THE UNIVERSITY OF CALGARY

IS-95 Cellular Mobile Location Techniques

by

Geoffrey G. Messier

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF ELECTRICAL AND COMPUTER
ENGINEERING

CALGARY, ALBERTA
July, 1998

THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "IS-95 Cellular Mobile Location Techniques" submitted by Geoffrey G. Messier in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE.

Chairman and Co-Supervisor, Dr. Michel Fattouche
Department of Electrical and Computer Engineering

Co-Supervisor, Dr. Brent R. Petersen
Department of Electrical and Computer Engineering

Dr. Abu Sesay
Department of Electrical and Computer Engineering

Dr. John Nielsen
Nortel

Abstract

Wireless mobile phones are used a great deal for emergency phone calls. However, most cellular networks provide only the most basic 911 services. In response to this, the Federal Communications Commission passed a series of regulations in 1996 that require all cellular service providers to provide Enhanced 911 (E-911) service. When the final phase of E-911 is implemented in 2001, every cellular network in the United States will have to be able to determine the position of each mobile in their coverage areas to an accuracy of 125 meters, RMS.

This thesis addresses the problem of determining the position of IS-95 cellular mobiles. First, an analysis of the mobile's received signal level is performed in order to determine what location methods are practical. Then, two strategies of locating the mobile are investigated. First, a system locating the mobile using the time-of-arrival of its signal at several base stations is proposed and evaluated. Second, two systems that locate the mobile using measurements it makes on the base station pilots are presented and analyzed.

Acknowledgements

The author would like to thank Dr. Fattouche and Dr. Petersen for their supervision and guidance of the work presented in this thesis. The author would also like to thank Dr. Nielsen for all his help on this project. Finally, the author would like to thank the National Sciences and Engineering Research Council, TRLabs and the University of Calgary for their funding support.

Dedication

For Margie.

Table of Contents

List of Figures

Table of Abbreviations

FCC	Federal Communications Commission
E911	Enhanced 911
CDMA	Code Division Multiple Access
SIR	Signal to Interference Ratio
TOA	Time of Arrival
TDOA	Time Difference of Arrival
MUSIC	Multiple Signal Classification
PN	Pseudo-Random Noise
AOA	Angle of Arrival
QPSK	Quadrature Phase Shift Keying
PDF	Probability Density Function
AMPS	Advanced Mobile Phone System
DLL	Delay Locked Loop
FFT	Fast Fourier Transform
MAHO	Mobile Assisted Handoff
MSC	Mobile Switching Station
RMS	Root Mean Squared
m	meters

Chapter 1

Thesis Introduction

1.1 Thesis Overview

The topic for this thesis is IS-95 cellular mobile location. The focus of this work is the characterization of the IS-95 mobile location problem and the development of possible solutions for the problem.

The motivation behind this project is a series of regulations passed in 1996 by the Federal Communications Commission (FCC) in the United States. The intent of these regulations are to encourage cellular service providers to improve the quality of the 911 service available to cellular phone users. One of these regulations states that all cellular phone service providers must be able to determine the physical location of the phones in their system. This is to assist in providing a more effective emergency response when a cellular phone is used to make an emergency call. The regulations passed by the FCC are discussed in detail in Section 1.2. In addition to emergency call service, cellular mobile location would be useful for applications like fleet deployment and handoff management.

The IS-95 cellular system is gaining increasing popularity around the world. Section 1.3 discusses the recent growth of IS-95. With the increasing deployment of this system, developing a location system for IS-95 is a priority.

An overview of the contents of the thesis is given in Section 1.4. A brief introduction to the problem is given and then each of the chapters in the thesis are discussed.

1.2 Enhanced 911 Service for Cellular Systems

This section gives a brief overview of the Enhanced 911 (E911) regulation for wireless service providers introduced by the FCC [1].

The 911 emergency calling system was introduced in 1968 and has since become a very important factor in providing effective emergency service. In the United States, approximately 260,000 911 phone calls are made every day [1]. In response to the increasing demand on the system, the FCC started an initiative to make the 911 phone service more effective. This has come in the form of regulations requiring the telephone companies to add features to their phone networks to aid in the effective dispatch of emergency services.

E911 services have already been implemented to a large extent in the wireline phone networks. These services automatically provide a 911 dispatcher with information like the phone number and address of a 911 caller. This is a great help in situations where a 911 caller does not know his or her location, is disoriented or even becomes unconscious during the call.

With the rapidly increasing popularity of wireless phones, more and more emergency phone calls are being made over the wireless phone network. In 1994 in the United States, almost 18 million wireless 911 phone calls were made [1]. The total number of cellular subscribers in the United States currently exceeds 33 million and in a recent survey, 62% of cellular users said safety and security were their main reasons for getting a wireless phone [1].

Despite this, most cellular service providers only support the most basic 911 services. As a result, the FCC passed a series of regulations requiring all cellular networks operating in the United States to support certain E911 services. These services are

listed below.

1. By 1997, all cellular service providers must process any 911 call made by a wireless phone in their coverage area. These calls should not be subject to the usual user validation procedures used to detect cellular phones from different networks.

2. By 1997, all wireless 911 calls must be able to support special equipment used by disabled people. An example would be a Text Telephone Device used by a hearing impaired caller.

3. By April 1, 1998, the cellular network must be able to provide the 911 operator with the location of the base station handling the 911 call.

4. By October 1, 2001, all cellular service providers must be able to determine the longitude and latitude of a 911 caller within a radius of 125 m, RMS.

The focus of this thesis is the development of a system that satisfies the 4th requirement presented by the FCC.

There has been some debate on the interpretation of the 4th requirement. One possible approach to designing the location system would be to provide a location estimate with an error less than 125 m, 67% of the time. The rest of the time, no location estimate would be provided at all. This type of system would be unacceptable to the cellular subscribers. Users need to know that they will always be located, no matter where they are in the cellular coverage area. As a result, an effective location system must be able to provide some time of location estimate for the mobile at all times.

1.3 The IS-95 Cellular System

Code Division Multiple Access (CDMA) cellular telephones are becoming a very popular choice for providing digital cellular and PCS service. Not only are IS-95 systems being installed extensively in the North American domestic market, they are also being used in several international markets. Figure 1.1 shows a map produced by the CDMA Development Group [2]. The dark areas on the map indicate countries that either already provide CDMA cellular coverage or are doing trials on the technology.

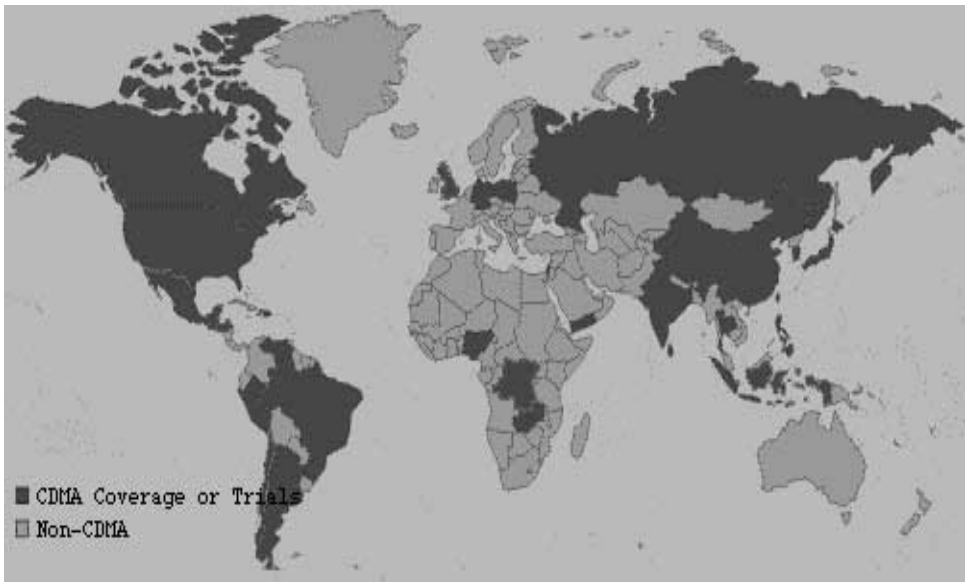

Figure 1.1: CDMA Cellular Technology Penetration

Recently, a survey of the market penetration of CDMA cellular systems was performed [3]. At the time of the survey, CDMA cellular systems were located in 25 countries, serving over 6 million subscribers. The production of CDMA mobiles is predicted to exceed 17 million units in 1998.

The growth in the number of CDMA subscribers is occurring in domestic and

international markets. In the North American market, service providers like AirTouch Communications who use CDMA networks, are reporting a 30% rise in the number of their subscribers that are switching from analog to digital phones [3]. Internationally, CDMA technology has been adopted by high growth markets like China, Japan and India. This will ensure a strong demand for CDMA cellular systems in the future.

CDMA cellular phone networks are becoming increasingly common. A large amount of money has already been invested in a CDMA cellular infrastructure with more development to come. A location system for that infrastructure that complies with the FCC's requirements is clearly a priority.

1.4 Thesis Overview

Determining the location of an IS-95 mobile is a very challenging problem. Most methods used for locating a mobile use some property of the signal the mobile is transmitting. For example, the time-of-arrival of the mobile's signal could be used to determine a fix on its location. This approach requires the mobile's signal to be received at more than one location. For practical reasons, these locations are usually the existing cellular base stations.

Under some conditions, the received signal to interference ratio (SIR) of the mobile's signal at base stations in neighbouring cells can be very low. This is due to a severe near-far effect at the neighbouring base stations. These low SIR levels can make it extremely difficult to locate an IS-95 mobile using its signal. Chapter 2 discusses this in detail and characterizes the severity of the problem using IS-95 system simulations.

There are two criteria that must be met in order to solve the IS-95 mobile location problem. The first is to deal with the problems described in Chapter 2. The second is to implement the location system with as little modification of existing mobiles and base stations as possible.

Rather than searching for a single, optimal solution to the IS-95 cellular location problem, this thesis proposes several sub-optimal solutions. Given the difficulty of the problem, it is unlikely that a single solution will be found that will locate the mobile in all circumstances. Instead, a combination of several different location techniques will likely be required to consistently produce a reliable position estimate.

The first type of location system proposed determines the mobile's location using the signal transmitted by the mobile. This approach is described in Chapter 3. The

technique used in this chapter is Time Difference of Arrival (TDOA).

The first step in the TDOA location approach is to determine the TOA of the mobile's signal at several different base stations. The difference of the TOA values at two different base stations defines a hyperbola. The intersection of several of these hyperbola gives the location of the phone. Chapter 3 also discusses some enhancements that can be made to a TDOA system using a TOA super-resolution technique based on the Multiple Signal Classification (MUSIC) algorithm.

Since minimizing the additional modifications required by the location system is important, any location scheme that uses existing features of the IS-95 standard to locate the mobile is attractive. Chapter 4 describes two location solutions that locate the mobile using measurements taken by the mobile on pilot signals transmitted by the cellular base stations.

The mobile makes two measurements on the pilot signals transmitted by the base station to aid in soft handoff. First it measures the received SIR of the pilots. In Chapter 4, a new approach is described that uses the ratio of the SIR measurements from two different pilots to determine the mobile's location. The mobile also measures the PN sequence offset of the spreading codes used by the pilot signals. These offsets are equivalent to coarsely quantized TOA measurements on the pilots and can also be used to determine the position of the mobile. The performance of this method is also presented in Chapter 4.

Finally, Chapter 5 makes some conclusions about the work presented in this thesis and some suggestions on future work for the project.

Chapter 2

IS-95 Cellular System Characterization

2.1 Introduction

Most traditional radio signal location methods determine a device's location by using some characteristic of a signal transmitted by that device. Usually, the signal has to be received at several receivers with known locations. The positions of these receivers are combined with some characteristic of the transmitted signal and used to solve for the location of the device. The signal characteristic used could be the time-of-arrival of the signal, the angle-of-arrival or even the strength of the signal.

A practical location system would have the receivers for the mobile's signal located in existing base stations. Building new base stations for the location system would be far too expensive. That means that in order to locate a mobile using its signal, it is necessary to receive that signal at the base station in its current cell as well as two or more base stations in neighbouring cells.

Unfortunately, in an IS-95 cellular network, situations can arise where the SIR of the mobile's signal at neighbouring base stations is extremely low. This is due to the near-far problem. IS-95 uses CDMA. This means that several users transmit on the same carrier and each user's signal is interference for every other user's signal. A near-far problem occurs when the mobile being communicated with is far from the base station and there are several interfering mobile's close to the base station. The signal of the desired mobile can be obscured by the interference from the nearby

mobiles to the point where the processing gain of its signal is no longer sufficient to extract its information.

Within a cell, this problem is solved using power control. The transmit powers of the mobiles are regulated so that they are all received at the same SIR. This means that the mobiles close to the base station reduce their power and the mobiles far away from the base station increase it. Once this condition is satisfied, the processing gain of the mobile's spreading codes can be used to overcome the interference from the other users.

However, this condition does not hold between cells. When receiving a mobile located outside a cell, the received power of the mobile's signal will be much less than the interference from the mobiles within the cell. This can result in extremely low SIR. The situation is especially bad when the mobile being located gets closer to the base station in its cell. When this happens, the power of its signal is reduced by power control which makes it even more difficult to extract at a neighbouring base station.

A quick calculation shows that it is possible to run into extremely low signal levels when trying to locate an IS-95 mobile. Consider the following example of a 2 base station IS-95 system, shown in Figure 2.1.

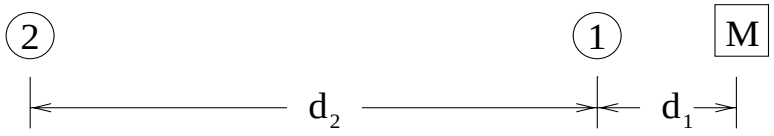

Figure 2.1: SIR Example

Base station 1 and 2 are represented by the circles. The mobile is represented by the square. The mobile is currently communicating with base station 1 but its signal

is also being received at base station 2 for location purposes. Each base station uses power control to make sure that they receive the signals from all the mobiles in their respective cells with a power of P.

This example illustrates a worst case situation for receiving the desired mobile's signal at base station 2. The mobile is as close as possible to base station 1, which means its power will be turned down very low by power control commands. For this example, d_1 is equal to 400 m and d_2 is equal to 2000 m.

The SIR of the desired mobile's signal at base station i is given by

$$SIR_i = \frac{P_i}{I} \tag{2.1}$$

where P_i is the received signal power of the mobile and I is the total interference power. A typical SIR value of a mobile's signal at the base station it communicates with is approximately -15 dB.

The ratio of the two received SIR values are given in Equation 2.2.

$$\frac{SIR_1}{SIR_2} = \frac{P_1/I_1}{P_2/I_2} = \frac{P_1}{P_2} \tag{2.2}$$

If the interference received at both base stations is the same, $I_1 = I_2 = I$, the ratio of the received SIR values at each base station is equal to the ratio of the average received signal power of the mobile at each of the base stations. This assumption holds approximately if each cell has the same number of uniformly distributed users.

An average value for the received signal power, S_i, from the desired mobile can be calculated approximately using the log distance path loss model, given in Equation 2.3

[4].

$$P_r = P_o \left(\frac{d_r}{d_o} \right)^{-n} \tag{2.3}$$

In this equation, P_r is the average received power at a distance d_r from the transmitter, P_o is a reference power received at a distance d_o from the transmitter and n is the path loss exponent of the channel. A conservative estimate for n in a typical cellular system is 4 [4].

Using the log-distance equation and expressing the values in dB, the average received SIR value at base station 2 is given by

$$\begin{aligned} SIR_2 &= SIR_1 - n10\log\left(\tfrac{d_1+d_2}{d_1}\right) \\ &= -15 - 4 \cdot 10\log\left(\tfrac{400+2000}{400}\right) \\ &\simeq \text{-46 dB} \end{aligned} \tag{2.4}$$

This calculation illustrates that while performing cellular mobile location in an IS-95 system, situations can arise where the SIR of the signal used to locate the mobile is extremely low. Clearly, dealing with signals at this level will make some location methods impractical. The received signal level of the mobile will greatly affect what location methods are feasible for this project. This means that it is important to have a very detailed picture of the signal levels that can be expected in an IS-95 system.

This chapter outlines an IS-95 system simulation that was performed in order to help characterize the received signal levels of IS-95 mobiles. Section 2.2 explains the components of the IS-95 cellular system that are implemented in the simulation, Section 2.3 describes the setup of the simulation, Section 2.4 presents the results of the simulation and Section 2.5 gives an interpretation of the simulation results.

2.2 The IS-95 Cellular System

The object of the system characterization simulation discussed in this chapter is to get an idea of what the SIR of a mobile's signal are at different locations in the cellular coverage area. In order to find this out, it is necessary to do a fairly detailed simulation of the reverse channel of an IS-95 system. This means that the signals of several IS-95 mobiles must be generated. In order to do this realistically, it is necessary to take a close look at how an IS-95 mobile works.

This section deals with the components of the IS-95 cellular system that are relevant to this simulation. It covers mainly how the mobile constructs its signal and how it functions in the cell.

2.2.1 The IS-95 Reverse Link

Figure 2.2 shows the blocks of the IS-95 mobile transmitter that are implemented in the simulation. The blocks previous to the Walsh Code Modulator block deal mostly with adding error correction coding and interleaving to the information bit stream. In this simulation, these blocks are approximated by a random bit stream used as an input to the Walsh Code Modulator block. Each of the blocks in Figure 2.2 are discussed below.

Orthogonal Walsh Code Modulation

The Walsh Code Modulation block allows the base station to perform non-coherent demodulation of the mobile's signal. In the forward channel, the base station transmits a pilot signal which allows the mobiles to coherently demodulate the signals they receive from the base station. However, the mobiles do not currently transmit pilot signals on the reverse channel. This means that the base station receiver cannot lock

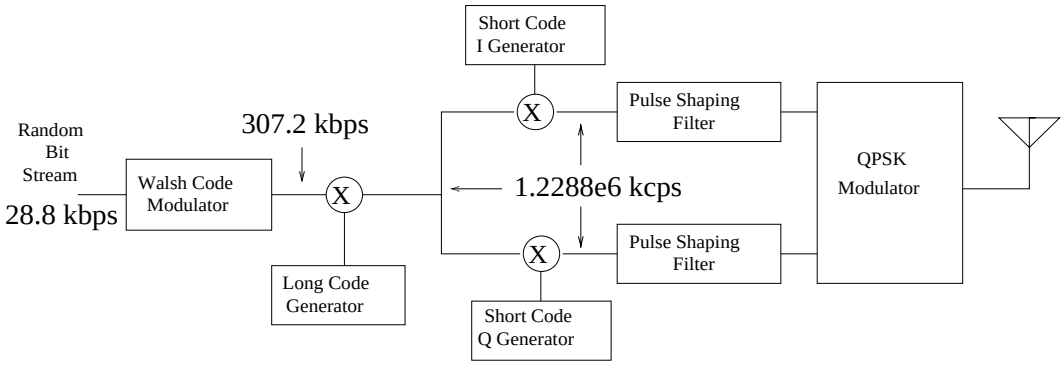

Figure 2.2: IS-95 Mobile Block Diagram

onto the phase reference of the mobile.

Since the base station does not have the phase reference of the mobile, it's necessary to use a non-coherent modulation scheme for the mobile's signal. There are several non-coherent modulation schemes, but the one used by the IS-95 mobiles is orthogonal modulation [5]. Orthogonal modulation means that the transmitter sends one of a set of K orthogonal symbols through the channel to represent the transmitted bits. Each symbol represents $\log_2(K)$ bits. The receiver has a bank of parallel correlators to correlate the received signal. Each correlator correlates the signal with a different symbol from the orthogonal set. The output of each correlator is then squared, to remove the unknown phase of the signal, and input into a decision device. The correlator with the largest output, and hence the largest energy, corresponds to the signal that was sent by the transmitter.

The orthogonal signaling used by the IS-95 mobile is slightly unconventional. Usually, orthogonal signaling is implemented using signals that are different frequencies. However, in this case, the orthogonal symbols are orthogonal bit sequences called Walsh sequences. The orthogonal symbol set consists of the rows of a 64x64 Walsh-Hadamard matrix [6]. This means that the orthogonal symbol set consists of 64 Walsh

symbols, each 64 bits long. Each symbol represents $\log_2(64) = 6$ bits. The orthogonal Walsh sequences are spread and then transmitted. At the receiver, they are despread and then passed through a bank of digital non-coherent correlators that determine which symbol has been received.

Long Code Spreading

The first stage of spreading applied to the Walsh sequences is the long code spreading stage. The primary purpose of this spreading stage is to keep the IS-95 mobiles from interfering with each other. Each mobile is given its own, unique long code spreading code. The set of long codes used by all the mobiles is made up of codes that are different shifts of the same PN sequence. The PN sequence is generated by a 42 bit shift register [6]. The period of the PN sequence is equal to $2^{42} - 1 \simeq 4.4 \cdot 10^{12}$ chips, which is why it is called the long code. At the IS-95 chip rate, this period is equal to just under 42 days.

The mobiles synchronize to the pilot signals transmitted by the base stations and the base stations are synchronized using GPS receivers. This makes it possible for each mobile to use different shifts of the same very long PN sequence and still remain orthogonal. The mobile's long code shift is generated using its electronic serial number, so each mobile will use the same long code for as long as it's in operation. The output bit rate of the Walsh Code Modulator is 307.2 kbps. The chip rate of the long code is 1.2288 Mcps. This means the spreading factor at this stage is 4.

Short Code Spreading

The final spreading stage is the short code spreading stage. A short code set also consists of different shifts of the same PN sequence. There are two short code sets,

one is used for the inphase and the other for the quadrature bit streams input to the QPSK modulator. The PN sequences used for the two short code sets are generated using 15 bit shift registers, which means their periods are only $2^{15} - 1 \simeq 3.2 \cdot 10^4$ chips long [6].

The short code spreading stage makes a somewhat unconventional use of QPSK modulation. Normally, QPSK is used to transmit two bits per symbol, one bit in the inphase stream and one bit in the quadrature stream. However, in the IS-95 mobile, the same output chip from the long code spreading block is input to the inphase and quadrature streams. Each of the streams are spread using a different short code and then they are pulse shaped and transmitted. This technique is used because it improves the multiple-access interference characteristics of the system and it aids in carrier recovery [5].

Since the chip rate of the short code generators are also 1.2288 Mcps, no additional bandwidth expansion occurs at this stage.

2.2.2 Power Control

Power control is the second aspect of an IS-95 mobile's operation that is important to the simulation. Power control is used by a base station to combat the near-far problem and to optimize capacity. The near-far problem occurs when the mobile a base station is communicating with is very far away and the interfering mobile or mobiles are very close. The distant mobile's signal can be attenuated to the point where even the processing gain of its spread spectrum signal will not be sufficient to overcome the interference from the nearby mobiles.

The solution to the near-far problem that IS-95 uses is power control. The base

station controls the transmit power of all the mobiles in its cell so that it receives their signals all at the same power. This means that mobiles that are far away from the base station are asked to turn up their power, while mobiles close to the base station are asked to turn their power down.

There are two kinds of IS-95 mobile power control: open loop and closed loop. For open loop power control, the mobile decides what its transmit power should be, based on measurements it takes on the forward channel. Closed loop power control is controlled by the base station. The base station sends the mobile commands telling it to adjust its power.

Equation 2.5 gives the relation used by the mobile to calculate its transmit power for open loop power control. This equation is sometimes called the 73 dB rule.

$$P_{out} = -P_{in} - 73 \tag{2.5}$$

The term P_{out} is the mobile's mean transmit power in dBm and P_{in} is the mean input power the mobile receives on the forward channel in dBm. The idea behind open loop power control is that the total power the mobile receives on the forward channel will be dominated by the signal power of the base station it is communicating with. As the mobile gets closer to the base station, the total forward channel power that it measures will increase and the mobile will decrease its transmit power accordingly.

Closed loop power control is controlled by the base station. The base station sends the mobile commands telling it to adjust its transmit power. Closed loop power control is used by the base station to fine tune the mobile's open loop estimate so that its signal is being received at a target SIR.

The base station uses closed loop power control to attempt to tune all the mobiles in its cell so that it receives them all at approximately the same power. This ensures that the base station will receive the signals from all the mobiles in its cell at the same SIR.

2.3 The Simulation

This section discusses the setup of the IS-95 system characterization simulation.

2.3.1 Simulation Topology

The simulation is a 7 cell simulation. The cell arrangement is shown in Figure 2.3. The numbered circles represent the base stations in the center of each cell. The radius of each cell is defined as the length of the dotted line labeled R in Figure 2.4. Interfering mobiles are uniformly scattered over each of the 7 cells. Each of the interfering mobiles remain in the same position for the entire simulation.

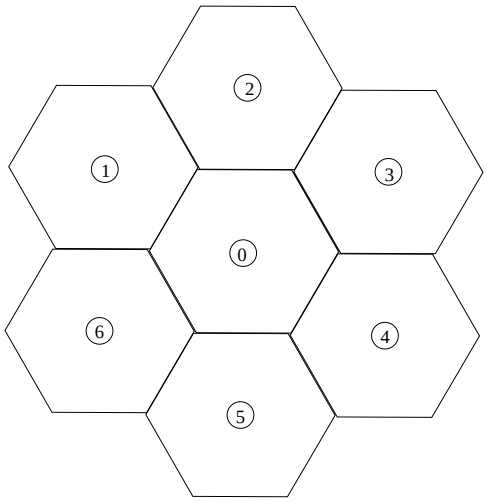

Figure 2.3: Cell Topology

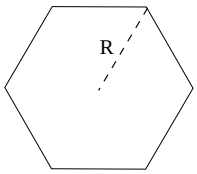

Figure 2.4: Cell Radius

It is assumed that the mobile the system is trying to locate is in the center cell.

This mobile is referred to as the quarry. The position of the quarry is moved in a grid pattern in the center cell. For each position, the SIR of the quarry's signal is measured at each of the 7 base stations in the system. The interference term in the SIR value includes interference from other mobiles as well as thermal noise. The result is several hexagonal grids of SIR data showing what the mobile's received SIR is at each of the neighbouring base stations for every region of the center cell.

The received SIR of the quarry is calculated only at the base station locations because of practical considerations for the cellular location system. In any cellular telephone location system, it is desirable to keep the cost of the system down by minimizing the extra infrastructure required to implement the system. This means that the receivers used to determine the mobile's position should be located in existing cellular base stations. Deploying receivers at additional locations would be very costly. With this in mind, only the received SIR values of the quarry's signal at the base station locations are important.

2.3.2 Simulation of Base Stations and Interfering Mobiles

The signals for the interfering mobiles and the quarry are generated separately and then combined. The mobiles' signals are generated using the transmitter structure discussed in Section 2.2.1. A random bit stream is used as an input to the structure and the output is a QPSK symbol stream. The spreading codes used are compliant with IS-95 specifications [6]. Each mobile signal is simulated at two samples per chip and pulse shaped using the filter specifications in the IS-95 standard [6].

Perfect power control is implemented for every mobile in the system. For each simulation run, the transmit power of each of the mobiles is adjusted so that they are

received at exactly the same power at their respective base stations. This simulates a system where the closed loop power control is quick enough to compensate for all variations of the mobile's received power from the mean.

In a real system, this type of power control would not be possible. Even though the mobile's signal is very wide band, there will be a bit of small scale variation in the received power of the mobile, caused by Rayleigh fading. A statistically independent small scale fade occurs at every half wavelength. In IS-95, a half wavelength is approximately 16 cm at cellular frequencies. If the mobile is moving, these types of channel variations would occur too quickly for power control to compensate. Power control has a finite response time for even large scale channel variations, like shadowing. As a result, the received SIR values of the mobiles will usually have a standard deviation of approximately 1-2 dB from the ideal.

Each base station in the simulation use 120 degree sectorized antennas. An ideal 120 degree sectorized cell with sector numbers is shown in Figure 2.5.

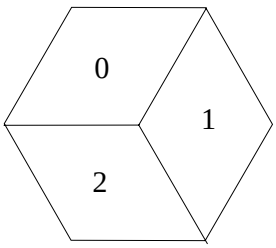

Figure 2.5: 120 Degree Sectorization

An approximate expression for the antenna gain pattern of a real 120 degree sectorizing antenna is used [7]. The gain, $G(\theta)$, as a function of angle from the

maximum gain line is given in Equation 2.6.

$$G(\theta) = \begin{cases} 1 - \frac{(1-b)}{(\pi/3)^2}\theta^2, & |\theta| \leq \sqrt{\frac{1-a}{1-b}}\frac{\pi}{3} \\ a, & \text{elsewhere} \end{cases} \tag{2.6}$$

The constant a is equal to $3.162 \cdot 10^{-2}$ and the constant b is equal to $3.981 \cdot 10^{-1}$.

In addition to the interference of other mobiles, thermal noise is also added to the waveforms received by the base stations. The formula for thermal noise power is given by

$$P_n = k_B T W \tag{2.7}$$

where k_B is Boltzmann's constant, $1.38 \cdot 10^{-23}$Ws/K, T is the operating temperature, 300 K, and W is the bandwidth, 1.25 MHz.

2.3.3 Channel Model

The channel model used in this simulation is a log-distance path loss model with log-normal shadowing added for some of the simulation runs. The log distance path loss model is given in Equation 2.8 [4].

$$\begin{aligned} P_{r,ld} &= c(d)^{-n} \\ c &= \frac{G_b G_m \lambda^2 P_t d_o^n}{(4\pi)^2 d_o^2} \end{aligned} \tag{2.8}$$

$P_{r,ld} \equiv$ received power calculated using the log-distance equation (W)

$d \equiv$ distance of the receiver from the transmitter (m)

$n \equiv$ path loss exponent

$G_b \equiv$ base station antenna gain

$G_m \equiv$ mobile antenna gain

$\lambda \equiv$ signal wavelength (m)

$d_o \equiv$ reference distance (m)

$P_t \equiv$ transmit power (W)

For this simulation, an omni-directional antenna is assumed for the mobile ($G_m = 1$). A transmit frequency of 900 MHz is assumed for the base station and a transmit frequency of 800 MHz is assumed for the mobile. A transmit power of 10 W is used for the base station and a distance of 50 m is used for d_o. The base station antenna gain, G_b, is given in Equation 2.6.

Once the received power of the mobile's signal is calculated using the log-distance model, log-normal shadowing is added to that value. The final received power in dBW, P_r, is given by Equation 2.9 [4].

$$P_r = P_{r,ld} + \Omega_i \tag{2.9}$$

The term $P_{r,ld}$ is the power calculated using Equation 2.8, expressed in dBW, and Ω_i is the shadowing value in the direction of base station i, expressed in dB.

Each mobile receives a different shadow in the direction of each of the 7 base stations in the simulation. It is assumed that the shadows experienced by the mobile in direction of each of the towers are correlated. The shadowing values in the direction

of base station i, Ω_i, are calculated using Equation 2.10 [8].

$$\Omega_{k+1} = \xi\Omega_k + (1-\xi)X_\sigma \tag{2.10}$$

In this equation, the shadow value in the direction of base station $k+1$, is calculated using the shadow value in the direction of base station k and a new shadowing value, X_σ. The value X_σ is a log-normal random process with a standard deviation of 10 dB [4]. The correlation factor between shadows is ξ. For this simulation, ξ was equal to 0.5.

One problem with is equation is that it produces a discontinuity in the correlation between the first and last shadowing values that are calculated. However, for these simulations it is still a good approximation.

Rayleigh fading was not implemented in this simulation. This was partly for simplification of the simulation and partly because the fluctuations in the signal due to small scale channel effects are quite small compared to the shadowing process described above. A study done by Wu, *et. al.* supports this assumption [9].

The paper by Wu, *et. al.* performs an analysis of received power fading in a RAKE receiver due to small-scale channel effects. The RAKE receiver is the receiver style used by IS-95 mobiles and base stations and consists of several parallel components called fingers. Each finger tracks one of the multipath components of the signal being received over the radio channel. Those multipath components are combined in the receiver in order to utilize as much of the energy transmitted through the channel as possible. Based on measurements taken on cellular radio channels, Wu performs an analysis to determine the probability distribution function (PDF) of the magnitude

of the multipath signals resolved by the RAKE fingers. The paper shows that the results didn't quite fit a Rayleigh or Ricean distribution. However, it does show that 92% of the time, the small-scale fades of the finger following the strongest multipath component vary less than 6 dB from the mean. This error is small compared to the log-normal shadowing process described above.

2.3.4 Calculation of Signal-to-Interference Ratio

Once the signals from the interfering mobiles and the quarry are simulated, it is necessary to determine the SIR of the quarry's signal. The SIR of the quarry's signal, SIR_q, is defined in Equation 2.11 where A is the received amplitude of the quarry's signal and σ_I^2 is the variance of the two sided, complex process that represents the interfering users plus the thermal noise in the channel. Since the interference from other IS-95 users can be approximated as a Gaussian process [5], the power of the total interference process can be characterized using the variance σ_I^2.

$$SIR_q = \frac{A^2}{\sigma_I^2} \tag{2.11}$$

The SIR of the quarry's signal is determined using a correlation. Note that since a stationary channel is assumed, a coherent correlation is performed. The derivation below assumes a coherent correlation. The received signal from the quarry plus the signals from the interfering mobiles are correlated with a clean copy of the quarry's signal. The quarry's signal is essentially a long PN sequence because of the spreading codes used by the mobile. This means that when the quarry's signal is correlated with itself, a spike is produced at the correlator output. If the correlation is long enough, that spike will rise out of the interference caused by the other mobiles. The

SIR of that peak, SIR_p, can be used to calculate SIR_q.

The SIR of the correlation peak is given in Equation 2.12.

$$SIR_p = \frac{A_p^2}{\sigma_{I,p}^2} \tag{2.12}$$

Equation 2.12 can be related to SIR_q. The term A_p is the amplitude of the correlation peak and $\sigma_{I,p}^2$ is the variance of the two-sided, complex Gaussian process representing the interfering users and channel noise at the output of the correlator. The peak of a PN sequence correlation is N times the amplitude of each chip, A, where N is the number of chips in the correlation. Spreading a signal with an N chip PN sequence expands the required bandwidth of the system by N. This increases the interference power by a factor of N, if a Gaussian PSD is assumed. This also assumes that the interference in each sample is independent. This gives Equation 2.13.

$$SIR_p = \frac{(NA)^2}{N\sigma_o^2} = \frac{NA^2}{\sigma_o^2} \tag{2.13}$$

Using the results in Equation 2.13, Equation 2.14 shows the SIR of the quarry's signal expressed as a function of the SIR of its correlation peak. Both SIR values are shown in dB.

$$SIR_q = SIR_p - 10\log(N) \tag{2.14}$$

The signals of the mobiles are complex, since each of the mobiles use QPSK modulation. This means that it is necessary to perform a complex correlation and take the magnitude of the correlation output in order to extract the quarry's signal.

Since the magnitude of the complex correlation is taken, it is necessary to use some property of that magnitude waveform to calculate $\sigma_{I,p}^2$. The multi-user interference

and thermal noise in the system can be approximated as a complex Gaussian process. The magnitude of a complex Gaussian process is a Rayleigh random variable. The mean of a Rayleigh variable can be used to calculate the variance of the Gaussian process used to form it [4]. The relation is given in Equation 2.15 and is used to calculate the denominator of the SIR_p term.

$$E[R] = \sqrt{\frac{\pi}{2}}\sigma_{o,p} \tag{2.15}$$

2.3.5 Types of Coverage Areas Simulated

Three different types of coverage areas are simulated: urban, suburban and rural. The cell radii, channel path loss exponent and number of users per cell for each of the area types are given in Table 2.1.

Coverage Area	Radius (m)	Path Loss Exponent	Number of Users per Cell
Rural	5000	2.7	15
Suburban	1500	2.7	20
Urban	750	4	25

Table 2.1: Simulated Coverage Areas

The cell radii values and the number of users are estimates of what typical values for each coverage area type might be. A path loss exponent of 4 was used for the urban simulations since it is a typical conservative estimate for cellular systems [4]. However, in 1992, Rappaport published a paper specifically discussing the path loss exponent for wide band CDMA signals [11]. In this paper, he suggests that a path loss value of 2.7 is more realistic. In the paper, the variance of the measurements used to calculate this value is very large. However, it is still considered a good estimate and was used for the rural an suburban simulations.

2.4 Simulation Results

This section presents the results from the simulation runs. The SIR data of the quarry is presented in two forms: contour plots and histograms. Contour plots are used because they show what the SIR levels of the quarry's signal are in different regions of the center cell. The histograms give a good breakdown of the fraction of quarry positions falling into a specific range of SIR values. The contour plots are given in Section 2.4.1 and the histogram results are in Section 2.4.2.

2.4.1 Contour Plots

The contour plots generated for the three coverage area types are shown below. Each contour plot is a plot of the area in the center cell. The position of the center base station is indicated by the X. The contour lines show the regions in the center cell where the quarry's received signal power is at a certain level. Contour plots were generated showing the SIR of the quarry's signal when the receiver is located at base station 1. The location of base station 1 is shown in Figure 2.3. Initially, contour plots were generated showing the SIR levels for all 6 neighbouring cells. However, due to the symmetric nature of the cell topology, each of the neighbouring base stations had very similar plots. As a result, only the contour plots for the SIR levels at base station 1 are given. Figure 2.6 shows the contour plot for the urban case, Figure 2.7 shows the suburban case and Figure 2.8 shows the rural case.

Contour plots were only generated for the simulations without shadowing. When shadowing was added, it caused the received SIR values of adjacent quarry positions to vary considerably. This made it impossible to pick out any specific trends using contours. However, since shadowing is a zero mean process, the average trends in the

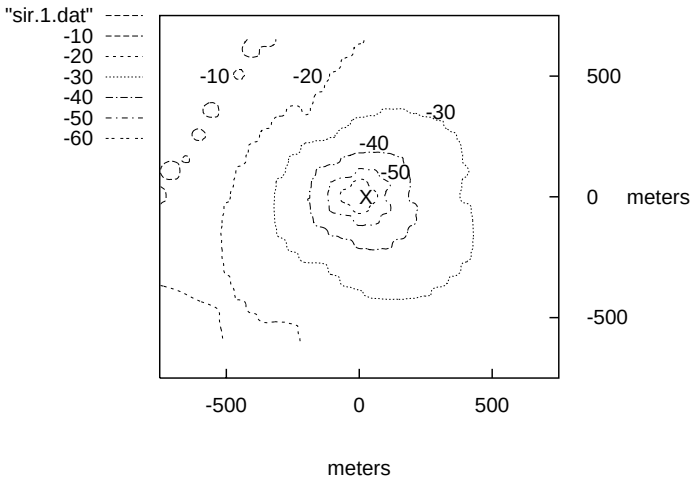

Figure 2.6: Urban SIR Contour Plot

results with shadowing will tend to be the same as the results without shadowing.

All the contour plots show that the region of lowest SIR is in the center of the cell. This is due to power control. When the quarry is very close to the base station in the center of the cell, power control commands from that base station turn the quarry's power down to a very low level. This reduces the SIR level of the quarry's signal at base station 1. The approximately symmetric rings around the center of the cell show that power control has an even greater effect than attenuation due to distance, when the mobile is close to the center of the cell. In the areas further away from the center base station, the regions of low SIR are larger on the far side of the base station due to the extra attenuation due to distance.

The same contour regions were present on each of the histograms. This means

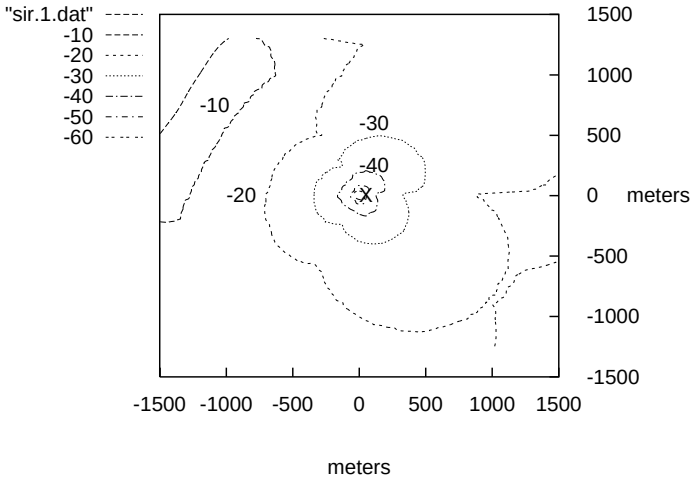

Figure 2.7: Suburban SIR Contour Plot

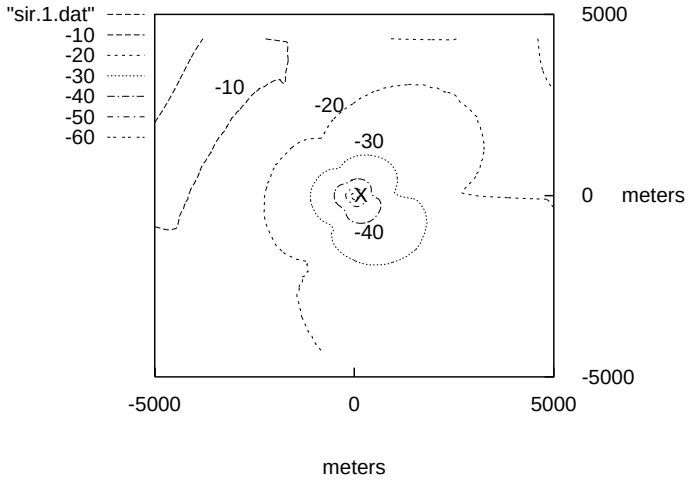

Figure 2.8: Rural SIR Contour Plot

that it is possible that a location system might have to deal with the same extreme signal levels in each of the three cases. However, the areas of very low SIR occupy a progressively smaller proportion of the total coverage area in the less populated coverage area types. This is due to two factors. First, the total number of interfering mobiles is decreased, which reduces the interference term. Secondly, the channel path loss exponent is also reduced, which means the mobile's signal is attenuated less as it travels to the neighbouring base station. The signal attenuation is increased by the increasing cell size, however this has less of an effect than the reduced path loss exponent and interference. The net result is an overall increase in the proportion of the cell covered by regions of favorable SIR for the suburban and rural cases.

2.4.2 Histograms

The histograms generated for the 3 simulation runs are given in this section. Quarry positions were generated in a 50 m grid in the center cell. The received SIR of the quarry's signal was recorded in a histogram for each of those positions. The histograms show the distribution of the SIR levels of the quarry's signal for all the positions in the center cell.

Initially, separate histograms were generated showing the distribution of the SIR of the quarry's signal for receivers at each of the neighbouring base stations. However, like the contour plots, the histograms for different neighbouring base stations were all very similar due to the symmetric arrangement of the cells in the simulation. As a result, the received SIR levels recorded at each neighbouring base station were combined into a single histogram. The histograms for the different coverage areas are given below. When each histogram was generated, it was assumed that a mobile

could get no closer than 75 m to the base station in the center cell. As a result, SIR measurements for all quarry positions inside a 75 m square in the middle of the cell were not included in the histogram plots.

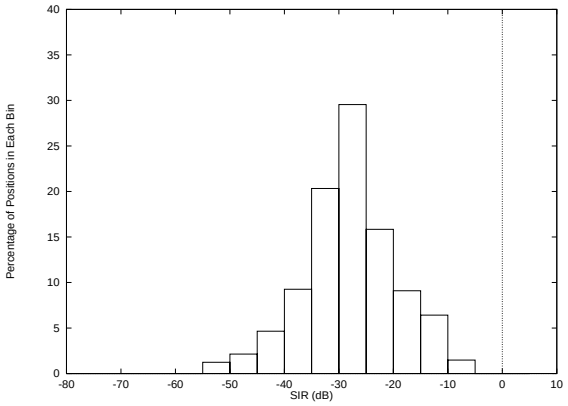

Figure 2.9: Urban Coverage Area (No Shadowing)

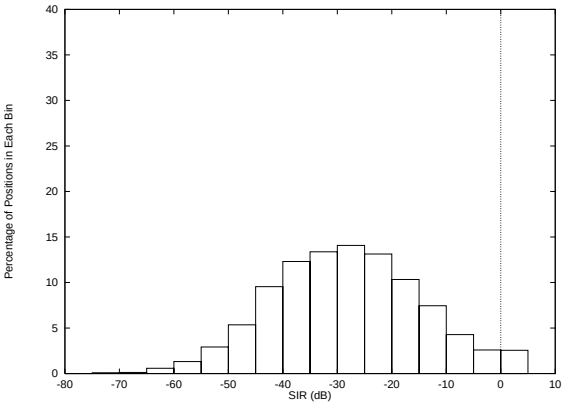

Figure 2.12: Urban Coverage Area (Shadowing)

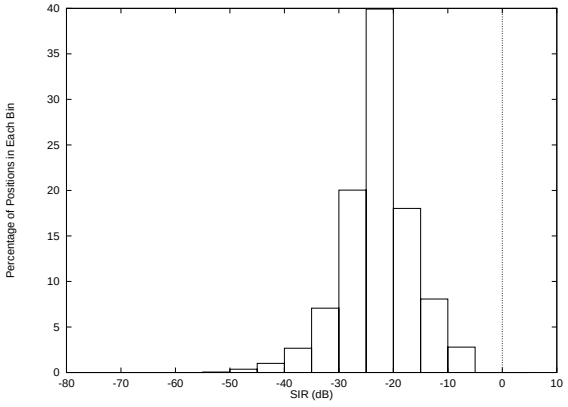

Figure 2.10: Suburban Coverage Area (No Shadowing)

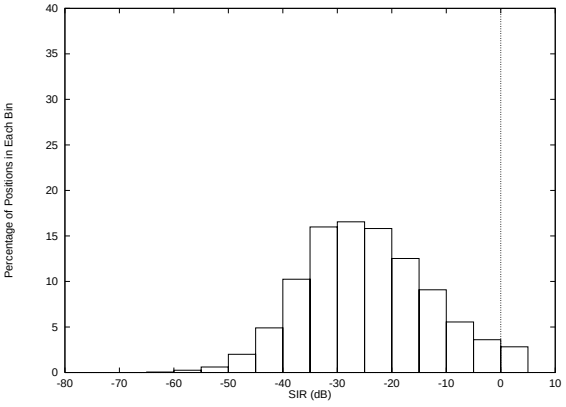

Figure 2.13: Suburban Coverage Area (Shadowing)

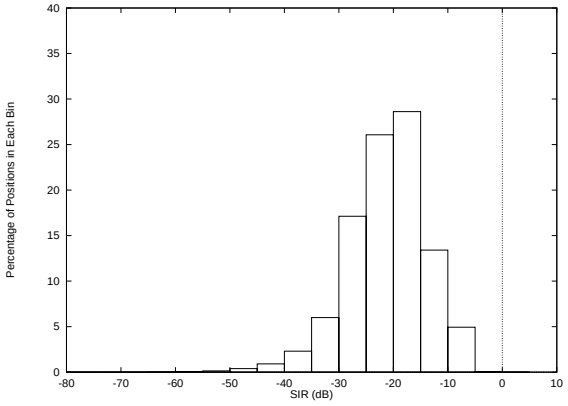

Figure 2.11: Rural Coverage Area (No Shadowing)

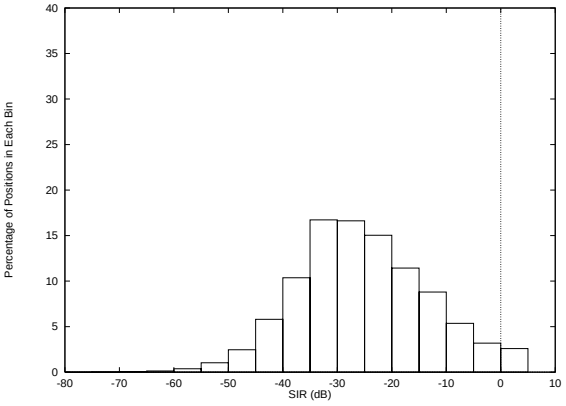

Figure 2.14: Rural Coverage Area (Shadowing)

The histograms allow the SIR distribution for a specific coverage area to be compared for the shadowing and non-shadowing cases. When shadowing is added, the histograms become slightly more spread out. This was to be expected, since the additive log-normal shadowing process increases the spread of the received SIR levels.

The histograms give an approximate picture of the probability density function (PDF) of the received power of the mobile's signal at a neighbouring base station. All histograms are approximately Gaussian with the shadowing results having a larger variance than the results without shadowing.

While no detailed statistical analysis is performed here, the plots show that extremely low SIR levels are possible. In order for a location system to be effective, it must be designed to handle these types of signal levels.

The histogram plots also confirm the information shown in the contour plots. For both shadowing and non-shadowing cases, as the coverage area became more sparse, a higher proportion of the SIR values moved into higher valued bins. This means that a location system that functions at higher SIR values would be able to come up with a larger number of position fixes in less populated areas, like suburban or rural areas, as opposed to denser urban areas.

2.5 Conclusion

The system characterization simulations presented in this chapter clearly indicate that very low SIR levels must be dealt with in order to consistently extract the signal of an IS-95 mobile at base stations in neighbouring cells. The contour plots show that the main reason for this is power control. As the mobile gets closer to the base station it is communicating with, its transmit power is reduced considerably.

The simulations also show that the proportion of the cell coverage area at low SIR values increases for a denser coverage area. The urban simulations had the highest proportion of quarry positions at low SIR values while the rural simulation had the highest proportion of quarry positions at high SIR values.

The histograms generated give an approximate picture of the PDF of the received SIR level of a mobile at a neighbouring base station. They indicate that it is possible to experience extremely low SIR levels, in all three coverage are types. A system designed to provide reliable position estimates must be prepared to deal with these types of received signal levels.

Performing a simulation of an entire cellular network is a very complex process. As with any simulation, assumptions were made that affect the simulation outcome. This means that these simulation results should not be considered definitive values for all IS-95 cellular systems. Instead, they should be treated as estimates that are a good place to start a location system design. If a higher degree of certainty is required, the simulations should be complemented by measurements performed on an actual IS-95 cellular system.

However, these results do show that the signal levels in a real IS-95 system are

likely to be extreme. Designing a location system with such small signal values will clearly make some location techniques impractical. The results presented in this section are a valuable guideline for deciding which location techniques are practical solutions for finding an IS-95 mobile.

Chapter 3

IS-95 Mobile Signal Location

3.1 Introduction

Most traditional techniques used to locate a device transmitting a radio signal use some characteristic of that signal. Usually the signal transmitted by the device must be received at several known locations. After reception, the signal characteristic is combined with the known receiver locations and used to solve for the location of the signal source.

The location system presented in this chapter uses the signal transmitted by the mobile to determine its location. The location estimates are calculated using the time difference of arrival (TDOA) of the mobile's signal at different base stations. Section 3.2 describes how TDOA works and why it was chosen for this system. Previous work in IS-95 mobile time-of-arrival based location is discussed in Section 3.3. Section 3.4 discusses how TDOA measurements are used to solve for an estimate of the coordinate position of the mobile. Section 3.5 addresses the problem of determining the time-of-arrival of an IS-95 mobile's signal. Section 3.6 presents a simulation that evaluates the accuracy of this approach. Section 3.7 describes how a super-resolution approach, based on the Root-MUSIC algorithm, can be used to enhance the accuracy of the system. Section 3.8 presents some conclusions.

3.2 Time Difference of Arrival Location

3.2.1 Benefits of a Time-of-Arrival Based Location Scheme

There are three basic characteristics of a mobile's signal that can be used to determine its location. The first characteristic is the received strength of the signal, the second is the angle-of-arrival of the signal and the third is the time-of-arrival of the signal.

A signal strength location based scheme takes advantage of the fact that the average received power of a radio signal decays in a known fashion with distance. The received power of a signal can be recorded at several known locations and used to determine the distance of the mobile from each of those locations. Those distances can be used to determine the coordinate position of the phone. The main disadvantage with a signal strength location system is the large random fluctuations from the mean received signal power due to multipath reflections in the wireless channel. Small scale fading in the channel, combined with shadowing due to large obstructions, can cause errors in the received signal power of 10 dB or more. This large amount of error generally makes a signal strength location scheme worth considering only when other location techniques are not possible.

An angle-of-arrival (AOA) approach determines the mobile's location based on the direction of the mobile's signal. The angle-of-arrival of the mobile's signal can be determined at several known locations. These angles define the directions of lines that start from each receiver and intersect the mobile's location.

There are three disadvantages with the angle-of-arrival approach. The first disadvantage is that a large antenna array is usually required to determine the angle of arrival of a signal. This adds a great deal to the complexity and expense of the

location system. The second disadvantage is that this technique can be confused by multipath reflections. In a radio channel, several reflected copies of the mobile's signal will arrive at the base station receiver. It's possible for the antenna array to lock onto one of the signal reflections rather than the line-of-sight component of the signal. The third disadvantage is that the error in angle-of-arrival systems increases with distance [12].

The time-of-arrival (TOA) location approach uses the arrival time of a device's signal to determine its position. The time-of-arrival of a mobile's signal is determined at several known locations. The relative times of arrival at two known locations or the absolute times of arrival at each location can be used to solve for the position of the phone.

Using the time-of-arrival of a signal to determine the location of its source has some advantages over the methods described above. The first advantage is that a time-of-arrival location system can be implemented using the antennas already present in the base station. No antenna array is required. This dramatically cuts down on the cost and complexity of the location system. The second advantage is that this technique is fairly robust when used in a multipath channel. Generally, the first rising edge of the signal is used to determine its arrival time. Since the multipath reflections of the radio channel arrive after the line-of-sight component of the radio signal, they usually do not corrupt the rising edge of the signal. This means that the multipath radio channel does not cause as much error in this method as it does in angle-of-arrival or signal strength location. Because of these advantages, time-of-arrival location was the technique chosen for the location system presented in this chapter.

One possible disadvantage to a time-of-arrival based scheme is that the receivers

in the system have to have precise timing synchronization. This synchronization has to be accurate to the order of less than 10^{-7} s. However, this problem can be solved by synchronizing using the clock reference of a GPS receiver. GPS receivers are able to synchronize their clocks to an accuracy well below this level [13]. Even with the extra error due to Selective Availability, the time synchronization accuracy is sufficient.

3.2.2 Overview of Time Difference of Arrival Location

The time-of-arrival technique used in this location system is called relative time-of-arrival or time difference of arrival (TDOA) location. The basic operation of the system can be illustrated using Figure 3.1. The numbered circles indicate three base stations. Each base station is connected to a location processor by a land-based network connection. In the middle of the base stations is the mobile that the system is trying to locate. The first step is for each of the base stations to determine the TOA of the mobile's signal at each of their locations. Section 3.5 discusses in detail how this can be done for an IS-95 mobile. The TOAs of the mobile's signal at each base station are then sent to the location processor. The processor uses those TOA estimates to determine the location of the mobile.

The location processor uses the difference of the TOAs of the mobile's signal at two base stations to define a hyperbola. The TOA of the mobile's signal at base station i is denoted as t_i and the TOA of the mobile's signal at base station j is denoted as t_j. The difference between the two measurements is given by

$$t_i - t_j = \frac{d_i - d_j}{c} \tag{3.1}$$

where d_i and d_j are the distances of the mobile from base station i and j, respectively.

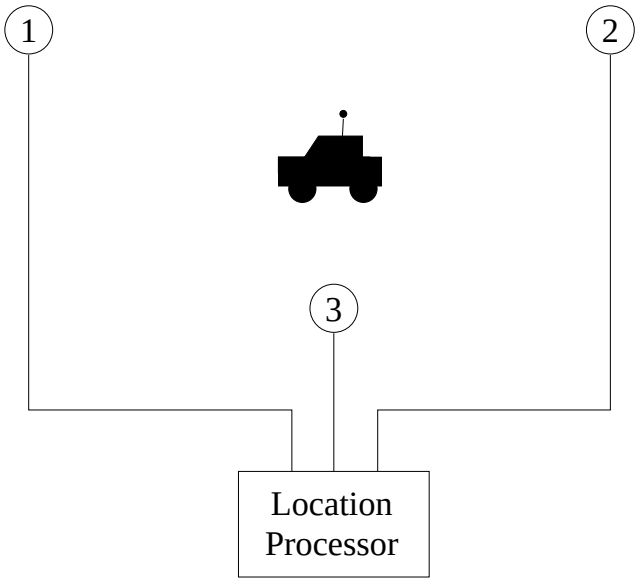

Figure 3.1: Simple Cellular System

The c term denotes the speed of light. Equation 3.2 shows Equation 3.1 when the distances are expressed as a function of the coordinate position of base station i, (x_i, y_i), the coordinate position of base station j, (x_j, y_j) and the coordinate position of the mobile (x, y).

$$t_i - t_j = \frac{\sqrt{(x_i - x)^2 + (y_i - y)^2} - \sqrt{(x_j - x)^2 + (y_j - y)^2}}{c} \tag{3.2}$$

The mobile's position can be located anywhere on the hyperbola defined by this equation.

TOA estimates from several base stations define several hyperbola. The location of the mobile is determined by solving for the intersection of the hyperbola. For example, the TOA estimates from the base stations in Figure 3.1 define two hyperbola, as shown in Figure 3.2. The position of the mobile is given by the intersection of the hyperbola. It is possible for the hyperbola to intersect at more than one location, giving several

location solutions. This can be dealt with by using approximate knowledge of the mobile's position to discard unreasonable location estimates.

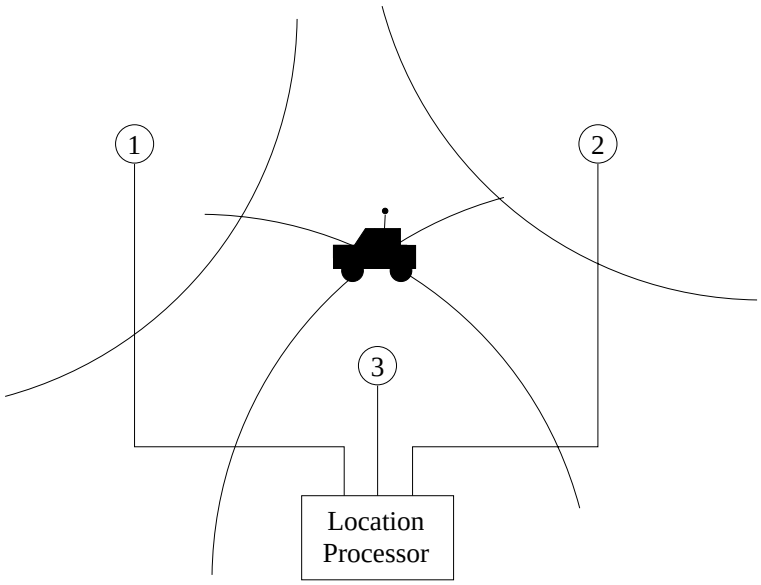

Figure 3.2: Relative Time-of-Arrival Hyperbola

3.3 Previous Work

Locating a mobile using the time-of-arrival of its signal is a very popular technique. TOA location schemes have appeared in patents as well as other technical papers. While there is some variation in how the technology is applied, each of the systems below all use the TOA of the mobile's signal in some way to determine its location.

The system disclosed by Schilling is designed for a system very similar to IS-95 but with some modification [14]. In this patent, the base station transmits a pilot signal that carries a generic spreading code that does not carry any information. The mobile synchronizes its timing reference to that pilot signal. Once it has established a time reference, the mobile transmits back a pilot signal of its own, consisting of another generic spreading code without information. The base station receives the mobile's pilot and calculates the difference in timing between the mobile's pilot and the pilot it transmits. That difference is approximately equal to the round trip propagation time of the signal. That value is then used to calculate the distance from the base station to the mobile. This system has the advantage of providing location information, even when the mobile's signal is only being received by one base station. The disadvantage of this system is that it requires considerable modification to the mobile.

The location system described by Mallinckrodt uses a round-time TOA [15]. The base station sends a command to the mobile. The mobile transmits a reply back to the base station and the TOA of that reply signal is used to determine the round trip travel time of a signal between the mobile and base station. That time is then used to determine the distance between the mobile and base station. While this approach is simple, trying to calibrate for the response time of the mobile results in large errors in the position estimate [16].

The system described by Tong *et. al.* uses the relative time-of-arrival of a reference signal received from the mobile at several base stations to determine its location [17]. In this patent, the base station queries the mobile for a spread spectrum reference signal. The mobile transmits the signal and the base station demodulates it, re-encodes it and sends the signal over a land-based communications link to the other base stations. The base station then queries the mobile again and the mobile retransmits the reference signal. Using the copy of the signal they received from the base station communicating with the mobile, the other base stations determine the TOA of the signal when it is retransmitted by the mobile. Those TOA values are then used to determine the mobile's location. This is a good approach that could likely be implemented with very little modification of the cellular system and almost no modification to the mobile.

TOA based location schemes have also been proposed for other cellular systems. In a paper by Klukas *et. al.*, a TOA system is presented that determines the position of AMPS cellular telephones [18]. In this system, the TOA of an 11 chip Barker sequence is determined using a sliding correlator approach. The estimate is further refined using a super-resolution technique. While this location system is designed for a different cellular system than IS-95, many of the techniques used in this paper can be applied to the IS-95 location problem. In particular, the super-resolution technique for IS-95 TOA estimates is discussed in Section 3.7. The location system proposed by Borsodi *et. al.* is for the IS-54 digital system, but it is very similar to the system discussed by Klukas *et. al.* [19].

Some analysis has already been done for using TOA location schemes for IS-95 mobiles [20]. In this paper, Aatique *et. al.* investigates the accuracy of a TOA based

location scheme for IS-95 mobiles with power control. This paper concludes that due to power control reducing the transmit power of the mobile, there will be areas where the mobile's signal will be too weak for TOA location to work. The paper suggests having the mobile turn its power up as a solution for the problem. This solution and some other solutions for this problem are discussed in Section 3.5.

A TOA location system using a Delay Locked Loop (DLL) has also been proposed [8]. The use of the DLL provides improved TOA estimates. However, this paper does not deal at all with the problem low received SIR from the mobile presents to the coverage of a TOA based location system.

3.4 Solution Method for Time Difference of Arrival Location Equations

In a real location system, the TOA estimates of the mobile's signal will have some error. This means that the hyperbola defined by the time difference measurements will not all intersect at exactly the same point. This makes it necessary to find an estimate the mobile's position that best satisfies all the TOA equations.

A common technique for solving TDOA equations is the least-squares estimation of linearized versions of the equations [21]. This is the technique used for the location systems in this chapter because it is a straight-forward method to implement and it produces fairly good results. There are other approaches to solving the TDOA equations that can give slightly better accuracy. For example, maximum-likelihood estimation has been shown to improve location estimates over least-squares approach [22].

A least squares estimator is used to solve for the position of the mobile. The least squares solution for data in the form of the matrix equation

$$\mathbf{Ax} = \mathbf{b} \tag{3.3}$$

is given in Equation 3.4 [23].

$$\hat{\mathbf{x}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}. \tag{3.4}$$

In order to find the least squares solution for the relative time-of-arrival location equations, it is necessary to put them in the form of Equation 3.3. The TDOA equa-

tions are linearized by using the first two terms of their Taylor series [21]. The TDOA of the mobile's signal at base station i and j is denoted by τ_k which is a function, $f_k(x, y)$, of the unknown mobile position, given by Equation 3.2. The linearization of this equation is given by

$$\tau_k = f_k(x, y) = f_k(x_o, y_o) + a_{k,1}\delta_x + a_{k,2}\delta_y \tag{3.5}$$

where

$$\begin{aligned} a_{k,1} &= \left.\frac{\partial f_k}{\partial x}\right|_{x_o,y_o} = \frac{x_i - x_o}{c\sqrt{(x_i-x_o)^2+(y_i-y_o)^2}} - \frac{x_j - x_o}{c\sqrt{(x_j-x_o)^2+(y_j-y_o)^2}} \\ a_{k,2} &= \left.\frac{\partial f_k}{\partial y}\right|_{x_o,y_o} = \frac{y_i - y_o}{c\sqrt{(x_i-x_o)^2+(y_i-y_o)^2}} - \frac{y_j - y_o}{c\sqrt{(x_j-x_o)^2+(y_j-y_o)^2}} \\ \delta_x &= (x - x_o) \quad \delta_y = (y - y_o). \end{aligned} \tag{3.6}$$

The values x_o and y_o are initial guess values for the mobile position. The initial guess can be some previous solution for the mobile position or it could also be obtained from a prediction algorithm. This derivation assumes that TOA estimates are available from N base stations. This means that $K = N - 1$ linearly independent TDOA equations can be formed.

Equation 3.7 expresses the set of linearized equations in matrix form.

$$\mathbf{z} = \mathbf{A}\boldsymbol{\Delta} \tag{3.7}$$

where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ \vdots & \vdots \\ a_{K1} & a_{K2} \end{bmatrix} \quad \boldsymbol{\Delta} = \begin{bmatrix} \delta_x \\ \delta_y \end{bmatrix} \quad \mathbf{z} = \begin{bmatrix} \tau_1 - f_1(x_o, y_o) \\ \tau_2 - f_2(x_o, y_o) \\ \vdots \\ \tau_K - f_K(x_o, y_o) \end{bmatrix} \tag{3.8}$$

Equation 3.7 is in the same form as Equation 3.3 and can be solved for a least squares solution. The least squares solution to Equation 3.7 is

$$\hat{\mathbf{\Delta}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{z}. \tag{3.9}$$

The matrix $\hat{\mathbf{\Delta}}$ is a least squares estimate of the difference between x_o and y_o and the actual solution of the equation. This difference can be added to x_o and y_o to give an estimate of the mobile position.

$$\begin{aligned} \hat{x} &= \hat{\delta_x} + x_o \\ \hat{y} &= \hat{\delta_y} + y_o \end{aligned} \tag{3.10}$$

However, unless x_o and y_o are already close to the actual solution of the equation, this estimate will not be very accurate. It will be more accurate than the initial guess, but it still could be quite different from the actual position of the mobile. To increase the accuracy of the estimate, the values for $\hat{x}$ and $\hat{y}$ are substituted for x_o and y_o in Equation 3.7 and a new least-squares solution is found. For this second solution, the Taylor series approximation is being evaluated closer to the actual solution of the equations, which improves the accuracy of the linearization. The result is a new estimate that is closer to the actual solution than the previous one.

This process of feeding back the new estimates into the algorithm continues until the matrix $\hat{\mathbf{\Delta}}$ is equal to zero. At this point,

$$\hat{\delta_x} = 0 \simeq (x - x_o) \quad \hat{\delta_y} = 0 \simeq (y - y_o) \tag{3.11}$$

and the estimates currently being used for x_o and y_o represent the least squared

estimate of the true position of the mobile.

It is possible for the technique to converge on a local minimum. It is also possible for the hyperbolic equations to have more than one solution and for this technique to find the incorrect one. If the algorithm does converge on an incorrect solution or diverge completely, this can be detected using some *a priori* knowledge that gives a coarse estimate of the phone's position. For example, knowledge of the cell the mobile is operating in could be used for this purpose. If an incorrect solution is detected, it can then be discarded.

3.5 Determining Signal Time-of-Arrival

Section 3.4 presents the method used to determine the position of the mobile using the relative time-of-arrival of the mobile's signal. This section discusses how the time-of-arrival values of an IS-95 mobile's signal are determined.

Section 3.5.1 illustrates how a correlator can be used to determine the TOA of the mobile's signal. Section 3.5.2 discusses some solutions for overcoming the very low SIR levels seen in Chapter 2.

3.5.1 Determining Time-of-Arrival using Correlation

Figure 3.3 shows the basic structure used to determine the TOA of a mobile's signal. The discrete time signal transmitted by the mobile is $m(n)$. The signal that arrives at the receiver is $r(n)$. In a simplified radio channel with no multi-path reflections, $r(n)$ consists of the mobile's signal $m(n)$ delayed by some amount k and multiplied by an attenuation factor A. The delay k is the TOA of the signal that is to be estimated. The received signal $r(n)$ is passed into a sliding correlator. As it arrives over the channel, the correlator correlates $r(n)$ with a copy of the signal transmitted by the mobile.

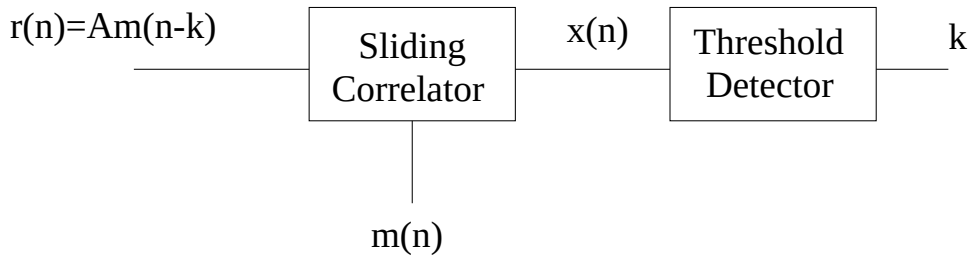

Figure 3.3: Sliding Correlator TOA Detector

The signal transmitted by the IS-95 mobile over the voice channel will consist

either of some information, multiplied by a PN sequence spreading code, or some PN spreading sequence without information that is known *a priori* by the base stations. In both cases, the correlation properties of the mobile's signal, $m(n)$, will be the same as the correlation properties of a simple PN sequence. This means that when a copy of the mobile's signal is correlated with the signal received from the mobile over the channel, a spike will be produced when the signal received from the mobile lines up with its copy in the correlation. This spike in the correlator output $x(n)$ occurs at the TOA of the signal, k.

The correlator output, $x(n)$, is passed into a threshold detector. This detector has a clock that is synchronized to the clocks in all the other TOA detectors in the location system. The threshold detector records the time the correlation output spike occurs. That time is forwarded to a processing center and used to determine the mobile's location as described in Section 3.4.

If the mobile transmits a known spreading sequence when being located, the base stations can continually correlate a copy of that signal with what they receive over the channel. However, if the mobile's signal contains voice data, it is impossible to know the content of the signal beforehand. In this case, the TOA detector shown in Figure 3.3 has to be modified. The modifications are shown in Figure 3.4.

For this structure, when a location request is submitted for a mobile, all the base stations in the area would buffer the signals they were receiving over the voice channel. Then, the base station closest to the mobile would demodulate the mobile's signal and send the information content of the signal, $v(n)$, to the other base stations in the area. That signal would be used to reconstruct a copy of the signal transmitted by the mobile, $m_v(n)$, at each of the base stations. That copy is then correlated with

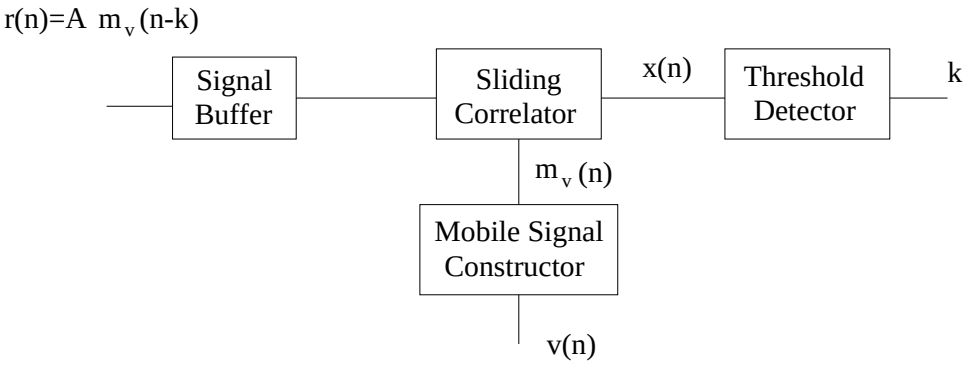

Figure 3.4: Sliding Correlator TOA Detector Modifications

the buffered signal received over the voice channel. In this case, the spike produced at the correlator output indicates the length of time after the base station started recording that the mobile's signal arrived. This TOA value is then forwarded to a central location and used to solve for the position of the mobile.

Before either of the structures described above can be used, the very low SIR levels seen in Chapter 2 have to be addressed. Some possible solutions to this problem are presented in the next section.

3.5.2 Solutions for Low Mobile SIR

The received SIR of a mobile's signal at a neighbouring base station can be extremely low. The SIR levels are especially severe when the mobile is close to the base station it is communicating with and its transmit power is reduced due to power control. One possible solution to this problem is for the mobile to violate its power control commands and transmit at maximum power until it is located.

The disadvantage of this approach is that increasing the transmit power of a mobile also increases the interference experienced by all the other mobiles in the area. Recently, an analysis was performed by Gosh *et. al.* on the effects of mobile

power increase for location purposes [24]. The conclusion was that the effects of mobile power increases can be tolerated by the system, which makes this approach a viable solution.

Another approach to solve the problem of very low received SIR is to use very long correlations. The height of the spike produced by the sliding correlator is equal the number of chips in the correlation. This is called the processing gain of the correlation, which is expressed as $10\log(N)$ where N is the number of chips in the correlation.

It is necessary to ensure that the processing gain is large enough to overcome the low SIR of the signal. The results in Chapter 2 show that SIR levels on the order of -50 dB are quite possible. If the correlation peak must be 10 dB above the noise in order to be detected, the processing gain must be 60 dB. This means $N \simeq 10^6$ chips.

The large processing gain required by the system means the TOA detectors have to perform very long correlations in order to extract the TOA of the mobile's signal. This must be taken into account when designing the hardware of the detectors. The first consideration is the buffer for storing the received signal. If signals on the order of 10^6 must be stored, assuming 2 samples per chip and at least 8 bits per sample, each buffer has to be able to store at least 2 Mbytes of data. Since this is a small amount of memory even for commercial personal computers, this buffer size should not present a problem.

The second hardware consideration is actually performing the correlation operation. Correlating signals that are several million samples long can be a fairly computationally expensive task. The problem with correlations is that their operation count is $O(N^2)$. However, this operation count can be reduced by using some *a priori* knowledge of the mobile's position. A reasonable assumption is that the system

will know at least what cell the mobile is in. This gives a position estimate with an accuracy of ± 1000 m for a typical urban cell size. This means that the TOA detector can estimate the TOA of the mobile's signal to somewhere in a window that is (2000 m)/$(3 \cdot 10^8$ m/s) $= 6.67 \mu s$ or approximately 7 chips wide. Since the TOA of the signal can be narrowed down to somewhere in this window, it's only necessary for the correlator to generate this window to determine the TOA of the signal. In the case of a 7 chip window at two samples per chip, the operation count is reduced to $O(14N)$ when only the window is generated.

Once the operation count is reduced, the correlation is well within the abilities of modern digital signal processors. The TMS320C62 is a 200 MHz DSP chip from Texas Instruments. Due to the highly pipelined nature of this chip, it is possible for it to execute two memory accesses, two multiply operations and an addition operation per clock cycle [25]. This means that the chip can execute the correlation operation at 1 billion operations per second.

With these long correlations, it might be undesirable for the phone to interrupt communication while it transmits a known spreading sequence with no voice information content. If that is the case, the technique described in the previous section for correlating a signal with voice content would be well suited for these long correlations.

When the base station near to the mobile transfers the mobile's signal to the base stations that are further away, it sends only the unknown information content of the signal. All the redundant information is removed. This idea can be considered a very effective form of compression. The chip rate of the IS-95 mobile is 1.2288 Mcps. If the signal was sampled at 2 samples/chip and 8 bits/sample, transferring a 10^6 chip sequence would mean approximately 16 million bits would have to be transferred

over the network. Assuming links of 64 kbps between each of the base stations, this would mean a transfer time of slightly under 5 minutes. This is obviously too long to wait. However, sending only the information content of the signal greatly reduces the number of bits to be transferred. A typical voice information bit rate is 9600 bps [6]. That means that the number of information bits that make up a 10^6 chip sequence is approximately 9600. This reduces the number of bits from 16 million to 9600, which is a compression ratio of 1667:1. Over a 64 kbps link, this means the transfer time is reduced to 150 ms.

The major problem with long correlations is not the correlation operation itself, but the Doppler spread in the channel. The coherence time of the channel, T_c, is a measure of approximately how long the radio channel will remain stationary. Coherence time can be calculated using

$$T_c \approx \frac{9}{16\pi f_m} \tag{3.12}$$

where f_m is the maximum Doppler shift of the channel [4]. The Doppler shift is equal to v/λ where v is equal to the velocity of the user and λ is the wavelength of the carrier. For a velocity of 100 km/h and a carrier frequency of 900 MHz, the Doppler shift is 83.3 Hz. This gives a coherence time of approximately 2.1 ms.

The correlation time necessary to extract the mobile's signal is much longer than the coherence time of the channel. This means that the random phase and amplitude changes in the channel will severely degrade the processing gain of the correlation.

The only way correlation can be used without the channel reducing the processing gain is to make sure the correlation time is less than the coherence time of the channel. Therefore, in order to use this type of TOA detection method, the long correlation

must be divided into several short correlations. The duration of each correlation must be less than the coherence time of the channel. The output of each of the short correlations can be combined on a power basis. That means each correlation output is squared and then added together. As more and more correlation peaks are added together, the interference will start to average out and the peak caused by the signal arrival will appear.

3.6 Time Difference of Arrival Location Simulations

This section presents the simulations used to evaluate the accuracy of a sliding correlator TOA location system that operates on the signal transmitted by the mobile. These simulations assume that by using one of the methods described in Section 3.5.2, the correlation operation in the detector is able to raise the correlation peak above the interference waveform and use it to estimate the TOA of the mobile's signal. If successfully implemented, either of the two techniques presented in the previous section will produce this same result. This means that the simulations presented below evaluate the accuracy of both approaches.

3.6.1 Simulation Setup

The simulation used the 7 cell topology, shown in Figure 3.5. The base stations in the center of each cell are indicated by the numbered circles. For each iteration, the mobile being located was placed at a random location in the center cell. This mobile is referred to as the quarry. The quarry's signal was generated using the IS-95 mobile structure described in Section 2.2. The spreading codes used by the mobile complied with the codes specified in [6]. Once the mobile's signal was generated, it was convolved with a channel model to determine the signal received at each of the 7 base stations. For the initial simulations, the mobile's signal was generated at two samples per chip, and the pulse shaping specified in the IS-95 standard [6] was used.

The simulation assumes that each base station has a sliding correlator TOA detector that implements one of the approaches described in Section 3.5.2. The simulated received signal from quarry was correlated with a clean copy of itself at each of the base station locations. The peak from that correlation was used to determine the

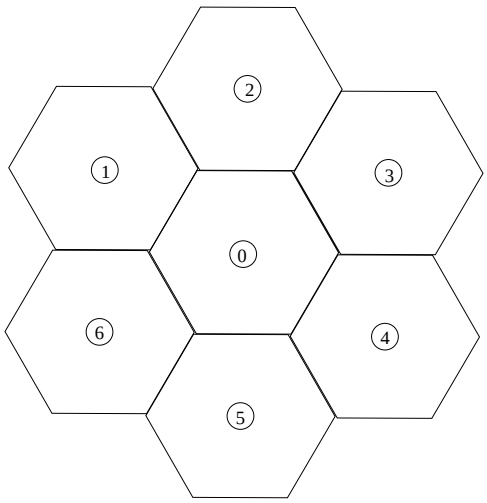

Figure 3.5: Cell Topology

TOA of the quarry's signal. The TOA of the signal at each of the base station locations was then used to calculate the quarry's position using the method described in Section 3.4. For each iteration, statistics were recorded on the error between the quarry's actual position and the position solved for by the system.

A simulation was also performed for a sampling rate of 8 samples/chip. The results were compared with the position error for the 2 samples/chip simulations to observe the effect of oversampling the signal.

The Cramer-Rao bound was plotted with the simulation results so the simulation error figures could be compared with a lower bound. The bound was generated using the Cramer-Rao bound for estimation of the TOA of a rectangular pulse [26]. This gives a lower bound on the variance of the TOA estimates. TOA estimates with this error variance were used in the position solving method described in Section 3.4 to solve for a number of mobile positions. The error distances of these position estimates give the Cramer-Rao bound in terms of radial error distance.

Channel Models

The first channel model used in this simulation was a simple 1-ray model. This channel model assumes a single, line-of-sight signal component. The impulse response of the channel is shown in Figure 3.6. Since a line-of-sight path for the signal is assumed, the time delay of the impulse, $T0$, is equal to d_i/c where d_i is the distance of the mobile from base station i and c is the speed of light.

The ray has a constant amplitude, $A0$. Since an IS-95 signal is wideband, Rayleigh fading for this channel model will be much less severe than the fading found in narrow band cellular systems [9]. As a result, this channel model assumes a constant amplitude ray, rather than a Rayleigh faded one. The value of $A0$ is calculated using the log-distance path loss model given in Equation 2.8. This equation is used to calculate the average received power of the quarry's signal, $P_{r,ld}$, at a distance d from the base station. The amplitude, $A0$ is then calculated using

$$A0 = \sqrt{P_{r,ld} T_c} \tag{3.13}$$

where T_c is the duration of a single chip.

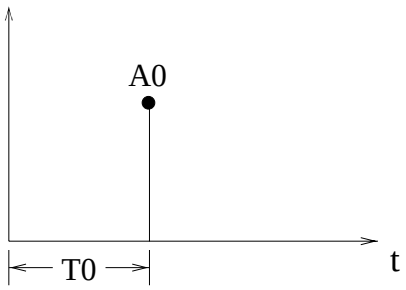

Figure 3.6: One Ray Channel Impulse Response

The second channel model used was a 4-ray model. This model assumes that the first ray in the channel impulse response is a line-of-sight ray and the three that come after it are multipath reflections. The impulse response of this channel model is given in Figure 3.7.

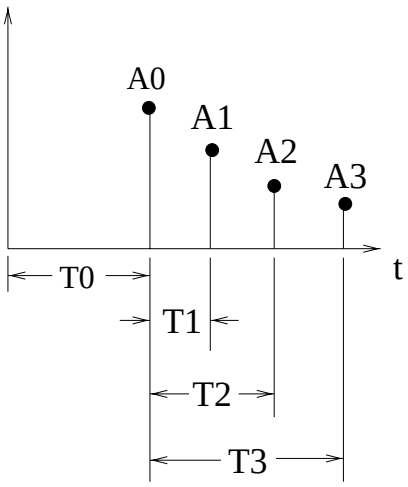

Figure 3.7: Four Ray Channel Impulse Response

Since the first ray is line-of-sight, $T0 = d_i/c$, where d_i is the distance from the quarry to base station i and c is the speed of light. The arrival times, $T1$, $T2$ and $T3$ are exponentially distributed after $T0$. Each of these arrival times are generated using the function shown in Figure 3.8. First, a value, y_i, is generated from a uniform distribution between 0 and 1. The value y_i is then mapped to a value x_i using the function $y = f(x) = 1 - e^{-x}$. The number x_i is a unitless value that is multiplied by a value t_{dmax}. This value is called the maximum delay spread of the channel and has units of seconds. For this simulation, the maximum delay spread was 10 μs. When 400,000 impulse responses were simulated with this value, the average RMS delay spread of the generated impulse responses was $1.26 \cdot 10^{-6}$ seconds. This is a typical average RMS delay spread value found in an outdoor channel [4].

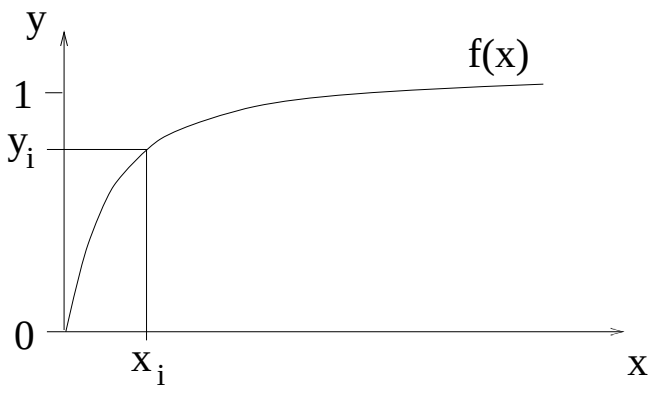

Figure 3.8: Exponential Distribution Function

The amplitude of each of the rays were also calculated using the log-distance path loss model. The amplitude of ray, Ai, is calculated using

$$Ai = \sqrt{P'_{r,ld} T_c} \tag{3.14}$$

where $P'_{r,ld}$ is the average received power of the mobile calculated at a distance d'_i using the log distance path loss model. The distance d'_i is equal to $d + (Ti - T0)c$ where d is the line-of-sight distance between the mobile and base station, Ti is the TOA of ray i and $T0$ is the TOA of the line-of-sight ray. This accounts for the additional attenuation of the multipath interferers due to the extra distance they travel.

Peak Detection Methods

Two simple approaches were used to determine the TOA of the spike at the correlator output in the TOA detector. Both techniques select a sample that corresponds to the start of the spike. The TOA of that sample is then used as the TOA of the mobile's signal.

Both methods assume that the TOA detector knows what cell the mobile is in.

This allows the detector to define a time window that should contain the TOA of the mobile's signal. The peak detection algorithm is then applied to only those samples in the window.

The first approach simply selected the sample with the largest amplitude. This technique is called max sample detection.

The second approach attempted to select a sample on the rising edge of the peak. This technique will give more accurate results if one of the multipath reflections in the channel has a larger amplitude than the line-of-sight component. This technique is called rising edge detection. This method involves finding the amplitude of the maximum sample in the window, A_{max}, and the standard deviation of the amplitude of the samples in the window, σ. An edge detection amplitude, A_{edge}, is defined as

$$A_{edge} = A_{max} - \gamma(A_{max} - \sigma) \tag{3.15}$$

where γ is a constant between 0 and 1. The first sample in the TOA window that exceeds A_{edge} is considered the rising edge of the correlation output spike and used as the TOA of the signal. For these simulations, γ was set to 0.6.

3.6.2 Simulation Results

Plots were generated showing the mean and standard deviation of the radial error distance of each of the position estimates of the algorithm. Radial error distance is defined as the two dimensional Euclidean distance between the actual position of the mobile and the position estimated by the algorithm. The radial error distance statistics are plotted vs. the SIR of the correlation peak.

Each plot shows 6 curves. The first five curves show the accuracy when TOA

estimates were available from a specific number of base stations. The number of base stations able to give TOA estimates was varied from 3 to 7. In all cases, the closest base stations to the mobile were used to provide TOA estimates. The sixth curve is the Cramer-Rao bound for the system.

Curves were generated showing the position error for each of the channel models and correlation spike detection methods discussed in the previous chapter. The figure numbers are given in Table 3.1 and the figures are given at the end of the section.

Detection Method	Channel Model	Mean Error Figure Number	Standard Deviation Figure Number
Max Sample	1 Ray	3.11	3.15
Max Sample	4 Ray	3.12	3.16
Rising Edge	1 Ray	3.13	3.17
Rising Edge	4 Ray	3.14	3.18

Table 3.1: Error Plot Figure Numbers

The plots show that the SIR of the correlation peak affects the accuracy of the simulations. All 4 plots show an improvement as the SIR of the correlation peak is increased. For the 1 ray channel, the error levels off at an SIR level of 15 dB. For the 4 ray channel, the error levels off at approximately 20 dB. This error bias is caused by bias in the TOA detection methods. It could be reduced by determining this bias and removing it.

The simulations show that for high SIR levels, increasing the number of base stations also increases the accuracy of the simulations. However, for low SIR levels, increasing the number of base stations actually increases the position error in some cases. The explanation for this is simply the large errors in the TOA estimates at these signal levels. The TOA estimates from each base station have such a large level of error that adding another estimate increases the overall error of the solution.

The plots also compare the effectiveness of maximum sample and rising edge detection. For the 1 ray channel, both methods have very similar accuracy. However, the max sample technique has much larger error values than the rising edge method for the 4 ray channel. This was expected, since it is possible for the max sample method to select one of the channel multipath rays by mistake. However, the rising edge of the line-of-sight ray is relatively unaffected. These results show that TOA detection methods that operate on the rising edge of the signal are much more reliable in a multipath environment.

Some simulations were also performed to investigate the effect of oversampling the signal. A simulation run was performed using the 4 ray channel model, 7 base stations per cell, rising edge detection and 8 samples per chip.

The mean error distance of this simulation is shown in Figure 3.19 and the standard deviation of the error distance is shown in Figure 3.20. The error statistics for the same simulation performed at 2 samples/chip are also shown on these figures for comparison.

The results show that increasing the sampling rate does reduce the error in the position estimates at high SIR values. This is because choosing a sample time as the signal TOA at 8 samples/chip has much lower quantization error than choosing a sample time at 2 samples/chip.

In these simulations, the rising edge detection method is able to satisfy the accuracy requirements of the FCC. When the SIR of the correlation peak is sufficient and rising edge detection is used, the mean radial error distance of the techniques is below 125 m for both channel types, regardless of the number of base stations.

However, even though the technique satisfies the FCC's requirements in these simulations, situations could arise in a real system where the system's accuracy would not satisfy the FCC. These situations relate to the geometry of the location system.

The accuracy of the position estimate depends a great deal on the geometry of the base stations determining the TOA of the mobile. Figure 3.9 shows an example of a good geometry situation. Here the mobile is surrounded by nearby base stations that are determining the TOA of its signal. In this situation, the location estimate calculated by the system is likely to be very good. Figure 3.10 shows an example of bad geometry. In this case, the mobile is located a fair distance outside the cluster of base stations and the estimate of its position in this case is likely to be very bad.

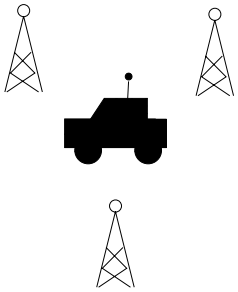

Figure 3.9: Example of Good Geometry

Figure 3.10: Example of Bad Geometry

A quantitative parameter of geometry is the Dilution-of-Precision (DOP) factor [27]. It relates the standard deviation of the position estimates, σ_p, to the standard

deviation of the TOA estimates, σ_{TOA}, as shown in

$$\sigma_p = (DOP)(\sigma_{TOA}). \tag{3.16}$$

A small DOP value indicates a strong geometry and a large DOP value indicates a weak geometry. In many cases, situations can arise where the DOP factor is much larger than 1. In this case, in order to keep the error of the position estimates below 125 m, the error of the TOA estimates would have to be well below 125 m. This means that it's quite likely that a more accurate method of estimating the TOA of the signal will be required. One method of enhancing the TOA estimate accuracy is presented in Section 3.7.

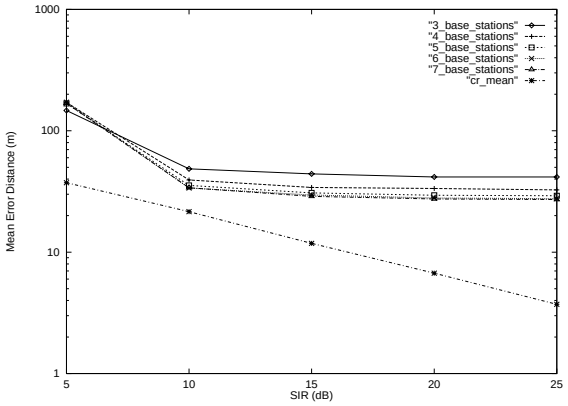

Figure 3.11: Mean Error Distance (Max Sample Detection, 1 Ray Channel)

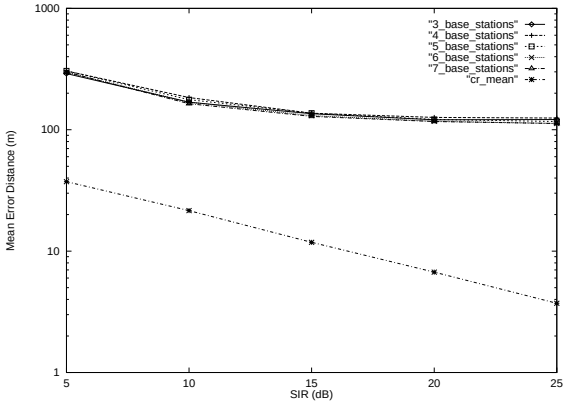

Figure 3.12: Mean Error Distance (Max Sample Detection, 4 Ray Channel)

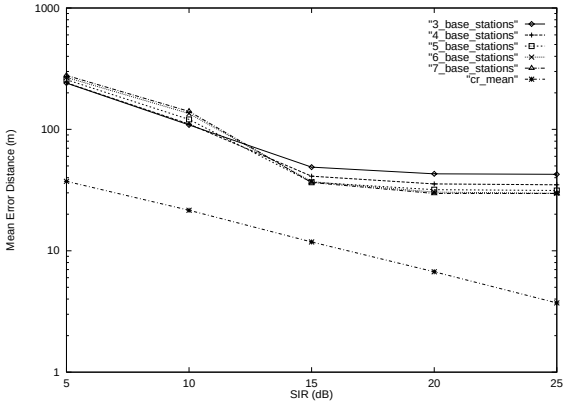

Figure 3.13: Mean Error Distance (Rising Edge Detection, 1 Ray Channel)

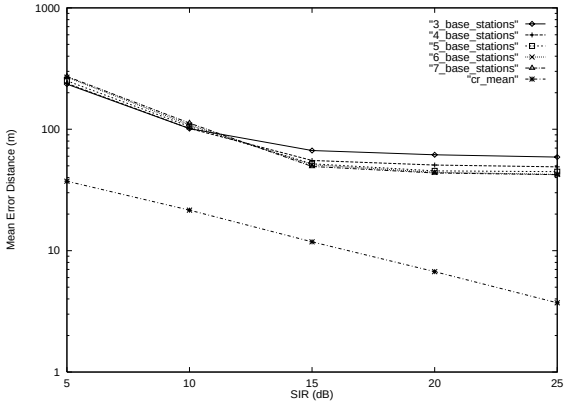

Figure 3.14: Mean Error Distance (Rising Edge Detection, 4 Ray Channel)

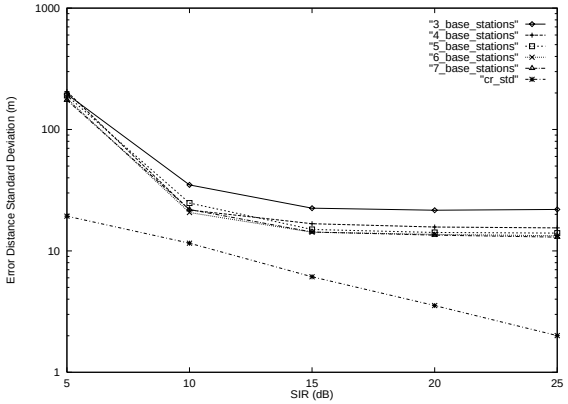

Figure 3.15: Error Distance Standard Deviation (Max Sample Detection, 1 Ray Channel)

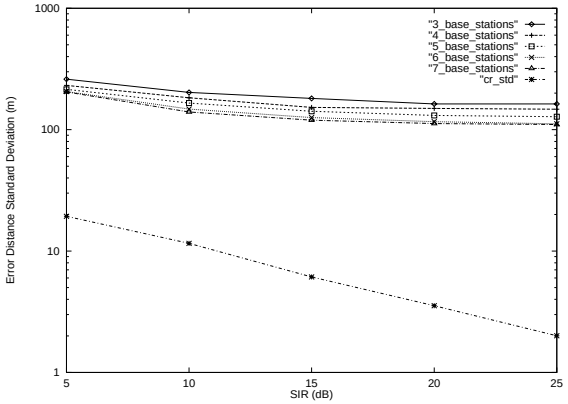

Figure 3.16: Error Distance Standard Deviation (Max Sample Detection, 4 Ray Channel)

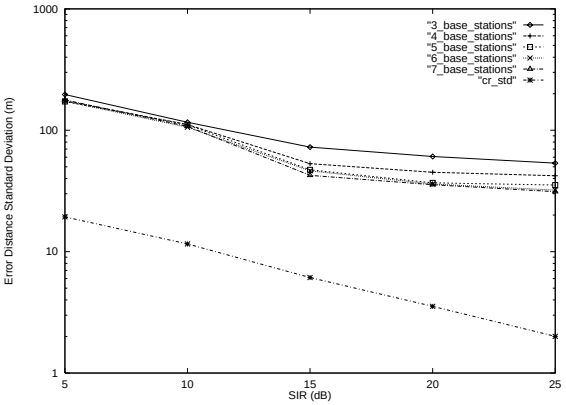

Figure 3.18: Error Distance Standard Deviation (Rising Edge Detection, 4 Ray Channel)

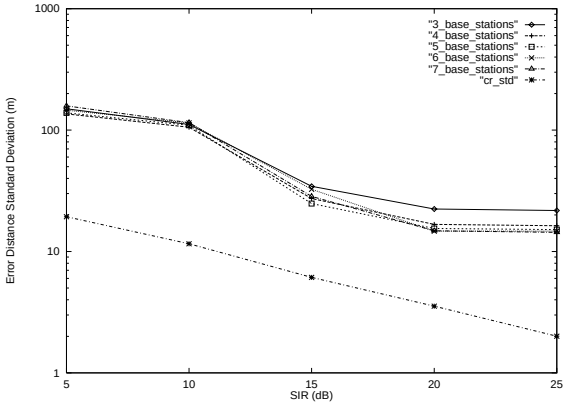

Figure 3.17: Error Distance Standard Deviation (Rising Edge Detection, 1 Ray Channel)

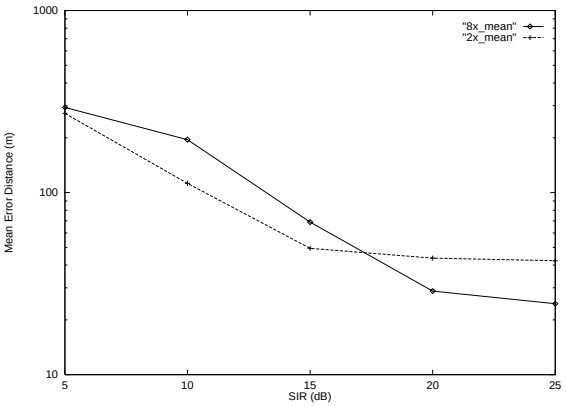

Figure 3.19: Mean Error Distance - 8 samples/chip

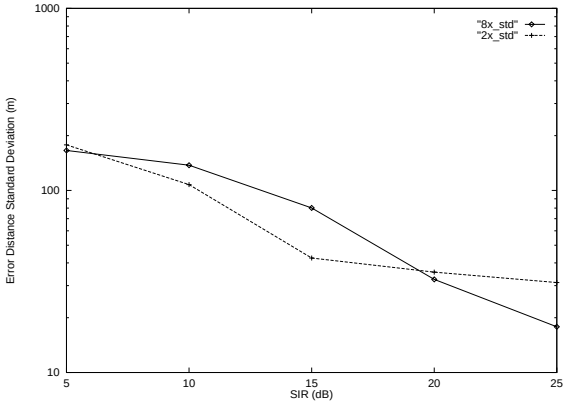

Figure 3.20: Error Distance Standard Deviation - 8 samples/chip

3.7 System Enhancement using a Super-Resolution Algorithm

This section discusses the use of super-resolution techniques for enhancing the performance of the TOA based location system presented in this chapter. The discussion is centered on the Multiple Signal Classification (MUSIC) super-resolution algorithm, which is well suited for improving TOA estimates [28]. Originally, super-resolution algorithms were used to calculate the frequency spectrum of a time domain signal with a much finer resolution that would be possible with a Fast-Fourier Transform (FFT). Section 3.7.1 introduces the MUSIC algorithm in this context. It turns out that super-resolution algorithms can also be used to determine a finely resolved picture of a time domain signal using frequency spectrum data. This gives a more accurate estimate of the TOA of a signal. This application is presented in Section 3.7.2. Section 3.7.3 illustrates the accuracy improvements possible when these techniques are incorporated into the location simulations presented in Section 3.6. Finally, Section 3.7.4 discusses some of the limitations of this technique.

3.7.1 Frequency Spectrum Super-Resolution

The MUSIC Algorithm

Super-resolution algorithms are specialized algorithms for estimating the frequencies of complex sinusoids in additive noise. It is assumed that the input signal consists of a sum of L uncorrelated, zero mean complex sinusoids given by

$$u(n) = \sum_{l=1}^{L} \alpha_l e^{j\omega_l n} + z(n) \tag{3.17}$$

where $z(n)$ is a complex additive noise component with variance σ^2. Unlike an FFT, the frequencies of the sinusoids do not have to be harmonically related.

The algorithm presented in this section is the MUSIC algorithm. This section is an overview of the discussion on MUSIC presented in [29]. The algorithm attempts to estimate the angular frequencies of the sinusoids in Equation 3.17. To do this, the algorithm uses the (M+1) by (M+1) ensemble-averaged correlation matrix $\mathbf{R}$ of the input sequence $u(n)$. The algorithm is based on the expression for $\mathbf{R}$ given in Equation 3.18.

$$\mathbf{R} = \mathbf{SDS}^H + \sigma^2\mathbf{I} \tag{3.18}$$

The matrix $\mathbf{I}$ is an (M+1) by (M+1) identity matrix and the rectangular matrix $\mathbf{S}$ is called the frequency matrix. It is defined in Equation 3.19.

$$\begin{aligned} \mathbf{S} &= [\mathbf{s_1}, \mathbf{s_2}, \ldots, \mathbf{s_L}] \\ &= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-j\omega_1} & e^{-j\omega_2} & \ldots & e^{-j\omega_L} \\ e^{-j2\omega_1} & e^{-j2\omega_2} & \ldots & e^{-j2\omega_L} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-jM\omega_1} & e^{-jM\omega_2} & \ldots & e^{-jM\omega_L} \end{bmatrix} \end{aligned} \tag{3.19}$$

The diagonal matrix $\mathbf{D}$ is defined as $\mathbf{D} = \text{diag}(P_1, P_2, \ldots, P_L)$ where P_i is the average power of the ith sinusoid.

The MUSIC algorithm is an eigenvector based method. The eigenvalues of $\mathbf{R}$ are denoted by $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_{M+1}$ and the eigenvalues of $\mathbf{SDS}^H$ are denoted by $\nu_1 \geq \nu_2 \cdots \geq \nu_{M+1}$. If the L sinusoids in $\mathbf{S}$ have distinct frequencies, the columns of

$\mathbf{S}$ will be linearly independent. This means that the $(M+1-L)$ smallest eigenvalues of the matrix $\mathbf{SDS}^H$ are equal to zero. Based on this assumption and Equation 3.18, we can relate the eigenvalues of $\mathbf{R}$ and $\mathbf{SDS}^H$ in Equation 3.20.

$$\lambda_i = \begin{cases} \nu_i + \sigma^2, & i = 1, 2, \ldots, L \\ \sigma^2, & i = L+1, \ldots, M+1 \end{cases} \tag{3.20}$$

A plot of λ_i is called an eigenspectrum. This plot shows that as the index i increases, the eigenvalues, λ_i, decay into a noise floor with amplitude σ^2. The eigenspectrum can be used to divide the eigenspace spanned by the eigenvectors into two subspaces. The eigenvectors corresponding to the first L eigenvalues form the signal plus noise subspace. The remaining eigenvectors form the noise subspace.

If $\mathbf{q_1}, \mathbf{q_2}, \ldots, \mathbf{q_{M+1}}$ denote the eigenvectors of $\mathbf{R}$, all the eigenvectors that form the noise subspace satisfy the relations

$$\begin{aligned} \mathbf{Rq_i} &= \sigma^2 \mathbf{q_i}, \quad i = L+1, \ldots, M+1 \\ (\mathbf{R} - \sigma^2 \mathbf{I})\mathbf{q_i} &= \mathbf{0}, \quad i = L+1, \ldots, M+1. \end{aligned} \tag{3.21}$$

Using Equation 3.18, we can express this relation as

$$\mathbf{SDS}^H \mathbf{q_i} = \mathbf{0}. \tag{3.22}$$

Since the columns of $\mathbf{S}$ are linearly independent and all the diagonal entries of $\mathbf{D}$ are non-zero, the only possible solution to Equation 3.22 is

$$\mathbf{S}^H \mathbf{q_i} = \mathbf{0} \tag{3.23}$$

The eigenvectors of a correlation matrix are orthogonal with each other. This means that the noise and the signal plus noise subspaces defined above are orthogonal complements. Since the multiplication of the columns of **S** and the noise subspace eigenvectors in Equation 3.23 equals zero, the subspace spanned by the columns of **S** must be equal to the span of the signal plus noise subspace. This important observation is the basis for the MUSIC algorithm. The algorithm determines the frequencies of the sinusoids given in Equation 3.17 by searching for the sinusoidal signal vectors in **S** that span the signal plus noise subspace and are therefore orthogonal to the noise subspace.

Estimating the Correlation Matrix

In a practical system, the ensemble-averaged correlation matrix **R** must be estimated. The estimated correlation matrix $\hat{\mathbf{R}}$ is calculated using

$$\hat{\mathbf{R}} = \frac{1}{2(N-M)}\mathbf{A}^H\mathbf{A} \tag{3.24}$$

where N is the number of samples used [29].

The data matrix **A** is formed using the samples $u(n)$ of the input waveform. The matrix is formed using a technique called temporal smoothing. This method helps to decorrelate the sinusoids contained in the signal. It is the analog of a method called spatial smoothing which helps to decorrelate plane wave sources that are incident on an antenna array. Uncorrelated sinusoids are essential for the operation of the MUSIC algorithm.

The operation is performed on a sequence of N samples. The smoothing is formed using windows of samples that are $M+1$ samples long. The Hermitian transpose of

the data matrix is given by $\mathbf{A}^H = [\mathbf{F}\mathbf{B}^*]$. The columns of the $\mathbf{F}$ matrix are formed using forward temporal smoothing. This means that the window is positioned at the start of the sample sequence and moved sample by sample towards the end of the sequence. The samples inside the window at each step compose a column of the matrix $\mathbf{F}$. The columns of matrix $\mathbf{B}$ are formed using backwards temporal smoothing. In this case, the window is positioned at the end of the sequence and stepped towards the beginning. The samples in the window at each step are put in the columns of $\mathbf{B}$. The expanded form of matrix $\mathbf{A}^H$ is shown in Equation 3.25.

$$\mathbf{A}^H = \begin{bmatrix} u(M) & \cdots & u(N-1) & u^*(0) & \cdots & u^*(N-(M+1)) \\ u(M-1) & \cdots & u(N-2) & u^*(1) & \cdots & u^*(N-(M+2)) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ u(0) & \cdots & u(N-(M+1)) & u^*(M) & \cdots & u^*(N-1) \end{bmatrix} \tag{3.25}$$

Calculation of the MUSIC Spectrum

The MUSIC spectrum is a way of plotting the frequency estimates the MUSIC algorithm makes for the input signal. The first step in generating the MUSIC spectrum is to calculate the eigenvalues and eigenvectors of $\hat{\mathbf{R}}$. The eigenvalues are sorted in descending order. The eigenvectors corresponding to the L largest eigenvalues are denoted $\mathbf{v}_1, \ldots, \mathbf{v}_L$. The eigenvectors corresponding to the remaining $(M+1-L)$ smallest eigenvalues are denoted $\mathbf{v}_{L+1}, \ldots, \mathbf{v}_{M+1}$.

These eigenvectors can be put in matrix form, as shown in Equation 3.26.

$$\begin{aligned} \mathbf{V}_S &= [\mathbf{v}_1, \ldots, \mathbf{v}_L] \\ \mathbf{V}_N &= [\mathbf{v}_{L+1}, \ldots, \mathbf{v}_{M+1}] \end{aligned} \tag{3.26}$$

Since the noise and the signal plus noise subspaces are orthogonal

$$\mathbf{V}_N^H \mathbf{V}_S = \mathbf{0} \tag{3.27}$$

The MUSIC spectrum is plotted using a variable frequency scanning vector, $\mathbf{s}(\omega)$ defined as

$$\mathbf{s}^T(\omega) = [1, e^{-j\omega}, \ldots, e^{-j\omega M}]. \tag{3.28}$$

The frequency scanning vector has the same form as the columns of the $\mathbf{S}$ matrix defined in Equation 3.19. Equation 3.23 illustrates that each of the columns in $\mathbf{S}$ are orthogonal to the noise subspace. This means that if the frequency of the scanning vector ω is equal to one of the frequencies of the sinusoids expressed in the columns of $\mathbf{S}$, then the scanning vector will also be orthogonal to the noise subspace. Based on this principle, the MUSIC spectrum, $S_{MUSIC}(\omega)$, is defined as

$$S_{MUSIC} = \frac{1}{\mathbf{s}^H(\omega)\mathbf{V}_N\mathbf{V}_N^H\mathbf{s}(\omega)}. \tag{3.29}$$

The angular frequency ω in the scanning vector is a continuous variable that can be adjusted over the frequency range $-\pi \leq \omega \leq \pi$. When the frequency of the scanning vector equals the frequency of one of the sinusoids in the input signal, the scanning vector becomes orthogonal to $\mathbf{V}_N$. At that point, the product $\mathbf{s}^H(\omega)\mathbf{V}_N\mathbf{V}_N^H\mathbf{s}(\omega)$ is equal to zero and an impulse is produced in $S_{MUSIC}(\omega)$. These impulses indicate the resolved estimates of the frequencies of the sinusoids.

Root MUSIC

The Root MUSIC algorithm is based on finding the poles of the MUSIC spectrum in the z-domain, rather than impulses in a continuous angular frequency plot. The complex exponential $e^{j\omega}$ in Equation 3.29 is replaced by the complex variable z. The result is given in Equation 3.30.

$$S_{MUSIC}(z) = \frac{1}{D(z)} \tag{3.30}$$

The term $D(z)$ represents the polynomial that results from the substitution of z for $e^{j\omega}$. The zeros of the polynomial $D(z)$ exhibit inverse symmetry with the unit circle. The position of the poles of $S_{MUSIC}(z)$ on the unit circle indicate the resolved frequencies of the sinusoids in the input signal.

3.7.2 Time Domain Super-Resolution

The benefit of super-resolution techniques for TOA location applications is not from calculating finely resolved frequency spectra using time signals. Instead, the benefit is being able to calculate a finely resolved time domain picture from frequency data.

The first researchers to use super-resolution to resolve a time domain signal presented their results in a paper by Yamada *et. al.* [30]. They used the MUSIC algorithm on frequency data from a spectrum analyzer to help improve the accuracy of antenna measurements.

The MUSIC algorithm can be used to resolve the TOA of a PN sequence correlation peak. The procedure used is shown in Figure 3.21. A small window of samples that includes the correlation peak is converted to frequency data using an FFT. That

data is then inverse transformed back to the time domain using the Root-MUSIC algorithm.

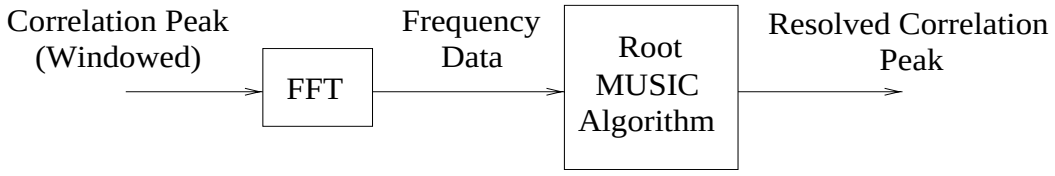

Figure 3.21: Time Domain Super-Resolution

It's possible to use the Root-MUSIC algorithm for this application because the signal being transformed to the time domain matches the model presented in Section 3.7.2. In an infinite bandwidth system, the correlation peaks can be modeled as impulses. An impulse, $\delta[n - n_o]$, is equal to a sinusoid, $e^{-j\omega n_o}$, in the frequency domain [31]. Several multipath arrivals result in several delayed copies of the correlation peak. This is equivalent to a sum of several sinusoids in the frequency domain. The delay information of each of the correlation peaks is contained in the frequency term of one of the sinusoids. This sum of sinusoids is equivalent to the sum of sinusoids in Equation 3.17 that the MUSIC algorithm operates on. This means that the MUSIC algorithm can be used to estimate the frequency terms of the sinusoids and use them to calculate a finely resolved estimate of delays of each of the correlation peaks.

The MUSIC algorithm is able to perform a frequency to time conversion with only a minor modification. The frequency scanning vector, $\mathbf{s}(\omega)$, in Equation 3.28 is changed to the scanning vector shown in Equation 3.31 [28].

$$\mathbf{s}^T(\omega) = [1, e^{j\omega}, \ldots, e^{j\omega M}] \tag{3.31}$$

For this system, the Root-MUSIC algorithm was chosen to resolve the TOA's of

the correlation peaks. The Root-MUSIC algorithm was used because it has worked well for other TOA resolution applications [28, 18, 19]. The algorithm produces several poles. For a multipath channel, each multipath arrival will be represented by one of the poles. Usually, there are more poles produced by the algorithm than there are multipath arrivals. The other poles are aliased copies of the multipath arrival poles [30]. When the Root-MUSIC algorithm is used to resolve a frequency spectrum, it provides a repeating spectrum. When it is used to resolve a time domain signal, it gives a repeating picture of the time domain. The pole that represents the the line-of-sight multipath arrival component must be selected from the other poles. The position of that pole on the unit circle is then used to calculate the resolved TOA of the mobile's signal. As discussed later in Section 3.7.4, selecting the correct pole is a problem.

A deconvolution operation is also required to improve the performance of the technique. The discussion above assumes that the correlation peaks are ideal impulses. However, due to the finite bandwidth of the system and the response of non-ideal components, the correlation peaks will be spread out over several samples. In order to improve performance, it is necessary to deconvolve the effects of all the non-ideal components in the system [28]. This can be performed after the FFT operation in Figure 3.21. The frequency response of the correlation peak is divided by the measured frequency response of the filters and components in the location system. The resulting spectrum is closer to one that would be generated by an ideal impulse train, which improves the performance of the algorithm.

The correlation based TOA detectors discussed in Section 3.5 can be modified to use Root-MUSIC by adding the blocks shown in Figure 3.22 to the output of their correlator blocks.

The Root-MUSIC modifications consist of several blocks. The first block in Figure 3.22 after the output of the correlator is a windowing block. This block takes a small window of samples centered around the correlation spike and passes that window on to the FFT block. This ensures that only the samples with correlation spike information are submitted to the super-resolution algorithm. The next block is an FFT that converts the correlation spike to the frequency domain. The block following the FFT operation is the system deconvolution block. This block contains a stored copy of the frequency spectrum of the impulse response of all the components in the mobile and the TOA detection system. The spectrum of the correlator spike is divided by this stored spectrum. The system impulse response can be measured or approximated using theoretical component values. Once the system response is removed, the correlation frequency spectrum is input to the Root-MUSIC algorithm. The Root-MUSIC algorithm performs the inverse transform to convert the correlation peak back to the time domain in the form of poles around the unit circle. Those poles are input into block that selects the pole corresponding the the line-of-sight arrival of the signal. The position of that pole is used to calculate the resolved TOA of the mobile's signal.

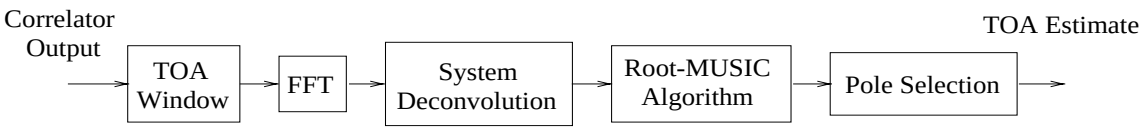

Figure 3.22: TOA Detector Root-MUSIC Modifications

3.7.3 Simulation

The simulations used to investigate the accuracy improvements possible from super-resolution were based on the simulations presented in Section 3.6. Exactly the same cell topology and channel models are used. For each iteration of the simulation, the mobile's position was randomly generated and the received signal of the mobile at each of the base stations was simulated. The mobile waveforms were simulated at 8 samples per chip and IS-95 compliant pulse shaping was used.

The TOA detectors in each of the 7 base stations were enhanced using the Root-MUSIC blocks shown in Figure 3.22. For this simulation, the pole closest to the actual TOA of the signal was always selected. This means that the simulation results show the best possible case. A real pole selection algorithm would periodically make mistakes and select the wrong pole. This would increase the error of the simulations. The TOA's calculated using the position of the Root-MUSIC poles were used to solve for the coordinate position of the mobile.

Deconvolution of the non-ideal components in the simulation was used to improve the accuracy of the simulation. Before the spectrum of the correlation peak was passed into the Root-MUSIC algorithm, the response of the IS-95 pulse shaping filter and the non-ideal shape of the impulse was deconvolved from the spectrum.

Plots were generated showing the mean and standard deviation of the radial error distance of the position estimates returned by the algorithm. The mean and standard deviation of the error distance for the 1 ray channel model are shown in Figures 3.23 and 3.24. The mean and standard deviation plots for the 4 ray model are shown in Figures 3.25 and 3.26. The Cramer-Rao bound is also plotted on each figure.

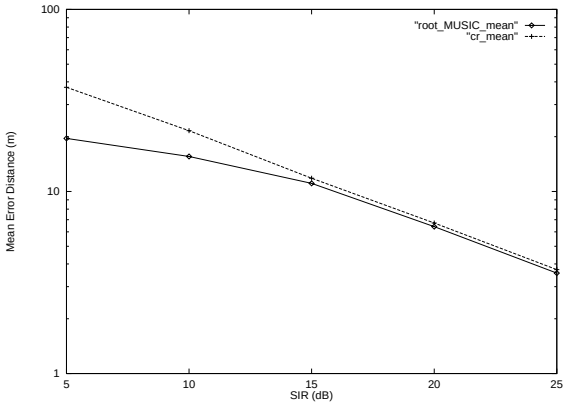

Figure 3.23: Mean Error Distance (1 Ray Channel)

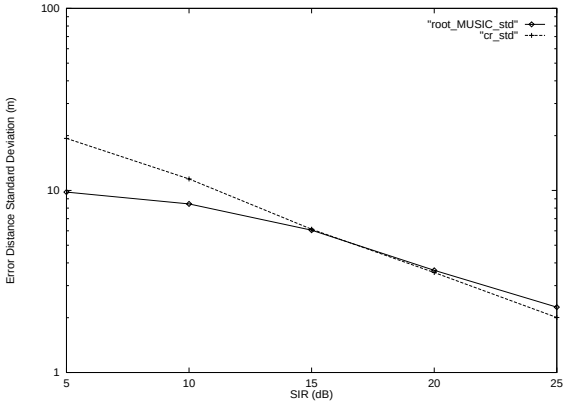

Figure 3.24: Error Distance Standard Deviation (1 Ray Channel)

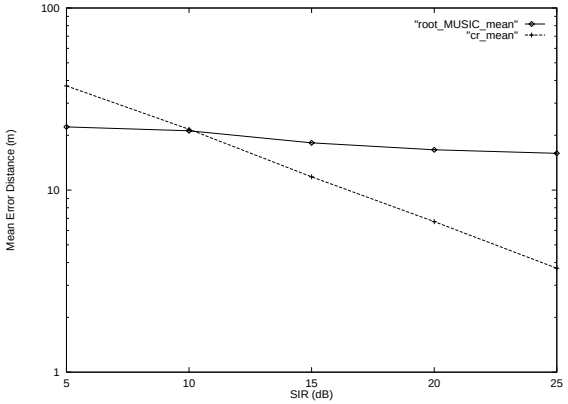

Figure 3.25: Mean Error Distance (4 Ray Channel)

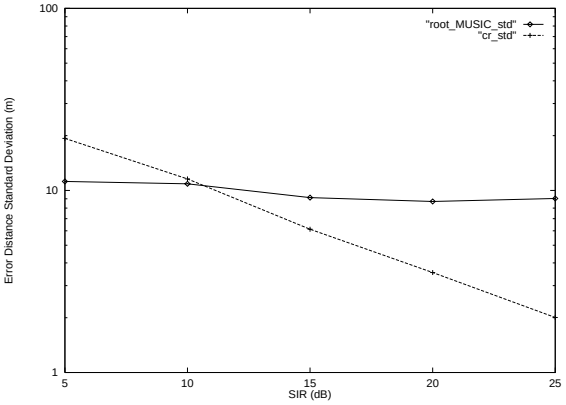

Figure 3.26: Error Distance Standard Deviation (4 Ray Channel)

The 1 ray channel model plots show a significant accuracy improvement over the threshold methods used in Section 3.6. In fact, the error of the method achieves the Cramer-Rao bound. This proves that, in some situations, the Root-MUSIC technique is optimal.

The accuracy of the scheme is reduced somewhat when the 4 ray channel model is used. The degradation in accuracy with the 4 ray channel model is a result of the additional multipath components interfering with the line-of-sight components. While the resolution of the Root-MUSIC algorithm is very good, there is a point where it cannot resolve two arrivals that are very close together [28]. When this occurs, the error in the Root-MUSIC TOA estimate is increased.

It should be noted that at low SIR, it appears that Root-MUSIC actually performs better than the Cramer-Rao bound. This is because the pole closest to the actual TOA of the signal is selected each time. The Root-MUSIC algorithm produces a number of poles, some representing the multipath arrivals of the signal and some that are aliased copies of the multipath poles. The result is a number of poles that are approximately uniformly spread around the unit circle. This means that there will always be a pole reasonably close to the actual signal TOA, even when the algorithm is no longer producing accurate estimates due to low SNR. As long as the pole closest to the actual TOA is selected, there will be an upper limit on the position error. This is simply due to the fact that the algorithm produces a lot of poles that are spread over the entire unit circle. This error ceiling is reached by the simulations at low SIR values.

3.7.4 Limitations of Time Domain Super-Resolution

The accuracy of the resolved TOAs returned by the Root-MUSIC algorithm depends on the selection of the correct pole. The Root-MUSIC algorithm generates several poles on the unit circle. Each of the multipath reflections in the correlation peak are represented by one of these poles. The position of the poles on the unit circle is used to determine their resolved TOA.

However, the poles representing the multipath reflections are surrounded by a number of other poles. Yamada *et. al.* describes these poles as aliased copies of the multipath poles [30]. A typical pole arrangement is shown around the unit circle in Figure 3.27. The actual multipath poles are shown in black while the aliased poles are shown in white. This time domain aliasing is caused by performing Root-MUSIC on a discrete frequency spectrum that doesn't have sufficient resolution. This problem of aliasing the time domain picture was solved in the paper by increasing the frequency resolution of the data taken by their spectrum analyzer.

In this system, the frequency data that the algorithm operates on comes from an FFT operation. The frequency resolution of an FFT can be increased by increasing the number of time domain points it operates on. This can be done in three ways.

The first way to do this is to increase the window size. However, increasing the window size means that a larger amount of noise is passed into the algorithm, along with the correlation peak. This reduces the accuracy of the algorithm to an unacceptable level.

The second possible technique is to zero pad the time domain signal. This causes the algorithm to interpolate the missing frequency samples between the real frequency data points. When zero padding was used in the simulations in Section 3.7.3, the

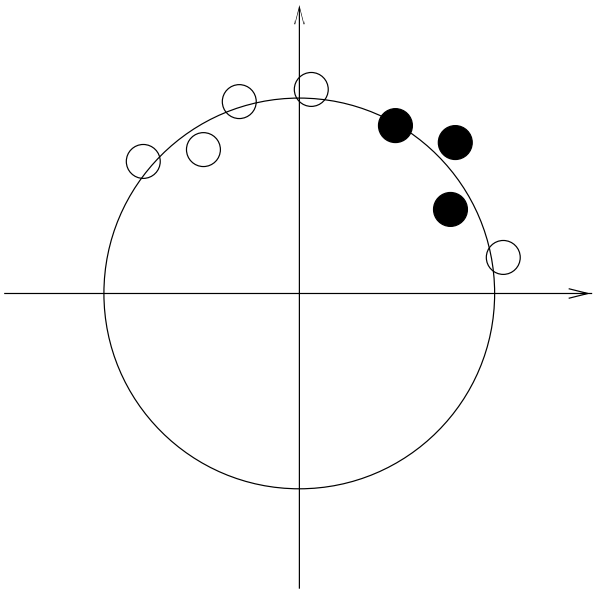

Figure 3.27: Root-MUSIC Pole Constellation

aliasing problem was eliminated and the Root-MUSIC pole corresponding to the line-of-sight arrival had the lowest angle on the unit circle every time. However, when zero padding is used, the majority of the points passed into the Root-MUSIC algorithm are interpolated points, rather than actual data points. This means that even a small amount of noise in the system causes a very large amount of error. As a result, zero padding is not a practical solution for the aliasing problem.

The third possible solution is to oversample the correlation peak. If this is performed after the lowpass filtering of the signal, the amount of noise input into the algorithm should not be increased significantly. This is the technique with the most promise for resolving the aliasing problem.

Another technique for selecting the line-of-sight pole makes use of a coarser estimate of the TOA. The algorithm can first estimate the TOA of the mobile's signal using a coarse threshold-based method. Then, the Root-MUSIC pole with an arrival

time closest to this coarse estimate is chosen as the resolved TOA of the signal. This method works well at high SIR values, when the correlation peak is well out of the noise and the accuracy of the threshold estimate is quite good. However, at low SIR values, the error in the threshold TOA estimates increases and the algorithm starts to choose incorrect poles. When this occurs, the errors of the resolved estimates have approximately the same value as the coarse estimates.

The accuracy advantage of super-resolution is the most significant at low levels of SIR. Since the method of using coarse TOA estimates only works reliably at high levels of SIR, the advantage of supplementing those estimates with super-resolution at that signal level is less significant. For Root-MUSIC to be most effective, a scheme for selecting the correct pole at low SIR values should be found.

3.8 Conclusion

At the beginning of this chapter, three main advantages of time-of-arrival location were given. A time-of-arrival location system is cheap to implement since it can use existing base station antennas, it is accurate and it is resistant to multipath. The first advantage is intuitive and the second advantages were confirmed with the simulations in Section 3.6.

Even with the crude TOA detection techniques used in these simulations, with a sufficient SIR at the correlator output and number of base stations, the position error in the simulated accuracy estimates was low enough to satisfy the FCC requirements. This shows that TOA location has a lot of potential for providing accurate location estimates. However, these results are very geometry specific. The geometry in the simulation was very favorable. It's possible in a real cellular system that there would be many cases where the geometry would be much worse. In these cases, a technique that finds more accurate TOA estimates would be required.

One possible technique for enhancing TOA estimates is the use of super-resolution algorithms. In this chapter, the Root-MUSIC algorithm was used to improve the accuracy of the TOA simulations. The TOA of the line-of-sight component of the signals was determined using one of the poles returned by the algorithm. The simulation results showed the accuracy possible when the correct pole was selected each time. These results indicate that there is a significant advantage to using the Root-MUSIC algorithm. However, before Root-MUSIC can be used, an algorithm for reliably selecting the correct TOA pole must be developed.

The simulations also confirm the robustness of the TOA location technique to multipath. When multipath reflections were introduced to the system, the accuracy

of the threshold technique that operated on the rising edge of the signal did not degrade significantly. This shows that as long as the TOA detector operates on the rising edge of the correlation peak, multipath reflections do not significantly degrade the accuracy of the system.

Chapter 4

IS-95 Mobile Handoff Measurement Location

4.1 Introduction

Locating an IS-95 mobile using the signal from the mobile can be very difficult. The simulation in Chapter 2 shows that the received signal level of an IS-95 mobile at a base station in a neighbouring cell can be extremely low. In Chapter 3, a system was presented that locates the mobile using its signal. While the accuracy levels of the system can potentially satisfy the FCC requirements, the system is somewhat complex to implement and still has some problems that must be addressed.

This makes it worthwhile to take a close look at the IS-95 cellular system to see if there are any features already implemented in the system that could assist in locating the mobile. Taking advantage of features already present in the IS-95 network would result in a location system that is both cheaper and less complex than one that is based entirely on extra equipment that must be added to the cellular system.

This chapter presents two location schemes that attempt to locate the mobile using features that are already present in the IS-95 cellular system. They both locate the mobile using measurements the mobile is already taking for another purpose. Each mobile takes measurements on pilot signals that are transmitted by each of the base stations in the system. These measurements are used to assist in the soft handoff operation. Section 4.2 describes the soft handoff operation and the measurements the mobile takes to assist soft handoff. Section 4.3 and 4.4 present two location systems

that are based on these measurements. Section 4.5 presents some conclusions about the two approaches.

4.2 IS-95 Soft Handoff Overview

Most cellular systems manage mobiles moving between cells using hard handoff. Hard handoff means that as soon as a mobile makes the transition into a new cell, the base station in the old cell will drop the mobile's call. At approximately the same instant, the base station in the new cell assigns the mobile a new channel and the call continues through the new base station. This transition sometimes results in an audible click during the call. With this handoff scheme, there is a danger of dropping the mobile's call. A dropped call occurs when the base station in the cell the mobile moves into does not have any channels available for the mobile or if noise or interference causes the base station to lose the mobile.

IS-95 attempts to address some of the problems with hard handoff by using the soft handoff scheme. When an IS-95 mobile approaches the boundary between two cells, it starts to communicate with the base station in the new cell. However, rather than immediately halting communication with the base station in the old cell, the mobile continues to communicate with it as well. This means that in the region around the border between two cells, the mobile communicates simultaneously with the base stations in both cells. It is possible that the mobile could even communicate with more than two base stations during soft handoff. This scheme makes sure that the mobile has been communicating with the base station in its new cell for some time before it stops communicating with the base station in its old cell. This "make before break" approach reduces the probability of dropped calls.

Soft handoff is a Mobile Assisted Handoff (MAHO) scheme. This means that the mobile plays a part in determining when it should go into handoff. It does this by

taking measurements on the pilot signals transmitted by the base stations. These pilot signals are used for MAHO as well as for timing and phase synchronization for the mobiles in the system. An IS-95 mobile takes measurements on the pilot signals from all nearby base stations and stores them. Periodically, the mobile returns these measurements to the Mobile Switching Office (MSC). The MSC uses these measurements to determine when the phone should go into handoff.

The first measurement taken by the mobile on a base station pilot signal is the PN sequence phase of the pilot signal short code. The set of short code spreading codes used by the base stations are made up of different offsets or phases of the same PN sequence. Each base station uses a different offset. A base station's pilot signal consists of a repeating short code that doesn't carry any information. The mobile measures the PN phase of a pilot from a nearby base station relative to its own timing reference. The mobile's time reference is the pilot signal of the base station it is currently communicating with. The mobile uses the PN offset of each pilot signal to identify which base stations the pilot signals are coming from.

The second measurement taken by the mobile on each base station pilot signal is its SIR. The SIR of a pilot signal is defined as the received signal power of the signal divided by the total power the mobile is receiving on the forward channel. The mobile uses the SIR of each pilot to determine how close it is to the base station transmitting the pilot and whether or not it should enter soft handoff with any of them. Based on their SIR values, the mobile organizes all the pilots it measures into several sets.

The two main pilot signal sets are called the Active and Candidate sets. The Active Set contains all the base stations the mobile is currently communicating with. The mobile will start to communicate with a base station when it receives a command

from the MSC. The Candidate set contains all the base stations that are candidates for the mobile to go into soft handoff with. When the MSC directs the mobile to start communicating with a new base station, that base station must be a member of the Candidate Set. When a mobile starts communicating with a new base station, that base station is taken out of the Candidate Set and placed into the Active Set. When the mobile stops communicating with a base station, it is removed from the Active Set and returned to the Candidate Set. A mobile will add a pilot to the Candidate Set when its SIR exceeds a certain threshold. It drops a pilot from the Candidate Set if that pilot's SIR drops below a certain threshold for a certain length of time.

The mobile returns the short code PN offsets and the SIR levels of each pilot in the Active and Candidate Set to the MSC. Measurements for a maximum of 11 pilots can be returned at one time. The location schemes discussed in this chapter use the measurements taken on the pilots in the Active and Candidate Sets to determine the location of the mobile taking the measurements. The scheme in Section 4.3 uses the SIR readings to locate the mobile. The scheme in Section 4.4 uses the short code PN offset measurements to locate the mobile.

These two location schemes share one major advantage: they are incredibly easy to implement. The measurements these schemes use are already present in the software of the MSC in numerical form. A system that uses these measurements to determine mobile locations could be implemented entirely in the software of the MSC. This type of location system would be very inexpensive and could be added to an existing IS-95 system very easily.

4.3 Location Using Pilot Signal Strength Measurements

This section describes a scheme for locating the mobile based on the SIR measurements it takes on the base station pilot signals. Section 4.3.1 describes some other cellular location systems that also use signal strength. Section 4.3.2 describes the derivation of equations for the SIR measurements that can be solved for the location of the mobile. Section 4.3.3 discusses the algorithm used to solve the location equations. Section 4.3.4 contains a simulation showing the accuracy of the signal strength location scheme and Section 4.3.5 offers some conclusions.

4.3.1 Background

Radio signal strength location is not a technique typically used to determine a device's location. The reason is the very large, random errors found in the wireless channel. However, there still has been a considerable amount of work done in using signal strength measurements to determine the location of cellular mobiles.

In the system disclosed by Dent, each base station transmits location information to the mobiles [32]. This information consists of a table or contour of the received strength of the signal from that base station versus distance. The mobile measures the received signal strength from a base station and then uses the information it receives from that station to determine its distance from it. The mobiles also transmit back the signal strength measurements they take to the base stations who use them to update their contours. The disadvantage of this system is that it requires extensive modifications of the mobile and the base stations.

A different type of signal strength location system is disclosed by Song [33]. This system uses a location device placed in a vehicle that monitors the strengths of the

signals received from the surrounding base stations. The strength of a radio signal decays with the distance from the transmitter, according to an exponential factor. The system assumes that this path loss exponent and the transmit power for each base station is known before-hand from measurements or calculation. That information is stored in the memory of each location device. The device uses the log-distance channel equation to solve for the distances from the base stations to the mobile. It then uses the distances to triangulate the mobile's location. There are two disadvantages with this system. First, additional equipment or some modification of the mobile is required. Second, the path loss exponent for each base station can vary considerably, depending on location in the cell. Approximating this with a single value will result in errors in the system. Also, if there is any change in the path loss exponent values, the memory in every one of the location devices will have to be updated.

Unlike the first two systems, the system presented by Dufour does not require the modification of the mobile [34]. It uses signal strength measurements taken by the mobile that are used for another function. The system forces the mobile to sequentially go into handoff with each of its neighbouring cells. It then uses the signal measurements the mobile takes of the base stations as a result of the handoffs to determine the location of the phone. The distances from each of the base stations are calculated and the mobile position is found using triangulation. The disadvantage of this system is the added complexity of forcing the cellular phone to perform several unnecessary handoffs. This method could also cause degradation in call quality and would not work if the mobile had only a small number of nearby cells. It also increases the probability of a dropped call.

The system disclosed by Doner also does not require modification of the mobile

[35]. It uses the signal strength readings the mobile takes of its surrounding base stations for MAHO. It also uses transmission time advance information for each mobile to determine location. This system divides the cell into large contour regions based on signal strength measurements taken in the area. The MAHO measurements taken by the mobile are used to map it into a specific location using the contour regions. The location is further refined by constraining the position estimate to lie on known service areas, like roads. The disadvantage of this system is that it requires extensive signal strength measurements to be made in the area. Another disadvantage is that in order to constrain the position estimate to a known service area, the topology of each cell must be surveyed in detail and stored.

4.3.2 Derivation of SIR Location Equations

A signal strength location system takes advantage of the fact that the average received power of a signal decays in a known fashion with distance. The log-distance channel model is given in Equation 4.1 and states that the average received power of a signal, P_r decays proportionally with the distance, d between the transmitter and receiver to a negative path loss exponent n [4]. The constant k depends on the channel between the transmitter and receiver, the transmit frequency, the transmitted power and the antennas on the mobile and base station.

$$P_r = k(d)^{-n} \tag{4.1}$$

This equation means that since the transmit powers of the base station pilot signals are known, it would be possible to use the received power measurements the mobile makes on the pilots to determine how far the phone is from each of the base stations.

Those distances could be combined with the known positions of the base stations to determine a coordinate (x, y) position for the mobile.

The problem with the measurements taken by the phone is that they are SIR measurements and not pure received signal power measurements. The SIR measurement of the pilot signal from base station i is defined in the IS-95 standard as the ratio of the received power of the pilot from base station i, P_i, divided by the total power the mobile is receiving on the forward channel, P_{tot} [6]. The value of P_{tot} changes depending on where the mobile is located in the cell. Since P_i is divided by a non-constant denominator, the SIR measurements cannot just be substituted into Equation 2.8 and solved for distance.

The mobile continually updates the SIR measurements it makes on the pilot signals. In addition to this, the IS-95 signal is wideband and will not experience the severe small scale fades that are typical in a narrowband system [9]. That means that the SIR measurements of the pilots in the Active and Candidate sets returned to the MSC can be considered as being taken in approximately the same location. All SIR measurements taken in the same location will have the same value for P_{tot} in their denominators. That means that if the SIR measurement from base station i, SIR_i, is divided by the SIR measurement from base station j, SIR_j, their denominators will cancel and the result is simply the ratio of the received signal powers from each of the base station pilots, given in Equation 4.2.

$$\frac{SIR_i}{SIR_j} = \frac{P_i}{P_j} \tag{4.2}$$

Equation 4.2 can be written as a function of the distances from base station i and

j and the path loss exponent by using the log distance channel model.

$$\frac{SIR_i}{SIR_j} = \frac{k_i(d_i)^{-n}}{k_j(d_j)^{-n}} \tag{4.3}$$

If the channels between the mobile and base station i and j are assumed to be statistically identical and the mobile and base stations are using omindirectional antennas, $k_i = k_j$ and Equation 4.3 becomes

$$\frac{SIR_i}{SIR_j} = \left(\frac{d_i}{d_j}\right)^{-n}. \tag{4.4}$$

In this form, the ratio of the SIR measurements of the pilots from two different base stations simplifies to a function of the distances from the mobile to the two base stations and the path loss exponent. The coordinate position of base station i is (x_i, y_i), the coordinate position of base station j is (x_j, y_j) and the coordinate position of the mobile is (x, y). Equation 4.4 can be written as a function of these positions.

$$\frac{SIR_i}{SIR_j} = \left(\frac{(x-x_i)^2 + (y-y_i)^2}{(x-x_j)^2 + (y-y_j)^2}\right)^{-\frac{n}{2}}. \tag{4.5}$$

This equation expresses the pilot SIR measurements from base station i and j as a function of three unknowns: the x coordinate position of the mobile, the y coordinate position of the mobile and the path loss exponent n of the channel. In this form, the equations could be solved for the mobile's position.

One additional correction has to be added before the location equations can be used. Most modern IS-95 base stations use directional sectorizing antennas, so the assumption of omnidirectional base station antennas is poor. The constant k in

Equation 2.8 is proportional to the base station antenna gain, G_b. When a directional sectorized antenna is used by the base station, that antenna gain will be a function of where the mobile is located in the cell. However, this antenna gain pattern can be determined using the antenna gain specifications from the manufacturer and the orientation of the antenna in the base station.

This means that the directional gain of the base station can be compensated for in the location equations. Using the antenna gain specifications and its orientation, an approximate expression for the base station antenna gain in each cell, $G'_b(x, y, x_i, y_i)$, can be written that is a function of the mobile's position in the cell, (x, y), and the base station's position (x_i, y_i). When this correction factor is added, Equation 4.5 becomes Equation 4.6.

$$\frac{SIR_i}{SIR_j} = \left(\frac{G'_b(x, y, x_i, y_i)}{G'_b(x, y, x_j, y_j)}\right)\left(\frac{(x - x_i)^2 + (y - y_i)^2}{(x - x_j)^2 + (y - y_j)^2}\right)^{-\frac{n}{2}} \tag{4.6}$$

4.3.3 Algorithm for Solving the SIR Location Equations

This signal strength location method is somewhat unconventional in its approach to solving the signal strength location equations. Specifically, it is unconventional in its approach to the path loss exponent factor in its equations. Many signal strength location schemes attempt to measure or characterize the path loss exponent of the channel and then use that value in their signal strength location equations [33, 35]. However, Section 4.2 states that the mobile can return signal strength measurements from up to 11 pilot signals. With this many pilot measurements, it would be possible to form enough equations to solve for the path loss exponent, in addition to the (x, y) position of the mobile. It is possible that there will be regions in the cellular coverage

area where there will not be enough pilot measurements to solve for n as well as the mobile position. In this case, the system would just use the most recently solved for value of n until enough pilot measurements are available for it to be updated.

Solving for the path loss exponent greatly simplifies the implementation of the system. Extensive channel measurements required to characterize the path loss exponent in each cell are no longer needed. The same system will be able to function in areas with different path loss exponents with no additional modification.

The method used to solve the signal strength location equations is similar to the method used to solve the TDOA equations in Chapter 3. The method is based on solving for a least squares solution to linearized versions of the signal strength equations.

The ratio of the SIR values of the pilot signals from base station i and j can be written as a function of the mobile coordinate position (x, y) and the path loss exponent of the channel n.

$$m_k = \frac{SIR_i}{SIR_j} = f_k(x, y, n) \tag{4.7}$$

The ratio of two SIR measurements is denoted m_k. The function $f_k(x, y, n)$ is the equation given in Equation 4.5 where the appropriate coordinate positions have been substituted in for base station i and j.

The function $f_k(x, y, n)$ is nonlinear. In order to solve the equations using a linear least-squares method, the function is expanded out using the first two terms of its

Taylor Series. The expansion is

$$m_k = f_k(x, y, n) = f_k(x_o, y_o, n_o) + a_{k,1}\delta_x + a_{k,2}\delta_y + a_{k,3}\delta n \tag{4.8}$$

where

$$\begin{gathered} a_{k,1} = \left.\frac{\partial f_k}{\partial x}\right|_{x_o,y_o,n_o} = \frac{10n(x-x_j)}{(x-x_i)^2+(y-y_i)^2} - \frac{10n(x-x_i)}{(x-x_i)^2+(y-y_i)^2} \\ a_{k,2} = \left.\frac{\partial f_k}{\partial y}\right|_{x_o,y_o,n_o} = \frac{10n(y-y_j)}{(x-x_i)^2+(y-y_i)^2} - \frac{10n(y-y_i)}{(x-x_i)^2+(y-y_i)^2} \\ a_{k,3} = \left.\frac{\partial f_k}{\partial n}\right|_{x_o,y_o,n_o} = 5\ln\left(\frac{(x-x_j)^2+(y-y_j)^2}{(x-x_i)^2+(y-y_i)^2}\right) \\ \delta_x = (x - x_o) \quad \delta_y = (y - y_o) \quad \delta_n = (n - n_o) \end{gathered} \tag{4.9}$$

The values x_o,y_o and n_o are initial guess values for the mobile position and path loss exponent.

The total number of equations used to find a solution is given by K. If SIR readings for N pilots are returned by the mobile, $N-1$ linearly independant SIR ratio equations can be formed.

The K equations can be written in matrix form

$$\mathbf{A\Delta} = \mathbf{z} \tag{4.10}$$

where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \vdots & \vdots & \vdots \\ a_{K1} & a_{K2} & a_{K3} \end{bmatrix} \quad \mathbf{\Delta} = \begin{bmatrix} \delta_x \\ \delta_y \\ \delta_n \end{bmatrix}$$
$$\mathbf{z} = \begin{bmatrix} m_1 - f_1(x,y,n) \\ m_2 - f_2(x,y,n) \\ \vdots \\ m_K - f_K(x,y,n) \end{bmatrix} \tag{4.11}$$

The K location equations form an inconsistant system which means there are more equations than unknowns and there is no unique solution. However a best fit solution, in the least squares sense, can still be found for the equations. The least squares solution for $\mathbf{\Delta}$ is given in Equation 4.12 [23].

$$\hat{\mathbf{\Delta}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{z} \tag{4.12}$$

The $\hat{\mathbf{\Delta}}$ matrix solved for is the least squares solution for the difference between x_o, y_o and n_o and the actual solution. This difference is added to x_o, y_o and n_o to give an approximation of x, y and n and shown in Equation 4.13

$$\begin{aligned} \hat{x} &= \hat{\delta_x} + x_o \\ \hat{y} &= \hat{\delta_y} + y_o \\ \hat{n} &= \hat{\delta_n} + n_o \end{aligned} \tag{4.13}$$

As described in Chapter 3, these new estimates are substituted back into Equation 4.10 and a new solution is found. This iteration continues until $\hat{\mathbf{\Delta}}$ converges to

zero. At this point, the estimates $\hat{x}$, $\hat{y}$ and $\hat{n}$ represent the least squares solutions for the signal strength location equations.

The conventional least squares technique described above was modified slightly to improve its performance for signal strength location. In early simulations, this technique had convergence problems when log-normal shadowing was added to the channel model. Examination of the algorithm showed that it would often diverge because of the large amount of error in the first few calculations for $\hat{x}$, $\hat{y}$ and $\hat{n}$. The first few solutions for $\hat{\mathbf{\Delta}}$ would contain such a large amount of error that the calculated values for $\hat{x}$,$\hat{y}$ and $\hat{n}$ would be much further from the solution than the initial guesses, rather than closer. When this occurred, the algorithm was unable to recover and it continued to diverge.

This problem was solved by adding a step size parameter. The relation shown in Equation 3.10 was replaced by

$$\begin{aligned} \hat{x} &= \rho\hat{\delta_x} + x_o \\ \hat{y} &= \rho\hat{\delta_y} + y_o \\ \hat{n} &= \rho\hat{\delta_n} + n_o \end{aligned} \tag{4.14}$$

where ρ is a step size subject to the constraint $0 < \rho \leq 1$. A small step size allows the algorithm to recover from a few initial bad calculations and still converge on the correct solution.

4.3.4 SIR Location Simulation

Simulation Setup

Simulation was used to determine the accuracy possible with the location scheme described in this section. The simulation used 7 cells, arranged as shown in Figure 4.1. The numbered circles represent the base stations in the center of each of the cells. The simulation assumes the mobile being located is somewhere in the center cell. That mobile is referred to as the quarry.

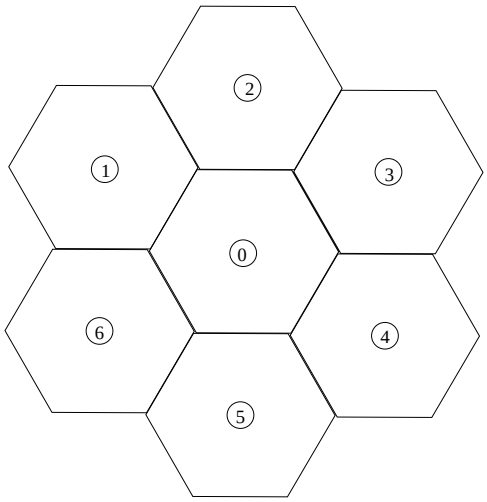

Figure 4.1: Cell Topology

Several random quarry positions were generated in the center cell. For each position, the SIR measurements of the pilot signals from the base stations closest to the quarry were simulated. The number of pilot signals measured by the mobile was a parameter of the simulation and was varied between 4 and 7.

The channel model used was the log distance channel model with log-normal shadowing. As discussed in Chapter 2, the errors due to the small-scale channel effects are small compared to the shadowing. As a result, they were not implemented in the simulation.

Each base station had directional antennas that divided the cells into 120 degree sectors. The gain pattern formula used is given in Equation 2.6.

Simulation Results

Initially, the simulations described in the previous section were performed without the step size parameter. However, when a log-normal shadowing of 10 dB was used, which is a typical value for cellular systems [4], the algorithm experienced extreme convergence problems. This was due to the large errors in the received signal strength caused by the shadowing. This is illustrated using a plot of the percentage of attempted solutions that converged. Figure 4.2 shows the convergence percentages for a log-normal shadowing standard deviation of 10 dB when no step size parameter was used for the trilateration algorithm.

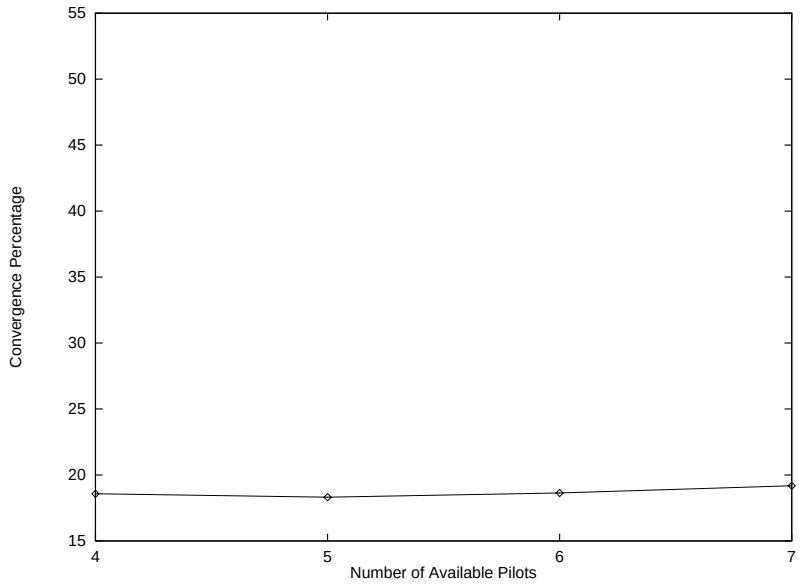

Figure 4.2: Algorithm Convergence without Step Size Parameter

Clearly, these convergence percentages are unacceptable. In order to correct this problem, the step size parameter ρ described in Section 4.3.3 was added. A simulation

was performed with a ρ value of 0.001 and a log-normal shadowing standard deviation of 10 dB. The convergence results are shown in Figure 4.3. This plot shows that

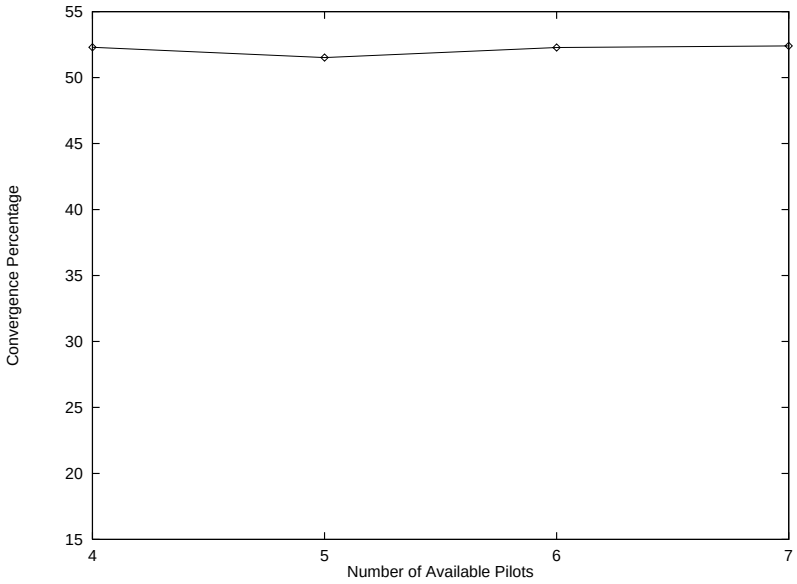

Figure 4.3: Algorithm Convergence with Step Size Parameter

the addition of a step size parameter resulted in a significant improvement in the convergence of the algorithm.

Once the step size parameter was added to improve algorithm convergence, several simulations were performed to investigate its accuracy. For each simulation, 5000 random mobile positions were generated. For each random position, SIR pilot measurements were simulated and used to solve for the mobile position. The radial error distance for each solution was recorded. Radial error distance is defined as the two dimensional Euclidean distance between the mobile's actual coordinate position and the coordinate position solved for by the system. Figure 4.4 shows the mean radial error distance and Figure 4.5 shows the standard deviation of the radial error distance. Each plot contains 3 curves, generated by 3 simulation runs that used a shadowing

standard deviation of 10 dB, 5 dB and 2 dB.

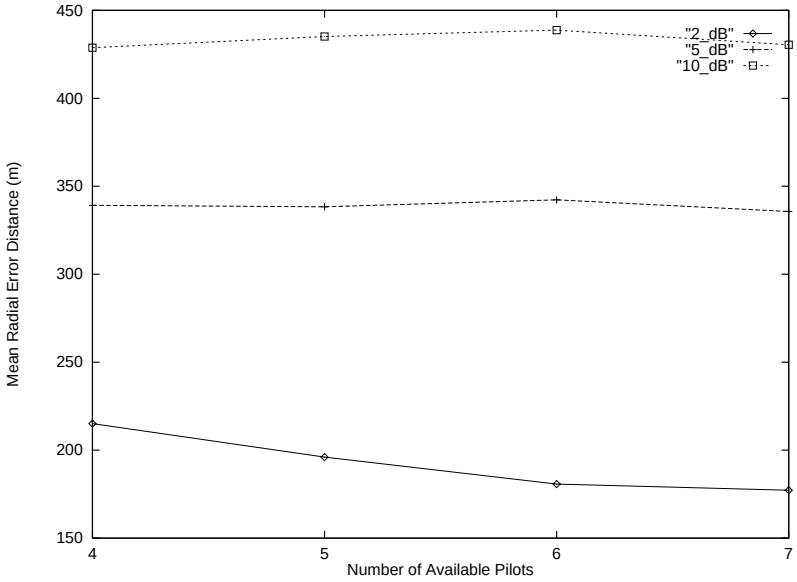

Figure 4.4: Mean Radial Error Distance

The errors in these plots are quite large compared to the FCC accuracy requirements of 125 m. A typical value for shadowing standard deviation is approximately 10 dB. For this shadowing value, the average error for this location method is approximately 400 m, which is much too large for the FCC requirements.

These graphs show that for large shadowing standard deviations, the error seems to increase if measurements are available from a larger number of pilots. This can be explained by examining the convergence percentage figures for each of the simulations. Figure 4.6 shows the percentage of solution attempts that converged for the 10 dB, 5 dB and 2 dB shadowing simulations. These plots show that an increased number of measurements increases the percentage of solution attempts that converge on an answer. When a small number of pilot measurements are available, the algorithm will only converge for locations where the shadowing error is small for all the measurements. This is a fairly small percentage of the total number of quarry

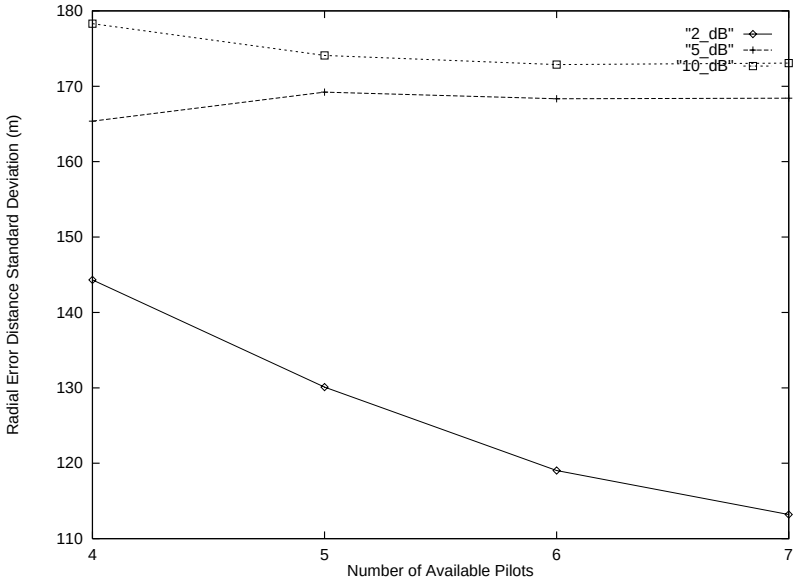

Figure 4.5: Radial Error Distance Standard Deviation

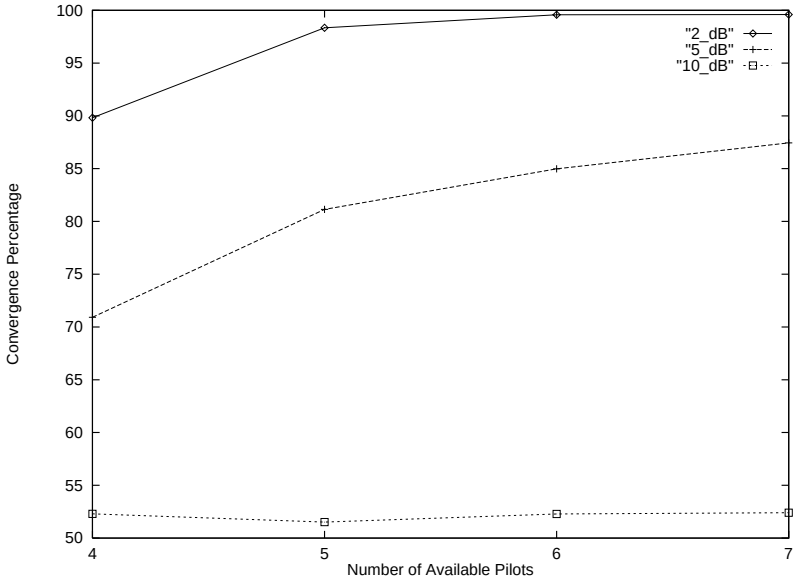

Figure 4.6: Simulation Convergence Percentages

locations. However, when a larger number of pilot measurements are available, the algorithm is able to converge for higher error levels. Since the algorithm now converges for measurements with higher error levels, the solutions found also have, on average, more error.

One penalty of adding a step size parameter to the algorithm is that it increases the number of iterations required to reach convergence. Here an iteration is defined as finding a single solution for $\hat{\mathbf{\Delta}}$. Figure 4.7 shows the average number of iterations for convergence for each solution attempt for the 10 dB, 5 dB and 2 dB shadowing simulations. These plots show that while the number of iterations required is large,

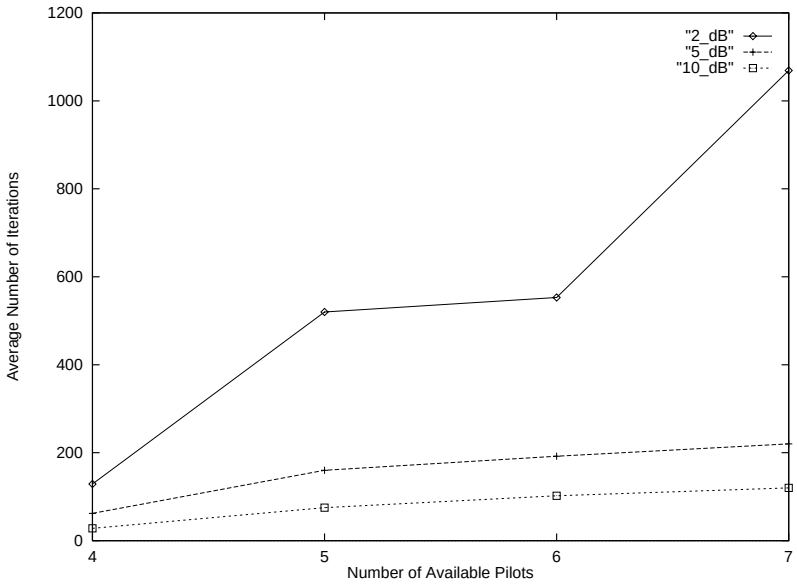

Figure 4.7: Average Number of Iterations to Convergence

it is well within the capabilities of modern digital processors. Also, the iterations increase when more pilot signal strength measurements are available. This is because a larger number of measurements allows the trilateration algorithm to converge for measurements with larger error. This increased amount of error means that the

algorithm requires more iterations to find a solution.

4.3.5 Conclusion

The simulation results show that this location technique is able to solve for the location of the mobile using the SIR measurements it makes on the pilot signals of nearby base stations. However, an environment with very low shadowing levels combined with a large number of available measurements is required before the technique will produce location estimates with a high degree of reliability. Even at the point where the technique consistantly converges to a solution, the radial error distance is still larger than the maximum allowed by the FCC.

Clearly, this technique will not satisfy the FCC requirements on its own. However, it could be used to supplement another type of location technique. Since these SIR measurements are already available within the IS-95 cellular system, using this location device to supplement another system would not add a great deal to the total cost. This location approach would work well in low multipath environments, like rural areas or on highways. Using signal strength location in these areas would reduce the amount of extra infrastructure that would have to be added to implement cellular location.

This technique would also complement a location technique that was more accurate but took more time to find a solution. The signal strength location system would give a quick approximation of the mobile's position. This would allow 911 dispatchers to send emergency vehicles towards the general area of the accident while they wait for a more accurate location estimate.

4.4 Mobile Location using Pilot Short Code Phase Measurements

4.4.1 The Mobile's Short Code Phase Calculation

Each base station is assigned its own unique short code spreading code offset. These codes are used to keep signals in neighbouring cells from interfering with each other. Each code in the short code set is generated using a different shift of the same PN sequence. These shifts are also called the phase of the short code. Since all the base stations in the cellular system are synchronized using the time reference transmitted by the GPS satellites, it is possible for them to use different shifts of the same PN sequence for their spreading codes.

The pilot signals transmitted by the base stations consist of a repeating copy of their short codes that are not multiplied by any information. The IS-95 mobile measures the short code phase of the pilots from all the base stations that are nearby. Since each base station uses a unique short code phase, the mobile does this to identify which base stations it is close to.

The calculation used by the mobile to determine these short code phases is specified in the IS-95 standard and is given in Equation 4.15 [6].

$$\texttt{PILOT\_PN\_PHASE} = \texttt{PILOT\_ARRIVAL} + (64 \cdot \texttt{PILOT\_PN}) \bmod 2^{15} \qquad (4.15)$$

The PILOT_PN field is the short code phase of the pilot from the base station in the cell the mobile is currently in. The PILOT_ARRIVAL field is defined as the time-of-arrival of the first usable multipath component of the pilot signal being measured, relative to the timing reference of the phone. The mobile gets its timing reference from the

pilot of the base station in the mobile's cell. The PILOT_ARRIVAL field is quantized to one spreading chip.

4.4.2 Short Code Phase Measurement Location

The short code phase measurement described in the previous section can be used to locate the mobile. The PILOT_ARRIVAL field is essentially a time-of-arrival measurement, similar to the measurements used in Chapter 4. In Chapter 4, the mobile's signal was received at several locations. The relative time-of-arrival of the signal between two receiver locations defined a hyperbola. If the mobile's signal was received at 3 or more locations, 2 or more hyperbola can be defined. The mobile's location is given by the intersection of these hyperbola.

A similar idea is used to locate the mobile using its short code phase measurements. In this case, synchronized pilot signals are transmitted from several known locations. The mobile measures the relative time-of-arrival of these signals in the form of the PILOT_ARRIVAL field in Equation 4.15. These measurements can also be used to define several hyperbola, with the position of the mobile being given by their intersection.

The short code phase measurements can be put into relative time-of-arrival form by subtracting the measured short code phase from base station i from the measured short code phase from base station j, shown in Equation 4.16. Since the PILOT_PN field is the same for both measurements, it cancels out.

$$\tau_{ij} = \texttt{PILOT\_PN\_PHASE}_i - \texttt{PILOT\_PN\_PHASE}_j = \texttt{PILOT\_ARRIVAL}_i - \texttt{PILOT\_ARRIVAL}_j \tag{4.16}$$

The term τ_{ij} denotes the relative time-of-arrival of the pilot signals from base station i and j at the mobile.

Since the short code phase measurements taken by all the mobiles are returned to the Mobile Switching Center (MSC), this location system can be completely implemented in the MSC software. This makes this system very inexpensive to implement.

In addition to the normal errors that time-of-arrival measurements are subject to, this location method has some additional error due to quantization. The short code phase measurements are quantized to one chip. The chip duration, T_c, for and IS-95 mobile is approximately 1 μs. Assuming the PDF of the arrival times are uniformly distributed between $\pm T_c/2$, the variance of the error due to the chip quantization, σ_e^2 can be calculated using Equation 4.17.

$$\begin{aligned} \sigma_e^2 &= \tfrac{2}{T_c} \textstyle\int_0^{T_c/2} x^2 dx \\ &= \tfrac{2}{T_c} \left.\tfrac{x^3}{3}\right|_0^{T_c/2} \\ &= \tfrac{T_c^2}{12} \end{aligned} \tag{4.17}$$

This means that the standard deviation of the error is $T_c/\sqrt{12}$ or approximately 0.2 μs. This translates into a standard deviation of approximately 70 m.

There are two other potential sources of error for this technique. It is possible for the timing reference of the mobile to have an additional error of $\pm$ 1 μs relative to the time of arrival of pilot signal from the base station in its cell [6]. However, this error in the absolute timing reference of the mobile does not affect the location system since we use the difference between two short code phase measurements to locate the mobile. A second source of error is the synchronization of the pilot signals between

base stations. It is possible for the short code radiated by each of the pilot signals to have a timing error of $\pm$ 3 μs relative to IS-95 system time [6]. This problem can be addressed in two ways. The first is to modify the base stations to improve the synchronization of the pilot signals. The second, perhaps cheaper, solution would be to add a device to the base station that measures the exact timing reference of the pilot signal. That reference can be sent back to MSC and used as a correction factor in Equation 4.16.

4.4.3 Previous Work

The idea of using the timing of the base station pilot signals to locate the mobile is a very popular one. Many different versions of this idea have been proposed in patents and in technical papers.

A system is proposed by Sood that is very similar to the system discussed in this section [36]. A cellular network is described where the base stations transmit synchronized timing references. The mobile measures the relative time-of-arrival of the pilots from all the base stations nearby. Those measurements are then returned to one of the base stations and used to determine the mobile's location. While this patent does not specifically mention short code phase measurements, that may be the intent of the author.

A system very similar to the one described by Sood is disclosed Khan *et. al.* [37]. This system also measures the relative time-of-arrival of synchronized pilot signals. However, instead of returning the measurements to the base station, the mobile keeps the measurements and calculates its position itself. This way, the position of the phone is available to the user as well as to the system. A system is described by Sawada

[38] that is almost identical to the one described by Khan *et. al.*.

An analysis of the performance of pilot signal location is given by Hepsaydir *et. al.* [39]. The paper gives a brief explanation of TOA location and simulates the accuracy expected from pilot TOA estimates in a multipath channel. However, in the paper, they assumed that the correlator used to estimate the TOA of the pilot oversampled the signal by 32 times the chip rate. Since the short code PN phases are only returned at a quantization level equal to 1 chip, these results are not valid for the particular location scheme considered in this section.

4.4.4 Short Code Phase Location Simulation

This section presents the simulation used to investigate the accuracy of the pilot short code phase location method.

Simulation Setup

The simulation was a 7 cell simulation, as shown in Figure 4.8. The numbered circles represent base stations in the center of each of the cells. The radius of each cell was 1000 m.

The mobile being located was placed at random locations in the center cell. For each location, the short code phase measurements made by the phone on the pilot signals from the base stations close to the mobile were simulated. Those measurements were substituted into Equation 4.16 and used to find τ_{ij}. The time difference measurements were then used in the least squares algorithm discussed in Chapter 3 and used to solve for the coordinate position of the phone. It is assumed that the mobile has a line of sight path to each of the base stations.

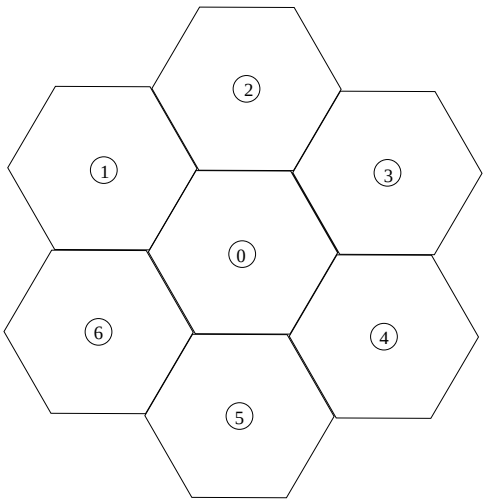

Figure 4.8: Cell Topology

Simulation Results

The first simulations performed assumed that nothing was done to correct the $\pm$ 3 μs error in the base station pilot signal synchronization. Each run of the simulation generated solutions for 5000 random mobile positions in the center cell. For each position, the radial error distance was recorded. Several runs were performed and the number of pilot signals received by the mobile was varied. The results are shown below. Figure 4.9 shows the mean radial error distance and Figure 4.10 shows the standard deviation of the error distance.

The one chip quantization of the arrival times and the pilot synchronization error result in significant position error. Figures 4.9 and 4.10 show that the system will not satisfy the FCC's requirements if nothing is done about the pilot synchronization error.

The next simulations assumed that a device was added to the antenna inputs in each base station to measure the exact timing of the pilots. These timing measurements

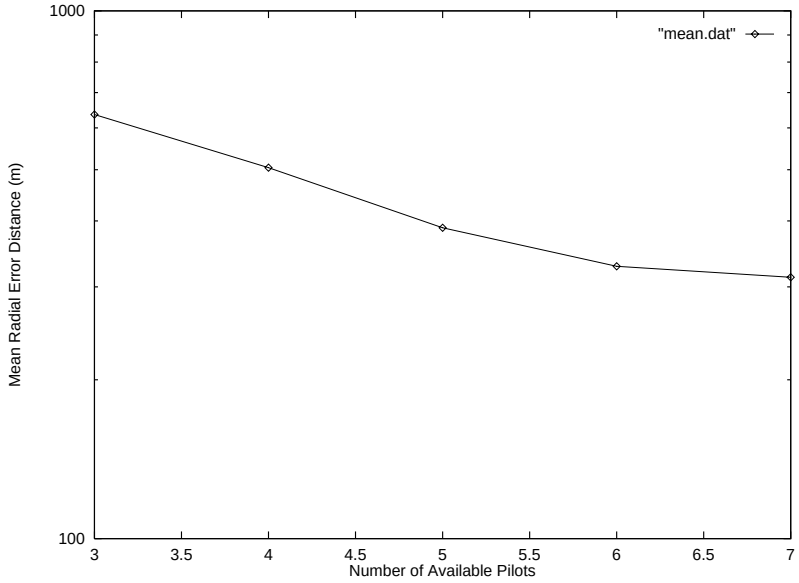

Figure 4.9: Mean Radial Error Distance

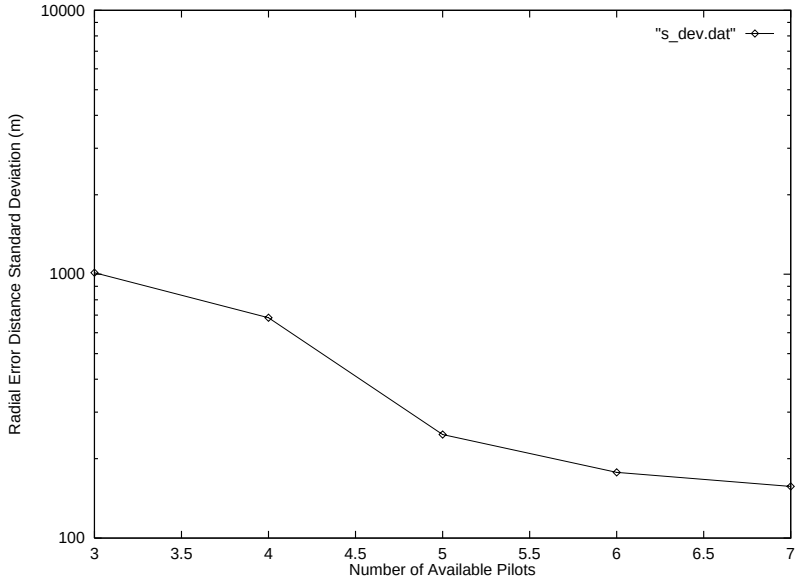

Figure 4.10: Standard Deviation of the Radial Error Distance

are returned to the MSC and used as a correction factor in Equation 4.16. This eliminates the error due to pilot synchronization. Figure 4.11 shows the mean radial error distance assuming the pilot synchronization error has been compensated for. Figure 4.12 shows the standard deviation of the radial error distance.

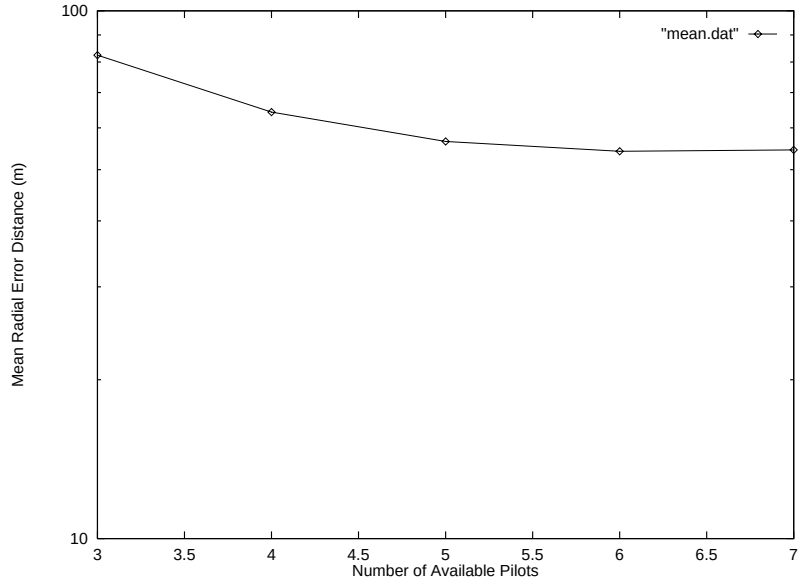

Figure 4.11: Mean Radial Error Distance

These plots show that once the pilot synchronization error is addressed, short code pilot phase location does satisfy the accuracy requirements of the FCC. The accuracy is better than 125 m even when short code signal phase measurements are available from only three pilot signals. Of course, these simulation results are very optimistic. Like the TOA position estimates in Chapter 3, the geometry of the simulation was very good and a LOS path from the mobile to the base stations was assumed every time.

The plots also show the error of the simulation leveling off as the number of available pilots is increased. This is due to the quantization error of the TOA estimates from the short code pilot phase measurements.

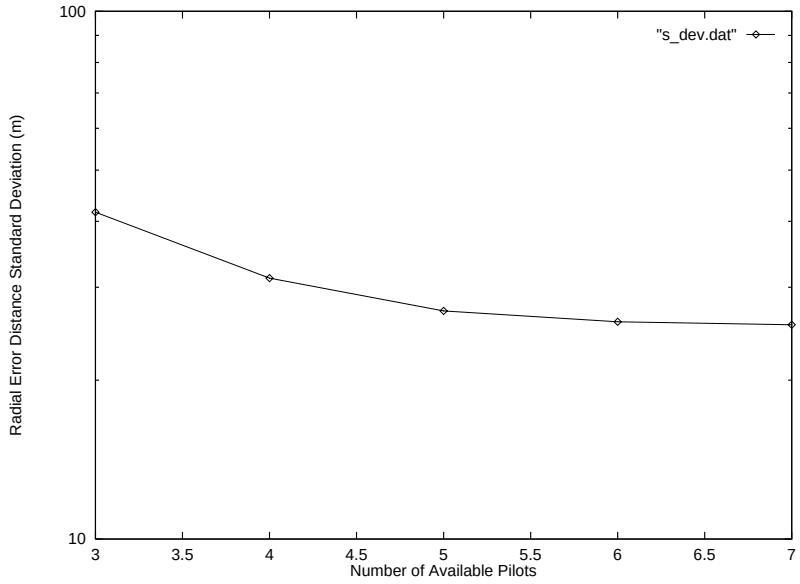

Figure 4.12: Standard Deviation of the Radial Error Distance

4.4.5 Conclusion

Short code signal phase location is a very good solution for the IS-95 mobile location problem. It uses measurements that the phone is already taking so no hardware modifications of the mobile are required. The short code pilot phase measurements are already available at the MSC, so the system can be implemented almost entirely in software. The only additional hardware that has to be added to the system would be a device in each base station to measure the timing of the base station pilot and return that timing to the MSC.

An additional modification of the IS-95 system might be required to provide location estimates on request. The IS-95 standard states that the mobile only returns base station pilot measurements when some change occurs in its Active or Candidate sets. The standard does not mention any command the base station can use to request these measurements from the mobile. If a location system that uses pilot

measurements is going to be useful, it must be able to determine the location of any mobile at any time. In order to do this, a command must be added to the system to allow the base stations to request the pilot measurements the mobile is taking. However, this modification would not be a major one and could be implemented entirely in the software of the base stations and the phones. This command may even exist already in some cellular systems for troubleshooting purposes.

4.5 Conclusion

This chapter presents two systems able to determine the position of the mobile using measurements it makes on the pilot signals transmitted by nearby base stations. The first system uses the pilot signal strength measurements taken by the mobile. While the accuracy of this approach does not satisfy the requirements of the FCC, it could still be a valuable supplement to another location system. The second approach uses the short code phase measurements taken by the mobile. Assuming that the pilot synchronization error is compensated for, this approach was able to satisfy the accuracy requirements of the FCC. Short code phase location is a good solution to the IS-95 mobile location problem. Both of the techniques presented in this chapter are attractive because they require little extra equipment or modification of the existing cellular system and can be implemented almost entirely in software.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

IS-95 mobile location is a challenging problem. There are two main criteria that must be met when developing an IS-95 mobile location system. The first is to develop a system that can reliably produce mobile location estimates under all circumstances. The second criteria is the location system should require little modification of the base stations and preferably no modification of the mobiles. The reason for this is the large investment that has already been made in existing IS-95 infrastructure. There are already hundreds of base stations and thousands of mobiles in use. Even a small modification of this equipment would be expensive. If a change was made to the mobile, eventually the old ones would be phased out. However, it is still very likely that many unmodified mobiles will still be in use by 2001.

The first step in this thesis project was to characterize conditions in an IS-95 system. This work was presented in Chapter 2. The object of the simulations was to provide an estimate of what the received SIR levels of a mobile would be at neighbouring base stations. The results showed that power control has the largest effect on reducing the SIR levels of the mobile's signal. They also showed that regardless of the type of coverage area, very low SIR levels can be expected. However, for less populated coverage areas, a larger proportion of the cell could be serviced by location systems that require a better SIR.

Of course, these simulations can only approximate the conditions found in a real system. As a result, the SIR values given in Chapter 2 should not be considered exact values. However, they do provide a good estimate of the very weak signal levels that can be expected from the mobile. This estimate is a good starting point for the design of a location system.

There are many possible solutions for the IS-95 location problem. This thesis investigates the performance of some conventional techniques and also proposes some new techniques for locating a mobile.

Chapter 3 investigates the performance of location systems that use the time-of-arrival of a mobile's signal at several base stations to determine its position. The simulations showed that this technique is very robust when multipath is introduced in the channel. The results also showed that TOA location has a lot of potential for providing accurate position estimates. Even with the crude threshold TOA estimation techniques used in the simulation, the accuracy of the position estimates satisfied the FCC's requirements for favorable geometry situations. Further accuracy enhancement is possible by using super-resolution techniques, but more work has to be done before they are reliable.

While the work in Chapter 3 shows a lot of promise for providing accurate, reliable position estimates, implementing that type of location system would require extra equipment to be added to the base stations. Since one of the criteria is to minimize the required modification of existing infrastructure, Chapter 4 discusses some location ideas that use measurements already being taken by the mobile for soft handoff purposes. The first system uses signal strength measurements the mobile makes on base station pilot signals for location and the second system uses PN phase

measurements made by the mobile on the pilot signals. The accuracy of the signal strength location system does not satisfy the FCC requirements. The PN phase location system will satisfy FCC specifications but only when a large number of pilot readings are available and the geometry of the base stations is quite good. The real advantage of these systems is that they are extremely inexpensive and easy to implement. The measurements used by both systems are already available in the software of the cellular switching office. Even if the accuracy of these approaches does not satisfy the FCC requirements under all conditions, they would still be an inexpensive supplement to another location system.

In addition to the solutions presented in this thesis, there are many other possibilities for solving the IS-95 location problem. Angle-of-arrival using antenna arrays is being seriously considered by some researchers. Putting a GPS receiver in each of the mobiles is another solution with the potential for providing very accurate location estimates. However with the hostile nature of the wireless channel and the demands for location estimates in increasingly congested urban areas, it is unlikely that a single technique will be found that will provide reliable location estimates for all situations.

As a result, the final solution will likely be the combination of several location systems. These systems would be complementary and each one would provide location estimates in situations where the other systems could not. The accuracy of the position estimates could be further improved by post-processing of the position estimates. The use of Kalman filtering or curve fitting techniques has been shown to improve the accuracy of location estimates [40] [16].

The work presented in this thesis is a good foundation for developing this type of solution. The analysis and simulation of the IS-95 system presented here does a

good job of characterizing the IS-95 mobile location problem. In addition, the several location systems presented have the potential to play a major part in the solution to that problem.

5.2 Future Work

All the results presented in this thesis are generated using simulations. Simulations are an excellent tool for estimating the performance of real systems. In the early stages of the design process, they show which techniques have promise and which techniques will not work. However, in order to have a detailed picture of how a location system will perform, measurements should be done on the physical system itself. If building the entire system is impractical, often it is good enough to take field measurements and insert the results from those measurements into key sections of the simulations.

Field measurements should be performed to supplement the system simulations presented in Chapter 2. Received signal levels are a key component in deciding what location methods are practical. Given their importance, some field measurements should be performed to confirm the simulation results. Since very low signal levels are expected when the mobile's signal is being received at neighbouring base stations, it would likely not be possible to accurately measure received signal levels at those points. However, it would be relatively easy to measure the mobile's received signal power at the base station it is communicating with. The simulation could be modified to generate these same signal levels and the two results could be compared. If the simulation results are close to the measured signal levels at the base station in the mobile's cell, the simulation can be considered accurate for the levels in the neighbouring cells.

Field measurements can also be used to improve the performance estimates of the TOA location approaches presented in Chapter 3. Specifically, measurements made in a real radio channel should be incorporated into the simulations. The simulated signal

of a mobile could be convolved with an actual radio channel, or an actual mobile signal received through a radio channel could be recorded and used in the simulations. In either case, the measurements should be performed at known locations so the accuracy of the TOA position techniques can be investigated.

It would be very easy to investigate the performance of the location techniques described in Chapter 4 in a real system. The measurements are contained in the phone and the Mobile Switching Center (MSC). A field trial would consist of moving a mobile to different known locations. For each location, the measurement taken by the mobile on the base station pilots could be recorded from the mobile or the MSC. Those measurements could then be incorporated directly into the simulations.

Additional work could be done on the SIR level simulations. A statistical analysis could be performed to determine the probability of coverage for a certain required SIR level. A confidence analysis could then be performed for different location systems. The simulations could also be redone for different base station configurations.

Clearly, geometry has a large effect on the final accuracy of the location system. Some simulations that calculate the DOP factor for arbitrary base station geometries should be performed. This would provide a better idea of what type of accuracy is required from the TOA estimates.

The final location system will likely consist of many different location techniques. Each technique provides another input to the location equation. An optimal way of combining these inputs should be investigated for different coverage area types.

Finally, post-processing algorithms on the location estimates should be investigated. Kalman filtering or some type of curve fitting algorithm could be used on the individual position estimates for a specific mobile to give a more realistic esti-

mate of its path. These techniques could potentially provide a significant accuracy improvement without any additional infrastructure cost.

Bibliography

[1] Federal Communications Commission, "Revision to the commission's rules to ensure compatibility with enhanced 911 emergency calling systems," July 1996, CC Docket No. 94-102.

[2] "CDMA development group website," www.cdg.org.

[3] P. LaForge, "1997 the year of globalization," *CDMA Spectrum*, December 1997.

[4] T. S. Rappaport, *Wireless Communications, Principles and Practice*, Prentice Hall, 1996.

[5] A. J. Viterbi, *CDMA: Principles of Spread Spectrum Communication*, Addison-Wesley Wireless Communications Series. Addison-Wesley Publishing Co., 1995.

[6] TIA/EIA/IS-95-A, *TIA/EIA Interim Standard: Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System*, Telecommunications Industry Association, 1995.

[7] S. Wang and I. Wang, "Effects of soft handoff, frequency reuse and non-ideal antenna sectorization on CDMA system capacity," in *IEEE 43rd Vehicular Technology Conference (VTC '93)*, 1993, pp. 850–4.

[8] J. J. Caffery and G. L. Stuber, "Radio location in urban CDMA microcells," in *Sixth IEEE International Symposium on Personal, Indoor, and Mobile Communication; PIMRC '95*, 1995, pp. 858–62.

[9] G. Wu, A. Jalali, and P. Mermelstein, "On channel model parameters for microcellular CDMA systems," *IEEE Transactions on Vehicular Technology*, vol. 44, no. 3, pp. 706–11, August 1995.

[10] B. R. Petersen, "Wireless communications course notes," ENEL 673, unpublished, fall term, University of Calgary, Calgary, AB, Canada, 1996.

[11] T. S. Rappaport and L. B. Milstein, "Effects of radio propagation path loss on DS-CDMA cellular frequency reuse efficiency for the reverse channel," *IEEE Transactions on Vehicular Technology*, vol. 41, no. 3, pp. 231–42, August 1992.

[12] J. J. Caffery and G. L. Stuber, "Overview of radiolocation in CDMA cellular systems," *IEEE Communications Magazine*, pp. 38–44, April 1998.

[13] G. Lachapelle, M. E. Cannon, and J. Ray, "GPS theory and applications," Course Notes, July 1998.

[14] D. L. Schilling, "CDMA communications and geolocation system and method," U.S. Patent Number 5506864, Filed: September, 1994, Assignee: InterDigital Technology Co.

[15] A. J. Mallinckrodt, "Position determination in an integrated cellular communications system," U.S. Patent Number 5612703, Filed: May, 1995, Assignee: Cellsat America Inc.

[16] M. Pent, M. A. Spirito, and E. Turco, "Method for positioning GSM mobile stations using absolute time delay measurements," *Electronics Letters*, vol. 33, no. 24, pp. 2019–20, November 1997.

[17] L. Tong and F. Ling, "Method and apparatus for locating a mobile station in a spread spectrum communications system," U.S. Patent Number 5675344, Filed: June, 1996, Assignee: Motorola Inc.

[18] R. Klukas, A. Borsodi, M. Asteridge, M. Fattouche, and G. Lachapelle, "A system to position analog cellular telephones," in *Wireless '96*, 1996, pp. 17–28.

[19] A. Borsodi and M. Fattouche, "Super resolution of discrete arrivals in a position location cellular system," in *Wireless '95*, 1995, pp. 727–39.

[20] M. Aatique, G. Mizusawa, and B. D. Woerner, "Performance of hyperbolic position location techniques for code division multiple access," in *Wireless '97*, 1997, pp. 261–74.

[21] W. H. Foy, "Position-location solutions by taylor-series estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-12, no. 2, pp. 187–93, March 1976.

[22] G. D. Morley and W. D. Grover, "Improved location estimation for automatic vehicle location with multipath-corrupted range estimates," in *Wireless '95*, 1995, pp. 711–25.

[23] G. Strang, *Linear Algebra and its Applications*, Harcourt, Brace & Company, 3 edition, 1988.

[24] A. Gosh and R. Love, "Mobile station location in a DS-CDMA system," in *IEEE Conference on Vehicular Technology, VTC '98*, 1998, pp. 254–8.

[25] Texas Instruments, *TMS320C62xx CPU and Instruction Set Reference Guide*, July 1997, Literature Number: SPRU189B.

[26] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice hall Signal Processing Series. Prentice-Hall Inc., 1993.

[27] R. W. Klukas, *A Superresolution Based Cellular Positioning System Using GPS Time Synchronization*, Ph.D. thesis, University of Calgary, 1997.

[28] L. R. Dumont, "Super-resolution of discrete arrivals in a spread spectrum system," M.S. thesis, University of Calgary, 1994.

[29] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall Inc., 2 edition, 1991.

[30] H. Yamada, M. Ohmiya, Y. Ogawa, and K. Itoh, "Superresolution techniques for time-domain measurements with a network analyzer," *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 2, pp. 177–83, February 1991.

[31] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, Prentice-Hall Inc., 1989.

[32] P. W. Dent, "Navigation assistance for call handling in mobile telephone systems," U.S. Patent Number 5404376, Filed: September, 1993, Assignee: Ericsson-GE Mobile Communications Inc.

[33] H. L. Song, "Vehicle locating and navigating system," U.S. Patent Number 5208756, Filed: January, 1991.

[34] D. Dufour, "System and method of locating a mobile terminal within the service area of a cellular telecommunication system," U.S. Patent Number 5613205, Filed: March, 1995, Assignee: Telefonaktiebolaget LM Ericsson.

[35] J. R. Doner, "Mobile telephone location process making use of handoff data," U.S. Patent Number 5657487, Filed: June, 1995, Assignee: AirNet Communications Co.

[36] P. L. Sood, "Apparatus and method for locating mobile and portable radio terminals in a radio network," U.S. Patent Number 5293645, Filed: October, 1991, Assignee: Sharp Microelectronics Technology Inc.

[37] M. H. Khan and J. Boccuzzi, "Method and apparatus for a portable communications device to identify its own location," U.S. Patent Number 5646632, Filed: November, 1994, Assignee: Lucent Technologies Inc.

[38] K. Sawada, "Apparatus for acquiring the position of a mobile station," U.S. Patent Number 5629710, Filed: January, 1995, Assignee: Fujitsu Ltd.

[39] E. Hepsaydir and W. Yates, "Performance analysis of mobile positioning using existing CDMA network," in *IEEE Third International Symposium on Spread Spectrum Techniques and Applications; ISSSTA '94*, 1994, pp. 190–2.

[40] M. Hellebrandt, R. Mathar, and M. Scheibenbogen, "Estimating position and velocity of mobiles in a cellular radio network," *IEEE Transactions on Vehicular Technology*, vol. 46, no. 1, pp. 65–71, February 1997.