability follows a Gaussian curve, while the compound cumulative probability reveals an exponential
curve. Because the probabilities exhibited by the simulation data also follow an exponential curve, as shown in the central range results of Figure 7.10, the compound probability mechanism inherently provides the better model, independently of the value of $C$.

The performance of the CP mechanism when hyperbolic areas are paired, as defined in Corollary 7.1 of Theorem 7.1, can be assessed by comparing the results of Figures 7.4 and 7.5 with those of Figures 7.11 and 7.12. Because the pairing algorithm exploits the natural symmetry between some of the hyperbolic area pairs, the correspondence between the CP distribution and simulation results is closer. For example, Figures 7.4 and 7.5 indicate a dip in the central range probability for 11 hyperbolic areas. Given that the grid points are centrally located, a significant portion of them are located in symmetric hyperbolic area pairs. As a result, their appearance in an odd number of hyperbolic areas is likely an anomaly. When hyperbolic areas are paired, as in Figures 7.11 and 7.12, this aberration disappears, and the simulation probabilities are more commonsensical and better fitted to the compound probabilities.

7.5 Summary

We presented a compound probability approach to combine together the confidence levels that the trusted receivers of a target message assign to the hyperbolic areas delineating an uncooperative device's probable position. Given that our scenarios are based on a closed world assumption necessitating the reduction of uncertainty rather than ignorance, the compound probability mechanism is based on a Bayesian conditioning approach. This technique enables the supportive aggregation of concurring evidence rather than its weakening through competitive probability redistribution.

Performance evaluation through simulation reveals that the compounding probability paradigm constitutes a better probabilistic model of the HPB simulation results than the simple multiplicative probability model. While the latter results in a maximum point-wise difference of up to 60% for the higher confidence levels when compared to the simulation data, our compound probability method never exceeds a 15% difference. Our model clearly yields the better probability distribution for the location estimation simulation results.
Figure 7.7: Cumulative Probability Distributions for $C = 0.95$

Figure 7.8: Cumulative Probability Distributions for $C = 0.90$
Figure 7.9: Compound and Multiplicative Cumulative Probability

Figure 7.10: Central Range Cumulative Probability
Figure 7.11: Paired Hyperbolic Area Probability Distributions for $C = 0.95$

Figure 7.12: Paired Hyperbolic Area Probability Distributions for $C = 0.90$
Chapter 8

Hyperbolic Position Bounding in Mobile WiFi/802.11 Networks

In [81], we validate the HPB simulation results obtained in Chapter 6 and further evaluate the HPB algorithm’s suitability for localizing an uncooperative transmitter. For this purpose, we describe an outdoor experiment involving five WiFi/802.11 devices, as outlined in [113]. We report on the log-normal shadowing model path loss parameters, which we compute from measured signal losses, and assess them against Liechty’s results for WiFi/802.11 networks [88]. We evaluate the performance of HPB using experimental RSS values. We assess the usability of the annuli method described in Chapter 6 by comparing its localization performance, assuming a range of EIRP values, with that of HPB. We compare the performance of HPB on the experimental RSS values with the prior simulation results.

Figure 8.1: Outdoor Experiment at Carleton University
8.1 Experiment Configuration

In order to facilitate a direct comparison between the experimental results and the simulation described in Chapter 6, our experiment is configured to closely model the simulation scenario. It thus comprises one transmitter and four receivers, as illustrated in Figure 8.1. The receiver locations are denoted in blue, and the consecutive transmitter positions in red.

We set up four fixed desktop receivers, labeled $R_1$ through $R_4$ in Figure 8.2, each equipped with a Trendnet network interface card, enabling access to the RSS values of received packets. A laptop is configured as a mobile transmitter, broadcasting packets with a transmitting power of 17 dBm using an antenna with a 7 dBi gain. Given that each receiver is also equipped with a 7 dBi gain antenna, the total EIRP is 31 dBm. The transmitter broadcasts from five separate locations, labeled $T_1$ through $T_5$ in Figure 8.2. The distances from the transmitter locations to each receiver are recorded in order to plot the path loss parameters. The transmitter antenna is situated at 1.5 m above the ground and the receiver antennas at 2.5 m.

![Figure 8.2: Experimental Grid](image)

To analyze the experimental results against the simulation results from Chapter 6, we require both grids to be comparable in size with respect to the area of the central range, i.e. the area situated between all four receivers. In the simulation scenario, the central range constitutes 16% of the simulation grid area. We therefore set the
experimental grid to be $100 \times 180 \text{ m}^2$, so that its central range also comprises 16% of the total area.

All the packets simultaneously received by the four desktop monitors are separated into two equal-sized groups. The first group is used to compute the log-normal shadowing model path loss parameters for our experiment, as presented in Section 8.2. The second group is input to HPB for localization of each packet’s originating position, as discussed in Section 8.3.

8.2 Path Loss Parameters

Our experiment varies slightly from Liechty’s previous work at the 2.4 GHz frequency because of the shorter T-R distances used. While our scenario involves a $100 \times 180 \text{ m}^2$ grid, Liechty employed a larger $500 \times 600 \text{ m}^2$ test area. We compute the path loss exponent $\eta$ and signal shadowing standard deviation $\sigma$ for our experiment and compare our values with those in [88].

![Figure 8.3: Experimental Path Loss Parameters](image)

The path loss for each packet, as the difference between the EIRP and RSS, is plotted in Figure 8.3 as a function of the logarithm of the T-R distance. We find that the best-fit path loss exponent $\eta$ equals 1.67 (depicted as a dashed line). The standard deviation associated with this value of $\eta$ is $\sigma = 3.49$. Liechty’s results yielded $\eta = 2.76$ and $\sigma = 5.62$. Although a path loss exponent of $\eta = 2$ is typically associated
with free space propagation, lower values for $\eta$ have been observed in experiments at short distances close to 100 meters [116]. Given that all the T-R distances in our experiment fall below 80 meters, our findings are consistent with previous research. A smaller shadowing standard deviation is also consistent with shorter T-R distances.

8.3 Experimental Results

The HPB success rate for every transmitter location in our experiment is computed for each of four confidence levels $C \in \{0.95, 0.90, 0.85, 0.80\}$. For $C = 0.95$, 96% of HPB executions are successful; for $C = 0.90$, the success rate is 89%; for $C = 0.85$, it is 76%; and for $C = 0.80$, it is 72%. Overall, the HPB success rate maps very closely to the corresponding confidence level, as previously observed with simulated RSS values in Chapter 6.

The success rates at individual transmitter locations, for confidence level $C = 0.95$, are illustrated in Figure 8.4. While success rates are quite high on the outer edges of the experimental grid, the successful localization of the middle transmitter position $T_2$ may suffer slightly from a large number of reflected packets, as it was within range of surrounding campus buildings on all four sides. For example, with transmitter location $T_5$, packets reflected off the Tory Building, positioned as shown in Figure 8.2, may be out of range for receivers $R_2$ and $R_3$ and not be measured. Since packets not received by all four monitors are unusable for our scenarios, they are omitted from our evaluation. In contrast, packets originating from $T_2$ may be reflected from all surrounding buildings and received by all four monitors, but at RSS values that are not commensurate with $T_2$’s relative distance to each receiver. Consequently, the success rate for transmitter location $T_2$ may have been affected.

Experimental results for candidate area sizes, given the four confidence levels, are found in Figure 8.5 and are deemed accurate with $\pm 4\%$ within a 90% statistical confidence interval. As with the simulation results outlined in Chapter 6, the intersection of hyperbolic areas decreases in size as the confidence level drops, due to the reduced amounts of signal shadowing taken into account. The correspondingly lower values of the normal distribution constant $z$ thus yield smaller hyperbolic areas. Candidate area sizes also tend to decrease with the distance from the middle of the experimental
Figure 8.4: HPB Experimental Results Success Rate for $C = 0.95$

Figure 8.5: HPB Experimental Results Candidate Area Size
grid, because a higher percentage of a centrally located area necessarily lies within the grid. Peripheral candidate areas are truncated at the edges of the grid and thus constitute a smaller percentage of the overall grid area.

8.3.1 HPB vs. Annuli Method Experimental Results

With the existing annuli method described in Chapter 6, a candidate area comprises the intersection of the minimum and maximum distances annuli around all the receivers, computed using Lemma 6.3. However, as previously observed, this candidate area may be too large to suitably pinpoint the location of an uncooperative transmitter, especially when a range of EIRP values is considered. We use our experimental scenario and measured RSS values to examine the relative performance of the annuli method and HPB.

For example, Figure 8.6 illustrates the minimum and maximum distances, $d_k^-$ and $d_k^+$ respectively, from each receiver $R_k$ to transmitter $T_4$, computed using Lemma 6.3. The reference distance $d_0$ is set to seven meters, in keeping with the reference distance used for the experiment. The dynamic EIRP range is set to $P^-$ equals 30 dBm and $P^+$ equals 38 dBm. The RSS values measured at each receiver $R_1$ to $R_4$ are -54 dBm, -50 dBm, -42 dBm and -59 dBm, respectively. The path loss at $d_0$ measured in the experiment equals 73 dB. The path loss parameters $\eta = 1.67$ and $\sigma = 3.49$ are determined in Section 8.2. The transmitter is located within each annulus with confidence level $\mathcal{C} = 0.95$, and so the associated normal distribution constant $z$ equals 1.96. The signal shadowing is thus contained within the interval $[-1.96 \times 3.49 \text{dB}, +1.96 \times 3.49 \text{dB}] = [-7 \text{dB}, +7 \text{dB}]$ with probability 0.95. For $R_1$, the minimum and maximum distances to the transmitter are 12 m and 247 m; for $R_2$, 7 m and 142 m; for $R_3$, 2 m and 47 m; and for $R_4$, 25 m and 292 m. The corresponding annuli are depicted in Figure 8.6. The dotted area represents the intersection of all annuli within the $100 \times 180 \text{ m}^2$ grid and constitutes 33% of the experimental grid area.

Figure 8.7 represents the candidate area computed by HPB for the same example scenario as in Figure 8.6. The intersection of all hyperbolic areas constructed with Theorem 6.2 yields a candidate area consisting of only 12% of the experimental grid.
Figure 8.6: Example of Intersecting Annuli for $T_4$

Figure 8.7: Example of HPB Candidate Area for $T_4$
In this case, the HPB area size is approximately 36% of the annuli candidate area shown in Figure 8.6, illustrating a clear improvement over the annuli method when a range of EIRP values is used.

Overall, for each of the four confidence levels, Figure 8.8 depicts a success rate for the annuli method higher than the corresponding confidence level, and thus greater than the HPB success rate as well. This phenomenon is due to the significantly greater candidate area size achieved with the annuli method.

Figure 8.9 illustrates the magnitude of this problem for confidence levels $C \in \{0.95, 0.90\}$. While the HPB candidate area size for $C = 0.95$ averages 26% of the experimental grid, the corresponding annuli method average candidate area size is 59% of the grid, more than twice the size obtained with HPB. Our experimental results thus confirm the limited usability of the annuli method in localizing an uncooperative transmitter.

8.3.2 HPB Experimental vs. Simulation Results

Figure 8.10 compares the HPB experimental and simulation success rates for the four confidence levels. While the two sets of results are comparable, the experimental results yield a slightly higher success rate.

Experimental and simulation candidate area sizes for $C \in \{0.95, 0.90\}$ are depicted in Figure 8.11. Inter-receiver distances for the experimental results are much shorter than for the simulation results, so fine-grained comparisons are difficult. However, within the central range, the average experimental candidate area size for $C = 0.95$ is nearly identical to the simulation results, at 26% and 27% of the total grid, respectively. For $C = 0.90$, the candidate area sizes are 20% and 21%; for $C = 0.85$, they are 15% and 16%; and for $C = 0.80$, they are 14% and 12%.

The performance of HPB along the success rate and candidate area size metrics yields consistent results in an experimental setting for T-R distances below 100 m, when compared with prior simulation results over longer distances.
Figure 8.8: HPB vs. Annuli Method Success Rate

Figure 8.9: HPB vs. Annuli Method Candidate Area Size
Figure 8.10: HPB Experimental vs. Simulation Success Rate

Figure 8.11: HPB Experimental vs. Simulation Candidate Area Size
8.4 Mobile Transmitter Tracking

Figure 8.12 illustrates an example set of HPB candidate areas computed from experimental RSS values, with each area associated with a transmitter location $T_1$ to $T_5$, for $C = 0.95$. The areas are staggered over time along the Z-axis to simulate mobility along a path from $T_1$ to $T_5$.

![Figure 8.12: Example of Candidate Areas for a Mobile Transmitter](image)

Given the size and central positioning of the $T_1$ to $T_3$ candidate areas, tracking a mobile transmitter from one of these areas to the next presents a significant challenge. However, as the transmitter moves away from the center of the experimental grid to positions $T_4$ and $T_5$, the corresponding candidate areas decrease in size. The shape and positioning of the $T_4$ and $T_5$ candidate areas, and the fact that they feature little overlap, unambiguously reveal the direction in which the transmitter has traveled from one location to the next. In this manner, HPB can provide a rudimentary means for tracking a mobile device broadcasting messages, especially when the candidate areas are small. It is expected that the inclusion of additional receivers computing more hyperbolic areas can further reduce the candidate area size and enable HPB to provide even finer-grained tracking capability. This enhancement is further investigated in Chapter 9.
8.5 Summary

We described an outdoor experiment for evaluating the hyperbolic position bounding of a mobile transmitting laptop emulating an uncooperative device in a WiFi/802.11 network, using RSS values harvested at four desktop receivers. This experiment provides a proof of concept scenario for the HPB algorithm in a practical setting, where a transmitter position is localized to a candidate area with a degree of confidence. The performance of HPB using experimental RSS values is assessed against simulation results obtained previously.

We find that the experimental results closely match the simulations along two tested metrics: the success rate in bounding a transmitter position to a candidate area, and the candidate area size. The success rate for the experimental results is found to be commensurate with the confidence level and slightly superior to that of the simulation results, especially in the experimental grid areas where the least signal reflection occurs. In terms of candidate area size, the experimental and simulation results average a candidate area of 26-27% of the total grid for confidence level $C = 0.95$, and 20-21% for confidence level $C = 0.90$. We also find that HPB can provide a coarse-grained tracking mechanism for a mobile transmitter as the computed candidate areas shift over time and space. The achievable level of tracking granularity is dependent upon the computation of sufficiently small candidate areas.

Our experiment confirms the HPB location estimation results obtained using RSS values simulated from the log-normal shadowing model. We thus conclude that this model is an appropriate tool for generating realistic simulated RSS values for use in the performance evaluation of our HPB algorithm.
Chapter 9

Hyperbolic Position Bounding in Vehicular Networks

Vehicle safety applications necessitate that each network device periodically broadcast position reports, or beacons [32]. A malicious insider generating false beacons whose digital signature is verifiable can cause serious accidents and possibly loss of life. Our HPB mechanism addresses this novel threat scenario in probabilistically delimiting the candidate location of an attack message’s originating device, assuming neither the cooperation of the attacker nor any knowledge of the EIRP. We extend the applicability of HPB to a vehicular network setting. We adapt our existing HPB algorithm for greater localization granularity and computational efficiency, as well as for the tracking of a mobile transmitter.

We demonstrate in [84] that by dynamically computing an EIRP range, we render the HPB mechanism impervious to varying power attacks, which are a known pitfall of RSS-based location estimation schemes. We propose three variations of HPB for computing sets of hyperbolic areas and the resulting candidate areas for the location of a transmitting attacker or otherwise uncooperative device. We extend HPB to include a mobile attacker tracking capability to estimate the mobility path of a transmitter in terms of location and direction of travel. We simulate a vehicular scenario with a variable number of receiving devices, and we evaluate the performance of HPB in both localizing and tracking a transmitting attacker, as a function of the number of receivers. We compare the HPB performance against existing location accuracy standards in related technologies, including the FCC guidelines for localizing a wireless handset in an emergency situation.

9.1 Mitigating Varying Power Attacks

The use of RSS reports has been criticized as a sub-optimal tool for estimating T-R distances due to their vulnerability to varying power attacks [38]. An attacker that
transmits at an EIRP other than the one expected by a receiver can appear to be
closer or farther simply by transmitting a stronger or weaker signal. Our HPB-based
algorithms are immune to such an exploit, since no fixed EIRP value is expected.
Instead, measured RSS values are leveraged to compute a likely EIRP range, as
demonstrated in Heuristic 9.1.

In order for HPB to compute a set of hyperbolic areas between pairs of receivers
upon detection of an attack message, a candidate range \([P_-, P_+]\) for the EIRP em­
ployed by the transmitting device must be dynamically estimated. We use the RSS
values registered at each receiver, as well as the log-normal shadowing model [115],
captured in Definition 6.1, for this purpose. The path loss \(L(d)\) is replaced with
its equivalent, the difference between the EIRP and the RSS measured at a given
receiver. Our strategy takes the receiver with the maximal RSS as an approximate
location for the transmitter and computes the EIRP range a device at those coordi­
nates would need to employ in order for a signal to reach the other receivers with the
RSS values measured for the attack message.

We begin by identifying the receiver measuring the maximal RSS for an attack
message. Given that this device is likely to be situated in nearest proximity to the
transmitter, we deem it the reference receiver. For every other receiving device \(R_k\),
we use the log-normal shadowing model to calculate the range of EIRP \([P_k^-, P_k^+]\) that
a transmitter would employ for a message to reach \(R_k\) with power \(RSS_k\), assuming
the transmitter is located at exactly the reference receiver coordinates. The global
EIRP range \([P^-, P^+]\) for the attack message is calculated as the intersection of all
receiver-computed ranges \([P_k^-, P_k^+]\).

**Heuristic 9.1. EIRP Range Computation.** Let \(R\) be the set of all receivers
within range of an attack message. Let \(\tilde{R}_m\) be the maximal RSS receiver and thus be
estimated as the closest receiver to the message transmitter, such that \(\tilde{R}_m \in R\) and
\(RSS_m \geq RSS_j\) for all \(R_j \in R\). Given that \(EIRP = \overline{L}(d_0) + 10\eta \log(d/d_0) + RSS + X_\sigma\)
from the log-normal shadowing model, let the EIRP range \([P_k^-, P_k^+]\) at any receiver
be determined, with confidence $C$, as:

$$P^-_k = \bar{L}(d_0) + 10\eta \log(d_{mk}/d_0) + RSS_k - z\sigma \tag{9.1}$$

$$P^+_k = \bar{L}(d_0) + 10\eta \log(d_{mk}/d_0) + RSS_k + z\sigma \tag{9.2}$$

where $d_{mk}$ is the Euclidian distance between $R_k$ and $\tilde{R}_m$, for any $R_k \in \mathbb{R} \setminus \{\tilde{R}_m\}$.

Then the estimated EIRP range $[P^-, P^+]$ employed by a transmitter is the intersection of receiver-computed EIRP intervals $[P^-_k, P^+_k]$ within which every receiver $R_k \in \mathbb{R} \setminus \{\tilde{R}_m\}$ can reach $\tilde{R}_m$. Since $P^-$ must be smaller than $P^+$, we iterate through the ascending ordered sets $\{P^-_k\}$ and $\{P^+_k\}$, for all $R_k \in \mathbb{R} \setminus \{\tilde{R}_m\}$, to find a supremum of EIRP values with minimal shadowing that is lower than an infimum of maximal shadowing EIRP values. Assuming the size of $\mathbb{R}$ is $n$, and thus the size of $\mathbb{R} \setminus \{\tilde{R}_m\}$ is $n - 1$, we compute the estimated EIRP range $[P^-, P^+]$ as follows:

1: $i \leftarrow n - 1$
2: $j \leftarrow 1$
3: while $i > 0$ and $j < n$ do
4: \hspace{1em} if $P^-_i < P^+_j$ then
5: \hspace{2em} $P^- \leftarrow P^-_i$
6: \hspace{2em} $P^+ \leftarrow P^+_j$
7: \hspace{2em} exit
8: \hspace{1em} end if
9: \hspace{1em} if $i > 1$ then
10: \hspace{2em} if $P^-_{i-1} < P^+_j$ then
11: \hspace{3em} $P^- \leftarrow P^-_{i-1}$
12: \hspace{3em} $P^+ \leftarrow P^+_j$
13: \hspace{2em} exit
14: \hspace{1em} end if
15: \hspace{1em} end if
16: $i \leftarrow i - 1$
17: $j \leftarrow j + 1$
18: end while

The only case where the pseudo-code above can fail is if every $P^-_i$ is greater than
every $\mathcal{P}_j^+$ for all $1 \leq i, j \leq n - 1$. This is impossible, since Equations (9.1) and (9.2) together indicate that for any $k$, $\mathcal{P}_k^-$ must be smaller than $\mathcal{P}_k^+$.

The log-normal shadowing model indicates that, for a fixed T-R distance, the expected path loss is constant, albeit subject to signal shadowing, regardless of the EIRP used by a transmitter. Any EIRP variation induced by an attacker translates into a corresponding change in the RSS values measured by all receivers within radio range. As a result, an EIRP range computed with Heuristic 9.1 incorporates an attacker’s power variation and is commensurate with the actual EIRP used, as are the measured RSS reports. The values cancel each other out when computing an HPB distance difference range, yielding constant values for the minimum and maximum bounds of this range, independently of EIRP variations.

**Lemma 9.1. Varying Power Effect.** Let $\mathbb{R}$ be the set of all receivers within range of an attack message. Let a probable EIRP range $[\mathcal{P}^-, \mathcal{P}^+]$ for this message be computed as set forth in Heuristic 9.1. Let the distance difference range $[\Delta d_{ij}^-, \Delta d_{ij}^+]$ between a transmitter and receiver pair $R_i, R_j$ be calculated according to Theorem 6.1. Then any increase (or decrease) in the EIRP of a subsequent message influences a corresponding proportional increase (or decrease) in RSS reports, effecting no measurable change in the range of distance differences $[\Delta d_{ij}^-, \Delta d_{ij}^+]$ estimated with a dynamically computed EIRP range.

**Proof.** Let an original EIRP range $[\mathcal{P}_{k}^-, \mathcal{P}_{k}^+]$ computed for all receivers $R_k \in \mathbb{R}$ yield an estimated global EIRP range $[\mathcal{P}^-, \mathcal{P}^+]$. Let a new varying power attack message be transmitted such that the EIRP includes a power increase (or decrease) of $\Delta \mathcal{P}$. Then for every $R_k \in \mathbb{R}$, the corresponding $\widehat{RSS}_k$ for the new attack message reflects the same change in value from the original $RSS_k$, for $\widehat{RSS}_k = RSS_k + \Delta \mathcal{P}$. Given new $\widehat{RSS}_k$ values for all $R_k \in \mathbb{R}$, the resulting EIRP range $[\widehat{\mathcal{P}}^-, \widehat{\mathcal{P}}^+]$ computed with
Heuristic 9.1 include the same change $\Delta \mathcal{P}$ over the original range of values $[\mathcal{P}^-, \mathcal{P}^+]$:

$$\hat{\mathcal{P}}^- = \sup\{ \hat{\mathcal{P}}_k^- \}$$

$$= \sup\{ \bar{L}(d_0) + 10\eta \log(d_{mk}/d_0) + \bar{R}\bar{S}\bar{S}_k - z\sigma \}$$

$$= \sup\{ \bar{L}(d_0) + 10\eta \log(d_{mk}/d_0) + R\bar{S}\bar{S}_k + \Delta \mathcal{P} - z\sigma \}$$

$$= \sup\{ \mathcal{P}_k^- + \Delta \mathcal{P} \}$$

$$= \mathcal{P}^- + \Delta \mathcal{P}$$

Conversely, we see that $\hat{\mathcal{P}}^+ = \mathcal{P}^+ + \Delta \mathcal{P}$.

As a result, the distance difference range $[\Delta d_{ij}^-, \Delta d_{ij}^+]$ for the new message is equal to the original range $[\Delta d_{ij}^-, \Delta d_{ij}^+]$:

$$\Delta \hat{d}_{ij}^- = (d_0 \times 10^{(\hat{\mathcal{P}}^- - \bar{R}\bar{S}\bar{S}_i - \bar{L}(d_0) - z\sigma)/10\eta})$$

$$- (d_0 \times 10^{(\hat{\mathcal{P}}^- - \bar{R}\bar{S}\bar{S}_j - \bar{L}(d_0) + z\sigma)/10\eta})$$

$$= (d_0 \times 10^{(\mathcal{P}^- + \Delta \mathcal{P} - \bar{R}\bar{S}\bar{S}_i - \bar{L}(d_0) - z\sigma)/10\eta})$$

$$- (d_0 \times 10^{(\mathcal{P}^- + \Delta \mathcal{P} - \bar{R}\bar{S}\bar{S}_j - \bar{L}(d_0) + z\sigma)/10\eta})$$

$$= (d_0 \times 10^{(\mathcal{P}^- - \bar{R}\bar{S}\bar{S}_j - \bar{L}(d_0) - z\sigma)/10\eta})$$

$$- (d_0 \times 10^{(\mathcal{P}^- - \bar{R}\bar{S}\bar{S}_i - \bar{L}(d_0) + z\sigma)/10\eta})$$

$$= \Delta d_{ij}^-$$

The same logic can be used to demonstrate that $\Delta \hat{d}_{ij}^+ = \Delta d_{ij}^+$.

A varying power attack is thus ineffective against HPB, as the placement of hyperbolic areas remains unchanged.

### 9.2 HPB Algorithm Variations

The HPB mechanism estimates the originating location of a single attack message from a static snapshot of a wireless network topology. Given sufficient computational efficiency, the algorithm executes in near real-time to bound a malicious insider’s position at the time of its transmission.

Hyperbolic areas constructed from Theorem 6.2 are used by HPB to compute a candidate area for the location of a malicious transmitter.
Definition 9.1. **Hyperbolic Area.** Let $\mathbb{G}$ be the set of all $(x,y)$ coordinates in the Euclidean space within radio range of a malicious transmitter. Let $\mathcal{H}_{ij}^-$ be the hyperbola computed from the minimum bound of the distance difference range between receivers $R_i$ and $R_j$ with confidence level $C$, as defined by Equation (6.3). Let $\mathcal{H}_{ij}^+$ be the hyperbola computed from the maximum bound of the distance difference range between $R_i$ and $R_j$ with the same confidence, as defined by Equation (6.4). Then we define the hyperbolic area $\mathcal{A}_{ij}$ as situated between the hyperbolas $\mathcal{H}_{ij}^-$ and $\mathcal{H}_{ij}^+$ with confidence level $C$. More formally, if $\delta(a,b)$ represents the Euclidian distance between any two points $a$ and $b$, then:

$$\mathcal{A}_{ij} = \{ p_k : \Delta d_{ij}^- \leq \delta(p_k, R_i) - \delta(p_k, R_j) \leq \Delta d_{ij}^+ \text{ for all } p_k \in \mathbb{G} \}$$

where $\Delta d_{ij}^-$ and $\Delta d_{ij}^+$ are defined in Theorem 6.1.

A set of hyperbolic areas may be computed according to three different algorithms, depending on the set of receiver pairs considered.

Definition 9.2. **Receiver Pair Set.** Let $\Omega$ be any set of unique receivers $R_k$. Then $S^\Omega$ is defined as the exhaustive set of unique, ordered receiver pairs in $\Omega$:

$$S^\Omega = \{ \{R_i, R_j\} : R_i, R_j \in \Omega \text{ and } i < j \}$$

where $s_h \neq s_k$ for all $s_h, s_k \in S^\Omega$ with $h \neq k$, and $|S^\Omega| = \binom{n}{2}$ where $n = |\Omega|$.

Our original HPB algorithm employs all possible combinations of receiver pairs to compute a set of hyperbolic areas. The intersecting space of the hyperbolic areas yields a probable candidate area for the location of a transmitter.

Algorithm 9.1. $A^\alpha$: **All-pairs Algorithm.** The all-pairs algorithm $A^\alpha$ computes hyperbolic areas between every possible pair of receivers. Let $\mathbb{R}$ be the set of all receivers within range of an attack message. Let $S^\mathbb{R}$ represent the set of all unique, ordered receiver pairs in $\mathbb{R}$, as put forth in Definition 9.2. Then the set of hyperbolic areas $\mathbb{H}^\alpha$ between all receiver pairs is stated as follows:

$$\mathbb{H}^\alpha = \{ \mathcal{A}_{ij}, \mathcal{A}_{ji} : \mathcal{A}_{ij}, \mathcal{A}_{ji} \text{ are computed as in Def. 9.1 for every } \{R_i, R_j\} \in S^\mathbb{R} \}$$
The $A^\alpha$ algorithm generates hyperbolic areas for every possible receiver pair, for a total of $\binom{n}{2}$ pairs given $n$ receivers, as put forth in Algorithm 9.1. While this approach works adequately for four receivers, additional receiving devices have the effect of dramatically increasing the computation time, as well as reducing the success rate due to the accumulated amount of signal shadowing excluded. The HPB execution time is based on the number of hyperbolic areas computed, which in turn is contingent upon the number of receivers. For $A^\alpha$, $n$ receivers locate a transmitter with a complexity of $O(n^2)$. 

An alternate algorithm $A^\beta$ aims to scale down the computational complexity by reducing the number of hyperbolic areas. We separate the set of all receivers into subsets of size $r$. Each receiver subset computes an intermediate candidate area as the intersection of the hyperbolic areas constructed from all receiver pair combinations within that subset. The final candidate area for a transmitter consists of the intersection of the intermediate candidate areas computed over all receiver subsets.

**Algorithm 9.2. $A^\beta$: r-pair Set Algorithm.** The r-pair set algorithm $A^\beta$ groups receivers in subsets of size $r$, computes intermediate candidate areas for each subset using the all-pairs approach within the subset, and yields an ultimate candidate area for a transmitter as the intersection of the receiver subset intermediate candidate areas. Let $R$ be the set of all receivers within range of an attack message. Let $\Psi$ represent the disjoint partition of $(m - 1)$ sets of $r$ receivers, with the $m^{th}$ element of $\Psi$ containing the remaining receivers:

$$\Psi = \{ \psi_k : \psi_k \subseteq R \text{ for } 1 \leq k \leq m \}$$

$$\text{and } |\psi_k| = r \text{ if } k < m$$

$$\text{and } 2 \leq |\psi_k| \leq r \text{ if } k = m$$

where $\psi_h \cap \psi_k = \emptyset$ for all $\psi_h, \psi_k \in \Psi$ with $h \neq k$. Let $S^{\psi_k}$ represent the set of all unique, ordered receiver pairs in a given set of receivers $\psi_k \in \Psi$, as put forth in Definition 9.2.
Then the set of hyperbolic areas $\mathbb{H}^\beta$ computed for sets of $r$ receivers is stated as follows:

$$
\mathbb{H}^\beta = \{ A_{ij}, A_{ji} : A_{ij}, A_{ji} \text{ are computed as in Def. 9.1 for every } \{R_i, R_j\} \in S^{\psi_k} \text{ for all } \psi_k \in \Psi \}
$$

For the $\mathbb{A}^\beta$ algorithm, the number of hyperbolic areas depends on the set size $r$, as well as the number of receivers $n$. Thus $\mathbb{A}^\beta$ locates a transmitter with a complexity of $\left(\frac{n}{r} + 1\right) \times \binom{r}{2} \approx O(n)$. For a small value of $r$, for example $r = 4$, the execution time is proportional to at most $\left(\frac{3n}{2} + 6\right)$.

A third HPB algorithm, the perimeter-pairs variation $\mathbb{A}^\gamma$, is proposed to bound the geographic extent of a candidate area within an approximated transmission range, based on the coordinates of the receivers situated farthest from a signal source. We establish a rudimentary perimeter around a transmitter’s estimated radio range, with the logical center of this range calculated as the centroid of all receiver coordinates. The range is partitioned into four quadrants from the center, along two perpendicular axes. Four perimeter receivers are identified as the farthest in each quadrant from the center. Hyperbolic areas are computed between all combinations of perimeter receiver pairs, as well as between every remaining non-perimeter receiver and the perimeter receivers in the other three quadrants.

**Definition 9.3. Receiver Quadrants.** Let $\mathbb{R}$ be the set of all receivers within range of an attack message. Let $R_X = (x_c, y_c)$ be the centroid of all $R_i \in \mathbb{R}$. Then we define the set of receiver quadrants $\mathbb{Q}$ as the disjoint set of all receivers $R_i \in \mathbb{R}$ partitioned into four quadrants from the centroid $R_X$:

$$
\mathbb{Q} = \{Q_k : Q_k = \{R_i : R_i \in \mathbb{R} \text{ and } R_i = (x_i, y_i) \text{ and } x_i \geq x_c \text{ and } y_i \geq y_c \text{ for } k = 1 \}
$$

$$
\text{x_i < x_c \text{ and } y_i \geq y_c \text{ for } k = 2} \}
$$

$$
\text{x_i < x_c \text{ and } y_i < y_c \text{ for } k = 3} \}
$$

$$
\text{x_i \geq x_c \text{ and } y_i < y_c \text{ for } k = 4} \}
$$

**Definition 9.4. Perimeter Receivers.** Let $\mathbb{Q}$ be the set of receiver quadrants associated with a set of receivers $\mathbb{R}$ with centroid $R_X$, as stated in Definition 9.3.
Then we define the set of perimeter receivers \( N \) as containing one receiver \( \rho_k \) for each of the four quadrants in \( Q \), such that \( \rho_k \) is the farthest receiver from the receiver centroid \( R_X \) in quadrant \( k \):

\[
N = \{\rho_k : \rho_k = q_i \text{ such that } q_i \in Q_k \text{ and } \\
\delta(q_i, R_X) \geq \delta(q_j, R_X) \text{ for all } q_j \in Q_k \\
\text{ for all } Q_k \in Q\}
\]

where \( \delta(a, b) \) represents the Euclidean distance between any two points \( a \) and \( b \).

**Algorithm 9.3. **\( A^\gamma \): Perimeter-pairs Algorithm. **The perimeter-pairs algorithm \( A^\gamma \) partitions a transmitter's radio range into four quadrants. Four perimeter receivers are determined. Hyperbolic areas are computed between all pairs of perimeter receivers, as well as between every perimeter receiver and the non-perimeter receivers of other quadrants. Let \( R \) be the set of all receivers within range of an attack message. Let \( Q \) be the associated set of receiver quadrants, as stated in Definition 9.3, and let \( N \) be the corresponding set of perimeter receivers, as found in Definition 9.4. Also let the set of non-perimeter receivers in a given quadrant be determined as all receivers in that quadrant other than the perimeter receiver:

\[
\overline{N} = \{\overline{\rho}_k : \overline{\rho}_k = \{Q_k \setminus \{\rho_k\}\} \text{ where } \rho_k \in N \text{ for every } Q_k \in Q\}
\]

Let \( S^N \) represent the set of all unique, ordered perimeter receiver pairs, as put forth in Definition 9.2.

Then the set of hyperbolic areas \( H^\gamma \) is stated as follows:

\[
H^\gamma = \left\{ A_{ij}, A_{ji} : A_{ij}, A_{ji} \text{ are computed as in Def. 9.1} \right\} \\
\text{for every } \{R_i, R_j\} \in \{S^N \cup \{(R_i, R_j) : R_i = \rho_k \text{ for every } \rho_k \in N \text{ and } R_j \in \overline{\rho}_m \} \right\}
\]

For example, Figure 9.1 illustrates a transmitter \( T \) and a set of receivers. The grid is partitioned into four quadrants from the computed receiver centroid. The set of perimeter receivers, as the farthest receivers from the centroid in each quadrant (I
to IV), form a rudimentary bounding area for the location of the transmitter. The $A^\gamma$ algorithm computes hyperbolic areas between all pairs of perimeter receivers, in this case between all possible pairs in $\mathcal{N} = \{R_3, R_4, R_7, R_5\}$. Additional receiver pairs are formed between the remaining non-perimeter receivers $\{R_1, R_2, R_6, R_8\}$ and the perimeter receivers of other quadrants. Receiver $R_6$, for instance, is situated in quadrant II, so it is included in a receiver pair with each perimeter receiver in $\{R_3, R_7, R_5\}$.

![Figure 9.1: Example of Perimeter Receivers](image)

In terms of complexity, the $A^\gamma$ algorithm is equivalent to $A^\beta$. Given $n$ receivers and four perimeter receivers such that $|\mathcal{N}| = 4$, $A^\gamma$ executes in time $\binom{4}{2} + 3(n - 4) = 3n - 6 \approx O(n)$.

The candidate area for the location of a malicious transmitter is computed as the intersection of a set of hyperbolic areas, $H^\alpha$, $H^\beta$, or $H^\gamma$, determined according to Algorithm 9.1, 9.2 or 9.3.

**Definition 9.5. Candidate Area.** Let $\mathbb{G}$ be the set of all $(x, y)$ coordinates in our sample Euclidean space. Let $\mathbb{V} \subseteq \mathbb{G}$ be the subset of all coordinates situated on the road layout of a vehicular scenario. Then the grid candidate area $GA^\ell$, where $\ell \in \{\alpha, \beta, \gamma\}$, is defined as the subset of grid points in $\mathbb{G}$ situated in the intersection
of every hyperbolic area computed according to Algorithm $A^\alpha$, $A^\beta$ or $A^\gamma$.

$$GA^\ell = \{ p_k : p_k \in G \text{ and } p_k \in \bigcap_{h=1}^{m} A_h \subset \mathbb{H}^\ell$$

where $\ell \in \{ \alpha, \beta, \gamma \}$ and $m = |\mathbb{H}^\ell| \}$

Similarly, the vehicular candidate area $VA^\ell$, where $\ell \in \{ \alpha, \beta, \gamma \}$, is defined as the subset of vehicular layout points in $V$ situated in the intersection of every hyperbolic area computed according to Algorithm $A^\alpha$, $A^\beta$ or $A^\gamma$.

$$VA^\ell = \{ p_k : p_k \in V \text{ and } p_k \in \bigcap_{h=1}^{m} A_h \subset \mathbb{H}^\ell$$

where $\ell \in \{ \alpha, \beta, \gamma \}$ and $m = |\mathbb{H}^\ell| \}$

While a candidate area contains a malicious transmitter with probability $C$, the tracking of a mobile device requires that a unique point in Euclidian space be deemed the likeliest position for the attacker. In free space, we can use the centroid of a candidate area, which is calculated as the average of all the $(x, y)$ coordinates in this area. In a vehicular scenario, we use the road location closest to the candidate area centroid.

**Definition 9.6. Centroids.** The grid centroid of a given $GA$, denoted as $G_X$, consists of the average $(x, y)$ coordinates of all points within the $GA$.

$$G_X = (x_G, y_G), \text{ such that } x_G = \frac{\sum_{i=1}^{GA} x_i}{|GA|} \text{ and } y_G = \frac{\sum_{i=1}^{GA} y_i}{|GA|},$$

for all $p_i = (x_i, y_i) \in GA$

The vehicular centroid of a given $VA$, represented as $V_X$, is the closest vehicular point to the average coordinates of all points within the $VA$.

$$V_X = v_k, \text{ such that } v_k \in V \text{ and } p_h = (x_V, y_V), \text{ where }$$

$$x_V = \frac{\sum_{i=1}^{VA} x_i}{|VA|} \text{ and } y_V = \frac{\sum_{i=1}^{VA} y_i}{|VA|}, \text{ for all } p_i = (x_i, y_i) \in VA,$$

and $\delta(p_h, v_k) \leq \delta(p_h, v_j)$, for all $v_j \in V$
Further work is required to ascertain whether the pure centroid of a HPB candidate area represents the best estimator for a transmitter position. While this approach constitutes a starting point, better performing schemes may be devised in future research.

9.3 Tracking a Mobile Attacker

We extend HPB to approximate the path followed by a mobile attacker, as it continues transmitting. By computing a new candidate area for each attack message received, a malicious node can be tracked using a set of consecutive candidate positions and the direction of travel inferred between these points. We establish a mobility path in our vehicular scenario as a sequence of vehicular layout \((x, y)\) coordinates over time, along with a mobile transmitter’s direction of travel at every point.

**Definition 9.7.** A mobility path \(P\) is defined as a set of consecutive coordinates \(p_i = (x_i, y_i)\) and angles of travel \(\theta_i\) over a time interval \(T\).

\[
P = \{ \{p_i, \theta_i\} : p_i = (x_i, y_i) \text{ is the transmitter location at } t_i \in T \text{ and } \theta_i = \text{atan2}(y_i - y_{i-1}, x_i - x_{i-1}) \}
\]

where \text{atan2} is an inverse tangent function returning values over the range \([-\pi, +\pi]\) to take direction into account.\(^1\)

In order to approximate the dynamically changing position of an attacker, we discretize the time domain \(T\) into a series of time intervals \(t_i\). At each discrete \(t_i\), we sample a snapshot of the vehicular network topology consisting of a set of receiving devices and their locations. Our approach is analogous to the discretization phase in digital signal processing, where a continuous analog radio signal is sampled periodically for conversion to digital form. We thus estimate the mobility path \(P\) taken by an attacker by executing an HPB algorithm for an attack message received at every interval \(t_i\) over a time period \(T\). The vehicular centroids of the resulting candidate areas constitute the estimated attacker positions, and the angle from one estimated point to the next determines the approximated direction of travel.

\(^1\)As first defined for the Fortran 77 programming language [3]
Algorithm 9.4. Mobile Attacker Tracking. Let $\mathcal{M}$ be the set of consecutive attack messages received over a time interval. Then the estimated mobility path $\hat{\mathcal{P}}$ of a transmitter over the message base $\mathcal{M}$ is computed as follows:

\[
\hat{\mathcal{P}} = \left\{ (\hat{\mathbf{p}}_i, \hat{\theta}_i) : \hat{\mathbf{p}}_i = (\hat{x}_i, \hat{y}_i) = Vx_i \text{ for } m_i \in \mathcal{M} \right. \\
\left. \quad \text{and } \hat{\theta}_i = \arctan2(\hat{y}_i - \hat{y}_{i-1}, \hat{x}_i - \hat{x}_{i-1}) \right\}
\]

For every attack message $m_i \in \mathcal{M}$, an estimated transmitter location $\hat{\mathbf{p}}_i$ must be determined. An execution of HPB using the RSS values corresponding to $m_i$ yields a vehicular candidate area $VA_i$, as put forth in Definition 9.5. The road centroid of $VA_i$ is computed as $Vx_i$, according to Definition 9.6. It is by definition the closest point in the vehicular layout to the averaged center of the $VA_i$, and thus the natural choice for an estimated value $\hat{\mathbf{p}}_i$ of the true transmitter location $\mathbf{p}_i$. The direction of travel of a transmitter is stated in Definition 9.7 as the angle between consecutive positions in Euclidean space. We follow the same logic to compute the estimated direction of travel $\hat{\theta}_i$ between transmitted messages $m_{i-1}$ and $m_i$ as the angle between the corresponding estimated positions $\hat{\mathbf{p}}_{i-1}$ and $\hat{\mathbf{p}}_i$.

Example. Figure 9.2 depicts an example mobility path of a malicious insider, with consecutive traveled points labeled from 1 to 20. The transmitter broadcasts an attack message at every fourth location, labeled as points 4, 8, 12, 16 and 20.

For each attack message, we execute the $A^\gamma$ HPB variation, for confidence level $C = 0.95$, using eight randomly positioned receivers, and a vehicular candidate area $VA^\gamma$ is computed. The estimated locations and directions of travel are depicted in Figure 9.3. The initial point’s direction of travel cannot be estimated, as there is no previous point from which to ascertain a traveled path. In this example, point 4 is localized at 100 meters from its true position, points 8, 16 and 20 at 25 meters, while point 12 is found in its exact location.

9.4 Performance Evaluation

We describe a simulated vehicular scenario to evaluate the localization and tracking performance of the extended HPB mechanisms described in Section 9.2. In order to
Figure 9.2: Example of Attacker Mobility Path

Figure 9.3: Example of Mobile Attacker Localization
model a mobile attacker transmitting at 2.4 GHz, we employ the log-normal shadowing model [115] to generate simulated RSS values at a set of receivers, taking into account an independently random amount of signal shadowing experienced at each receiving device. According to Rappaport [115], the log-normal shadowing model has been used extensively in experimental settings to capture radio signal propagation characteristics, in both indoor and outdoor channels, including in mobility scenarios. In Chapters 6 and 8, we evaluated HPB results with both log-normal shadowing simulated RSS values and RSS reports harvested from an outdoor field experiment at 2.4 GHz. We found that the simulated and experimental location estimation results are nearly identical, indicating that at this frequency, the log-normal shadowing model is an appropriate tool for generating realistic RSS values.

We compare the success rates of the A$^\alpha$, A$^\beta$ and A$^\gamma$ algorithms at estimating a malicious transmitter’s location within a candidate area, as well as the relative sizes of the grid and vehicular candidate areas. We model a mobile transmitter’s path through a vehicular scenario and assess the success in tracking it by measuring the distance between the actual and estimated positions, in addition to the difference between the approximated direction of travel and the real one.

Figure 9.4: Vehicular Scenario – Richmond, Ontario
9.4.1 Hyperbolic Position Bounding of Vehicular Devices

Our simulation uses a one square kilometer urban grid, as depicted in Figure 9.4. We evaluate the all-pairs $A^\alpha$, 4-pair set $A^\beta$ and perimeter-pairs $A^\gamma$ HPB algorithms with four, eight, 16 and 32 receivers. In each HPB execution, four of the receivers are fixed Road-Side Units (RSUs) stationed at intersections. The remaining receivers are randomly positioned On-Board Units (OBUs), distributed uniformly on the grid streets. Every HPB execution also sees a transmitter placed at a random road position within the inner square of the simulation grid. We assume that in a sufficiently dense urban setting, RSUs are positioned at most intersections. As a result, any transmitter location is geographically surrounded by four RSUs within radio range. For each defined number of receivers and two separate confidence levels $C \in \{0.95, 0.90\}$, the HPB algorithms, $A^\alpha$, $A^\beta$ and $A^\gamma$, are executed 1000 times. For every execution, RSS values are generated for each receiver from the log-normal shadowing model. We adopt existing experimental path loss parameter values from large scale measurements gathered at 2.4 GHz by Liechty et al. [88, 89]. From $\eta = 2.76$ and a signal shadowing standard deviation $\sigma = 5.62$, we augment the simulated RSS values with an independently generated amount of random shadowing to every receiver in a given HPB execution. Since the EIRP used by a malicious transmitter is unknown, a probable range is computed according to Heuristic 9.1.

For every HPB execution, whether the $A^\alpha$, $A^\beta$ or $A^\gamma$ algorithm is used, we gather three metrics: the success rate in localizing the transmitter within a computed candidate area $GA$; the size of the unconstrained candidate area $GA$ as a percentage of the one square kilometer grid; and the size of the candidate area restricted to the vehicular layout $VA$ as a percentage of the grid. The success rate and candidate area size results we obtain are deemed accurate $\pm 2\%$ and $\pm 0.8\%$ respectively, within a 90% statistical confidence interval. The average HPB execution times for each algorithm on an HP Pavilion laptop with an AMD Turion 64 x 2 dual-core processor are shown in Table 9.1. As expected from our complexity analysis, the $A^\alpha$ variation is markedly slower, and the computational costs increase as additional receivers participate in the location estimation effort. For example in the case of eight receivers, a single execution of $A^\gamma$ takes 23 ms, while $A^\alpha$ requires over 100 ms.
Table 9.1: Average HPB Variations Execution Time (seconds)

<table>
<thead>
<tr>
<th># Rcvrs</th>
<th>$A^\gamma$</th>
<th>$A^\beta$</th>
<th>$A^\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std Dev.</td>
<td>Mean</td>
</tr>
<tr>
<td>4</td>
<td>0.005</td>
<td>0.000</td>
<td>0.023</td>
</tr>
<tr>
<td>8</td>
<td>0.023</td>
<td>0.001</td>
<td>0.045</td>
</tr>
<tr>
<td>16</td>
<td>0.075</td>
<td>0.001</td>
<td>0.090</td>
</tr>
<tr>
<td>32</td>
<td>0.215</td>
<td>0.059</td>
<td>0.195</td>
</tr>
</tbody>
</table>

The comparative success rates of the $A^\alpha$, $A^\beta$ and $A^\gamma$ approaches are illustrated in Figure 9.5, for confidence level $C = 0.95$. While $A^\gamma$ exhibits the best localization success rate, every algorithm sees its performance degrade as more receivers are included. With four receivers for example, all three variations successfully localize a transmitter 94-95% of the time. However with 32 receivers, $A^\gamma$ succeeds in 79% of the cases, while $A^\beta$ and $A^\alpha$ do so in 71% and 50% of executions. Given that each receiver pair takes into account an amount of signal shadowing based on the confidence level $C$, it also probabilistically ignores a portion $(1 - C)$ of the shadowing. As more receivers and thus more receiver pairs are added, the error due to excluded shadowing accumulates. The results obtained for confidence level $C = 0.90$ follow the same trend, although the success rates are slightly lower.

![Figure 9.5: HPB Variations Success Rate for $C = 0.95$](image)

Figures 9.6 and 9.7 show the grid and vehicular candidate area sizes associated
with our simulation scenario, as computed with algorithms $A^\alpha$, $A^\beta$ and $A^\gamma$, for confidence level $C = 0.95$. The size of the grid candidate area $GA$ corresponds to 21% of the simulation grid, with four receivers, for both $A^\beta$ and $A^\alpha$, while $A^\gamma$ narrows the area to only 7%. In fact, the $A^\gamma$ approach exhibits a $GA$ size that is independent of the number of receivers. Yet for $A^\beta$ and $A^\alpha$, the $GA$ size is noticeably lower with additional receivers. This finding reflects the use of perimeter receivers with $A^\gamma$. These specialized receivers serve to restrict the GA to a particular portion of the simulation grid, even with few receivers. However, this variation does not fully exploit the presence of additional receiving devices, as these only support the GA determined by the perimeter receivers. The size of the vehicular candidate area $VA$ follows the same trend, with a near constant size of 0.64% to 1% of the grid for $A^\gamma$, corresponding to a localization granularity within an area less than 100 m x 100 m, assuming the transmitter is aboard a vehicle traveling on a road. The $A^\beta$ and $A^\alpha$ algorithms compute vehicular candidate area sizes that decrease as more receivers are taken into account, with $A^\alpha$ yielding the best localization granularity. But even with four receivers, $A^\beta$ and $A^\alpha$ localize a transmitter within a vehicular layout area of 1.6% of the grid, or 125 m x 125 m.

Generally, both the $GA$ and $VA$ sizes decrease as the number of receivers increase, since additional hyperbolic areas pose a higher number of constraints on a candidate area, thus decreasing its extent. We see in Figures 9.6 and 9.7 that $A^\beta$ consistently yields larger candidate areas than $A^\alpha$ for the same reason, as $A^\alpha$ generates a significantly greater number of hyperbolic areas. For example, while $A^\alpha$ computes an average $GA^\alpha$ of 10% and 3% of the simulation grid with eight and 16 receivers, $A^\beta$ yields areas of 15% and 9% respectively. By contrast, $A^\gamma$ yields a $GA$ size of 5-6%, but its reliability is greater, as demonstrated by the higher success rates achieved. The nearly constant 5% $GA$ size computed with $A^\gamma$ has an average success rate of 81% for 16 receivers, while the 9% $GA$ generated by $A^\beta$ is 79% reliable and the 3% $GA$ obtained with $A^\alpha$ features a dismal 68% success rate. Indeed, Figures 9.5 and 9.6 taken together indicate that smaller candidate areas provide increased granularity at the cost of lower success rates, and thus decreased reliability. This phenomenon is consistent with the intuitive expectation that a smaller area is less likely to contain
the transmitter.

9.4.2 Tracking a Vehicular Device

We generate 1000 attacker mobility paths $P$, as stipulated in Definition 9.7, of 20 consecutive points evenly spaced at every 25 meters. Each path begins at a random start location along the central square of the simulation grid depicted in Figure 9.4. We keep the simulated transmitter location within the area covered by four fixed RSUs, assuming that an infinite grid features at least four RSUs within radio range of a transmitter. The direction of travel for the start location is determined randomly. Each subsequent point in the mobile path is contiguous to the previous point, along the direction of travel. Upon reaching an intersection in the simulation grid, a direction of travel is chosen randomly among the ones available from the current position, excluding the reverse direction.

The $A^a$, $A^b$ and $A^g$ algorithms are executed at every fourth point $p_i$ of each mobility path $P$, corresponding to a transmitted attack signal at every 100 meters. The algorithms are executed for confidence levels $C \in \{0.95, 0.90\}$, with each of four, eight, 16 and 32 receivers. In every case, the receivers consist of four static RSUs, and the remainder are OBUs randomly placed at any point on the simulated roads.

For each execution of $A^a$, $A^b$ and $A^g$, a vehicular candidate area $VA$ is computed, and its centroid $V \chi$ is taken as the probable location of the transmitter, as described in Algorithm 9.4. Two metrics are aggregated over the executions: the root mean square location error, as the distance in meters between the actual transmitter location $p_i$ and its estimated position $\hat{p}_i = V \chi_i$; and the root mean square angle error between the angle of travel $\theta_i$ for each consecutive actual transmitter location and the angle $\hat{\theta}_i$ computed for the approximated locations.

The location error for the $A^a$, $A^b$ and $A^g$ algorithms, given confidence level $C = 0.95$, is illustrated in Figure 9.8. As expected, the smaller $VA$ sizes achieved with a greater number of receivers for $A^a$ and $A^b$ correspond to a more precise transmitter localization. The location error associated with the $A^a$ algorithm is smaller, compared to $A^b$, for the same reason. Correspondingly, the nearly constant $VA$ size obtained with $A^g$ yields a similar result for the location error. For instance with confidence
Figure 9.6: HPB Variations Grid Candidate Area Size for $C = 0.95$

Figure 9.7: HPB Variations Vehicular Candidate Area Size for $C = 0.95$
level $C = 0.95$, eight and 16 receivers produce a location error of 114 and 79 meters respectively with $A^\alpha$, but of 121 and 102 meters with $A^\beta$. The location error with $A^\gamma$ is once more nearly constant, at 96 and 91 meters. The use of all receiver pairs to compute a VA with $A^\alpha$ allows for localization that is up to 40-50% more precise than grouping the receivers in sets of four or relying on perimeter receivers when 16 or 32 receiving devices are present. Despite its granular localization performance, the $A^\alpha$ approach works best with large numbers of receivers, which may not consistently be realistic in a practical setting. Another important disadvantage of the $A^\alpha$ approach lies in its large complexity of $O(n^2)$ for $n$ receivers, when compared to $A^\beta$ and $A^\gamma$ with a complexity of $O(n)$, as discussed in Section 9.2.

Figure 9.9 plots the root mean square location error in terms of VA size for the three algorithms. While $A^\alpha$ and $A^\beta$ yield smaller VAs for a large number of receivers, the VAs computed with $A^\gamma$ offer more precise localization with respect to their size. For example, a 0.7% VA size obtained with $A^\gamma$ features a 96 meter location error, while a similar size VA computed with $A^\beta$ and $A^\alpha$ generate a 102 and 114 meter location error, respectively.

The error in estimating the direction of travel exhibits little variation in terms of number of receivers and choice of HPB algorithm, as shown in Figure 9.10. With eight and 16 receivers, for confidence level $C = 0.95$, $A^\beta$ approximates the angle of travel between two consecutive points within 77° and 71° respectively, whereas $A^\alpha$ estimates it within 76° and 63°. $A^\gamma$ exhibits a slightly higher direction error at 76° and 77°. It should be noted that for all three algorithms, for all numbers of receivers, the range of angle errors only spans 14°. So while the granularity of localization is contingent upon the HPB methodology used and the number of receivers, the three variations perform similarly in estimating the general direction of travel.

### 9.4.3 Discussion

The location error results of Figure 9.8 shed an interesting light on the HPB success rates discussed in Section 9.4.1. For example in the presence of 32 receivers, for confidence level $C = 0.95$, only 50% of $A^\alpha$ executions yield a candidate area containing a malicious transmitter, as shown in Figure 9.5. Yet the same scenario localizes a
Figure 9.8: HPB Variations Location Error for $C = 0.95$

Figure 9.9: HPB Variations Location Error for Vehicular Candidate Area Size
transmitter with a root mean square location error of 45 meters of its true location, whether it lies within the corresponding candidate area or not. This indicates that while a candidate area may be computed in the wrong position, it is in fact rarely far from the correct transmitter location. This may be a result of our strict definition of a successful execution, where only a candidate area in the intersection of all hyperbolic areas is considered. We have observed in our simulations that a candidate area may be erroneous solely because of a single misplaced hyperbolic area, which results in either a wrong location or an empty candidate area. In our simulations tracking a mobile attacker, we notice that while $A^\gamma$ and $A^\beta$ generate an empty $VA$ for 10% and 14% of executions, $A^\alpha$ does so in 31% of the cases. This phenomenon is likely due to the greater number of hyperbolic areas generated with the $A^\alpha$ approach and the subsequent greater likelihood of erroneously situated hyperbolic areas. While the success rates depicted in Figure 9.5 omit the executions yielding empty candidate areas as inconclusive, future work includes devising a heuristic to re-compute a set of hyperbolic areas in the case where their common intersection is empty.

Additional simulations have been carried out to assess HPB performance at 5.9 GHz, the frequency associated with DSRC/WAVE vehicular communications [8]. For this purpose, we use the average path loss exponent and signal shadowing standard deviation determined by Kunisch and Pamp [75] for V2V communications at 5.9 GHz, with $\eta = 1.75$ and $\sigma = 3.3$. These path loss parameter values are consistent with the ones measured by Paier et al. [106] for 5.2 GHz. We find that overall, the HPB success rates at 5.9 GHz are within 5% of those depicted for the 2.4 GHz frequency, while the grid candidate area sizes are within less than 1%. We thus surmise that HPB yields similar results in the gigahertz frequency band common to emerging wireless platforms.

In comparing the location accuracy of HPB with related technologies, we find that, for example, differential GPS devices can achieve less than 10 meter accuracy. However, this technology is better suited to self-localization efforts relying on a device’s assistance, and cannot be depended upon for the position estimation of a non-cooperative adversary. The FCC has set forth regulations for the network-based localization of wireless handsets in emergency 911 call situations. Service providers
are expected to locate a calling device within 100 meters 67% of the time and within 300 meters in 95% of cases [52]. In the minimalist case involving four receivers, the HPB perimeter-pairs variation \( A^7 \) localizes a transmitting device with a root mean square location error of 107 meters. This translates into a location accuracy of 210 meters in 95% of cases and of 104 meters in 67% of executions. While the former case is fully within FCC guidelines, the latter is very close. With a larger number of receivers, for example eight receiving devices, \( A^7 \) yields an accuracy of 188 meters 95% of the time and of 93 meters in 67% of cases. Although HPB is designed for the location estimation of an uncooperative transmitter such as a malicious insider, its use may be extended to additional applications such as 911 call origin localization, given that its performance closely matches the FCC requirements for emergency services.

9.5 Summary

We have extended our HPB mechanism to localize the originator of an attack signal within a vehicular network. We have devised two additional HPB-based approaches to compute hyperbolic areas between pairs of trusted receivers by grouping them in sets and establishing perimeter receivers. We demonstrate that due to the dynamic computation of a probable EIRP range utilized by an attacker, our HPB algorithms are impervious to varying power attacks. We have extended the HPB algorithms to track the location of a mobile attacker transmitting along a traveled path.

The performance of all three HPB variations is evaluated in a vehicular scenario. We find that the grouped receivers method yields a localization success rate up to 11% higher for a 6% increase in candidate area size over the all-pairs approach. We also observe that the perimeter-pairs algorithm provides a more constant candidate area size, independently of the number of receivers, for a success rate up to 13% higher for a 2% increase in candidate area size over the all-pairs variation. We conclude that the original HPB mechanism using all pairs of receivers produces a smaller localization error than the other two approaches, when a large number of receiving devices are available. We observe that for a confidence level of 95%, the former approach localizes a mobile transmitter with a granularity as low as 45 meters, up to 40-50% more
precisely than the grouped receivers and perimeter-pairs methods. However, the computational complexity of the all-pairs variation is significantly greater, and its performance with fewer receivers is less granular than the perimeter-pairs method. Of the two approaches with complexity $O(n)$, the perimeter-pairs method yields a success rate up to 8% higher for consistently smaller candidate area sizes, location and direction errors. In general, the $A^2$ variation produces superior results for better computational efficiency.

In a vehicular scenario, we achieve a root mean square location error of 107 meters with four receivers and of 96 meters with eight receiving devices. This granularity is sufficient to satisfy the FCC-mandated location accuracy regulations for emergency 911 services, thereby demonstrating the suitability of HPB for estimating the candidate location of a vehicular network malicious insider and for tracking such a device as it moves throughout the network. Our HPB mechanism may also be adaptable to a wide range of applications involving network-based device localization assuming neither target node cooperation nor knowledge of the EIRP.
Figure 9.10: HPB Variations Direction of Travel Angle Error for $C = 0.95$
Chapter 10

Relative Span Weighted Localization

As an open-form location estimation solution, HPB computes a candidate area, rather than a unique point, to determine the position of an uncooperative transmitter. While such approaches are sufficiently granular in a number of scenarios, other applications require a closed-form solution yielding a single point in Euclidian space for the location of a transmitter. In Chapter 9, we approximate an exact transmitter position from a HPB candidate area as the centroid of all the points situated in this area. However, we have observed through simulations that the centroid of a multiple hyperbolic area intersection is often not the best indicator for a transmitter position, as the hyperbolic area shapes may be highly influenced by a single receiver experiencing an amount of signal shadowing greatly different from the fluctuations at other receiving devices. We therefore investigate an additional, non-HPB location estimation approach that yields a single point solution for the location of an uncooperative transmitter. As with HPB, we assume that a single target message is received at a number of trusted nodes. We also presume neither the cooperation of the transmitting device, nor knowledge of the EIRP used to emit the target message.

In a sufficiently densely populated wireless network, the source location of a given message may be approximated from the coordinates of receiving devices, assuming an omnidirectional propagation pattern. We propose two localization algorithms [85] that estimate a transmitting node's position as the weighted average of receiver coordinates, assuming a single message is received from the target node. We compute a RSS span as the difference between the maximum and minimum RSS values for the transmitted message over all receivers. We assign greater weight to the receiver coordinates whose RSS value is closer to the maximum of the RSS span and thus closer to the transmitter. Conversely, lesser weight is ascribed to receivers with lower RSS values, as they are deemed farther from the transmitter.
We describe a *Relative Span Weighted Localization* (RWL) mechanism, where the concept of a weighted moving average is adapted to provide a linear mapping between the weight assigned to a receiver’s coordinates and the relative placement of its RSS value within the overall RSS span. We further propose an exponential variation of RWL, dubbed *Relative Span Exponential Weighted Localization* (REWL). This approach is conceptually related to an exponential moving average and relies on an exponential weight correspondence between a receiver’s coordinates and its relative situation within the RSS span. We evaluate the RWL and REWL algorithms using simulated RSS reports featuring a variety of node densities, number of receivers, and amount of signal shadowing representative of environment-based RSS fluctuations. We also test our localization mechanisms with RSS values harvested from an outdoor field experiment. We find that the exponentially weighted variation achieves better results and that, except for cases with a small number of receivers and a large amount of signal shadowing, our mechanism meets the E911 mandated location accuracy requirements.

10.1 Centroid Localization

We outline the centroid localization approaches on which our novel algorithms are based, and introduce the notation used throughout the description of our mechanisms.

**Notation.** The estimated coordinates of the transmitter we are striving to locate are denoted as $\hat{p} = (\hat{x}, \hat{y})$. Each receiver $R_i$ is situated at a point of known coordinates $p_i = (x_i, y_i)$. For the sake of simplicity in our algorithm descriptions, we depict operations on receiver points $p_i$. In fact, two separate calculations occur. The approximated $\hat{x}$ coordinate is computed from all the receiver $x_i$ coordinates, and $\hat{y}$ is calculated from the $y_i$ coordinates.

Given a set of known points $p_i$ in a Euclidian space, for example a number of receivers within radio range of a target transmitter to be localized, Bulusu *et al.* [30] approximate the location $\hat{p}$ of a node from the centroid of the known points $p_i$ as follows:

$$\hat{p} = \frac{1}{n} \times \sum_{i=1}^{n} p_i$$  \hspace{1cm} (10.1)
where \( n \) represents the number of points.

In the simple Centroid Localization (CL) approach, all points are assumed to be equally near the target node. Blumenthal et al. [23] argue that some points are more likely than others to be close to target node. Their Weighted Centroid Localization (WCL) scheme aims to improve localization accuracy by assigning greater weight to those points which are estimated to be closer to the target node and less weight to the farther points. The weighted centroid is thus computed as:

\[
\hat{p} = \frac{\sum_{i=1}^{n} (w_i \times p_i)}{\sum_{i=1}^{n} w_i}
\]

with

\[
w_i = \frac{1}{(d_i)^g}
\]

where \( d_i \) is the known distance between the target node and point \( p_i \), and the exponent \( g \) influences the degree to which remote points participate in estimating the target location \( \hat{p} \). Values of \( g \) are determined manually, with Blumenthal et al. and Schuhmann et al. [127] promoting different optimal values, depending on the experimental setting.

### 10.2 Localization Algorithms

Assuming an uncooperative node, we cannot presume to know \emph{a priori} the set of T-R distances \( d_i \) or the optimal value of \( g \) in a given outdoor environment. Further, we cannot estimate values of \( d_i \) from the log-normal shadowing model [115], as the transmitter EIRP may not be known. We therefore introduce the concept of relative span weighted localization in order to estimate the location of a transmitter with minimal information available at a set of receivers. Our approach adapts the concept of moving average from a weighting method over time and applies it to WCL in the space domain. But rather than ascribing weights according to known or approximated T-R distances, we weigh each receiver's coordinates according to the relative placement
of its RSS value within the span of all RSS reports for a given transmitted message. The receiver coordinates may be weighted linearly or exponentially.

**Definition 10.1. Minimal/Maximal RSS.** Let \( \mathcal{R} \) be the set of all receivers within range of a given message \( M_T \) originating from an uncooperative transmitter \( T \). Let \( \mathcal{T} \) denote the set of RSS values measured at each receiver \( R_i \in \mathcal{R} \) for message \( M_T \), such that:

\[
\mathcal{T} = \{ v_i : v_i \text{ is the RSS value for message } M_T \text{ at } R_i \text{ for all } R_i \in \mathcal{R} \}
\]

Then we define the minimal and maximal RSS values, \( \mathcal{V}_{\text{min}} \) and \( \mathcal{V}_{\text{max}} \), for message \( M_T \), as the smallest and largest RSS values in \( \mathcal{T} \):

\[
\mathcal{V}_{\text{min}} = \min \{ v_i \in \mathcal{T} \}
\]

\[
\mathcal{V}_{\text{max}} = \max \{ v_i \in \mathcal{T} \}
\]

**Definition 10.2. RSS Span.** Let the minimal and maximal RSS values for a message \( M_T \) be as stated in Definition 10.1. We define the RSS span \( \mathcal{V}^\Delta \) for this message at a set of receivers \( \mathcal{R} \) as the maximal range in RSS values over all receivers:

\[
\mathcal{V}^\Delta = \mathcal{V}_{\text{max}} - \mathcal{V}_{\text{min}}
\]

We describe two relative span weighted localization algorithms, both computing a weighted centroid as defined in Equation (10.2), but with novel approaches for computing the weights \( w_i \) assigned to each receiver’s coordinates \( p_i \).

### 10.2.1 Linearly Weighted Localization

The RWL algorithm computes a centroid of receiver coordinates, each weighted linearly according to the relative position of the receiver’s RSS value within the RSS span.

**Algorithm 10.1. RWL Algorithm.** The Relative Span Weighted Localization (RWL) algorithm estimates a transmitter’s coordinates \( \hat{p} \) as the weighted centroid of all receiver coordinates \( p_i \), as defined for WCL in Equation (10.2), but with a linearly increasing weight assigned to each receiver according to its presumed proximity to the
transmitter. Given the RSS values in $\mathbf{V}$, as found in Definition 10.1, and the RSS span $V_{\Delta}$ determined according to Definition 10.2, the weight $w_i$ of each receiver $R_i$ is computed from the relative placement of its RSS value $v_i$ in the RSS span, as follows:

$$w_i = \frac{v_i - V_{min}}{V_{\Delta}} \text{ for all } R_i \in \mathbb{R}$$

The relative span weighted centroid thus becomes:

$$\hat{p} = \frac{\sum_{i=1}^{n} [(v_i - V_{min}) \times p_i]}{\sum_{i=1}^{n} (v_i - V_{min})}$$

where $n = |\mathbb{R}|$.

10.2.2 Exponentially Weighted Localization

Exponentially weighted moving averages (EMAs) have been used for a variety of forecasting applications, for example in Muir [98], to predict future values based on past observations, with more weight exponentially ascribed to more recent data. A weighting factor $\lambda$ is used as a parameter to control the proportion of weight assigned to recent observations with respect to past ones.

According to [100], the EMA at time $t$ is stated as:

$$EMA_t = \lambda \times Z_t + (1 - \lambda) \times EMA_{t-1}$$

where $\lambda$ is the weighting factor, $Z_t$ is an observation at time $t$ and $EMA_0$ is the average of historical observation values.

Roberts [120] expands the EMA equation as follows:

$$EMA_t = \lambda \times \sum_{i=1}^{n} [(1 - \lambda)^{t-i} \times Z_i]$$

where $n$ is the number of observations.

We adapt the EMA concept from rating observations over time for the purpose of weighting receiver coordinates over the space domain. While EMA favors more recent observations in time with a weighting factor of $\lambda$, we bolster receivers that are likely
to be closer to a transmitter and thus feature higher RSS values. In addition, rather than increasing the weighting factor exponent by one for each observation in time, we correlate the exponent with the relative position of each receiver’s RSS value within the RSS span.

**Algorithm 10.2. REWL Algorithm.** The Relative Span Exponentially Weighted Localization (REWL) algorithm estimates a transmitter’s coordinates \( \hat{p} \) as the weighted centroid of all receiver coordinates \( p_i \), as defined for WCL in Equation (10.2), but with exponential weight assigned to each receiver according to a weighting factor \( \lambda \). Given the RSS values in \( T \) as found in Definition 10.1, the weight \( w_i \) of each receiver \( R_i \) is computed from the relative placement of its RSS value \( v_i \) in the RSS span, as follows:

\[
    w_i = (1 - \lambda)^{(v_{\text{max}} - v_i)} \quad \text{for all } R_i \in \mathbb{R}
\]

The relative span exponentially weighted centroid thus becomes:

\[
    \hat{p} = \frac{\sum_{i=1}^{n} [(1 - \lambda)^{(v_{\text{max}} - v_i)} \times p_i]}{\sum_{i=1}^{n} (1 - \lambda)^{(v_{\text{max}} - v_i)}} \quad \text{(10.4)}
\]

where \( n = |\mathbb{R}| \).

**10.2.3 Example**

Figure 10.1 compares the relative weights assigned to a set of receivers with a RSS span \( v^A \) of 15 dBm, given the RWL and REWL weight assignments, assuming three different values for the weighting factor \( \lambda \). The REWL algorithm with \( \lambda = 0.10 \), equivalent to a smoothing factor of 10%, represents the flattest curve and thus the closest to the constant weighting approach of simple CL. Less weight is assigned to receivers closest to the transmitter and more weight to those farthest, when compared to the linear RWL method. With \( \lambda = 0.20 \), the nearest receivers are ascribed far greater weight than under the RWL scheme, and the mid-RSS receivers are given much less importance. A weight factor of \( \lambda = 0.15 \) strikes a balance between the
two, with the highest RSS receivers assigned slightly more weight than with RWL, the mid-RSS receivers somewhat less, while the lowest still contribute marginally to estimating the transmitter location.

![Figure 10.1: Example of Relative Span Weights](image)

10.3 Performance Evaluation

We evaluate the performance of the RWL and REWL algorithms using simulated RSS values and experimental ones harvested from an outdoor field experiment.

10.3.1 Simulation Results

We ran the RWL and REWL mechanisms on simulations featuring a variety of node densities and number of receivers. For each of 10,000 executions, we generate a random transmitter position within a 1000 × 1000 m$^2$ simulation grid. We define our node densities as the number of nodes per 100 × 100 m$^2$. For every node density $d \in \{0.25, 0.50, 0.75, 1.00, 2.00, 3.00, 4.00, 5.00, 6.00, 7.00, 8.00, 9.00, 10.00\}$, we position $d$ nodes per 100 × 100 m$^2$ in uniformly distributed positions on our simulation grid. For each node, we compute a RSS value based on the log-normal shadowing model [115], with an independently random amount of signal shadowing generated along a Gaussian probability distribution. We assume two different radio propagation
environments with path loss constants obtained from outdoor experiments. For the 2.4 GHz WiFi/802.11g frequency, we use propagation values measured by Liechty et al. [88, 89], where a signal shadowing standard deviation is measured at nearly \( \sigma = 6 \) dB. For the 5.8 GHz frequency, licensed for vehicular networks [8], we make use of the constants determined by Durgin et al. [47], with a signal shadowing standard deviation close to \( \sigma = 8 \) dB. Similar experiments by Schwengler and Gilbert corroborate the amount of signal shadowing commonly experienced at this frequency [128]. Our setup allows us to gauge the performance of relative span weighted localization based on propagation environments featuring different amounts of signal fluctuations. Once our simulated nodes are positioned, we determine which ones can be used as receivers. We set all receiver sensitivity to -90 dBm, and the nodes that feature a RSS value above the sensitivity are deemed within range of the transmitter and thus become receivers. The non-receiver nodes are subsequently ignored as out of range.

Table 10.1 shows the average number of receivers for each node density, over all our simulated executions, given each radio propagation environment.

| Frequency (GHz) | \( \sigma \) (dB) | 0.25 | 0.50 | 0.75 | 1.00 | 2.00 | 3.00 | 4.00 | 5.00 | 6.00 | 7.00 | 8.00 | 9.00 | 10.00 |
|----------------|-----------------|------|------|------|------|------|------|------|------|------|------|------|------|
| 2.4            | 6               | 2    | 3    | 4    | 5    | 11   | 17   | 23   | 29   | 35   | 40   | 46   | 52   | 58   |
| 5.8            | 8               | 4    | 7    | 11   | 15   | 30   | 45   | 60   | 75   | 91   | 106  | 121  | 136  | 151  |

Figure 10.2 depicts an example simulation grid, with a transmitter at 2.4 GHz, a number of nodes generated with density one per \( 100 \times 100 \) m\(^2\), and receivers within range of the transmitter. It should be noted that some nodes may be located closer to the transmitter and yet be out of range. This is due to the different amounts of signal shadowing generated for each node. So while one node may be physically closer to the transmitter, if it experiences a large amount of negative shadowing, its RSS value may fall below the receiver sensitivity and thus be deemed undetectable.

For each execution, we use the known coordinates of all receivers to compute a possible position for the transmitter, according to four algorithms: the maximum RSS receiver method, where a transmitter is assumed to be at exactly the receiver position with the highest RSS value; the CL approach, as set out by Bulusu et al. in
Equation (10.1); the RWL algorithm using Equation (10.3); and the REWL algorithm as set forth in Equation (10.4), given three different values for the weighting factor \( \lambda \in \{0.10, 0.15, 0.20\} \). We assess the performance of each mechanism according to its location accuracy, computed as the Euclidean distance between the estimated position \( \hat{p} \) and the actual transmitter location, averaged over all executions. Our results are deemed accurate within ±3 meters in a 95% statistical confidence interval.

Figures 10.3 and 10.4 plot the average location error for each tested algorithm, given all defined node densities, for frequencies 2.4 GHz and 5.8 GHz respectively. We find that while higher densities consistently yield greater location accuracy, a larger amount of signal shadowing results in higher location errors. For example, for all densities, the REWL algorithm, with the 2.4 GHz frequency and \( \sigma = 6 \) dB, yields a location error consistently less than 75 meters, while the same mechanism at the 5.8 GHz frequency and \( \sigma = 8 \) dB reaches an error of 105 meters. For both frequencies and all node densities, the REWL algorithm with a weighting factor of 15% (\( \lambda = 0.15 \)) achieves optimal results.

The same observations can be made when nodes are generated by absolute numbers of receivers rather than node densities. Figures 10.5 and 10.6 demonstrate the location errors computed with each algorithm with fixed numbers of receivers, given the 2.4 GHz and 5.8 GHz frequencies respectively. Again, with similar numbers of receivers, the least shadowed environment produces lower location errors. As with the tests involving different node densities, the REWL mechanism with \( \lambda = 0.15 \)
Figure 10.3: Algorithm Location Error by Node Density (2.4 GHz)

Figure 10.4: Algorithm Location Error by Node Density (5.8 GHz)
performs better than the other algorithms for all numbers of receivers.

In order to gauge the performance of REWL ($\lambda = 0.15$) for a single frequency and different levels of environmental shadowing, we executed the algorithm at 2.4 GHz with three separate amounts of signal shadowing generated on the simulated RSS values: $\sigma \in \{6, 8, 10\}$ dB. As Figure 10.7 reveals, higher levels of shadowing have a significant impact on location error, with an error increase of roughly 50% for every 2 dB of additional signal shadowing standard deviation, for each node density.

We assessed the performance of each algorithm, and in particular the REWL ($\lambda = 0.15$) mechanism, when compared to the E911 regulations for location accuracy. Figures 10.8 and 10.9 show the location error cumulative probability distribution for each algorithm, given four receivers, for the 2.4 GHz and 5.8 GHz frequencies respectively. While every method evaluated meets the E911 requirements at 2.4 GHz with moderate signal shadowing ($\sigma = 6$ dB), none of the mechanisms succeed with 5.8 GHz and a larger amount of shadowing ($\sigma = 8$ dB). However, even in the latter case, the REWL approach with $\lambda = 0.15$ is nearly adequate.

The REWL algorithm, with $\lambda = 0.15$, was evaluated for different node densities, with the two different frequencies. Given the smaller amount of signal shadowing found at 2.4 GHz, REWL meets the E911 location accuracy requirements for every node density, as seen in Figure 10.10. For larger amounts of shadowing at 5.8 GHz, only the smallest node density of 0.25 per 100 $\times$ 100 m$^2$ fails to meet the E911 standard, as shown in Figure 10.11. Even in a heavily shadowed environment, higher node densities can accurately localize a transmitter within 100 meters 67% of the time and within 300 meters in 95% of cases.

Orooji and Abolhassani [105] simulate a cluster of seven cells, each featuring a base station with a one kilometer radius, in order to compute the location of a mobile station. A very small amount of signal shadowing $\sigma \in \{1, 2\}$ dB is taken into account. Even though their proposed T-R distance-weighted method assumes a known distance to one of the base stations, the mean location error is 48 meters, with 95% of executions resulting in a location error less than 103 meters. Our RWL and REWL ($\lambda = 0.15$) algorithms for 2.4 GHz with eight receivers yield an average 37 and 34 meter location error respectively. RWL locates a transmitter within 100 meters
Figure 10.5: Algorithm Location Error by Number of Receivers (2.4 GHz)

Figure 10.6: Algorithm Location Error by Number of Receivers (5.8 GHz)
98% of the time, while REWL does so in 99% of cases. Thus over a similarly sized simulation grid, our RWL and REWL mechanisms consistently yield more accurate results.

10.3.2 Experimental Results

We conducted an outdoor field experiment with four desktop receivers statically arranged in the corners of a rectangular area $80 \times 110$ m$^2$ in size. Each receiver collected the RSS values of packets transmitted by a laptop from each of ten separate locations. Only the messages simultaneously received by the four desktops were retained. The localization algorithms were executed on each message, and the average location errors for each transmitter location are depicted in Figure 10.12. The location error for each algorithm averaged over all transmitter locations can be found in Table 10.2. We find that the RWL and REWL mechanisms perform far better than the maximum RSS receiver and CL approaches, with a gain in location accuracy of up to 40%. On average, the RWL, REWL with $\lambda = 0.15$, and REWL with $\lambda = 0.20$ mechanisms perform equally well, with no algorithm emerging as clearly superior to the others. This may be due to our small experimental data set (approximately 400 messages), when compared to simulation results obtained over 10 000 executions. While our simulations also found consistently similar results between the RWL and REWL
Figure 10.8: Algorithm Location Error CDF for Four Receivers (2.4 GHz)

Figure 10.9: Algorithm Location Error CDF for Four Receivers (5.8 GHz)
Figure 10.10: REWL Location Error CDF by Node Density (2.4 GHz)

Figure 10.11: REWL Location Error CDF by Node Density (5.8 GHz)
mechanisms, the larger amount of simulated data allows us to draw more fine-tuned conclusions.

Figure 10.12: Algorithm Location Error for Experimental Data

Table 10.2: Average Location Error for Experimental Transmitter Locations

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Max RSS</th>
<th>CL</th>
<th>RWL</th>
<th>REWL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \lambda = 0.10 )</td>
</tr>
<tr>
<td>Avg Location Error (meters)</td>
<td>40</td>
<td>46</td>
<td>28</td>
<td>33</td>
</tr>
</tbody>
</table>

10.4 Summary

We have proposed a wireless network-based localization mechanism for estimating the position of an uncooperative transmitting device, whether it be a malfunctioning sensor, an attacker engaging in a security exploit or a low-battery cellphone in a critical emergency. We extend the concept of weighted centroid localization and
describe two additional receiver coordinate weighting mechanisms, one linear and the other exponential, that assume no knowledge of the T-R distances nor of the transmitter EIRP. We adapt the concept of moving averages based on observations over time to the space domain. We ascribe linear and exponential weights to each receiver coordinates, based on the relative positioning of the receiver’s RSS value relative to the RSS span over all receivers.

We tested our relative span weighted localization algorithms with simulated and experimental RSS values, using two frequencies featuring different amounts of signal shadowing. We found that our algorithms yield lower location errors than the existing centroid localization method. As expected, the location accuracy increases as more nodes participate in the localization effort. For example with REWL (λ = 0.15) at 2.4 GHz, one node per 100 × 100 m² localizes a transmitter within 44 meters, while ten nodes per 100 × 100 m² do so in less than ten meters. Yet the location accuracy decreases as the amount of signal shadowing between different receivers increases, with an average decrease of approximately 50% for every 2 dB of additional signal shadowing standard deviation. We conclude that, given our simulation parameters at the 2.4 GHz frequency, the exponential variation of our relative span weighted localization algorithm achieves a location accuracy that meets the FCC regulations for Enhanced 911, for all densities with moderate amounts of signal shadowing and for all but the smallest node densities with extensive shadowing.
Part IV

Conclusion
Chapter 11

Conclusions and Future Work

We set out to examine the extent to which security requirements in wireless mobile networks, especially in new and emerging technologies, could be harmonized with the reasonable degree of location privacy expected by their users. To this end, we establish a clear distinction between the privacy implications for two separate classes of network devices. In the case of honest nodes, whose behaviour conforms to network authority expectations, anonymity is of paramount importance, as the risk of location tracking may violate users' privacy and deter their adoption of new technologies. On the other hand, uncooperative devices should not retain any degree of location privacy if they engage in security exploits, since honest nodes require protection from attackers, whether they be outsiders to the network or malicious insiders. As a result, network authorities must ensure that such attackers are apprehended and the threats they pose mitigated.

We highlight our contributions in providing solutions to balance the location privacy of honest nodes in vehicular networks and the mitigation of threats posed by uncooperative, including malicious, devices. We discuss our principal conclusions, and we outline directions for future work.

11.1 Contributions

We have performed an analysis of the threats to the DSRC/WAVE vehicular network architecture. We have correlated the resulting assessed risks with those uncovered in a previous analysis of the WiMAX/802.16 wireless access technology. We have identified location tracking as a critical threat in vehicular networks, and insider attacks as an equally crucial risk to both DSRC/WAVE and WiMAX/802.16.

We have addressed the location tracking threat in vehicular networks by providing a Secure Anonymous Broadcasting (SAB) solution that maintains equivalent security
to DSRC/WAVE, while preserving driver anonymity. Our approach is found to generate smaller vehicle safety broadcast messages and employs more computationally efficient cryptographic operations.

We have approached the threat of insider attacks from a location estimation standpoint. We aim to attribute a target message to its originator by providing a localization solution that makes a novel and unique set of assumptions. In presuming neither assistance from a transmitter nor knowledge of the transmitting power (EIRP) it uses to broadcast its message, we can bound the probable position of any uncooperative device, whether it be malfunctioning or malicious. We have described the probabilistic Hyperbolic Position Bounding (HPB) mechanism, where the trusted receivers of a target message approximate a transmitter location from the RSS values registered for that message. As each receiver experiences an independently different amount of signal shadowing, this quantity is modeled as a random variable, subject to a level of confidence. HPB combines a range of signal shadowing with an interval of possible EIRP values, in order to compute a range of distance differences between a transmitter and a pair of receivers. Hyperbolas are constructed at the minimum and maximum bounds of this distance difference range and thus bound the location of the transmitter to a candidate area, with a given confidence level. We have evaluated the performance of HPB using simulated RSS values generated from the log-normal shadowing propagation model. We found that our HPB algorithm succeeds in locating a malicious transmitter to an area approximately $250 \times 250$ m$^2$ in size.

We have proposed an additional mechanism for combining the probability distributions assigned by multiple receiver pairs to the common intersection of their candidate hyperbolic areas. Our approach aggregates the transmitter position probability distribution within the common intersection, rather than redistributing it outside this intersection. We have concluded that our compound probability distribution closely relates to the simulated HPB results, within a 15% difference for higher confidence levels.

We have conducted an outdoor experiment with WiFi/802.11 devices in order to harvest measured RSS values to further evaluate HPB performance. We have found that the HPB results are consistent, whether simulated or experimental RSS values
are considered.

We have described three variations of the HPB algorithm, for the purpose of improving location accuracy and computational efficiency. We have also extended HPB for tracking a mobile attacker. We have concluded that HPB can be used to localize the position of a transmitter in a vehicular scenario within the requirements for Enhanced 911 (E911) location accuracy.

We have outlined two additional uncooperative transmitter location estimation algorithms to yield a single candidate point in Euclidian space. We extend the concept of Weighted Centroid Localization (WCL) to assign weights to receiver coordinates based on a moving average approach applied to the space domain, rather than over time. We describe two weighting mechanisms, one linear and another exponential. We have determined that our exponential algorithm can localize a transmitter within the E911 stipulated requirements.

11.2 Conclusions

Our work makes a number of hypotheses in tackling the location tracking threat against honest nodes and the location estimation of uncooperative transmitters. We outline the main propositions put forth in our research, and we describe how these have been addressed by our contributions.

1. We submit that it is possible for adequate network security to be maintained, while honest node anonymity is protected. We demonstrate this point with our SAB mechanism, where honest nodes in vehicular networks are permitted to participate in broadcasting vehicle safety messages on an anonymous basis, thus preserving their privacy from other non-authoritative devices. Network security is ensured by simultaneously providing the means for a network authority to identify a misbehaving node in the event that an attack message is detected.

2. We propose to deal with insider attacks by attributing the source of a malicious transmission to its originator. Given that a perpetrator is unlikely to use credentials linked to its true identity, we necessarily reduce this problem to one of estimating the probable location of a transmitting attacker. With an
approximated target position, network authorities can be equipped with additional mechanisms to link a physical location to a possible device identity, and to effect retribution (such as eviction) on the misbehaving node. The means to detect, identify and evict such devices remains an open issue.

3. We hypothesize that RSS values can be used to approximate a transmitter location, presuming that an attacker does not cooperate with localization efforts and that it may use an unexpected EIRP in order to obfuscate its position. We demonstrate that a malicious transmitter can be localized through a novel theoretical approach, despite these restrictive assumptions and the fact that no prior attempts have been made for localization under such circumstances. We determine that existing mechanisms, for example the annuli method, are unsuitable for localization when the EIRP is unknown. We further establish that our HPB algorithm does succeed under the same conditions, with a location granularity of approximately one quarter of a one square kilometer grid.

4. In assessing the performance of our HPB mechanism, we require a method to generate simulated RSS values based on T-R distances. We hypothesize that the log-normal shadowing model is appropriate for this purpose, given that it has been used in a number of experimental indoor and outdoor settings. We correlate the HPB results obtained using RSS values simulated with the log-normal shadowing model with RSS reports harvested from a practical outdoor experiment. Since both sets of results are consistent with each other, we determine that the log-normal shadowing model is a suitable tool for generating RSS values in simulation scenarios.

We acknowledge that RSS values can be subject to natural environment based fluctuations and have not previously been used without knowledge of the corresponding EIRP. We nonetheless conclude that RSS reports constitute valid estimators of T-R distances, and that an acceptable degree of localization accuracy is achievable, even without a known EIRP value as a basis for reference. E911 regulations require network-based location estimation mechanisms to meet an accuracy requirement of
100 meters in 67% of cases, and of 300 meters 95% of the time. Under these stipulations, our open-form HPB approach can successfully localize a transmitter in restricted conditions, for example in a vehicular scenario with a known navigational layout, assuming at least four receivers within radio range. Our closed-form REWL algorithm also meets the E911 guidelines with similar numbers of receivers, but in an unrestricted setting, and even in a highly shadowed environment.

11.3 Future Work

Our HPB mechanism takes into account a range of possible EIRP values when approximating the location of an uncooperative transmitter, but it assumes an omnidirectional signal propagation pattern. An attacker may thus employ a directional antenna to obfuscate its position. Directional signals can foil our attempts at localization using HPB, because such transmissions distort the relationship between the RSS values received by a pair of devices and their distance difference from the transmitter. We require an extension of the HPB algorithm to address directional attack signals.

Every type of antenna exhibits a particular radiation pattern represented by a 3D shape, where the signal strength is uniform on every surface point of this shape. For an omnidirectional antenna, the radiation pattern is roughly doughnut-shaped. The RSS is equal at an equal distance from a transmitter, in all directions. With directional antennas such as parabolic dishes or Yagi antennas, the transmitted signal is stronger in some directions than others. Barbeau and Kranakis [14] provide an introduction to this topic.

Given that the radiation pattern of an attacker's directional antenna is unknown, our contributions in this area will necessarily be two-fold:

1. We plan to investigate how the shape of an unknown antenna's radiation pattern can be approximated, based on the RSS values obtained at trusted devices and their coordinates.

2. The HPB mechanism can subsequently be refined to compensate for the directional antenna gain in the relative RSS values of an attack signal, once the antenna's radiation pattern is mapped.
For example, the scenario in Figure 11.1 shows an attacker (depicted as a solid circle) emitting a directional signal. The dotted lines represent the antenna's radiation pattern, and thus the receivers located on the dotted lines measure an equal RSS. Receiver $R_1$ registers a higher RSS value than $R_4$, even though $R_4$ is closer to the transmitter. However, most OBU's in $R_1$'s vicinity also receive the attack message, while some in close proximity to $R_4$ do and others do not. It may be possible to map out the shape and extent of an antenna's radiation pattern, given the coordinates of the devices that are within a signal's range and the ones that are out of range.

![Figure 11.1: Example of Directional Antenna Radiation Pattern](image)

Our radiation pattern matching mechanism can assume a finite set of known radiation patterns, for example Yagi and dipole patterns. The signal reception map can be matched to known radiation patterns using recognition techniques.

A new radiation pattern mapping scheme could be evaluated through the simulation of attack signals originating from antennas featuring a variety of patterns. The success rate can be gauged by comparing the similarity of an approximated radiation pattern to the actual one. With further simulations, the success rate of an extended HPB algorithm in bounding the position of an attacker can be assessed, given relative RSS values adjusted in function of the approximated radiation pattern. Receiving directional antennas may also be investigated and leveraged as an additional source of transmitter location information.
Other avenues for improvement of our HPB mechanism involve enhancements to the EIRP range estimation technique, described in Heuristic 9.1. One approach has been put forth by El Sayr [49], where receivers are assumed to experience correlated amounts of signal shadowing. Additional techniques may be devised to improve HPB localization performance when a transmitter is not physically situated between any pair of receivers, but is still within their radio range.
Bibliography


