2006

# Analysis of Relay-based Cellular Systems

Ansuya Negi
*Portland State University*

DISSERTATION APPROVAL


The abstract and dissertation of Ansuya Negi for the Doctor of Philosophy in Computer Science were presented December 5, 2006, and accepted by the dissertation committee and the doctoral program.

COMMITTEE APPROVALS:

_____
Suresh Singh, Chair


_____
Cynthia Brown


_____
Warren Harrison


_____
Su-Hui Chiang


_____
Douglas Hall
Representative of the Office of Graduate Studies


DOCTORAL PROGRAM APPROVAL: _____
Cynthia Brown, Director
Computer Science Ph.D. Program

ABSTRACT

An abstract of the dissertation of Ansuya Negi for the Doctor of Philosophy in Computer Science presented December 5, 2006

Title:   Analysis of Relay Based Cellular Systems

Relays can be used in cellular systems to increase coverage as well as reduce the total power consumed by mobiles in a cell. This latter benefit is particularly useful for mobiles operating on a depleted battery. The relay can be a mobile, a car or any other device with the appropriate communication capabilities. In thesis we analyze the impact of using relays under different situations. We first consider the problem of reducing total power consumed in the system by employing relays intelligently. We find that in a simulated, fully random, mobile cellular network for CDMA (Code Division Multiple Access), significant energy savings are possible ranging from 1.76 dB to 8.45 dB.

In addition to reducing power needs, relays can increase the coverage area of a cell by enabling mobiles located in dead spots to place relayed calls. We note that use of relays can increase the useful service area by about 10% with real life scenarios. We observe that in heavy building density areas there is more need of relays as compared to low building density areas. However, the chance of finding relays is greater in low building density areas. Indeed, having more available idle nodes helps in choosing relays, so we conclude that unlike present day implementations of cellular

networks, the base station should admit more mobiles (beyond the capacity of the cell) even if they are not placing calls since they can be used as relays.

One constraint of using relays is the potential to add interference in the same cell and in neighboring cells. This is particularly true if the relays are not under power control. Based on our analysis, we conclude that in interference limited systems like CDMA the relays have to be under power control otherwise we will reduce the total capacity by creating more dead spots. Thus, we believe that either the base station should be responsible to allocate relays or relays should be provided with enough intelligence to do power control of the downlink. Finally, we show how utility of data services can be increased by use of relays.

ANALYSIS OF RELAY BASED CELLULAR SYSTEMS

by

ANSUYA NEGI

A dissertation submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY
in
COMPUTER SCIENCE

Portland State University
2007

With HIS blessings,
Dedicated to my parents,
Sushila and Surendra Singh Chauhan

## Acknowledgement

I would like to take this opportunity to thank my advisor Prof. Suresh Singh and express my heartfelt gratitude for his constant encouragement and support. His guidance and editorial skills were essential in the completion of this thesis. I am thankful to department chair Prof. Cynthia Brown for her support these past years. I also thank my doctoral committee members, Prof. Cynthia Brown, Prof. Warren Harrison, Prof. Su-Hui Chiang and Prof. Douglas Hall.

My special appreciation and gratitude goes to Shashidhar Lakkavalli for fruitful discussions and for just being there. I am thankful to my lab colleagues, Maruti Gupta, Satyajit Grover, Dilip Sundarraj and Ravi Ghanta for their support. I am thankful to the faculty of Computer Science department and especially Prof. Jim Binkley for teaching finer aspects of networking. I also want to thank Computer Science staff especially Rene Remillard, Salinda Calder, JaNice Brewster, Kathy Lee and Beth Phelps.

All this would not have been possible without the help and support of my husband Jeewan, daughter Shivani and my mom. They accepted many busy evenings and weekends and I am eternally indebted for their encouragement and understanding. Finally, I am grateful to my parents for installing the values and patience that helped me immeasurably during this research.

**Table of Contents**

## List of Tables

# List of Figures

# 1. Background

In this research we investigate the use of relays in a Code Division Multiple Access (CDMA) based cellular system. Relays act as an intermediary between the user and the base station. The purported benefits of using relays include better coverage, lower energy consumption, and possibly increased capacity. However, given the nature of the cellular environment, we need to consider several side effects as well as implementation challenges. In this thesis, we introduce these constraints and then explore how relays can be used in the cellular environment. We first describe various components of a cellular system and the different multiple access technologies used. The evolution of air interfaces from analog to digital is then described. The emergence of CDMA as the technology of choice for cellular systems is discussed. Finally, relevant aspects of present day digital technology such as power control and the near far effect, are explained in the context of introducing relays.

## 1.1 Cellular System Architecture

A cellular system is a collection of many interconnected subsystems that communicate between themselves through network interfaces [Rappaport, 1996]. To illustrate a typical architecture consider the Global System for Mobile (GSM) cellular standard. The subsystems of GSM consist of the Base Station Subsystem (BSS), Network and Switching Subsystem (NSS), and the Operation Support Subsystem (OSS).
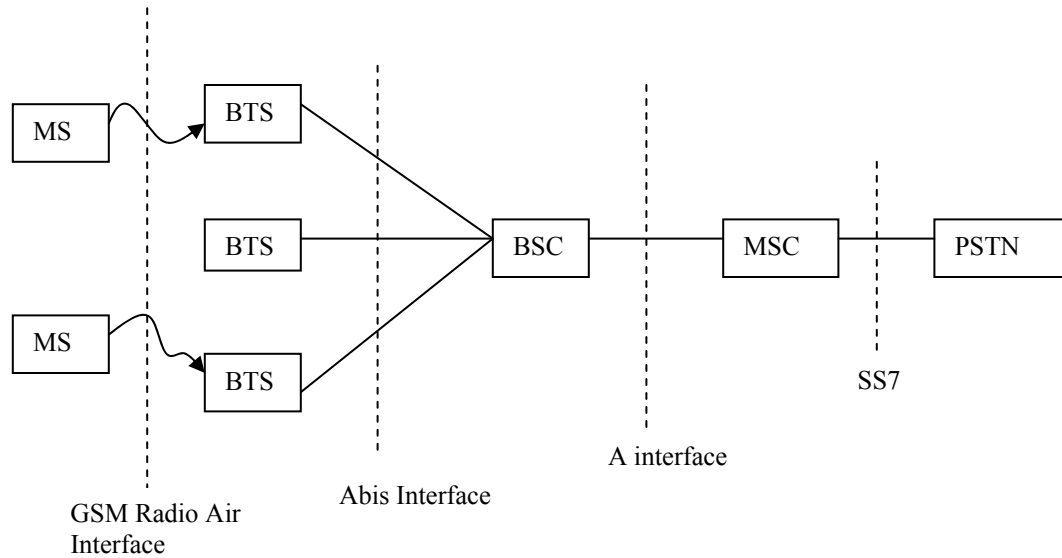
Figure 1. 1 Architecture of GSM

The Base Station Subsystem comprises of several Base Station Controllers (BSC), Base Transceiver Stations (BTS), and Mobile Stations (MS). Each BSC controls many BTSs and handoffs between two BTSs are controlled by the BSC. The interface between the MS and BTS is called an air interface, while the interface between the BTS and BSC is called the Abis interface. BTS and BSC are connected via dedicated leased lines or microwave links.

The Network and Switching Subsystem (NSS) consists of a Mobile Switching Center (MSC). The NSS manages the switching functions and also connects the cellular network with outer networks such as telephone networks or the Internet. BSCs are connected with the MSC via dedicated leased lines or microwave links. The Operation Support Subsystem is connected to the MSC to monitor and maintain the GSM system.
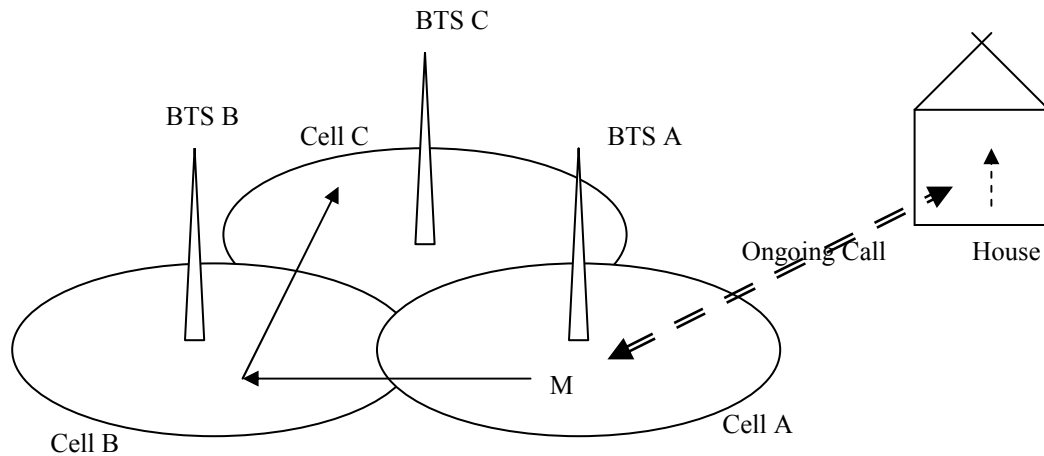
Figure 1. 2 Example illustrating the major functions in a GSM cellular system

In Figure 1.2, BTS A and BTS B are controlled by BSC 1 while BTS C is controlled by BSC 2. When the mobile user with an ongoing call moves from cell A to B, BSC 1 hands over the call from BTS A to BTS B. Handing over the call involves routing the traffic to BTS B instead of BTS A. BSC 1 also updates a location registry which keeps track of the mobile's latest cell. When the mobile moves from cell B to cell C, the call needs to be handed off to BSC 2. However, BSC 1 can not reach BSC 2 directly. Here, the MSC has to interact with both BSC's to hand off the call. As the call being shown in Figure 1-2 reaches a land phone, the MSC will direct the call to the appropriate Public Switched Telephone Network (PSTN) line.

## 1.2 Multiple Access Techniques

Mobiles in a cell have to share a finite radio spectrum and this means simultaneous allocation of the bandwidth to multiple users [Rappaport, 1996]. Over

the years, researchers have investigated several such multiple access techniques. Those relevant to cellular systems are summarized below.

1.2.1 Narrowband Systems

In these systems the available bandwidth is divided into many narrowband channels. Frequency Division Multiple Access (FDMA) is one such system where the frequency is divided into many channels. Each user is allocated one band. If it is a Frequency Division Duplex (FDD) system then different channels are allocated to the forward and reverse links. If it is Time Division Duplex (TDD) system then the same frequency bands are used for both the links, although separated by time. Time Division Multiple Access (TDMA) is another narrowband system where the spectrum is divided into time slots. Each user is allocated a time slot and slots are rotated cyclically between different users. TDMA again can operate in FDD or TDD mode.
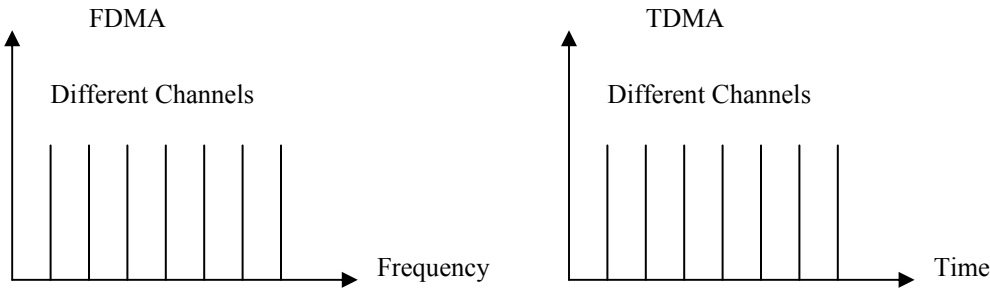


Figure 1. 3 Types of narrowband systems

1.2.2 Wideband Systems

In wideband systems the entire bandwidth is available to all the users. However, Code Division Multiple Access (CDMA) is used to provide different

channels. Thus, each user is assigned a unique code, which allows them to transmit simultaneously in the same channel.

To illustrate how CDMA works we take a signal denoted by $s_i(t)$ and wideband code signal denoted by $c_i(t)$ [Garg, 2002]. The signal is modulated and transmitted as $y_i(t)$, given by

$$y_i(t) = s_i(t) \bullet c_i(t) \tag{1.1}$$

where, '$\bullet$' denotes the multiplication operator. $c_i(t)$ has a very high rate as compared to $s_i(t)$ and therefore $y_i(t)$ is said to be spectrum spread. Since all mobiles transmit simultaneously, all the coded data streams can be represented as,

$$x(t) = \sum_i y_i(t) \tag{1.2}$$

This spread spectrum signal is modulated by a carrier and we get $z_i(t)$.

$$z_i(t) = A_i \left[ \sum_j c_j(t) \bullet s_j(t) \right] \cos(\omega_c t) \tag{1.3}$$

This signal is transmitted over the bandwidth and at the receiving end $z_i(t)$ plus noise is received. The original signal is recovered by despreading the signal at the receiver. In order for the receiver to recover each transmission, the codes $c_i(t)$ need to be orthogonal to each other. In addition, the power at which each signal is received needs to be carefully controlled to ensure that no one signal is drowned out by others. Indeed, unlike in pure TDMA or FDMA systems, power control is critical to the functioning of CDMA systems.

**1.3 Growth of Air Interfaces**

The initial cellular offerings were analog in nature. The signal was modulated and sent at a high frequency carrier. The standard is called the Advanced Mobile Phone System (AMPS). This technology of the early eighties came to be known as the first generation cellular technology. However there were many analog systems offered by different companies resulting in incompatibility. People couldn't use roaming to talk between different systems. To overcome such problems, the interfaces and protocols needed to be standardized. In Europe, the Global System for Mobile Communications (GSM) standardized the interfaces and protocols, making the network elements independent. In such an open architecture, any base station can communicate with any mobile switching center (MSC) and also can be independently modified. This heralded the second generation (2G) of cellular technologies. In America, the Joint Technical Committee (JTC) started standardization efforts in 1992. JTC established the Technical Adhoc Groups (TAG) to standardize the air interfaces. Some of these digital second generation standards [Garg, 2002] are summarized below.

- IS-95: This standard operates at 1.9GHz and uses CDMA access technology. This standard is interoperable with AMPS and thus can provide roaming where IS-95 is not available. The cells can have a range of upto 50 km. There are two transmission rates supported by different speech codecs. Rate set 1 (RS1) supports 9.6 kbps, 4.8 kbps, 2.4 kbps and 1.2 kbps. Rate set 2 (RS2) supports 14.4 kbps, 7.2 kbps, 3.6 kbps and 1.8 kbps. The transmission rate is variable and depends on

speech activity of the user. Not only voice but data services are also supported including asynchronous data, facsimile, packet data and short messaging.

- GSM1900: This standard operates at 1.9GHz and uses TDMA access technology. Gross transmission rate per traffic channel is 22.8 kbps. Both circuit switched and packet switched data can be provided. The cells have a range of 35km in rural areas and upto 1 km in urban areas. The standard provides a high level of security by means of Subscriber Identity Module (SIM) cards. SIM cards also help in roaming between different carriers of the GSM system. It is to be noted that GSM can support three times the system capacity of the first generation AMPS systems.

- Wideband CDMA (W-CDMA): This standard is based on wideband CDMA access technology and can support 5MHz, 10MHz and 15MHz channel spacing. It supports data rates of 16 kbps, 32 kbps and 64 kbps. The range of a cell can reach upto 5 km.

W-CDMA can support sixteen times the AMPS capacity, due to the use of CDMA as the access technology. The increase in system capacity is not only due to reuse of the spectrum but better coding gain/modulation schemes and the fact that CDMA uses a larger bandwidth. As a result the power is spread over a larger bandwidth, making the average power very low. The implication of this are reduced interference and a larger battery life at the mobiles. Another notable feature of CDMA is that it uses soft handoff, which means that a mobile can be connected to two or more base stations and can hand over the call to the base station with a better signal. This feature reduces the percentage of dropped calls and also avoids a ping-pong effect, which means that nearby base stations juggle with the mobile as the call is

rapidly switched. Finally, another advantage of CDMA technology is an improvement

in Quality of Service (QoS) in fading environments. This is a result of better

utilization of multipath propagation by the use of RAKE receivers. A RAKE receiver,

illustrated in Figure 1.4, consists of many integrators, one for every path, thus utilizing

the multipath environment. Each path is resolved independently and combined to

produce a net overall gain. The three strongest signals are selected by the RAKE

receiver and coherently combined to get an enhanced signal. This is not possible in

narrowband systems where fading is a major cause of signal degradation.
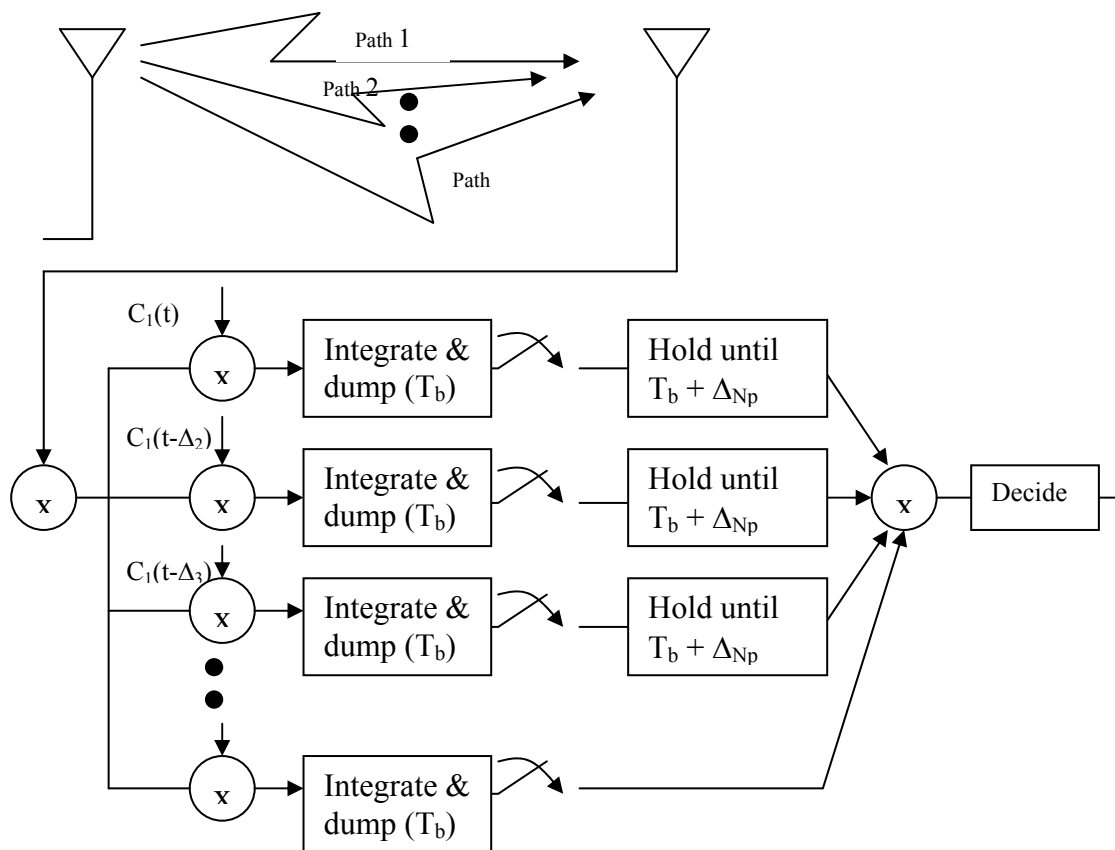


Figure 1. 4 RAKE receiver design

Third generation technologies (3G) are primarily CDMA based [Garg, 2002]. The International Telecommunication Union (ITU) has created a standard for 3G wireless systems called International Mobiles Telecommunications-2000 (IMT-2000). IMT-2000 standards are either the Universal Mobile Telecommunications System (UMTS) developed by the Third Generation Partnership Project (3GPP) [3GPP, 1998] or cdma2000 developed by the Third Generation Partnership Project 2 (3GPP2) [3GPP2, 2000]. Europe and Japan use UMTS based on W-CDMA while North America uses cdma2000, which is based on the existing IS-95 standard. The difference in these new standards over the 2G systems includes the provision of wireline voice quality and high data rates. They feature 144 kbps for mobile users and 2 Mbps for stationary users over a 2 GHz frequency band. The bit rate is thus higher than the 10-50 kb/s offered by 2G systems [Jamalipour and Yabusaki, 2003]. Another key difference between 2G and 3G systems is the use of a hierarchical cell structure that enables seamless transition to fixed data networks. This is accomplished by the use of hierarchical cells based on multi-layering picocells and microcells over macrocells.

## 1.4 Key Concepts in Cellular System Communications

This section provides an overview of the communication concepts in cellular systems that are relevant to the use of relays. Specifically, we describe path loss, interference, and power control.

1.4.1 Path loss

Path loss refers to the degradation in the transmitted signal as a function of the relative location of the receiver. As an example, assume that the transmitter and receiver are in free space. The transmitted signal may be viewed as an expanding sphere. Thus, the signal intensity degrades as distance squared (since the area of the sphere is $4\pi r^2$). The path loss exponent is thus said to be 2 and this is called free space path loss. For free space, the transmitted power $P_t$ as a function of distance, d, between transmitter and receiver, is given by the equation [Rappaport, 1996]

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

where, $P_r$ is the received power,

$G_t$ is the transmitter antenna gain,

$G_r$ is the receiver antenna gain,

$\lambda$ is the wavelength (meters),

L is the system loss factor.

We see that the received power as well as antenna gain for both transmitter and receiver can affect the transmitted power. We will use L =1 to indicate that there is no loss due to filter or antenna or transmission line.

In realistic situations, the signal suffers from additional effects including diffraction off rooftops, interference between multiple reflected components of the same signal, absorption in air and other material, and noise. The general form of the equation describing the received signal power is typically quite complex and depends

on these factors. Measured penetration loss into suburban homes is 4 to 7 dB at 800 MHz [Bertoni, 2000] and higher at higher frequencies. The transition loss through a row of houses is 4*n to 14*n dB, where n is the number of houses in between. The signal level drops about 15 dB when turning a corner and 30 dB when moving further down the street. Average building penetration loss is 12 dB with standard deviation of 8 dB according to UMTS standards.

[Hata, 1980] gave an empirical model, arrived by regressively fitting the curves reported by [Okumura, 1968]. Path loss in decibels in an urban area is given as,

$$L = 69.55 + 26.16\log f_M - 13.82\log h_{BS} - a(h_m) + 44.9 - 6.55\log h_{BS})\log R_k \qquad 1.5$$

where,

$f_M$ = frequency in megahertz, between 150 and 1500 MHz

$R_k$ = distance from the base station in kilometers, 1 to 20 Km

$h_{BS}$ = height of the base station antenna in meters, 30 to 200 m

$h_m$ = height of the subscriber antenna in meters, 1 to 10 m

The term a($h_m$) gives the dependence of path loss on subscriber antenna height and is defined such that a(1.5) = 0.

For regions classified as a large city,

$$a(h_m) = 8.29(\log 1.54 h_m)^2 - 1.10 \text{, for } f_M <= 200 \text{ MHz} \qquad 1.6$$

$$= 3.2(\log 11.75 h_m)^2 - 4.97 \text{, for } f_M >= 400 \text{ MHz} \qquad 1.7$$

For a small to medium sized city,

$$a(h_m) = (1.1\log f_M - 0.7)h_m - (1.56\log f_{M-0.8}) \qquad 1.8$$

The constant term of 69.6 dB in this model accounts for the building environment [Bertoni, 2000]. It is shown that this model corresponds to the one derived theoretically considering diffraction over rooftops. It shows that for $h_{BS}$ = 10m and $h_m$ = 1.5m, the theoretical and Hata models for path loss are,

Theory: $L = 53.7 + 21\log f_M + 38\log R_k$          1. 9

Hata:    $L = 49.2 + 26.2\log f_M + 35.2\log R_k$          1. 10

Another empirical formula that includes the effects of buildings on path loss is given by CCIR (Comite' Consultatif International des Radio-Communication, now ITU-R), which is given by,

$$L(dB) = 69.55 + 26.16\log_{10} f_{MHz} - 13.82\log_{10} h_1 - a(h_2) +$$
$$(44.9 - 6.55\log_{10} h_1)\log_{10} d_{km} - B \qquad \qquad 1.11$$

where,

$$a(h_2) = (1.1\log_{10} f_{MHz} - 0.7)h_2 - (1.56\log_{10} f_{MHz} - 0.8)$$
$$B = 30 - 25\log_{10}(\%buildings) \qquad \qquad 1.12$$

Here $h_1$ and $h_2$ are base station and mobile antenna heights in meters, $d_{km}$ is the link distance in kilometers, $f_{MHz}$ is the center frequency in megahertz. B is the correction factor and (% buildings) shows the percentage of area covered by buildings.

Introducing a relay into a cell changes the path loss seen by a transmitting mobile since the mobile-to-relay part of the connection will typically have a different set of propagation constraints as compared to a mobile to BS connection. For instance, if relays are also mobiles, then the antenna height of both connection end-points is the same at approximately 1 meter. We consider these issues later in this thesis.

1.4.2 Power Control

A mobile user nearer to the base station may send a higher power signal that may drown a weaker signal from a distant mobiles. This is called the near-far effect. In a CDMA cell, if users have to share the media they have to increase or decrease their power in a way that prevents the near-far effect. This is done via power control.

Power control is achieved in cellular systems in two ways. One is called open loop power control and the other is called closed loop power control [Garg, 2002]. In order for base station to properly distinguish each mobile signal, power from each mobile should be the same at the base station. This makes the signal contours look like circles around base station in a free space environment.

1.4.2.1 Open Loop Power Control

This is also known as autonomous power control since there is no feedback from the BS. This mechanism is used by mobiles that have not yet been assigned a traffic channel by the BS (e.g., new mobiles entering a cell). In reverse open loop power control (ROPC), the mobile changes its transmit power depending on the received power from all base stations. The received power includes power in all channels - pilot, paging, sync and traffic channels. If the received power is high the transmit power is reduced and vice versa. ROPC happens 50 times per second, i.e., every 20 ms.

1.4.2.2 Closed Loop Power Control

In closed loop power control (CLPC), the base station provides specific feedback to the mobile. The BS may ask the mobile to either reduce or increase its transmit power. Thus, unlike ROPC, which assumes identical forward and reverse link conditions, CLPC correctly interprets uplink channel conditions to provide accurate feedback to the mobile. Closed loop power control consists of an inner and outer loop power control called reverse inner loop power control (RILPC) and reverse outer loop power control (ROLPC). In order to understand this power control mechanism it is important to explore the frame structure. A frame has a length of 20 ms and it is composed of 16 time slots of equal duration, namely 1.25ms. These slots are called Power Control Groups (PCG). In the RIPLC mechanism, at the base station every 1.25 ms, the received signal strength is measured by measuring $E_b/I_t$, which is the ratio of energy per bit $E_b$ and total noise and interference power spectral density $I_t$. If the received signal strength exceeds a target value, a power down power control bit of 1 is sent, else a power control bit 0 is sent. Each power bit of 1, on being received by the mobile PCG produces a 1 dB change in mobile power. The ROPLC mechanism calculates a new set point $E_b/I_t$ at every frame interval, i.e., 20 ms, as against frequent checks for each frame i.e., 16 times in the RIPLC mechanism. The RIPLC helps the mobile to be as close as its target $(E_b/I_t)_{setpoint}$ and ROPLC adjusts the base station target $(E_b/I_t)_{setpoint}$ for a given mobile.

1.4.3 Interference

Introducing relays complicates power control and gives us new design choices. For example, should the relay act as a mini base station and provide closed loop power control to the mobile? How will other transmissions affect reception at the relay? If the base station does cell wide power control, by what mechanism can the base station ensure reduced interference at the different relays? We examine these questions in this thesis.

The signal to interference ratio for a user, i, at the base station is given as [Veeravalli and Sendonaris, 1999],

$$SIR_i = \frac{Signal_i}{\sum_{j \neq i} Signal_j + BackgroundNoise + OtherCellInterference} \qquad 1.\,13$$

The required SIR, denoted as SIR* is a function of target frame error rate (FER) or target SIR, given by SIR$^{target}$ and the multipath conditions [Veeravalli and Sendonaris, 1999] and is given by,

$$SIR_i^* = SIR_i^{t\,arg\,et}\delta_i \qquad 1.\,14$$

where, $\delta_i$ is the error in the power control algorithm. Each mobile user has the required SIR and power control algorithms have to satisfy this equation for all users.

When we introduce relays, however, we create a new set of problems – interference at relays in a cell and interference at mobiles in adjacent cells. As mentioned previously, we need to manage interference at relays from other mobiles in order for the relay to communicate with its mobile. A second problem created by

relays is increased interference in neighboring cells. This happens when the relay is located close to the cell boundary. The relay in one cell is not involved in power control in adjacent cell and thus we may create a race condition where increased interference due to the relay causes the neighboring BS to increase power which causes the relay to also increase power, and so on.

## 1.5 Roadmap to remainder of the thesis

Chapter 2 and 3 introduce the possible configurations in which relays can be a part of the cellular network. We explore the benefits and challenges that relays bring. We compare with Appendix C, where we look at work on using relays in cellular environments. Chapter 4 explores the benefit of saving power by the use of relays. A discrete event simulator to calculate total system energy savings is described in detail. Chapter 5 explores the increased reliability of coverage that is brought about by the use of relays. Chapter 6 explores the problem of interference due to relays. In chapter 7, a pricing mechanism of relay based cellular systems is discussed in terms of utility. The goal is to understand how users can be motivated to act as relays (since their batteries get depleted). In chapter 8, we conclude and try to assimilate the answers that have been gathered in the previous chapters.

## 2. Benefits and Challenges of Using Relays

As mentioned previously the CDMA cellular environment is a heavily power controlled, interference limited environment. In such a system one has to be very careful in introducing new components, such as relays. In this thesis, we explore the benefits of relays in cellular systems while addressing the problems caused by the introduction of relays. There are two mechanisms to incorporate a relay into a cellular system.

- Use of ad hoc networks as a separate channel in the cell giving us a hybrid network in which either only relays or all mobiles have two interfaces, one for the cellular system and one for the ad hoc system. An example is the Integrated Cellular and Ad-hoc Relay (iCAR) [Wu, Qiao, De and Tonguz, 2001] system, where each mobile has two interfaces, one cellular at 1.9GHz and another for ad hoc at 2.4 GHz.

- Use of the existing cellular technology but adding additional functionality to the mobiles.

The first mechanism involves introduction of new technology in current cellular systems. Even though ad hoc technology has been well studied, it brings about many of its own problems, namely, the hidden terminal problem, unlicensed band and interference with other same-band applications. For this reason, in this research we use the second approach and look at using mobile nodes as relays, in an on-demand basis.

## 2.1 Benefits of Relaying

To better appreciate the benefits of using relays in cellular networks, let us consider some scenarios. In the figures given below, M represents a mobile requiring its call to be carried (or relayed), R represents the carrier of the call (i.e., a relay) and BS represents the Base Station.

- Scenario I – Power Savings

  A shorter distance between a transmitter and receiver gives us a square of distance power law, where the received signal strength is inversely proportional to distance squared. For larger distances, however, the power law can be the third or fourth power of distance. This means that even halving the distance can bring about major savings in power. In the figure below M is at the cell periphery. By relaying the call through R, we reduce the transmission distance for M thus reducing its power consumption. In addition, in many cases, the total power used in M → R and R → BS is less than the direct M → BS transmission.
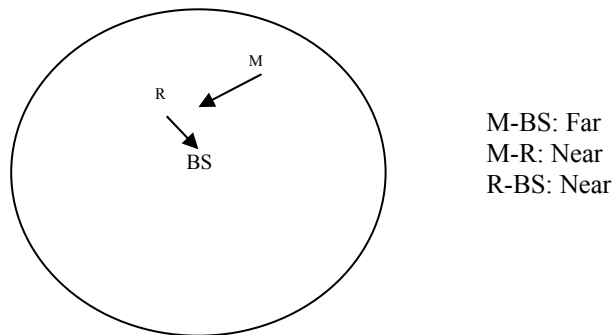


Figure 2. 1 Reduction in power by use of relays

- Scenario II – Power Savings

    Mobile is at Non Line of Sight (NLOS) and can significantly reduce total power consumed by using a Line Of Sight (LOS) transmission. As shown in the figure below, the buildings introduce a NLOS path to the mobile, which can be avoided if a relay is utilized.
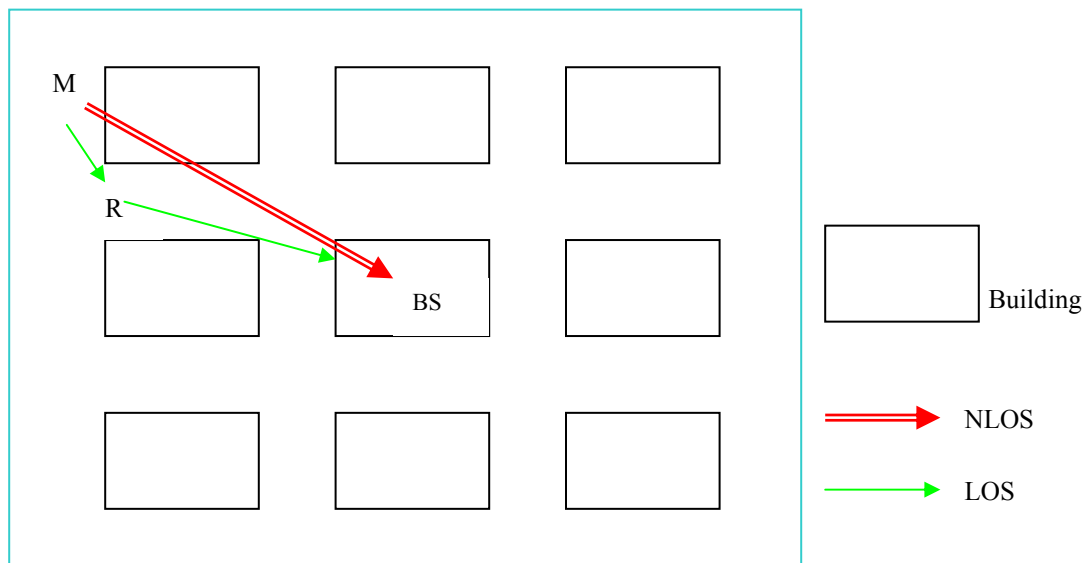


Figure 2. 2 Use of LOS path by relay to overcome NLOS path

- Scenario III – Power Savings

    Mobile needs to save its own battery power because it is running low. Thus, the relay's battery is consumed instead of the mobile's battery. This introduces an interesting economic dimension to the problem, where the relay is compensated for carrying the call.

- Scenario IV – Improvement in useful service area

    Mobile is at a dead spot where it cannot talk to the BS. For example, if we plot path loss contours (using CCIR path loss formulae) for an urban environment, we can see that local islands are introduced wherever buildings are present denoting dead spots. The mobiles at these areas can talk to the BS only if relays are present nearby.



M-BS: Signal unreachable
M-R: Signal reachable
R-BS: Signal reachable
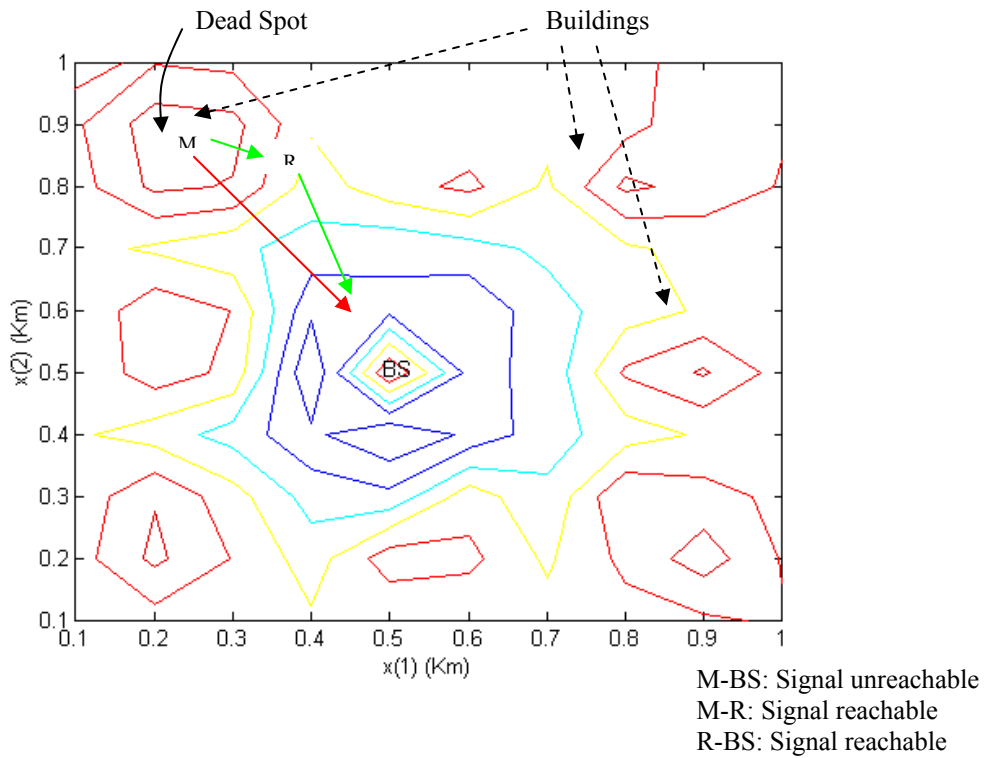
Figure 2. 3 Increasing coverage by eliminating dead spots

- Scenario V – Increase coverage

Mobile lying outside the cell needs to connect to BS.
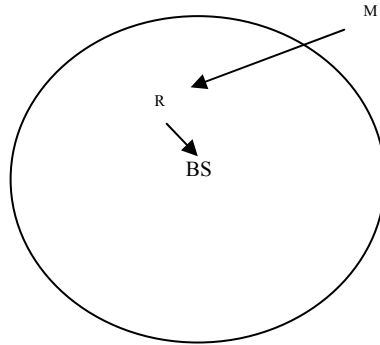


Figure 2. 4 Increasing coverage outside the cell

The coverage of a BS can be increased without increasing base station transmit power by using relays. A benefit of this approach is that inter-cell interference is reduced by keeping the BS transmit power low. This scenario is particularly useful in sparse rural areas, where relays are likely to be present. Thus instead of installing another expensive base station, relays can increase coverage.

## 2.2 Challenges of Relaying

Introduction of relays brings benefits but introduces challenges. This is due to the fact that interference is a big issue in interference limited CDMA systems. Normally the mobiles are in perfect power control with the base station, with mobiles near the base station using less power as compared to far away mobiles. With the introduction of relays, however, this balance can get affected. Any addition to capacity due to relays can increase the interference in the cell causing a reduction in capacity as well.

2.2.1 Interference due to relaying

Interference in the cell is composed of inter-cell interference, intra-cell interference and noise. Adding relays in the cell increases intra-cell interference. The Signal to Interference Ratio (SIR) is defined as the ratio of signal power, S, to the total interference power, I.

$$SIR = \frac{S}{\sum_{int\,erferers} I} \qquad\qquad 2.\,1$$

In the case of FDD-CDMA, interference consists of inter-cell interference and very little intra-cell interference as uplink and downlink use different frequencies. This scenario changes with TDD-CDMA, as the same frequency region is being utilized in downlink and uplink. The addition of relays creates extra interference in the cell as the power control may not be perfect for relay based calls. The base station (BS) tells the relay to transmit at a certain power to prevent the near-far effect. Now the relay to BS will be in perfect power control whereas the mobile to relay part of the connection may not be, unless the relay is capable of transmitting power control information. *We will study how much and how widespread relay-induced interference will be. This will enable us to properly design a CDMA cell. This also helps in locating regions where outages will most likely occur.*

2.2.2 Assigning relays to calls

Relays can be assigned to overcome dead spots or can be assigned to preserve battery power. But these objectives may contribute to a new source of interference in

the cell. Also, how a mobile discovers the presence of a relay – whether a relay broadcasts its presence - may add to overall interference in the cell.

In order for relays to be effective, there may be a need to change relays as the mobile moves. This handoff will depend on who allocates the relays. If the allocation is centralized, then hand off can be decided by the base station. However, if relays are to co-ordinate amongst themselves, then this again could add to interference in the cell.

2.2.3 Convincing users to participate as relays

A mobile that behaves as a relay experiences significant battery drain because it forwards calls for other users. Why would any user allow her mobile to participate as a relay? We believe that providing a financial incentive in the form of cash or free calling would persuade users to participate. Note that users in cars or in offices may have their phones attached to a charger and could thus serve as relays with no battery drain while still earning money or free minutes. One constraint we see, however, is that the payment needs to be tied in with the actual cost of forwarding, and mobiles whose calls are being forwarded need to pay more for the service. We study this pricing question later in this thesis.

## 3. Approaches for implementing relays

As mentioned previously, relays can be implemented either using ad hoc network technology or cellular technology itself. If non-cellular technology is used for relays, such as 802.11 wireless LANs, then the issue is more of setting up the handshake protocol and transferring calls between cellular and non-cellular technology. The problems and solutions from non-cellular technology are inherited while few changes need to be made for cellular technology. Some of the main challenges in implementation will include call setup and takedown, compensating ad hoc nodes for forwarding calls, and call handoff between cells and ad hoc relays. This is the option provided by RadioFrame Networks [RadioFrame Networks, 2003] in their RadioFrame Units (RFU). This RFU is a small capacity Base Transceiver Station (BTS) that has been modified to be IP literate. RFU contains interfaces for cellular systems called RadioBlade transceivers (RB). These transceivers are replaceable and can service GSM, CDMA, iDEN technologies through remote software downloads. The ad hoc interface in RFU is provided by WLAN – integrated RadioFrame Access Point (iRAP) and it supports 802.11 b/g radio frequency channel in the 2.4 GHz band, or 802.11 a/h radio frequency channel in the 5 GHz band. RFU also provides Base Band Blade (BBB) that provides start-up and authentication, BTS radio control, Quality of Service, IP tunneling, encryption, and management functions. Again it is to be noted that all radios are 'removable' thereby providing a flexible and modular solution for any future enhancements. RadioFrame Networks [RadioFrame Networks,

2003] provides different solutions for indoor coverage and hot-spot coverage in the form of B-series and S-series systems.

On the other hand, if relays are cellular based, then, depending on how they are implemented, we get different benefits and problems. Thus, relays may use the same frequency spectrum or a different one, and relays may either be other mobiles or they may be fixed special devices. If they use a different frequency spectrum then we can overcome interference issues. However parallel coverage of one more frequency is needed, making it a costly solution. Assuming that they use the same frequency, relays can be stationary special devices or mobile.

- If relays are stationary, the issue is then of planning ahead as to where they need to be placed. They are not as expensive as base stations, but due to immobility, they can prove costly if new buildings change the link budget calculations.

- If mobiles act as relays then the issue here is how they will affect interference and how carriers will solve security issues. Furthermore, there is the question of how relays can be advertised in the system for use. If relays are stationary special devices, the BS knows about their location and propagation environment. Thus, it can appropriately allocate relays to calls. However, if the relays are mobile, how do they get allocated? One way is that potential relays broadcast their presence to other mobiles. Thus any mobile in need of a relay can know the location of the relay. Another way is for each mobile to inform the BS about its location and propagation environment thus letting the BS make the allocation.

We consider the scenario of mobiles as relays and explore various questions associated with them:

1. How many relays are needed for a call?

2. How are relays allocated?

3. How are relays advertised in the system?

4. Are relays under power control?

5. How do relays affect interference and thus capacity in the system?

## 3.1 How many relays can be used for a call?

Since voice has stringent delay limit requirements, we can not have many relays between mobile to base station. A GSM frame length is 20ms and if one frame is required to forward a call, then we have added 20ms delay to the voice call. If we add two frames (i.e., two relays) then a 40ms delay occurs, which is unacceptable. So we choose one relay between mobile and base station.

## 3.2 How are relays allocated?

A fundamental problem we face in using relays within cellular networks is assigning relays to mobiles that need relays to connect to the base station. At a high level the problems can be summarized as:

- Assigning relays to mobiles: The criteria by which relays are assigned to mobiles can be diverse depending on the overall goals of the system. Thus, if the main goal is reducing power consumption cell-wide, relays should be selected based on the

propagation environment. On the other hand, if increasing coverage is the main goal (e.g., extending coverage to dead spots), then relays ought to be assigned in a way that gives priority to mobiles located in dead spots and other poor coverage areas. A related question here is who performs the actual assignment. One approach is to have the base station do all the work but requiring all mobiles to report details of the propagation environment such as interference, path loss to other mobiles, etc. Another approach may be to have a mobile-initiated procedure where mobiles select their own relays based on local monitoring of signal quality for the relays. In any case, any selected approach will need to be carefully evaluated using metrics such as signaling and computational overhead, quality of assignment, and complexity of the algorithm.

- Handoff among relays in a cell: If a call lasts a long time, the mobile to relay link may become sub-optimal (e.g., if the relay moves far away). In such a situation, we need to assign a new relay to the ongoing call and appropriately handoff the call without losing any data.

- Inter-cell handoff: When a mobile using a relay moves out of the cell and into another, we need to handoff the call from the relay to a new relay in the new cell or to the base station in the new cell. This requires additional overhead and participation by the two base stations as well as relay(s) if the relays are not under centralized control. Under centralized control, however, it is the new base station's responsibility to find a new relay for the mobile depending upon the availability and conditions in the new cell.

**3.3 How are relays advertised in the system?**

How can mobiles be made aware of the relays? This kind of scenario exists in ad hoc environments and they have only one option, to advertise themselves on the air. Each mobile floods other mobiles with the required information. However, in our cellular based scenario we have the advantage of a centralized base station. A base station is aware of each mobile's location and power requirements and so is the best option in scenarios where the base station is choosing the relay for a mobile. However, in scenarios where the base station is not able to hear a mobile and the mobile can only hear a relay, then either relay or mobile can initiate the handshake. If in such a situation relays advertise themselves, they consume a part of the bandwidth and create additional interference.

**3.4 Are relays under power control?**

In highly interference limited controlled environments any new addition of interference is a big issue. Relays being mobiles will be under power control. However the mobile-relay connection may or may not be under power control. If the mobile and relay can coordinate the power level of the mobile-relay connection, then there is no problem. However, if the mobile-relay connection is not under power control, it may contribute to additional interference. Another consideration will be how many frames are required to achieve power control. If more than one frame is required then we can not satisfy the delay limits for voice calls.

**3.5 How do relays affect interference and thus capacity in the system?**

If relays do add interference in the system, then we have to explore how much loss of capacity can occur.

In addition to these questions, it is helpful to understand how much hardware and software would be needed to implement a relay in a mobile. As a relay can transmit to either the mobile and base station, it should be equipped with two transmitters as can be seen in Figure 3.1. The receiver can be limited to only one at the relay since at any time it will be receiving only one call. Also, in a realistic scenario, the choice of relay should be made on a per frame basis. This requires that the relay should have the capability of buffering at least one frame. Whether this frame will be transmitted as such will depend upon the policies for power control and security. So if a frame needs to be demodulated and decoded then the relay needs the hardware in duplicate. The amount of intelligence needed at the relay also depends upon the type of calls to be forwarded. In voice services, the complexity is less since there is less variation in data rates. However, this is not the case in data based networks given a variety of modulation schemes and data rates.
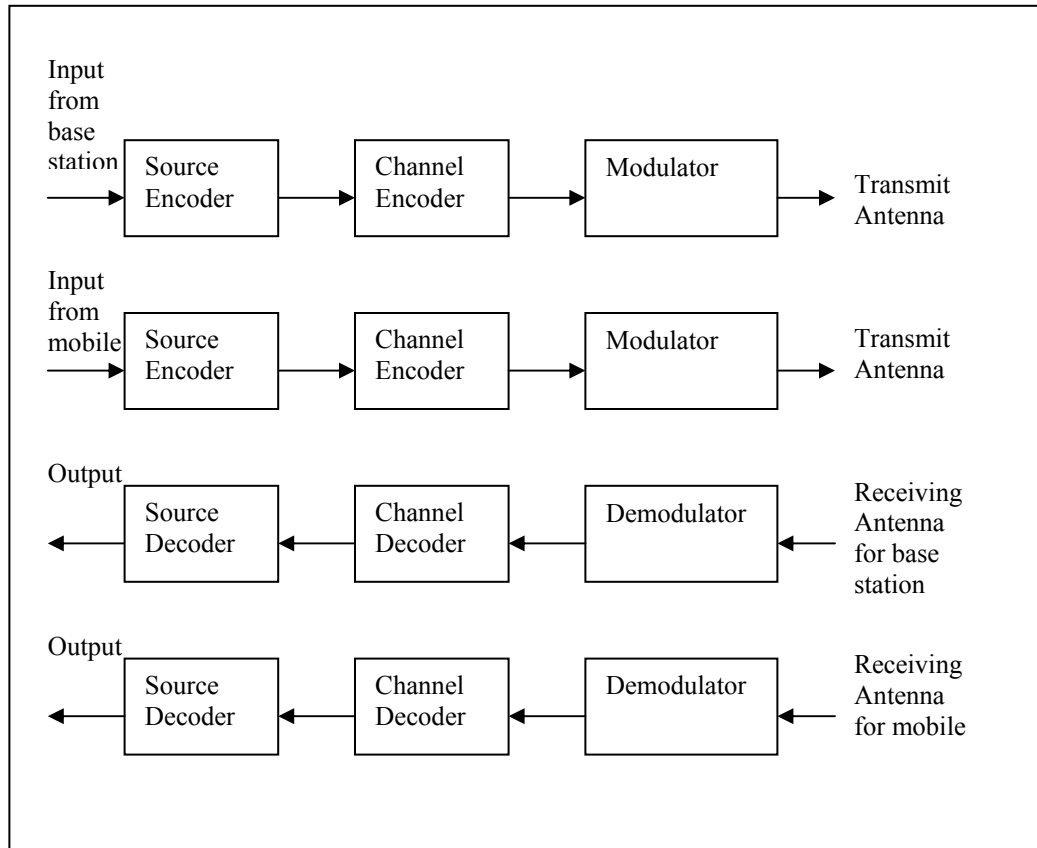
Figure 3. 1 Illustration of hardware requirements for a relay

We also note that if the relay based cellular system is centralized then most of the software complexity resides at the base station. However, if the relay based cellular system is decentralized, then the relay will bear the additional software complexity.

## 4. Energy consumption

Energy consumption is one of the key constraints in the design of any mobile device because these devices typically run on batteries. While battery technology has continued to improve over the years, the demands placed on them have grown at a faster rate. Thus, just a few years ago, cell phones used simple LCD screens that were far more energy efficient than the color screens (and the cameras) that are available in cell phones today. Indeed, we believe that mobile devices will continue to grow more feature rich in the future, thus placing greater demands on the battery. In other words, no matter by how much the battery capacity increases, the mobile devices will evolve to require even more energy.

In order to rein in this unquenchable demand for energy, hardware designers have started using energy-efficient designs for the hardware and software contained in mobile devices. Some examples of these techniques include:

- Dynamic Power Management: The idea is to shut down components of the device that are idle. For example, if the screen is idle, the intensity is reduced after some time and then it is turned off completely. Likewise, in a phone equipped with a camera, the camera is only powered on when it is in use. Energy is saved in the radio by powering it on and off appropriately. As an example, consider the pager, which is very energy efficient (one charge lasts two or more weeks). The pager's radio is off most of the time. It wakes up periodically to listen to the base station. The base station announces the IDs of the pagers that have a waiting message. All the other pagers turn off for the duration of the message transmissions. In the context of cell phones, a similar protocol is used. The cell phone monitors the base

station ID it is in and only contacts the base station when it moves to a new cell. Otherwise, the radio is powered off and only wakes up to hear a periodic base station initiated beacon.

- Dynamic Voltage Scaling: Here the clock speed (and hence voltage) of the hardware is changed based on demand and can result in huge savings. The relationship between voltage and clock frequency is given as [Gonzalez, Gordon and Horowitz, 1997],

$$T_g = K \frac{V}{(V - V_{th})^\alpha} \qquad \qquad 4.1$$

where $V$ is the peak voltage, $V_{th}$ the threshold operating voltage, $K$ is a constant, and $T_g$ is the gate delay, which increases with decreasing bandwidth. Here $\alpha$ depends on the technology used for the device; it is 2 for long-channel devices and is 1.3 for sub-micron CMOS devices.

The relationship between voltage, frequency and energy consumption E, is given as [Hadjiyiannis, Chandrakasan and Devdas, 1998],

$$E \propto C \times V^2 \times f \times t \qquad \qquad 4.2$$

where C is the effective switched capacitance, V is the peak voltage swing, f is the switching frequency and t is the length of the transmission. Let us compare energy consumption $E_0$ and $E_1$ for devices having different bandwidths $f_0$ and $f_1$, where

$$f_0 = f_1 / 2 \qquad \qquad 4.3$$

giving

$$t_0 = 2t_1 \qquad \qquad 4.4$$

The cycle time, t, is the sum of delays in the critical path. If we approximate $V_{th}$ to be zero, then $T_g^0$ and $T_g^1$, the gate times, are given as,

$$T_g^0 = \frac{K}{V_0^{\alpha-1}}$$

$$T_g^1 = \frac{K}{V_1^{\alpha-1}}$$

$$T_g^0 = 2T_g^1$$

$$V_0 = \frac{V_1}{2^{1/(\alpha-1)}}$$

4. 5

Thus,

$$E_0 = CV_0^2 f_0 t_0 = C\left(\frac{V_1}{2^{1/(\alpha-1)}}\right)^2 \left(\frac{f_1}{2}\right)(2t_1)$$

$$E_0 = E_1\left(\frac{1}{2^{1/(\alpha-1)}}\right)^2$$

4. 6

If $\alpha$ is 2, then energy consumption is a quarter of the original, when bandwidth is reduced by a half.

In addition to these techniques, newer cellular systems that use CDMA are more energy efficient than older systems based on TDMA or FDMA. Further savings are attained in these systems by combining most hardware functions into fewer chips. For example, the Intel PXA800F chip [Diasemi, 2003] combines power and audio management in one chip. Concurrently, batteries are being designed to fully exploit dynamic voltage scaling, for example the single cell lithium ion battery such as the Maxim 8890 [Maxim, 2001].

While all of these techniques to reduce energy consumption are beneficial, we believe that energy will always be a constraint and thus additional techniques to

reduce energy consumption are necessary. In this context, we believe that using relays in cellular networks will result in significant energy savings in the cell as a whole and at individual mobiles who have little remaining battery power. In this chapter we explore this assertion in detail.

## 4.1 The problem of assigning relays

Consider the simple example in Figure 4.1 (a) where we have a node, M, that wishes to place a call via a relay. All the cars shown are potential relays. The question is, which of these available relays is the "best" choice? Figure 4.1 (b) illustrates the path loss models from the location of the node wishing to use a relay. Thus, car 'i' is a poor choice because the mobile and 'i' are Non Line of Sight (NLOS) to each other, which means that the mobile will need to expend a great deal of energy getting to the relay. On the other hand, 'a' is a good choice because it is in Line of Sight (LOS) and thus requires the least energy. A second consideration here is the length of the call. Thus, if the call is of a long duration then the relay 'a' may become a poor choice when it becomes NLOS (e.g., if it moves to location 'i'). This means that either the call will need to be handed off between relays or the initial selection of the relay should be such that the probability of such a situation occuring is minimized. The mechanics of handing off calls between relays has been discussed previously in chapter 3.
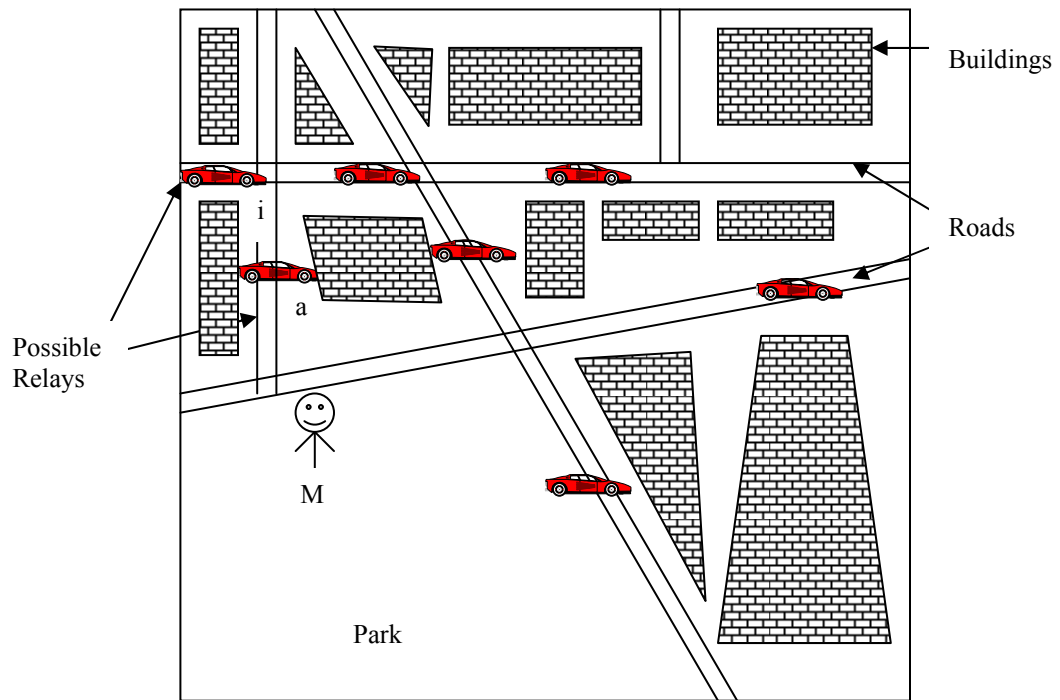
(a) Mobile, M, needs to set up call via a relay
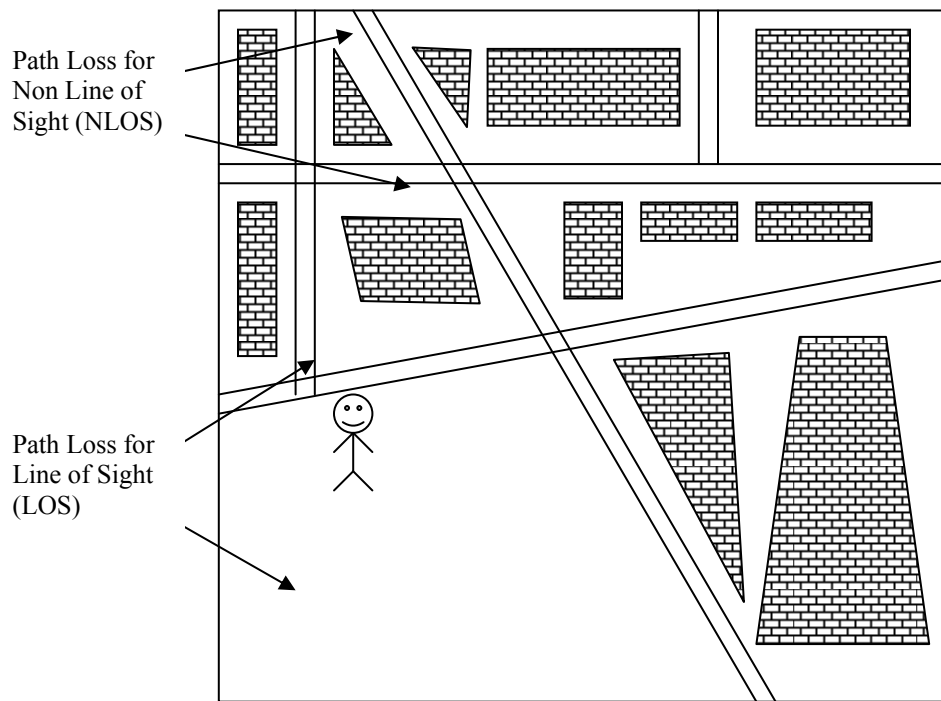


(b) Path Loss models from mobile point of view

Figure 4. 1 Example scenario for relay assignment

Consider now a more realistic problem where we have several mobiles in the cell, some of which can serve as relays and others that need relays. How do we now assign relays to mobiles taking into consideration the path loss models between each mobile and each potential relay as well as relative mobility of the mobiles and relays? What is the criterion to be used in making such as assignment? Since energy is our main concern here, we will use *minimization of total energy in the cell* as the primary criterion for allocating relays. Thus, in the case of a single mobile, if the energy in placing a direct call to the base station is greater than the total energy consumed in placing the call via a relay (i.e., the energy on the relay – base station hop plus mobile – relay hop) then the relay will be used. Similarly in the case of multiple mobiles, relay assignments that reduce the total energy cell–wide are the only feasible assignments and among these, the assignment resulting in the lowest energy is the best choice.

Given m mobiles wanting to place calls and k relays (say k >= m), we have k(k-1)(k-2) …. (k-m) possible assignments. If k and m are small, we can exhaustively explore the state but in any realistic system, this may prove impractical. Thus, we need more efficient algorithms to determine low-energy assignments. In this section, we explore three such assignment algorithms:

1. Greedy Approach: When a mobile requests a carrier, we only look at the idle mobiles (that agree to serve as relays) and select the first one that minimizes the total energy for that call. This method is not optimal because the assignment does not look at already assigned relays. For example, say relay R1 is being used by mobile M1 and mobile M2 now requests a relay. Our method assigns relay R2 to

M2 since R2 minimizes the energy among the set of idle relays. It may be the case that by reassigning R1 to M2 and assigning R2 to M1, more energy can be saved, but the greedy approach does not do this.

2. Intelligent Relay Approach: We try to use relays as listening devices rather than broadcasting devices. Figure 4.2 shows the attenuation of the transmission from various mobiles towards the base station. The curves are different for different mobiles because they may have different propagation environments. The mobiles further away from the base station start wth higher power to overcome near far effect. We can think of relays as being located between the mobile and the base station. Thus, relays will also receive transmissions from the mobiles and the assignment is now made based on the following algorithm. For a given relay and for a given mobile's location, assume that we have a value which denotes the expected received power from that location in the absence of obstacles such as buildings. If the received power from the mobile at that location is lower than this value, the relay is a good candidate to carry that mobile's call. Relay being nearer to the mobile can hear the mobile, whereas base station can not hear the mobile. The relay thus can save the mobile from attenuating the signal too quickly and guarantees that the signal will reach the base station. The mobile's signal has to reach only the relay instead of the base station, thus saving its battery power. An added benefit is that since the relay does not broadcast its presence, it avoids adding more interference to the cell.

Attenuated Signal Power on the y-axis
Distance from Base Station on x-axis
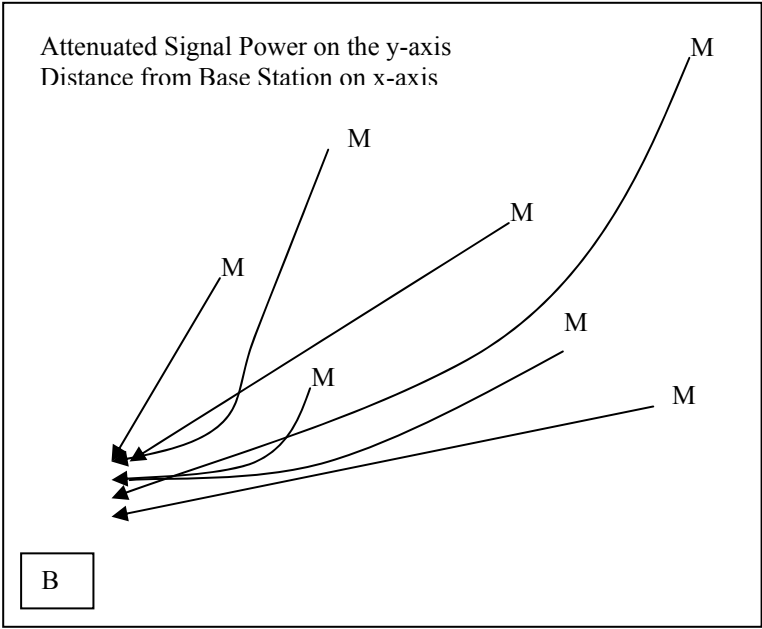
Figure 4. 2 Illustration of attenuation of signal strength
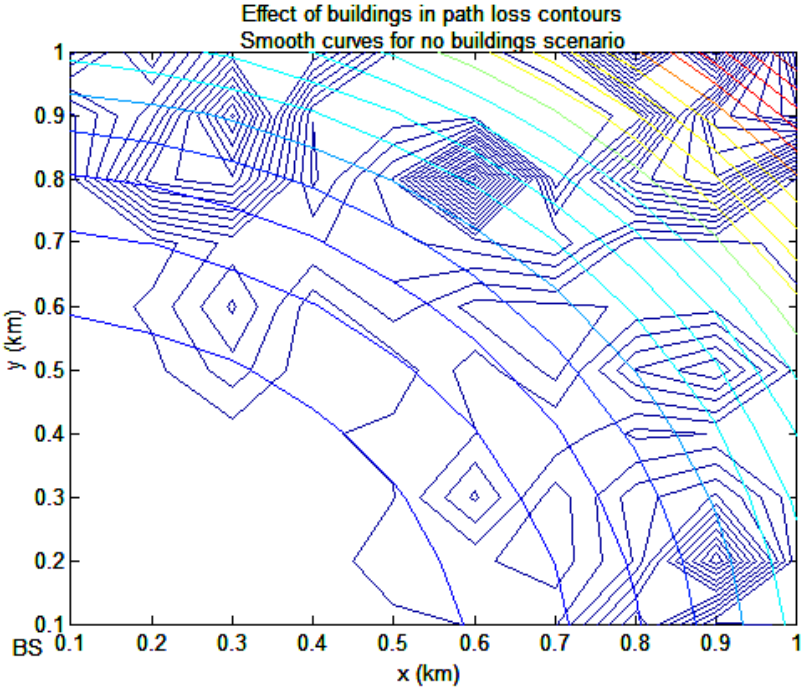


Figure 4. 3 Path loss contours with and without buildings

3. Optimum Approach - Assignment Problem: When a relay is selected to carry a call, it should lower the energy consumption of the cell. In the Greedy Approach it may happen that the choice of a particular relay for a mobile may force a costlier choice for another mobile. This is because the Greedy Approach is a local minimization approach. In order to minimize power requirements globally, each mobile should be allocated a fresh relay every time a request is made. There is a class of linear programming problems called transportation problems, where the goal is to minimize the cost of shipping a commodity from source to destination. The assignment model is a special case of the transportation problem in which a worker is assigned a job appropriate to his skill level. In this way the cost of employing workers is minimized. In our case we have to minimize total energy requirements. The Hungarian method is used to solve such assignment problems [Carpaneto and Toth, 1980] [Borlin, 1999]. We create a square matrix of calling mobiles and idle mobiles, which are suitable as carriers. If the matrix is not square, we can create dummy rows or columns.

Carrier

|        | | | |
|--------|--------|--------|--------|
| $C_{11}$ | $C_{12}$ | …. | $C_{1n}$ |
| $C_{21}$ | $C_{22}$ | …. | $C_{2n}$ |
| …. | …. | …. | …. |
| $C_{n1}$ | $C_{n2}$ | …. | $C_{nn}$ |

Mobile (label to the left of the matrix)

Figure 4. 4 Calculation of cost matrix

The element $C_{ij}$ of the matrix represents the cost (power) for mobile i when relay j is chosen. When i = j, it means the mobile is the carrier of itself, i.e., directly connected to the base station. If a mobile is not idle the cost of relaying is infinity, since it is not available.

As an example, let the mobile to relay power be shown in this matrix,

| Power | R1 | R2 | R3 | Row Min |
|-------|----|----|----|---------|
| M1 | 15 | 20 | 18 | 15 |
| M2 | 12 | 7 | 22 | 7 |
| M3 | 8 | 11 | 21 | 8 |

Subtract from each row, the minimum row value.

| Power | R1 | R2 | R3 |
|-------|----|----|----|
| M1 | 0 | 5 | 3 |
| M2 | 5 | 0 | 15 |
| M3 | 0 | 3 | 13 |
| Column Min | 0 | 0 | 3 |

Subtract from each column, the minimum column value.

| Power | R1 | R2 | R3 |
|-------|----|----|----|
| M1 | 0 | 5 | **0** |
| M2 | 5 | **0** | 12 |
| M3 | **0** | 3 | 10 |

The underlined zeros in each row give the assignments. M2 mobile will take R2 and M3 will take R1. For M1 mobile only relay R3 is left. The total system power here is (18 + 7 + 8), which is the optimum value.

4.1.1 Discrete Event Simulation

Stochastic simulation involves measuring performance of a system for some input. This involves generating samples of input processes and generating input/output relations in terms of events and states. When *events* we are interested in occur at discrete times, the simulation is a discrete event simulation. The *state* of the system is defined by variables. The future state of a system is dependent only on the current state of the system and not on any past state. The performance of the system is given by statistical variables.

As an example to illustrate the states and events, the call model of a mobile is illustrated in Figure 4.5. The model has two states, 'Idle' and 'Calling'. The change of state is triggered by an event. Event 'call starts' transitions the 'Idle' state to the 'Calling' state. Similarly event 'call ends' transitions the 'Calling' state to the 'Idle' state.
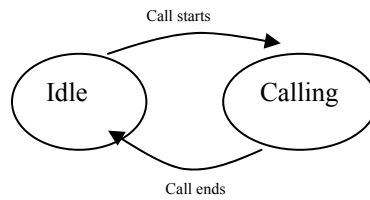
Figure 4. 5 Illustration of states and transitions

The state of the system changes only when events happen, as when a mobile starts calling. The events are ordered in an event list and they indicate when the next event will occur. As an example, when a call starts, the call termination event is put in the event list. If another call starts before the first call's termination, then it will be handled first. At each event, the state and statistical variables are updated.

In our simulator, events are generated by call initiation and termination and by the movement of mobiles. In addition, when a mobile is chosen as a relay, we can have another state, called "Carrying". We can see various states and transitions in Figure 4.6. The mobile is in an 'Idle' state if the mobile is not calling. When a call starts the new state of the mobile is 'Calling' to signify that the mobile is in direct contact with the base station. If a relay is found then the mobile changes its state to 'Stretched' and the relay changes its state to 'Carrying'.
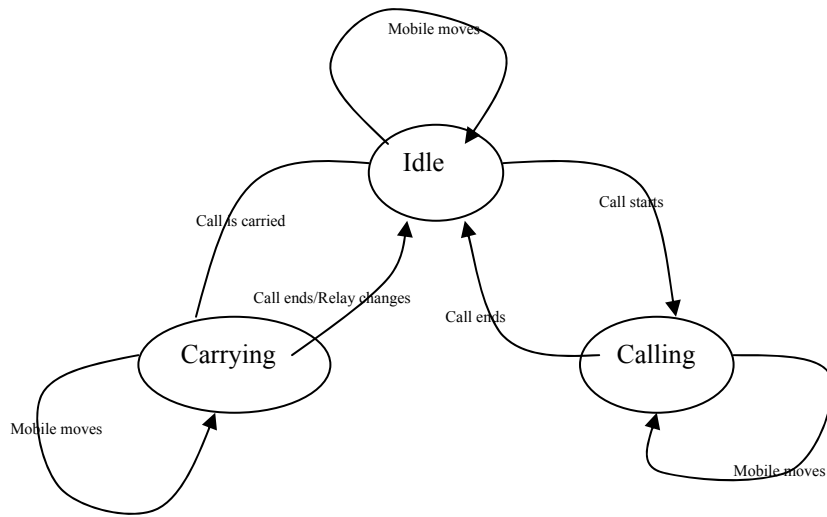
Figure 4. 6 Finite State Automata to choose mobiles as relays

## 4.1.2 Comparison of the energy savings by the three approaches

In order to compare the three approaches of assigning relays, we conducted simulations in Matlab using a simplified cellular representation. We use a 1 Km x 1 Km cell with the base station located in the center for Figures 4.7 and 4.8 and the base station located at a corner for Figure 4.9 (to simplify computation). The reverse link frequency is set at 1920 MHz. In order to model the path loss of an urban environment, we use the CCIR formula.

The other simulation parameters are summarized in Table 4.1 below.

| Variable | Value |
|---|---|
| Call rate | 1, 2 calls/hour |
| Call duration | 120 sec |
| Base station | Center, Corner |
| Number of Nodes | 5 to 50 |
| Idle to Calling Ratio | 0.5 to 1 |

Table 4. 1 Table showing variables used in simulation

Figure 4.7 plots the total energy consumed during a 10,000 second run for the Greedy approach and for the direct to base station approach (i.e., when no relays are used). The x-axis plots the idle/calling ratio of mobiles. Thus, going from left to right, we see an increase in the relative number of idle nodes allowing more of the calls to be relayed. At small values of idle/calling nodes, there is no difference between the Greedy approach and direct to base station approach. This is because few calls can be relayed. However, as more nodes become idle, the total energy cost for the Greedy approach is dramatically lower than the direct to base station cost. In fact, the Greedy approach shows 4x improvement over placing calls directly to the base station. Indeed, energy savings of up to 3.56 dB to 6.6 dB are possible as the number of idle mobiles increases.

Figure 4.8 plots the same parameters for the Intelligent relay approach. However, here the simulation time is 2,500 seconds (due to the computation cost of this approach we only ran the simulations for shorter periods). Again, as in the Greedy case we observe increasing benefits from relaying of up to almost 4x. However, unlike

the Greedy approach, the computational cost of this algorithm is very high. This is because the Intelligent relay approach requires an approximation of approaching signals at the relay location. We know that mobile signals originating at higher building density areas will require higher power to reach the base station. So at a relay location, if the power calculated with buildings between the mobile and the relay is higher than the expected power calculated with no buildings, then the relay chooses the mobile to be its carrier. The base station is taken to be at a corner in this simulation to simplify calculations.

Finally, Figure 4.9 plots the data for when we use the optimal approach (using the Hungarian approach) and run the simulation for 2,500 seconds. The energy savings here are 5-6x, as expected, more than in the Greedy case. As in the case of the Intelligent relay approach, the computational overhead here is also very high.

4.1.3 Computational overhead of the three approaches

| Approach | Complexity | Cellular Space |
|---|---|---|
| Greedy | $\Theta(n)$ | Half cell |
| Intelligent Relay | $\Theta(n^2)$ | Quadrant of cell |
| Optimum | $\Theta(n^2)$ | Full cell |

Table 4. 2 Table showing complexity of the approaches

Table 4.2 lists the computational complexity of the three approaches. In the greedy approach, the algorithm runs in linear time (in the number of relays) because all we do is find the maximum energy savings among the available n relays. Similarly, the intelligent approach in theory should be linear in n, as theoretically the relay can act as a receiver and thus compare just the arriving signals. However, here in the simulation the complexity is more as we approximate the arriving signals. The optimal approach examines the n x n array of costs and thus is proportional to $n^2$. It is to be noted that in realistic scenario the algorithms will be run on a per frame basis and the upper bound for the algorithm will be decided by the amount of time mobile is willing to take for searching.

4.1.4 Summary

The relay assignment problem is key to the idea of using relays in the cellular environment. In this section we presented three relay assignment algorithms and evaluated their performance in Matlab simulations. The results show that of the three relay assignment algorithms, the Greedy approach appears to deliver good energy savings at a small computational cost. Therefore, for the remainder of this chapter, we use the Greedy approach to study the question of using relays in more detail.
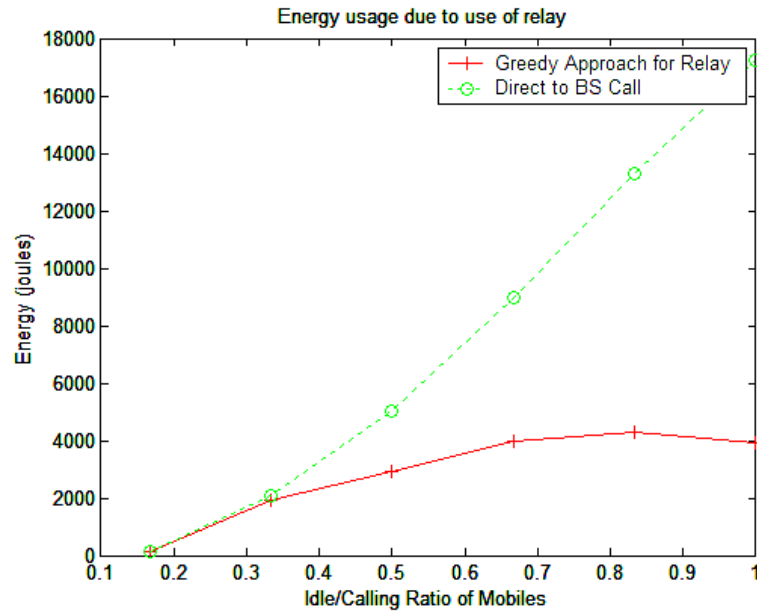
Figure 4. 7 Energy comparison between direct to BS and greedy approach
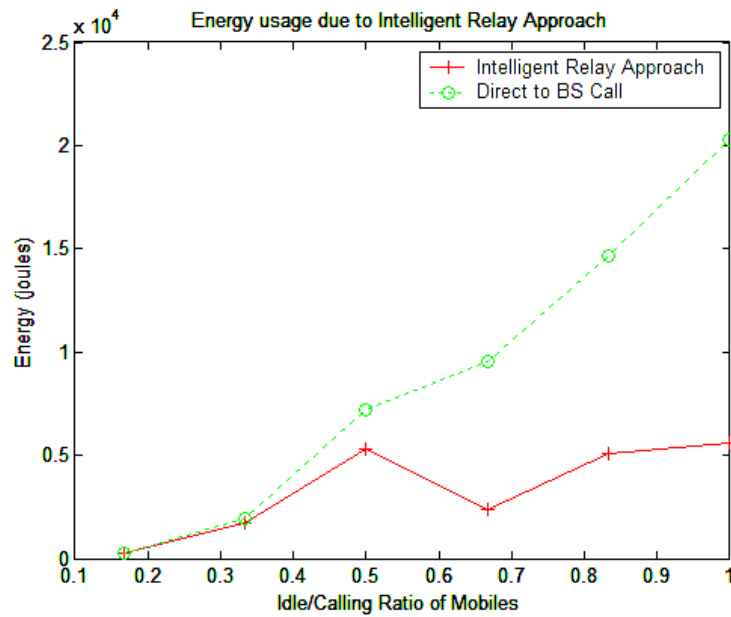


Figure 4. 8 Energy comparison between direct to BS and intelligent relay approach
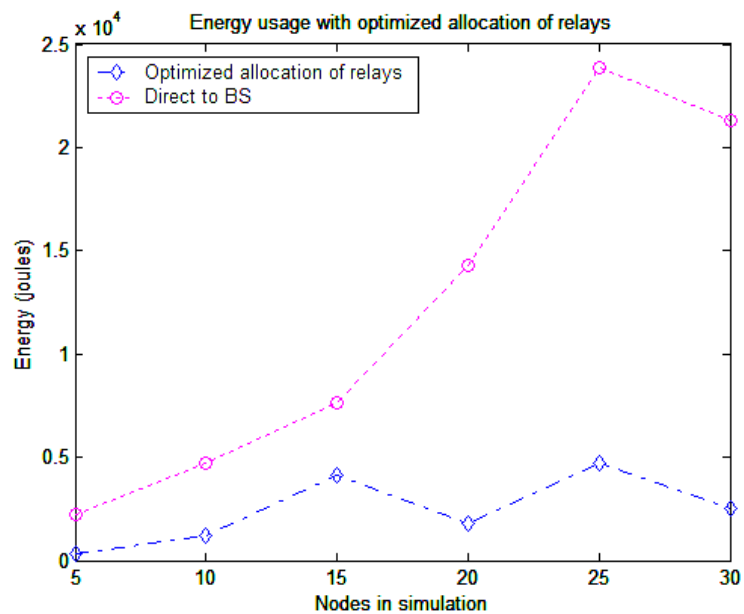
Figure 4. 9 Energy comparison between direct to BS and optimized approach

## 4.2 Overview of the Simulation Design

In order to quantify and evaluate the energy savings obtained by using relays, we rely on detailed simulations. The goal in designing these simulations is to ensure that they are complete and that they enable us to explore the problem in detail. By complete we mean that the simulations should allow us to examine a variety of cellular topologies as well as mobile and relay placements within the cell. Other variables include the number of mobiles, system load (i.e., call rate and call length), and movement patterns of the mobiles.

To document the energy savings possible using relays, we study the Greedy approach described previously using a representative urban topography for the cell. We decided to focus only on an urban topography because, in a rural setting, the propagation environment is simple with the base station and the mobiles being line of sight (LOS) of each other. Thus, relay allocation is trivial – the best relay for a mobile is the one that lies as close to the center as possible of the line connecting the mobile and the base station. On the other hand, in an urban setting (visualize a Manhattan type environment), the base station antenna is below rooftops and mobiles are frequently non line of sight (NLOS) of the base station. Indeed, these microcells are the most challenging environments from a coverage standpoint due to the complex propagation environment created by buildings. Thus, the propagation environment changes rapidly as a mobile travels, and, in the case of connections using relays, the change is twice as fast because both the mobile as well as the relay move. This change may prompt frequent reassignments of relays to mobiles if we are to minimize total energy consumed.

Allocating relays to mobiles is based on minimization of total energy. Thus, the simulator needs to be able to estimate the energy used by calls placed from any location directly to the base station as well as the energy of the calls being routed by relays that may be located anywhere in the cell. This means that we need expressions that give us the path loss between any two points in the cell (for the mobile – relay link) as well as between any point in the cell and the base station. Thus, one component of the simulator design was implementing algorithms that allow the computation of these path loss expressions in the course of the simulation.

**4.3 Detailed Simulator Design**

At the time this research got underway, there were no free cellular simulators available (there still are none that provide the level of detail we require). Thus, we designed and implemented the simulator in the first two years of this research. This implementation was done jointly with Shashidhar Lakkavalli. We use Java for implementation because of its object-oriented nature as well as the ease of visualizing the cell being simulated.

The cell is represented as a rectangular region that is divided into a grid. The resolution of the grid is 20 meters. Buildings are introduced into the cell interactively where the user identifies the set of grids that represent the building. Roads are also identified via the same user interface. The user can place the base station anywhere in the rectangle including the corners. Note that this models real cellular environments where base station placement is constrained by zoning and other considerations. The user can also specify the behavior of mobile nodes including a distribution for speed of travel, call length distribution, and call frequency.

4.3.1 Propagation Environment

To determine the path loss between two points in Manhattan style urban scenarios, we use a mathematical method for path loss calculations [Berg, 1995]. This method is suitable for buildings that are considerably taller than the height of the antennas. The method is recursive and is well suited for ray or path tracing techniques, which mimic paths taken by radio beams. This model accommodates the behavior of

the radio beams as they bend around corners and travel across different streets. It also includes the dual slope pattern of path loss when varied across distances.

The basic equation for free space path loss is written as,

$$L_{dB}^{(n)} = 20.\log\left(\frac{4.\pi.d_n}{\lambda}\right)$$
4. 7

where $\lambda$ is the wavelength, n is the number of sections between the transmitter antenna and receiving antenna and $d_n$ is the illusory distance between the two antennas and is calculated as [Berg, 1995],

$$k_j = k_{j-1} + d_{j-1}q_{j-1}$$
$$d_j = k_j s_{j-1} + d_{j-1}$$
4. 8

with the initial values, $k_0 = 1$ and $d_0 = 0$ and s is the actual distance.

The angle dependence is given as [Berg, 1995],

$$q_j(\theta_j) = \left(\theta_j \bullet \frac{q_{90}}{90}\right)^v$$
4. 9

The parameter values are $q_{90} = 0.5$ and v, determining the shape function is 1.5. A similar path loss equation is also applicable for dual slope behavior. Breakpoint distance is taken as 300 meters.

This can be illustrated by looking at the direction the beam takes as shown in Figure 4.10. Receiver R1 lies in direct line of sight of transmitter Tx. The receivers not directly in line of the transmitter, such as R2 and R3, get the rays diffracted at the street corners and the rays seem to originate from the street corners. Those receivers lying at different angles, $\theta$, receive a different share of the beam, making the path loss a function of $\theta$. Let us use this example to illustrate the algorithm in more detail. Let

s0 be the actual distance between Tx and the corner, and s1 is the distance between the

corner and the receivers. All receivers are equidistant from the corner.

If $s_0 = 100$ meters and $s_1 = 100$ meters, and $\lambda = 0.33$ meters (900 MHz),

$$\begin{aligned} k_0 &= 1, d_0 = 0 \\ k_1 &= k_0 + d_0 \bullet q_0 = 1 \\ d_1 &= k_1 \bullet s_0 + d_0 = 100 \end{aligned} \qquad 4.\ 10$$

1.  Path loss between Tx and R1 ($\theta = 0^0$, q(0) = 0)

$$\begin{aligned} k_2 &= k_1 + d_1 \bullet q_1(0) = 1 \\ d_2 &= k_2 \bullet s_1 + d_1 = 200 \\ L_{dB} &= 77.63 dB \end{aligned} \qquad 4.\ 11$$

2.  Path loss between Tx and R2 ($\theta = 45^0$, q(45) = 0.125)

$$\begin{aligned} k_2 &= k_1 + d_1 \bullet q_1(45) = 13.5 \\ d_2 &= k_2 \bullet s_1 + d_1 = 1450 \\ L_{dB} &= 98.84 dB \end{aligned} \qquad 4.\ 12$$

3.  Path loss between Tx and R3 ($\theta = 90^0$, q(90) = 0.3535)

$$\begin{aligned} k_2 &= k_1 + d_1 \bullet q_1(90) = 36.35 \\ d_2 &= k_2 \bullet s_1 + d_1 = 3735 \\ L_{dB} &= 103.06 dB \end{aligned} \qquad 4.\ 13$$

If $s_0 = 100$ meters and $s_1 = 100$ meters, and $\lambda = 0.15$ meters (2 GHz),

Path loss between Tx and R1 ($\theta = 0^0$, q(0) = 0), L = 84.48 dB

Path loss between Tx and R2 ($\theta = 45^0$, q(45) = 0.125), L = 101.69 dB

Path loss between Tx and R3 ($\theta = 90^0$, q(90) = 0.3535), L = 109.91 dB

If $s_0 = 20$ meters and $s_1 = 20$ meters, and $\lambda = 0.15$ meters (2 GHz),

Path loss between Tx and R1 ($\theta = 0^0$, q(0) = 0), L = 70.50 dB

Path loss between Tx and R2 ($\theta = 45^0$, q(45) = 0.125), L = 77.55 dB

Path loss between Tx and R3 ($\theta = 90^0$, q(90) = 0.3535), L = 83.63 dB
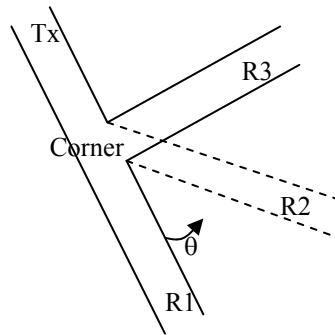


Figure 4. 10 Receivers at different angles from the transmitter [Berg, 1995]

In order to implement this algorithm in our simulator (at the level of grids), we use a concept from game programming, which is an approach to develop computer games.  Usually, in a simulated map of the game an object has to move around to reach a goal. The object has to avoid obstacles and find a path intelligently. Thus, in game programming, the shortest distance is often calculated. Since we need to find the best path between the mobile and base station, we need to find the shortest path.  Due to its simplicity, A* is the heuristic used in game programming, rather than breadth first or depth first search. The physical space can be divided into square shaped cells or grids. The shortest distance between the start cell and destination cell is calculated by starting from the start grid cell. The cost of traversing each grid cell is added until the destination is reached. Consider the example in Figure 4-11 [Macgill, 1999]. The dark grey grid cell (with 2.0 written on it) is the start grid and the light grey grid cell is

the destination. The black grid cell represents infinite cost. The cost of a white grid cell is one. As can be seen the cost of each grid cell starting from the start grid cell is added in each adjacent grid cell. As soon as the destination is reached, the heuristic traverses back as represented by the shaded path. The shaded path avoids the black obstacles and traverses the corners much as a ray of light would do.

The pseudocode of A* algorithm is given below from [Bourg and Seemann, 2003]. The open list is a list to keep track of which grid cells need to be searched. The closed list refers to grid cells that are already checked and no longer need to be examined. The cost here refers to the path loss cost.

```
add the starting node to the open list
while the open list is not empty
{
        curent node = node from open list with the lowest cost
        if current node = goal then
                path complete
        else
                move current node to the closed list
                examine each node adjacent to the current node
                for each adjacent node
                        if it isn't on the open list
                                and isn't on the closed list
                                        and it isn't an obstacle then
                                                move it to open list and calculate cost
}
```
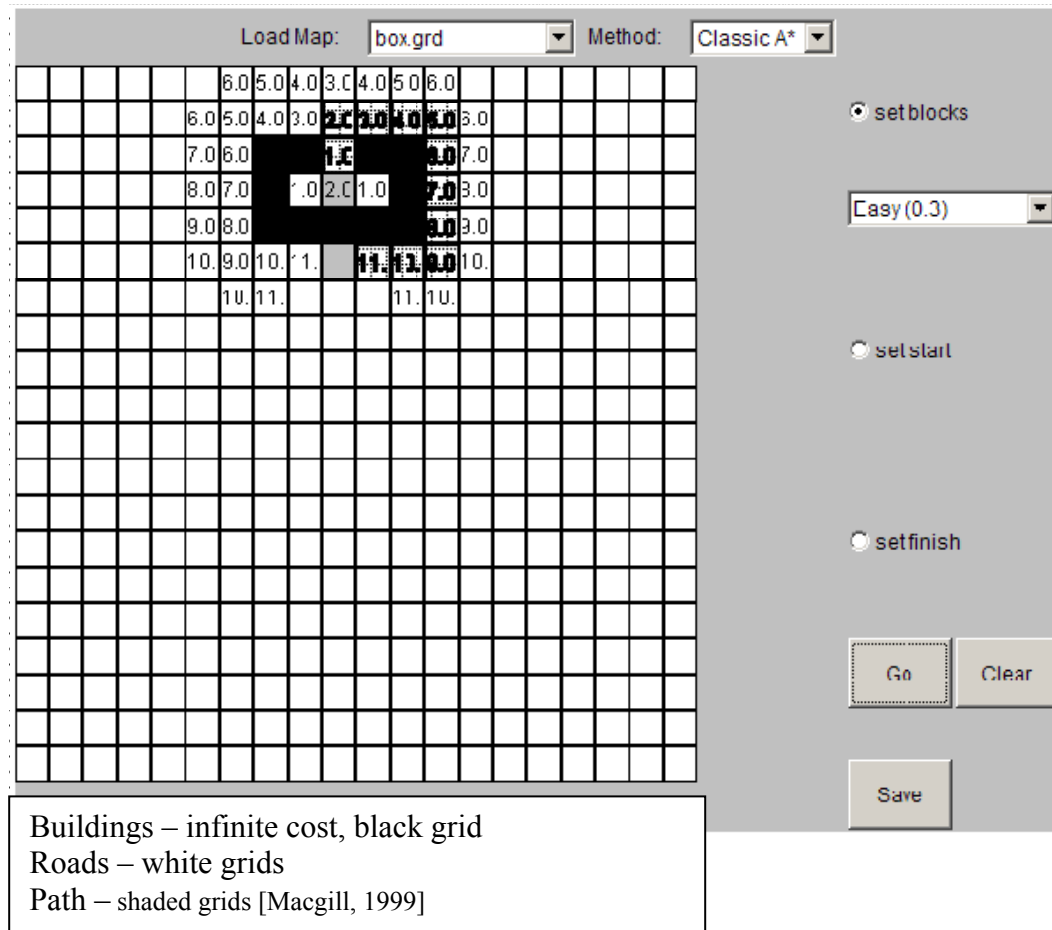
Figure 4. 11 Illustration of A* algorithm

The mobility state model is linked to each crossing of the grid cell, that is, when a mobile crosses a grid cell boundary, the event queue is updated. If the grid cell size is smaller, it means that we are unnecessarily adding to our computation. Also it may mean that we are adding to the ping-pong effect in the handoff between the intermediaries, where the relays are handed to one and again handed back quickly. On the other hand if the grid cell size is larger, the mobile may continue using a relay which is sub-optimal.

4.3.1 Call and Mobility Model

Call arrival is modeled by a Poisson distribution while the call length is modeled by an exponential distribution. A Poisson distribution is described by its mean and variance, $\lambda$, and is given as,

$$f(x) = \lambda^x \frac{e^{-\lambda}}{x!}$$ 
4. 14

Exponential distribution, with mean $a$, is given as,

$$f(x) = \frac{1}{a} e^{-x/a}$$
4. 15

In our case, $\lambda$ and $a$ are in units of seconds. Typical values from published studies for 'a' are 150 seconds [Lam, Cox and Widom, 1997] and for $\lambda$, it depends on the number of calls per hour and the call duration.

4.3.2 Modeling Node Behavior

For each mobile and relay in the cell, the base station maintains a finite state machine indicating the present state of that node. Figure 4.12 shows the state machine for a mobile. When a mobile needs to initiate a call, it moves from the Idle state to the Requested state where the mobile is waiting for the base station to accept the call. At this point, the base station may either find a relay or ask the mobile to place the call directly. These two situations correspond to the state transitioning to either Stretched (i.e., call is relayed) or Direct. If the call is being relayed, there may be a need to change the relay over time. Thus, the self loop in the Stretched state denotes the base

station changing the relay as propagation conditions change in relative node position. It is possible that there may be no available relay so the call can become Direct at some later time. Likewise, a Direct call in progress may become Stretched if a relay can be found.

Node mobility is modeled as a Gaussian distribution. The mobiles moves along the roads until it finds an intersection. At the intersection the node changes direction uniformly randomly and chooses a new velocity, which is 10% of the mean value.
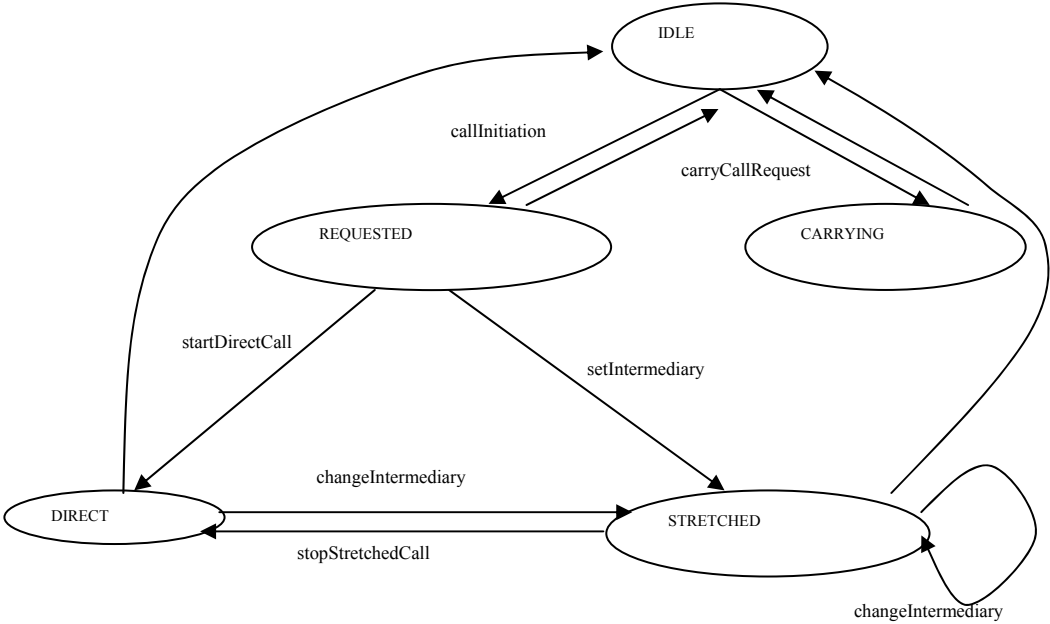


Figure 4. 12 Finite State Automaton of a Mobile
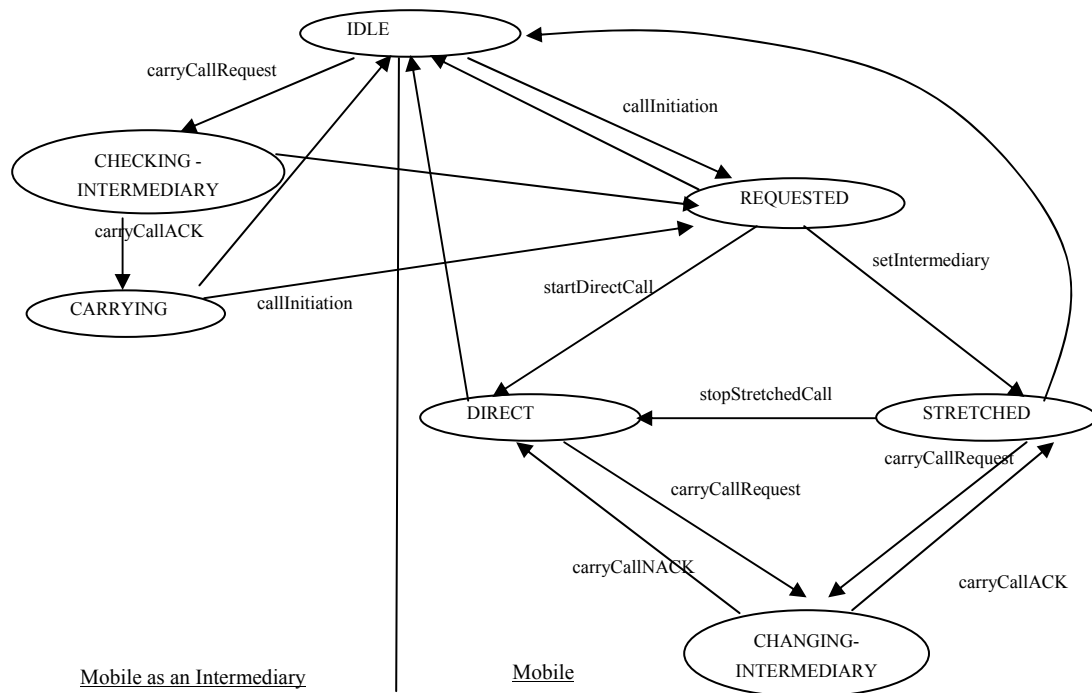
Figure 4. 13 Finite State Automaton of a Mobile maintained in the base station

An idle mobile can serve as a relay for another call. This situation is illustrated in Figure 4.13 where the mobile is contacted by the base station and if the mobile is willing, it becomes a relay. The transition from the Idle to CheckingIntermediary transition happens when the request is received from the base station. After the mobile agrees to become a relay, it responds with a carryCallACK and moves to the Carrying state. If a relay needs to place its own call, it informs the base station via the callInitiation transition which causes the base station to reassign the relay and to set up this new call.

4.3.3 Calculating Energy Consumption

To measure the energy consumption for a given amount of simulation time in the cellular environment we take into account the amount of power spent by the mobile. The power when multiplied by the time of call duration gives the total energy consumed by the mobile. We assume that an idle mobile does not consume energy. In reality, however, mobiles scan the pilot channel and thus need some energy all the time. We do not model this because it accounts for a very small value.
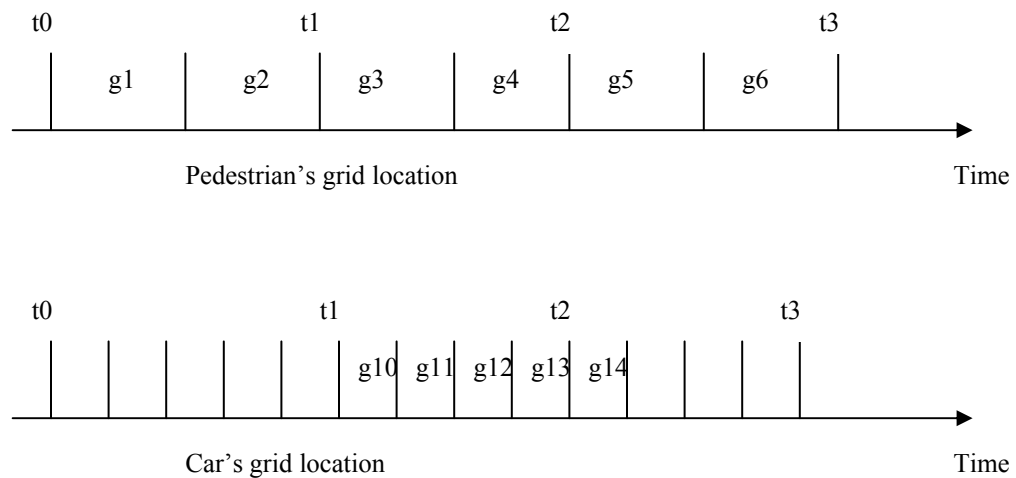


Figure 4. 14 Grid locations over time

Figure 4.1 (a) shows a pedestrian walking down a street while placing a call. The call is initially direct to the base station (t0 – t1) then it is relayed via the car for time t1 – t2 after which it is direct to the base station again until termination t2 – t3. To calculate the total energy consumed for this call, we need the path loss between the mobile and base station for time t0 – t1 and t2 – t3 plus the energy between the

pedestrian and the car for t1 –t2 as well as energy from the car to the base station for this same interval. The path loss changes as the mobiles move and in our simulation we have path loss expressions for all pairs and grid squares. Figure 4.14 shows the grid location of the pedestrian and relay (car) location for the time t0 – t3. To determine the total energy consumed, we calculate the energy for transmitting from grid cells g1 and g2 direct to the base station (transmit power multiplied with time in that grid cell) plus the energy consumed in transmitting between g3 – g10, g3 – g11, and g3 – g12 (since the car moves through three grid locations in the time it takes the pedestrian to move one grid location) plus the energy the relay expends in transmitting to the base station, and so on. The simulator makes these calculations for all ongoing calls and reports the total energy consumed for some given length of time.

**4.4 Simulator Implementation**

The simulator was developed in Java. The main components of this simulator include the event queue, comparison operators for events in the event queue, the state representation, data collection modules, and algorithms that make mobile move in a way that follows the topology of the cell. Each of these major components are described next.

4.4.1 Event Queue

The events are ordered in time in the event queue. As an example, if we add a call initiation event for some mobile X to start at time 101, then the call termination event should come after this time.  Similarly if another mobile Y has its call being

terminated at time 100, then this event should start before mobile X's call starts. One way to implement this ordering is to insert the events as they come into the sorted queue. Java JDK offers a 'TreeSet' collection to implement this, where elements are sorted by the ordering prescribed by the comparator, described below. 'TreeSet' takes log(n) time for 'add', 'remove' and 'contains' operation.

### 4.4.2 Comparable

Different objects representing different events populate the event queue. In order to sort the event queue, we need to define a way to compare the objects. Java JDK provides the interface 'Comparable', where the method 'compareTo' has to be specified. 'EventQueueObject' populates the event queue, named 'SimEventQueue'.

```
SimEventQueue = new TreeSet(
      new Comparator()
     {
      public int compare(Object a, Object b)
      {
       EventQueueObject itemA = (EventQueueObject) a;
       EventQueueObject itemB = (EventQueueObject) b;
       return itemA.compareTo(itemB);
      }
     }
    );
```

### 4.4.3 Observer and Observable

We decided to use the Model/View/Controller architecture [Sunsted, 1995] since many controllers can contribute to the model, as shown in Figure 4.15. This architecture is also useful when the same model/data has to be shown in different views and the binding between model and view occurs at run time. Class 'Observable'

provided by Java JDK represents the data that needs to be observed, which is the event queue in our design. The 'Observer' signifies the different views of the 'Observable'. 'Observable' notifies the 'Observers' of an event by the method 'notifyObservers', which calls the 'update' method of the 'Observers'. In this way the event queue notifies the nodes whose call has initiated or terminated or change of location has occurred. The nodes themselves act as controllers of the event queue as they add new events to the queue.
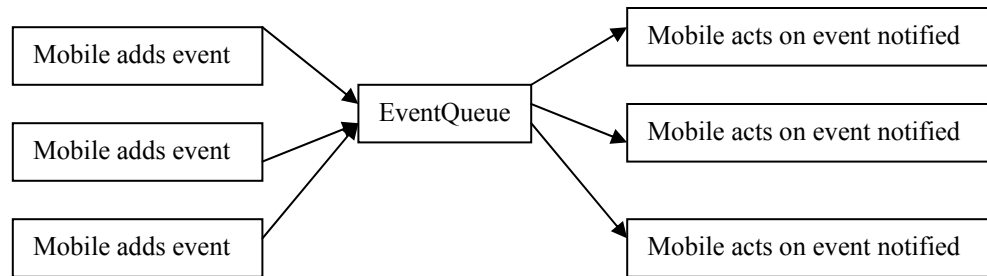


Figure 4. 15 Model/View/Controller architecture

4.4.4 Data Collection

The simulation is updated at each event, be it a mobility event or a call model event. As the mobile crosses a grid line, an update location event is sent. Every call termination creates another call creation event and vice versa. After a simulation is run for, say, 1000 seconds, it is repeated six times and the final average values of metrics are recorded.

## 4.5 Simulation Results

4.5.1 Factors and Metrics

The metrics that we use in this study include: total system energy, per mobile energy consumed, percentage of time a node uses relays, and number of relay – relay handoffs. Other metrics such as throughput and call blocking rate are irrelevant because under high load our system degenerates to a traditional cellular system (i.e., without relays).

The factors that can affect performance are many such as number of mobiles (5, 10, 15, 20, 25, 30), mean call rate (2, 3, 4 calls/hour), mean call duration (100, 200, 300 seconds), cell antenna gain (6 and 10 dBi), mean node velocity (1 or 5 m/s), and base station placement in the cell (center or corner). If we do a full factorial design with all these parameters, we will need to do 6x3x3x2x2x2 = 532 experiments (multiplied with the number of repetitions). Therefore we decided to do a fractional factorial design where we first determine if any of the factors can be eliminated. The number of mobiles, base station antenna gain, and base station placement are all necessary for our study. Therefore, we conducted an initial series of experiments to select among velocity, call rate and call duration. For these experiments, the number of mobiles is set to be 5.

Steps to determine the main effects [Jain, 1991]:

- Each factor has different values or levels
- For each call rate, the simulation is run for both velocity and call duration
- Repetition (five times) of simulation is done

- Results of system energy are put in matrix form

- Each row and column mean is calculated, giving mean for each factor

- A common mean is also calculated

- The difference of each factor (their mean) and the common mean gives the effect of each factor

| Main Factors | Level 1 | Level 2 | Level 3 | Effect of Factors |
|---|---|---|---|---|
| Velocity | 0.075129183 | -0.075129183 | | 0.17 |
| CallRate | -0.301627953 | 0.366536429 | -0.064908476 | 47.34 |
| CallDuration | -0.771235199 | 1.091740008 | -0.320504809 | 18.89 |

Table 4. 3 Effect of factors

We see from Table 4.3 that least effect factor is velocity, while call rate and call duration significantly affect our calculations.  Based on the above analysis, we therefore use the following reduced set of parameters for the final set of simulations using cellular layout as given in Figure 4.16 (blue grid as base station, black grid as building, yellow grid as intersection, white grid as road) and using Appendix A:

- Number of mobiles: 5, 10, 15, 20, 25, 30

- Base station position: Center or corner

- Call rate: 1, 2 calls/hour. We did not choose the values 3 and 4 because we want enough mobiles in an idle state to qualify as a relay.

- Call duration: 120 seconds. We took this factor to be a constant even though it was a significant factor, so as to have more idle nodes that can act as relays.

- Cell antenna gain: 6, 10 dBi

- Grid size is set at 20 meters

- Velocity = 10 m/s

64
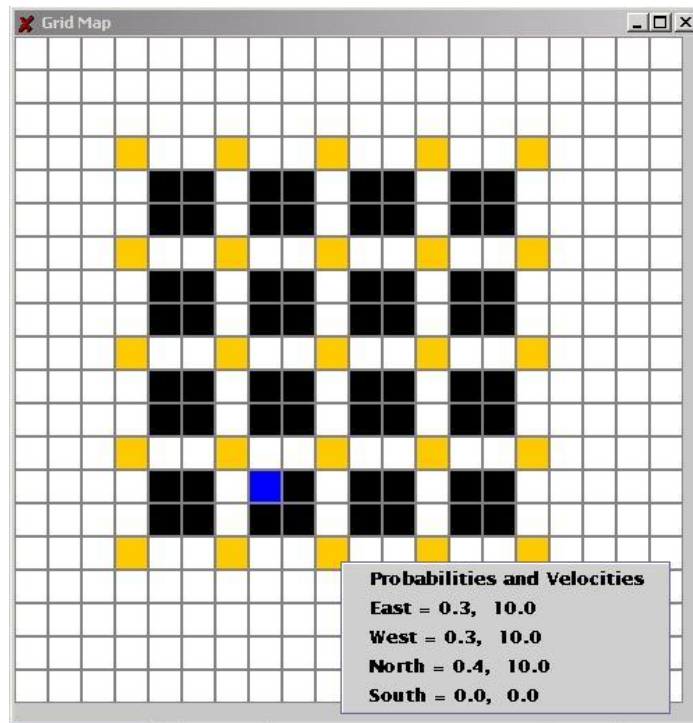
Figure 4. 16 Snapshot of cellular layout

The comparisons we performed were the following:

- Energy consumed (system wide as well as per node) for the case when all mobiles call the base station directly and when mobiles can use relays. This comparison seeks to demonstrate the energy saving benefit of using relays.

- Direct and stretched time which means how much time did mobiles talk directly to the base station and how much time were their calls relayed (we use the term stretched for relayed calls because in a sense the calls are stretched via the relay).

- Number of handoffs between relays in the case that we use relays.

4.5.2 Base station at center, Gain = 10dBi
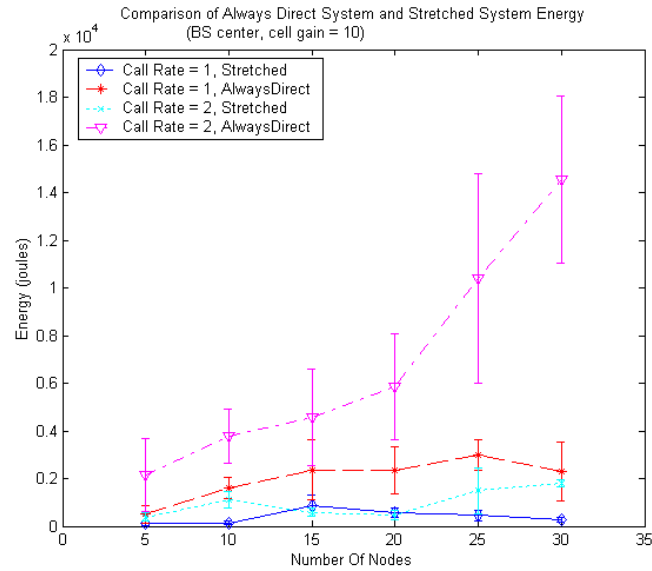


Figure 4. 17 Comparison of system energy for BS at center, cell gain =10 dBi

Figure 4.17 shows the total system energy for varying call rates. As can be seen, higher call rates consume more energy. However, irrespective of call rates the stretched system consumed less energy. For a lower call rate the savings are more than half, while for the higher call rate the savings vary from three to more than seven times. Also the spread in confidence interval is less for the stretched case. The reason is that relaying calls limits energy consumption by hand offs between relays.
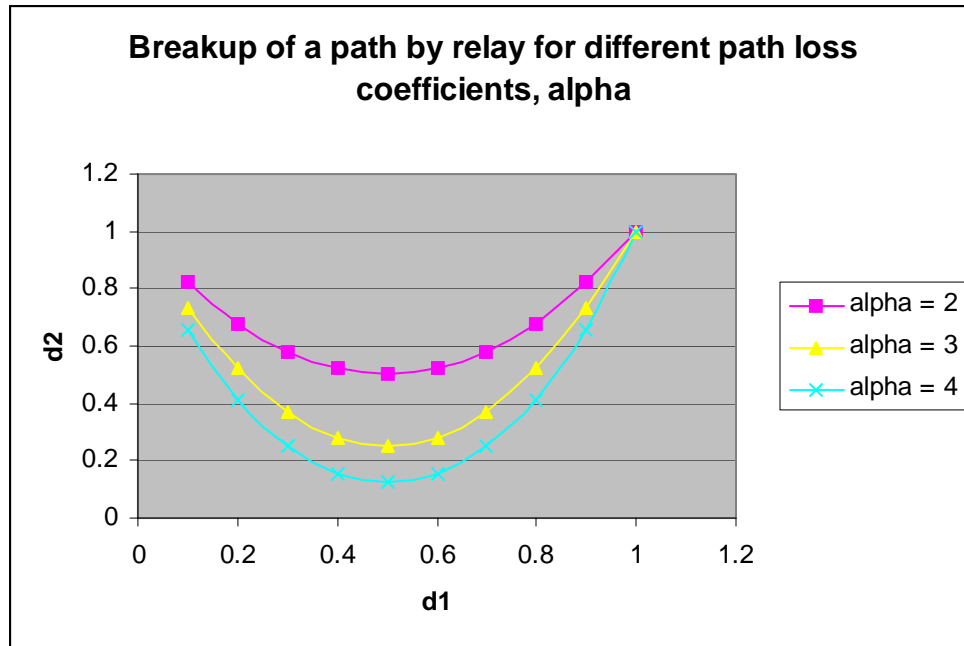
Figure 4. 18 Figure showing breakup of path with d =1

To see why relays are so effective at saving energy, assume that the path loss exponent is $\alpha$. Let d be the distance between the mobile and BS, then the transmit power in a direct call is proportional to $d^\alpha$. For the stretched call, say the distances between the mobile and relay is $d_1$ and between relay and BS is $d_2$. The total energy for the stretched call is then proportional to $d_1^\alpha + d_2^\alpha$. Figure 4.18 plots the values of $d_1$ and $d_2$ for which

$$d_1^\alpha + d_2^\alpha \leq d^\alpha$$

In the figure, d = 1 and $\alpha$ = 2, 3, 4 and $d^\alpha$ = 1. The values of $d_1^\alpha + d_2^\alpha$ are always less than 1. In our simulations there is a high probability of finding a relay in the region that reduces total energy and hence the energy savings as illustrated in Figure 4.17.

A problem may occur if all the mobiles have very high call rates, then at any given instance there is less chance of an finding an idle mobile. This results in less energy savings as well as higher number of handoffs. This is illustrated in Figure 4.19 where we see that at a call rate of 2, the number of relay to relay handoffs is more than double than for a call rate of 1.
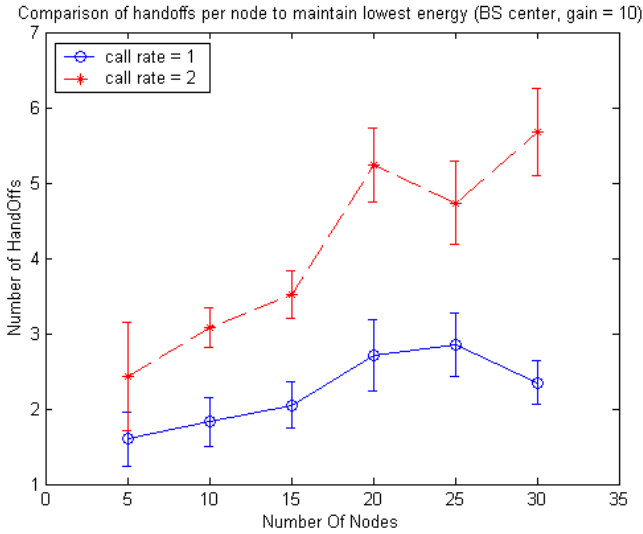


Figure 4. 19 Comparison of handoffs for BS at center, cell gain = 10 dBi

Figure 4. 20 Comparison of per node energy for BS at center, cell gain =10 dBi

Figure 4.20 shows the actual energy consumed per node for a run of 1000 seconds. As can be seen the direct mode always consumes more energy as both nearer ($d^2$) and further ($d^4$) paths are present. The higher call rate shows less than 0.5 joules consumption of energy. The stretched mode savings vary from three to ten times, with higher savings for a higher call rate.
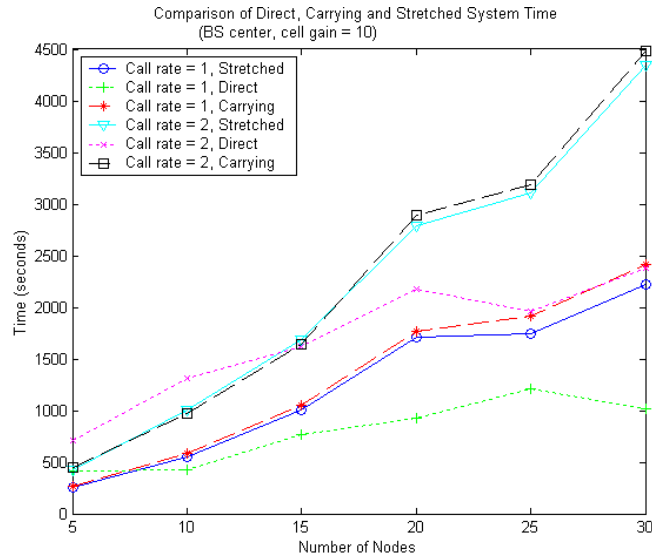
Figure 4. 21 Comparison of time for BS at center, cell gain = 10 dBi

Figure 4.21 plots the total time spent by all the mobiles in the system either in direct calls to the base station (label used in the figure is "Direct"), in stretched calls where the mobile is the originator of the call (label used is "Stretched"), and time spent by a mobile in acting as a relay for other calls (label here is "Carrying"). The first thing to note is that the curves for Stretched and Carrying (for a given call rate) are almost identical. This makes sense because for every stretched call, there is a node relaying that call. A second feature is that the total airtime in a Direct call is always below the total time in Stretched or Carrying. This is because a stretched call causes two mobiles to communicate thus doubling total call time. We note that for a given call rate the Direct curve is not exactly half of the Stretched (or Carrying) curve because our Greedy algorithm only relays a call if there are energy savings to be obtained by relaying; otherwise the call is placed directly to the base station. For a higher call rate, the total airtime goes up, as expected. Figure 4.22 plots the percentage

70

of time calls were placed directly to the base station or relayed. We see that as the number of nodes increases, the percentage of stretched call also increases. This makes sense since there are more candidates available to carry calls. This also explains the decrease in the Direct call percentage.
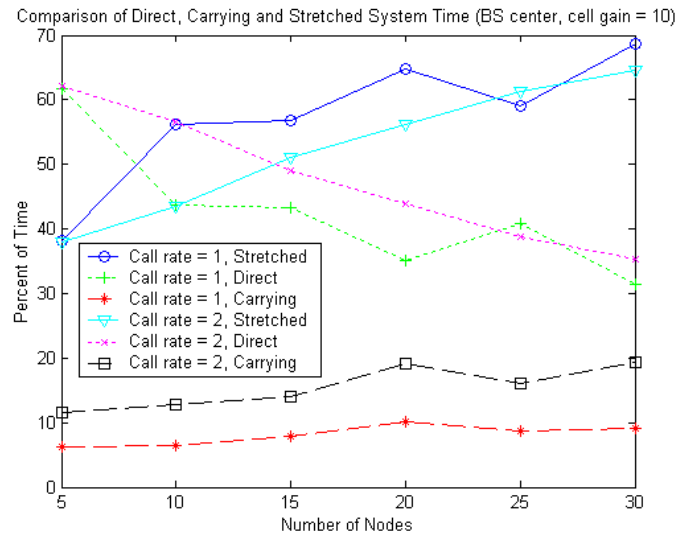


Figure 4. 22 Comparison of time for BS at center, cell gain = 10 dBi

4.5.3 Base station at center, Gain = 6 dBi
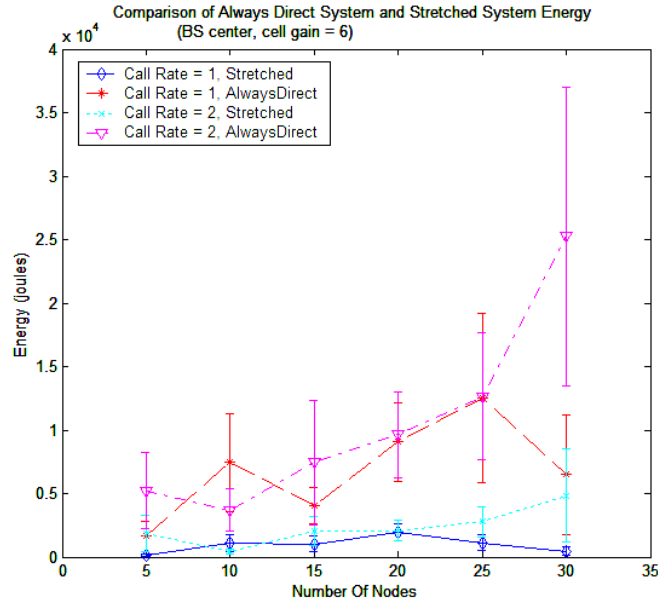


Figure 4. 23 Comparison of system energy for BS at center, cell gain = 6 dBi

Figure 4.23 illustrates the system energy with cell gain of 6 dBi. As expected it shows more energy consumption for higher call rates and lower energy for stretched connections. In comparing with Figure 4.17 with cell gain of 10 dBi, we see that the higher cell gain system consumes less energy. This is more evident for lower call rates.

Consider the case when the number of mobiles is 25. The energy consumed when placing direct calls using a cell gain of 6 dBi consumes almost five times the energy compared to the case when the cell gain is 10 dBi. This makes sense because a higher cell gain indicates a more directional antenna at the base station which requires a lower transmit power at the mobile. Interestingly, however, the total energy

72

consumed when calls are relayed does not change as much when we change cell gain. This is because when calls are relayed, each hop is far shorter than placing a direct call. Thus, the impact of higher cell gain is much lower.
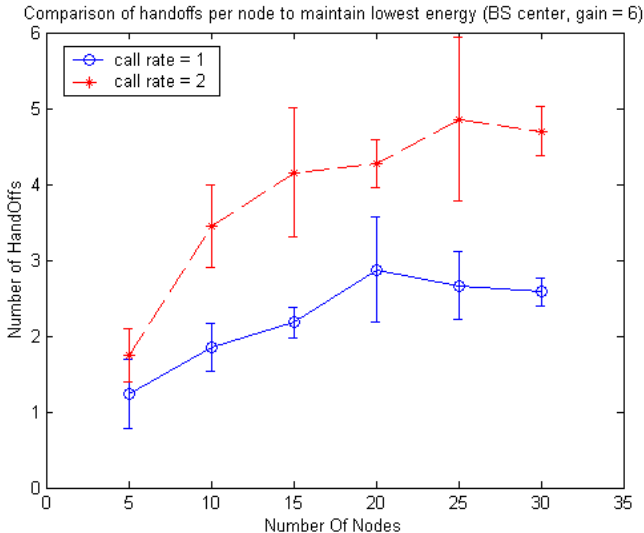


Figure 4. 24 Comparison of handoffs for BS at center, cell gain = 6 dBi

Figure 4.24 plots the number of handoffs as a function of the number of number of mobiles. We see that these numbers are not different from the case when the cell gain is 10 dBi. The reason is that in both cases the Greedy algorithm only relays a call when it is energy efficient to do so. Thus, cell gain does not affect this decision (though it does affect the total energy consumed).
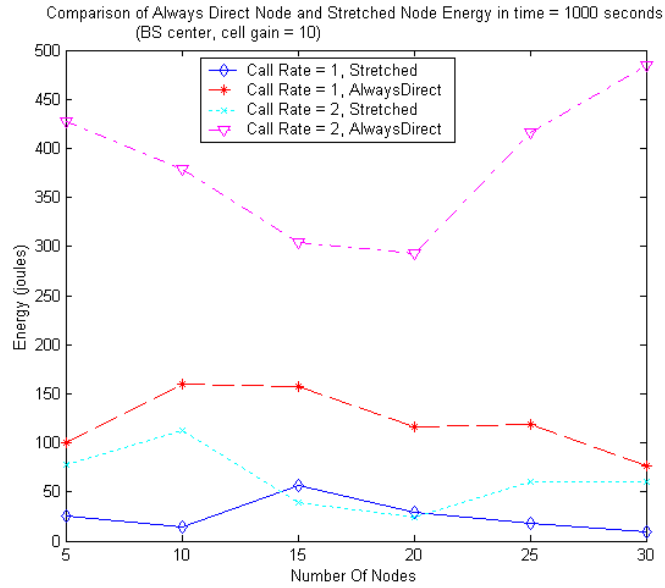
Figure 4. 25 Comparison of per node energy for BS at center, cell gain = 6 dBi

Figure 4.25 plots the per-node energy consumed for a run of 1000 seconds. The direct mode consumes an average of 0.6 joules as opposed to 0.5 joules for the case when the cell gain is 10 dBi. This makes sense since a higher gain implies lower transmit power. The per node energy for the stretched case is similar for both cell gain values. Finally, Figure 4.26 plots the Direct, Stretched, and Carrying time as a function of the number of mobiles. Cell gain does not affect this metric either because calls stretched with one cell gain are also stretched with the other.
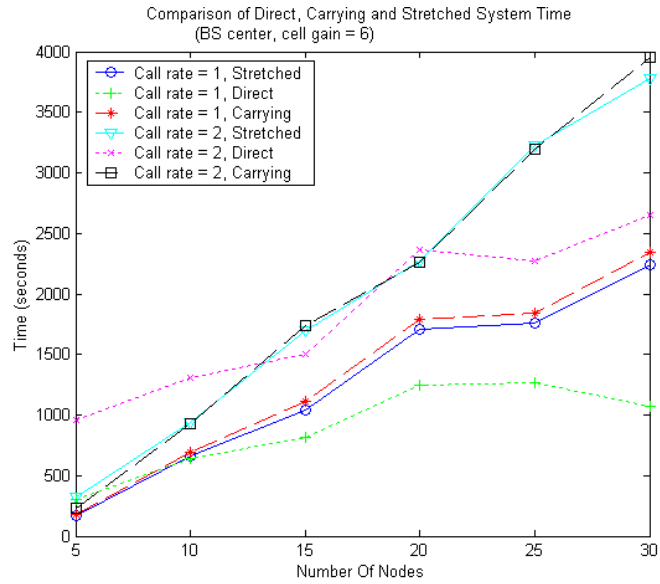
Figure 4. 26 Comparison of time for BS at center, cell gain = 6 dBi

### 4.5.4 Base station at corner, Gain = 10 dBi
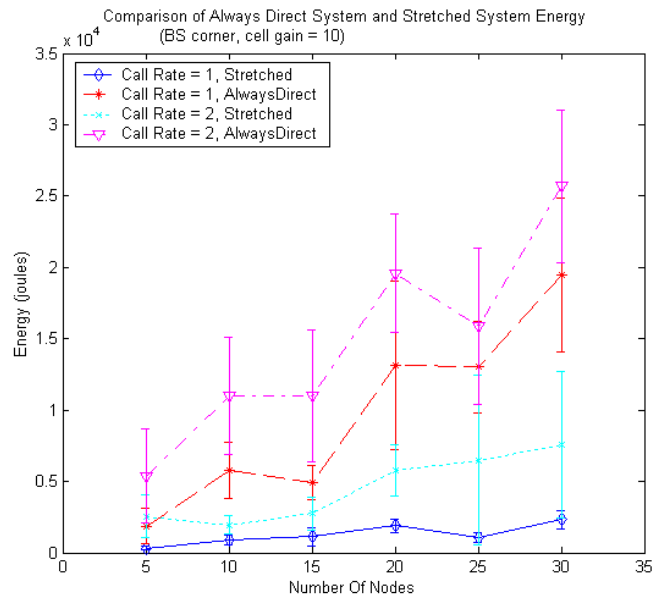


Figure 4. 27 Comparison of system energy for BS at corner, cell gain =10 dBi

We next look at the case when the base station is located at a corner of the cell. This scenario is not uncommon in actual deployments due to zoning and other constraints. In order to cover its cell with such a placement, the base stations use sectored antennas. Comparing Figure 4.27 with 4.17 (base station at center) we note some interesting differences. First, the direct calls have much higher energy consumption when the base station is at a corner. This is because the average distance between the mobile and base station is greater. We also see that when the base station is in a corner, the benefits of relaying the call become more evident. This is also because the distance to the base station is greater.
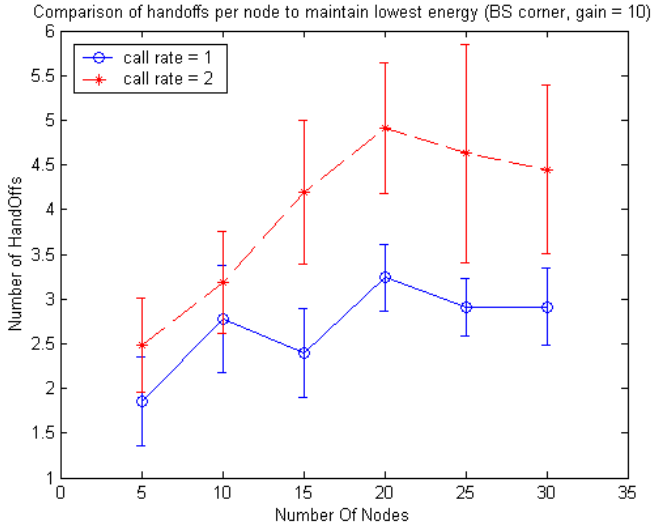


Figure 4. 28 Comparison of handoffs for BS at corner, cell gain = 10 dBi

Comparing Figure 4.28 and Figure 4.19, we see there is little difference in the number of handoffs between relays in either case. This is not unexpected because base

station placement should not really affect these values. Rather, the availability of free

mobiles and the load are the determining factors.
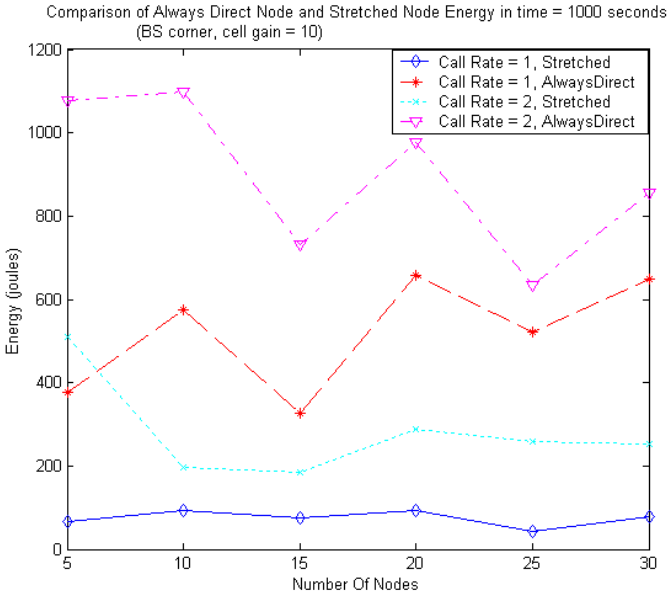


Figure 4. 29 Comparison of per node energy for BS at corner, cell gain =10 dBi


Figure 4.29 plots the per-node energy for a run of 1000 seconds. Comparing

this with Figure 4.20, we note that the trend is similar in both cases but the values for

the corner case are significantly higher. This is again due to the fact that since the base

station is at a corner, the total distances involved are larger. We observe that the

energy for the stretched case is fairly flat as the number of nodes increases indicating

that there are enough relays available at these call rates.
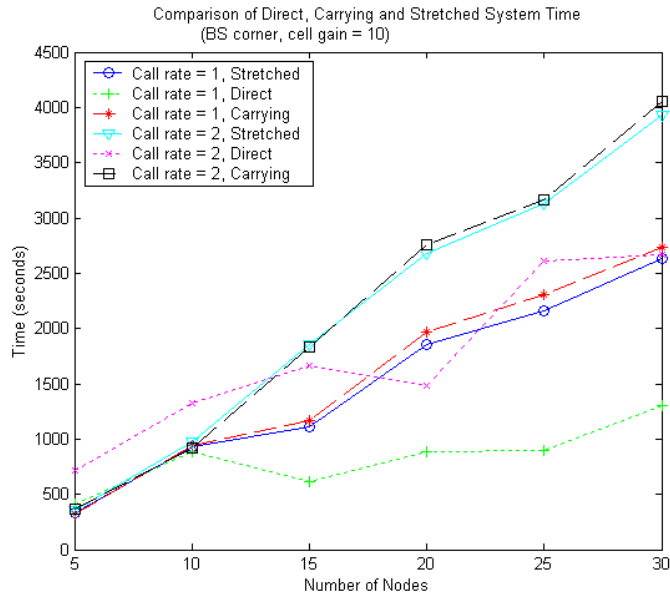
Figure 4. 30 Comparison of time for BS at corner, cell gain = 10 dBi

Figure 4.30 and 4.21 are almost identical. These figures plot the total amount of time that nodes in the system were placing Direct calls, Stretched calls, or Carrying other node's calls. The time values are almost the same regardless of base station location because in both cases, there are a sufficient number of idle relays available to stretch a call. That is, the probability of finding an idle relay in the feasible area shown in Figure 4.18 for a given node, is almost one in both cases. We note that for the same reason, Figures 4.31 and 4.22 are similar (these figures plot the percentage of time a node spends in these three modes).
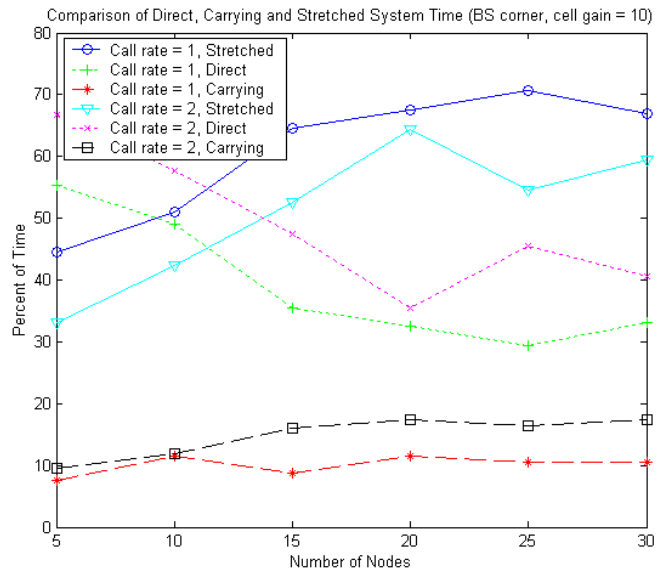
Figure 4. 31 Comparison of time for BS at corner, cell gain = 10 dBi

## 4.5.5 Base station at corner, Gain = 6 dBi
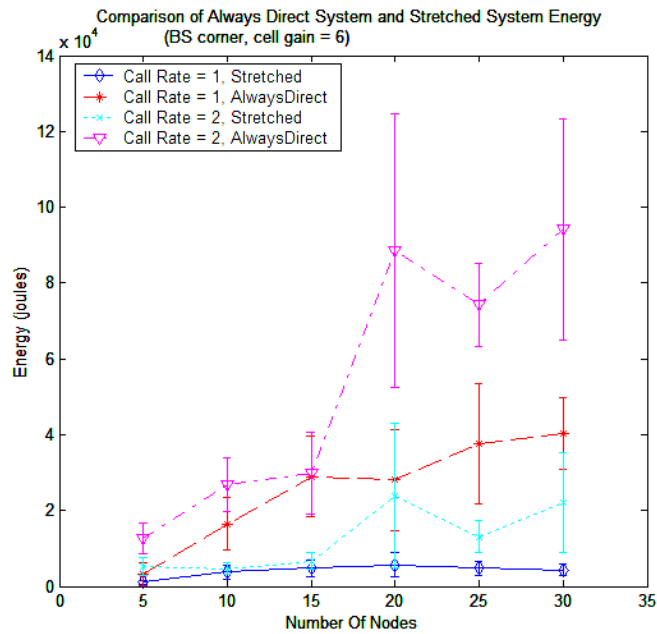


Figure 4. 32 Comparison of system energy for BS at corner, cell gain = 6 dBi

Figure 4.32 plots the total energy as a function of the number of nodes. As expected, because the cell gain is lower, the total energy is greater in all cases. Indeed, the increase compared to the 10 dBi case is at least twice. As compared with the base station in the center case with a cell gain of 6 dBi, we see an increase of four to six times. In other words, placing the base station in the corner of the cell and reducing the gain is the worst scenario in terms of energy consumption. We note the number of handoffs still follows the same trend as previously.
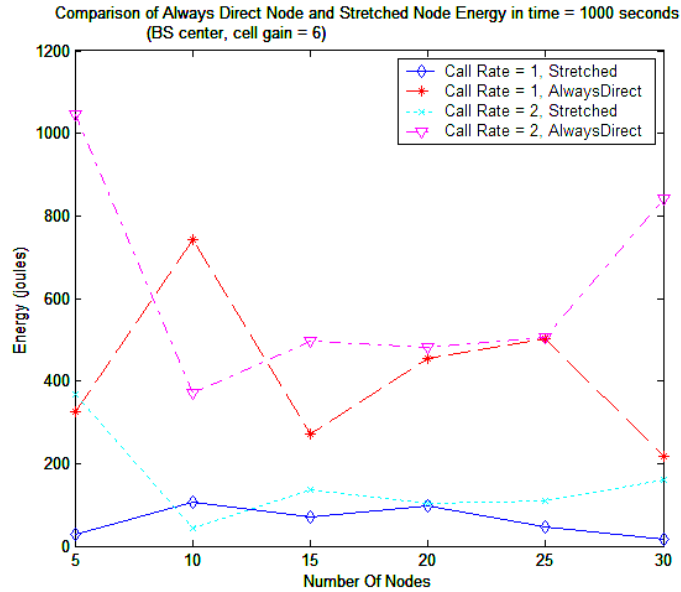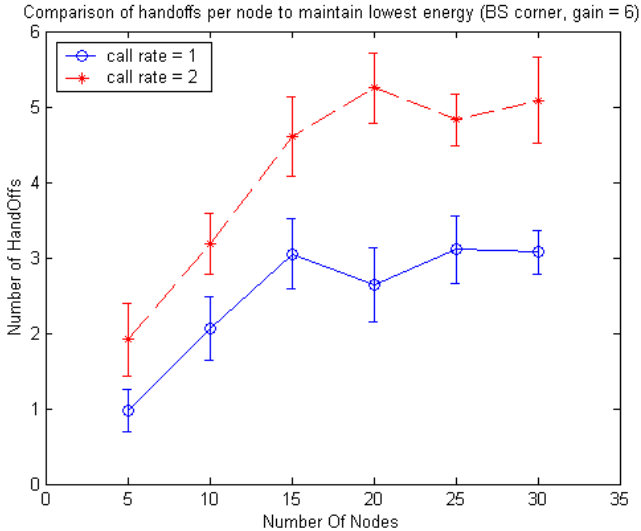


Figure 4. 33 Comparison of handoffs for BS at corner, cell gain = 6 dBi
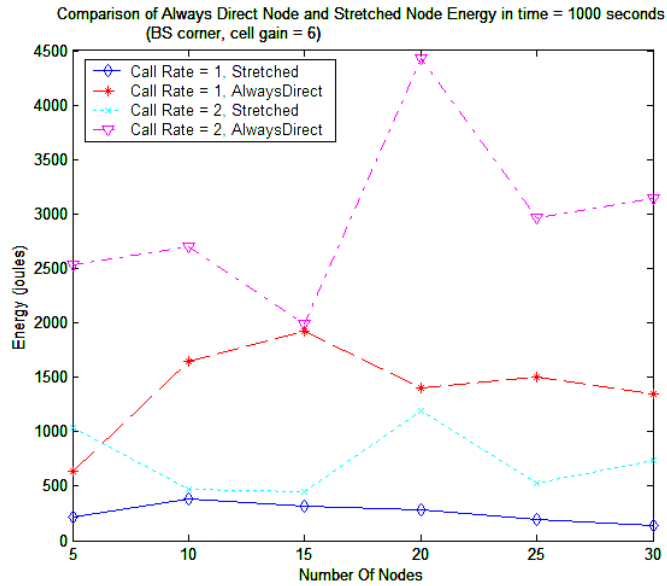
Figure 4. 34 Comparison of per node energy for BS at center, cell gain =10 dBi

Figure 4.34 plots the per node energy consumption. Compared with when the cell gain is 10 dBi, we see an increase of Direct and Stretched energy of at least three times. Compared with the 6 dBi case with the base station in the center, we see an increase of four to five times. Finally, note that the total time spent in various modes (Figure 4.35) is identical to all previous cases because the cell gain does not affect when a call is carried. The only determining factor is the presence of an idle node in the feasible region.
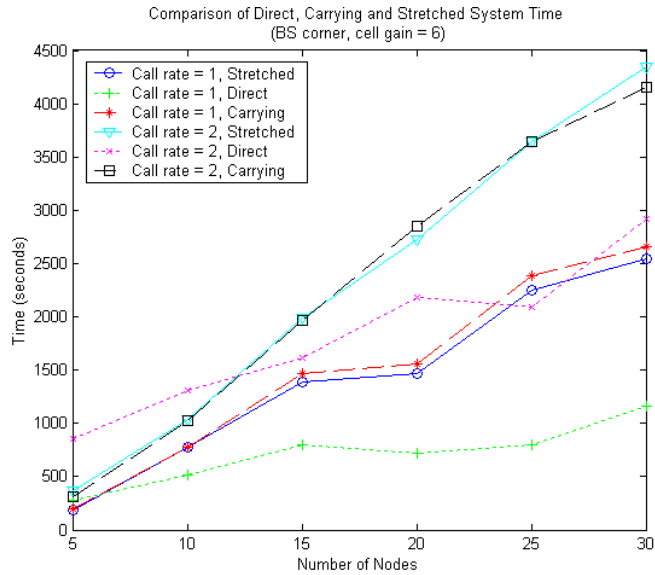
Figure 4. 35 Comparison of time for BS at corner, cell gain = 6 dBi

4.5.6 Summary of the Main Results

Based on the previous discussion, we can draw the following conclusions:

1.  Using relays is always beneficial with energy gains of between 1.76 dB to 8.45 dB.

2.  If the base station is in the corner of the cell, the benefits of using relays are even more significant.

3.  The energy savings are greater as the number of idle node increases.

4.  The number of relay-relay handoffs increases with an increase in the number of mobiles because the Greedy algorithm always attempts to find the lowest energy path for the connection.

5.  Using a higher cell gain reduces energy consumption of direct as well as stretched calls.

The base station at corner case has higher energy consumption when compared to the base station at the center, because of the fact that for the same area, higher distances are covered in the former case. In both the cases the stretched simulation mode achieves significant energy savings over time, which varies from 1.76 dB to almost 8.45 dB as the number of nodes is increased. The average energy per node shows savings in both simulation modes, showing the fairness in choosing relays over time. The amount of time a mobile carries for another is always more than the time for direct calls, even though direct calls have highest priority. This means that we have enough resources in terms of time to allot for relaying. The number of handoffs is similar in all cases.

## 5. Improving coverage with relays

The previous chapter discussed the improvements in system energy consumption when using relays. This chapter explores improvements in the reliability of coverage. Coverage is defined as the farthest distance from the base station where the signal has reliable signal strength [Rappaport, 1996], which is also the range of the cell. There is, however, the question of how uniform this coverage is. Using accurate ray tracing models, [Coinchon, Salovaara, and Wagen, 2001] report 12% rejections when 20dB building penetration loss is not taken into account and 20% rejections when the building penetration loss is taken into account. Intuitively the reason is that path loss contours in high building density areas make their own local minima. [Lee and Miller, 1998] term these as pockets in which the propagation can be very poor. [Lee and Miller, 1998] also shows that, with the CCIR propagation model, the impact of the building density on the cell radius is much more than the impact of the antenna heights. Further evidence is provided by [Neskovic, Neskovic and Paunovic, 2002], where they report coverage of Belgrade city. Areas identified as highly dense typically have lower signal strength when compared to semi-urban areas.

### 5.1 Gaps in Coverage

Radio signal penetration varies considerably due to variation in the terrain and due to the interference limited characteristic of CDMA system. [Jones and Skellern, 1995] report that contiguous cell coverage may not be possible in a microcell even if

the call loss rate is low. They analyze interference to noise ratio and show that interference from other mobiles is sufficient to produce gaps in coverage.

To illustrate the variation of signal due to buildings, contour maps are plotted using CCIR formula in Figure 5.1. Path loss contours here use 850 MHz for a typical urban area. The area of the cell is 1 km$^2$. The base station antenna height is 30m and mobile antenna height is 1.5m. The buildings lie in a grid formation all over the area. The base station is at the center of the cell. Only the area near the base station shows the circular curves as expected. The contours get distorted as we go further away from the base station due to the presence of buildings. In two dimensions, as we can see in Figure 5.1, several local minima occur (circular areas around buildings, away form the BS), which correspond to the high building density. In both cases, we see that several areas have signal strength well below the threshold, resulting in call dropping. Figure 5.2 plots the same contours for the case when the base station is in a corner. Figures 5.3 and 5.4 provide a 3-D plot of signal strength in the cell. The valleys that we see are all below threshold indicating a lack of coverage. We call these areas Below Threshold Areas (BTA).
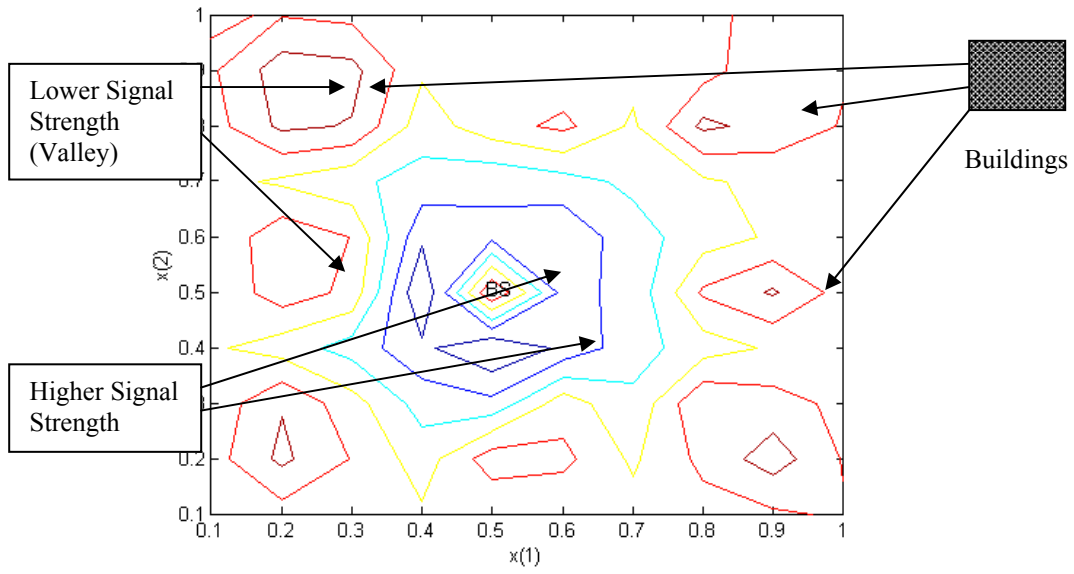
Figure 5. 1 Path Loss contours for base station at center
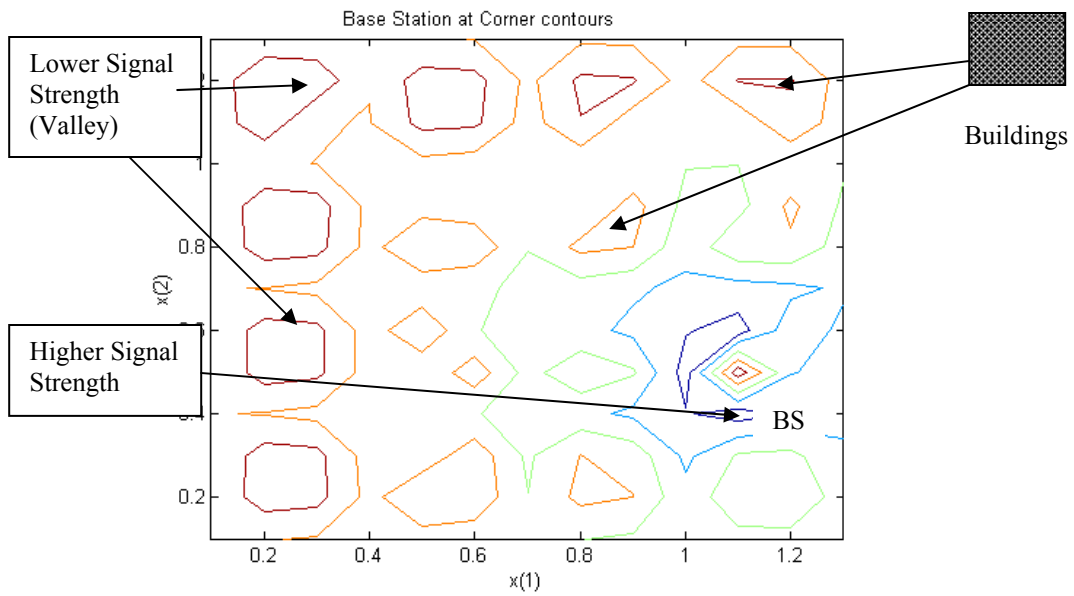


Figure 5. 2 Path Loss contours for base station at corner

Lower Signal
Strength
(Valley)

Buildings

Path Loss

Higher Signal
Strength
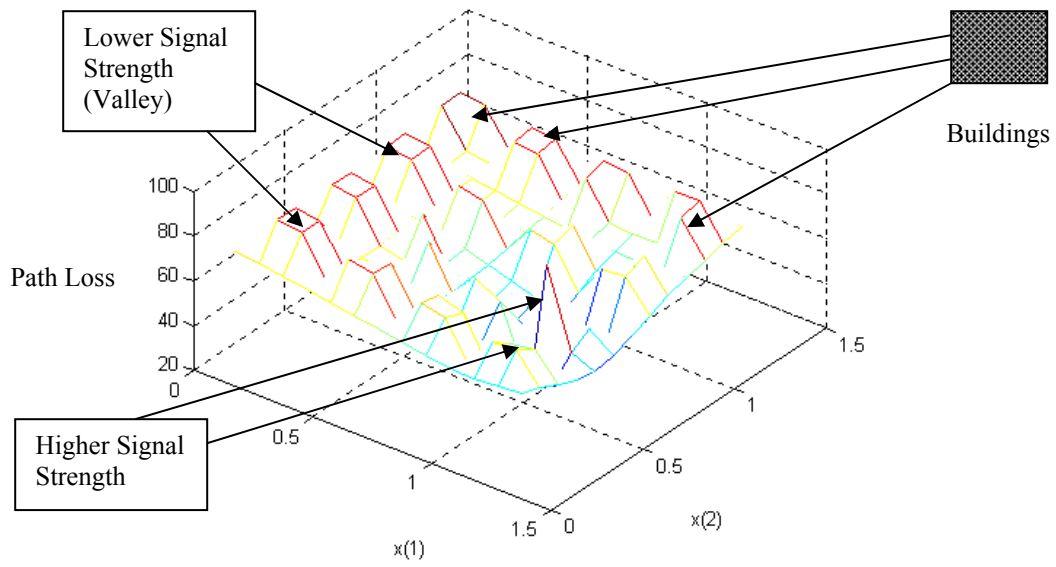
Figure 5. 3 Path Loss contours in 3D for base station at center

Lower Signal
Strength
(Valley)

Buildings

Path Loss

Higher Signal
Strength
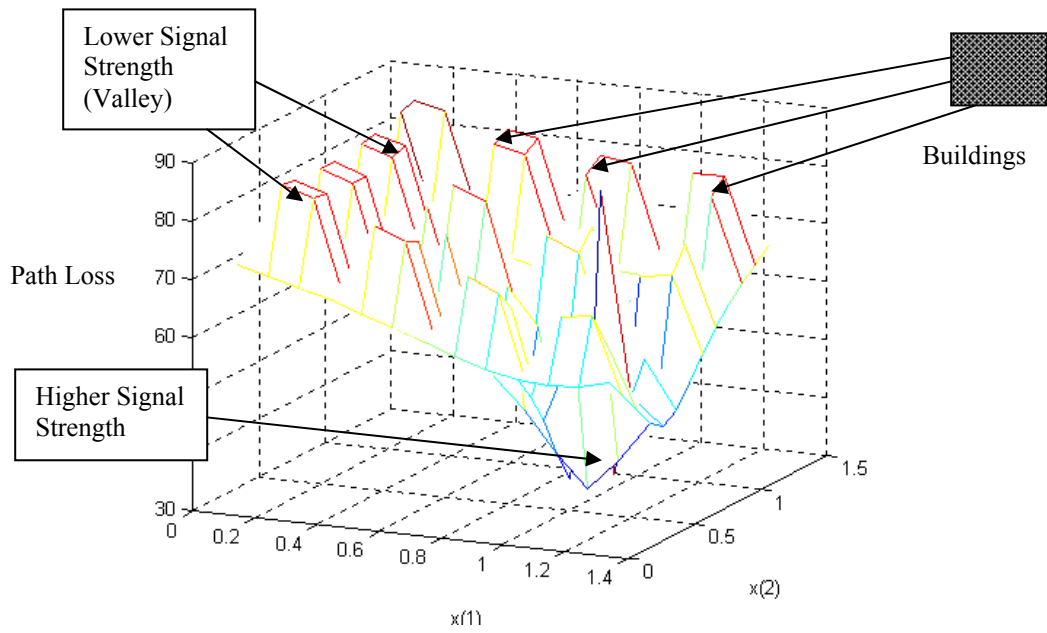
Figure 5. 4 Path Loss contours  in 3D for base station at corner

87

In order to deal with the problem of BTAs, two solutions are possible. The first is to add additional mini-base stations within the cell and the second solution is to have mobiles use a higher transmit power if they are in a BTA. The first solution is expensive and does not guarantee that all BTAs will be eliminated (unless we add a large number of mini-base stations). Furthermore, new building construction can invalidate some placements. The second solution is problematic in interference – limited CDMA systems because other mobiles will encounter greater interference resulting in lower system capacity.

We explore the potential of the relay – based approach to solve the above BTA problem. If a mobile is stuck in some below threshold area, it need not increase its power to reach the base station. It can instead search for another idle mobile, which can relay its call to the base station. This procedure is particularly useful in a dead spot [Aggelou and Tafazolli, 2001], as the base station may be out of reach for the mobile. Thus, in addition to saving battery resources since a low signal power is required to reach the relay, we can increase coverage.

## 5.2 Modeling Below Threshold Areas

In order to study the benefit of using relays to improve coverage, we first need to model a typical urban cell containing several BTAs. We develop a mathematical model to quantify the occurrence of Below Threshold Areas (BTA). The modeling challenge is that BTAs can vary in size dramatically, they can be only few meters wide or can stretch to whole building blocks. A cellular user thus experiences BTAs as small fades or whole block fades. To properly capture this behavior, we divide the cell

into grids. The high building density grids will have many BTAs, while sparse regions will have fewer BTAs.

In order to model the natural clustering behavior of BTA occurrence (e.g., a building causes several BTAs whereas open roads do not), we use the Generalized Poisson Distribution (GPD). The GPD was developed by Consul and Jain [Consul, 1989] where they noted that complete randomness is not realistic and we don't see it in nature. Thus living organisms tend to cluster more in some areas, maybe for food or for safety. They even lay eggs on wheat, seed or leaves, but only at those places where somebody else has already laid eggs. They even migrate in groups. So some type of clustering is inevitable. The symbols $\theta$ and $\lambda$ are called the first and second parameters of the GPD model. Parameter $\lambda$ is independent of $\theta$ and the lower limit is imposed to ensure that there are at least five classes with non–zero probability when $\lambda$ is negative. Variance of this GPD model is greater than, equal to, or less than the mean according to whether the second parameter $\lambda$ is positive, zero or negative. Both mean and variance tend to increase or decrease in value as $\theta$ increases or decreases. When $\lambda$ is positive, the mean and variance both increase in value as $\lambda$ increases, but variance increases faster than the mean. Let X be a discrete random variable (r.v.) defined over non - negative integral values, and let $P_x(\theta, \lambda)$ denotes the probability that r.v. 'X' takes the non – negative integral value, x.

The Generalized Poisson Distribution (GPD) [Consul, 1989] is,

$$P_X(\theta, \lambda) = \frac{\theta(\theta + x \cdot \lambda)^{x-1} e^{-\theta - x \cdot \lambda}}{x!}, x = 0,1,2,\ldots$$

$$P_X(\theta, \lambda) = 0, x > m, when, \lambda < 0 \hspace{3cm} 5.1$$

where $\theta > 0$ , max(-1, $\theta/m$ ) $< \lambda <= 1$ and m($>=4$) is the largest positive

integer for which $\theta + \lambda$ m > 0 when $\lambda$ is negative. By varying the two parameters we

can control the mean number of BTAs and the variation in each grid. In Figure 5.5, we

illustrate the use of GPD to generate BTAs in low and high building density areas.

Low Building Density Cell :   $\lambda$ = -0.1,   $\theta$ = 0.2, m = 1 => This gives almost eighty

percent of grids free of BTAs and the rest have only one BTA.

High Building Density Cell :    $\lambda$ = -0.2,   $\theta$ = 0.8, m = 2 => This gives a few highly

dense grids and lots of less dense grids. Fifty percent of grids still have no BTAs.
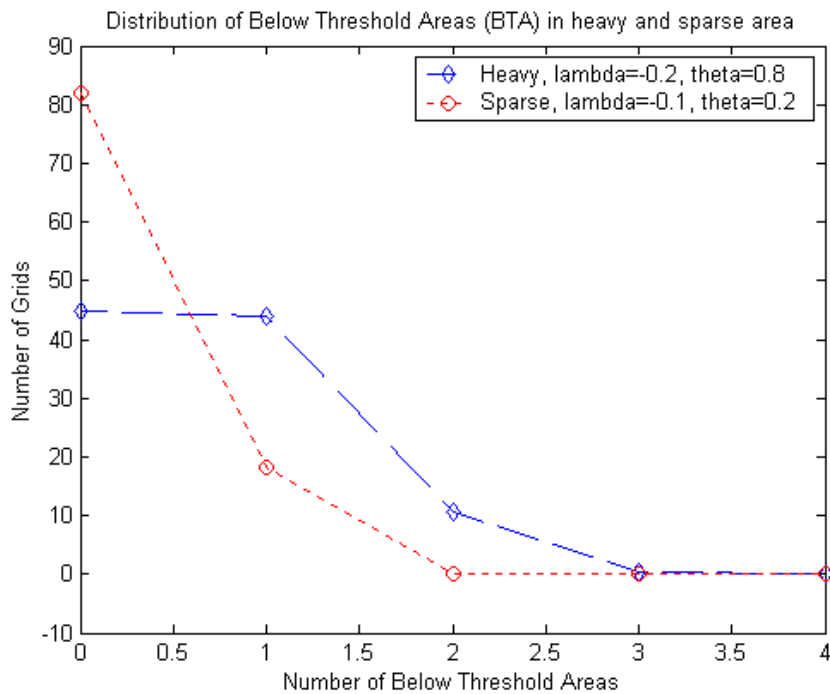


Figure 5. 5 Distribution of Below Threshold Areas

## 5.3 Reliability of Coverage

*We use the metric "useful service area" [Rappaport, 1996] to identify the total area where the signal level is above some threshold.* It can thus be seen as a measure of the reliability of coverage. This measure is quite relevant now as more and more people give up landlines in favor of cellular phones and expect good coverage. *Useful service area* (or useful coverage area) is given as,

$$\int_{area\ ,A} \Pr(\ \xi\ >\ \gamma\ )\,dA \qquad\qquad 5.\,2$$

where $\xi$ is the signal level and $\gamma$ is some threshold level. This gives the probability that a connection will not be lost. The integral is taken over the entire cellular area.

Improvement in the useful coverage area due to relayed calls is given by,

$$\int_{area\ ,A} \Pr(\ \xi\ >\ \gamma\ )\,dA\ +\ \int_{area\ ,A} \Pr(\ \xi\ <\ \gamma\ )\ *\ \delta\,(stretch\ )\,dA \qquad 5.\,3$$

where, $\delta$ (stretch) = 0 or 1 depending on the possibility of relaying the call.

The probability that the above delta function will be unity is equal to the probability of a calling mobile in the BTA finding another mobile to relay its call. The relaying mobile needs to be idle and it should not be in a BTA itself.
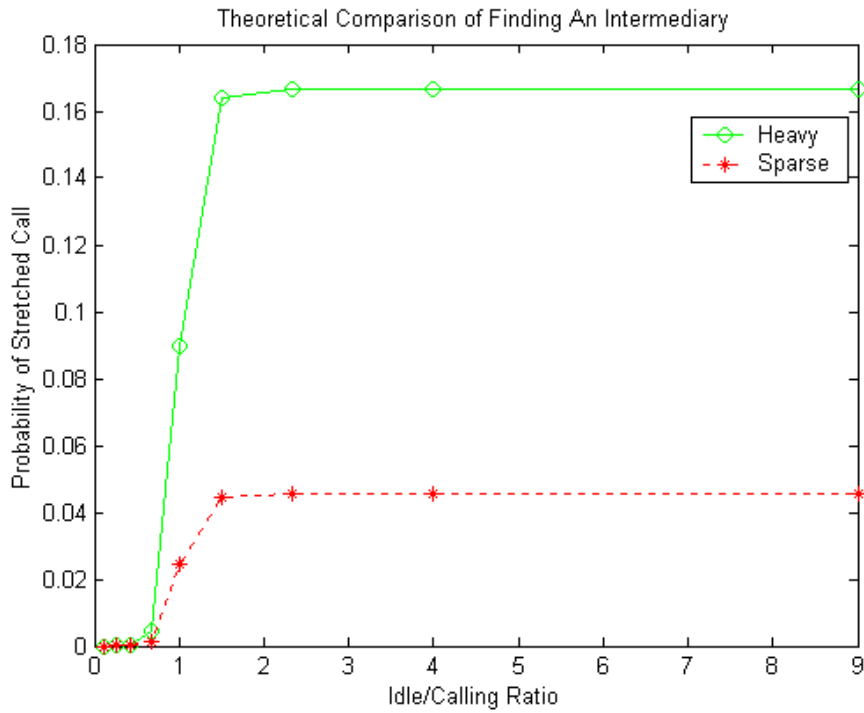
Figure 5. 6 Theoretical probability of finding an intermediary

To find the percentage of additional area covered by introducing a relay, we perform the following simple calculation. Let us denote by N the total number of grids in the cell. Out of these grids, let n lie below threshold. So there are N-n grids above threshold (without relays). Now out of the n grids we need to find the probability of placing a stretched call or finding a relay.

The probability of finding a relay, P, is proportional to the occurrence of BTAs as given by the GPD, which is dependent upon building density. In addition, P also depends upon the number of idle nodes and how far they are in order to save power and this factor is given by the probability of j stretched calls, $P(j \mid m)$. Here it is to be noted that if there are constant number of mobiles in the cell, the number of idle nodes

changes after one has been allocated. Here M is the total number of mobiles and m is the number of idle mobiles. So $P(j \mid m)$ follows hypergeometric distribution given as,

$$P(j \mid m) = \frac{\binom{m}{j}\binom{M-m}{m-j}}{\binom{M}{m}} \qquad 5.4$$

This is the probability of finding a mobile able to carry a call without being replaced. If we can replace the available set of mobiles, then it will follow the Bernoulli distribution. Thus the probability of finding a relay for a mobile is,

$$P = \sum_m P(j \mid m).P_X(\theta, \lambda) \qquad 5.5$$

In Figure 5.6 we plot the probability of a mobile using a stretched call due to poor coverage. As expected heavy density areas have a higher probability of using a stretched call. The area of interest is the rising curve for both heavy and sparse areas. This shows that saturation is achieved on reaching a certain idle to calling ratio of mobiles.

### 5.4 Methodology of Simulations

We use the above grid-based cell representation where each grid may have a number of BTAs. The high building density areas will have higher number of BTAs. The range of building density varies from 10-20% in sparse cells and 10-40% in heavy cells [Lee and Miller, 1998]. The mobiles are placed uniformly randomly in the cell. Thus mobiles inside buildings are easily represented here as against our previous representation where mobiles were only on the roads.

Monte Carlo simulations are run with 50 nodes on a 1km x 1km cellular area with the base station at the center. The path loss exponent was kept at 2 for distances less than a threshold distance of 300m and 4 for beyond the threshold distance. The results are a mean of 30-50 runs. Since we are using the stretched call model we choose the intermediary only when the direct path is costly and the stretched path is greedily chosen. The chosen intermediary also shouldn't fall in a local minima region, called the "constraint" here. With these constraints we consider four approaches.

5.4.1 Strategies/ Approaches for the system

1. "*No constraint and No replacement*" approach – The approach of "No constraint" means that any available mobile is considered for carrying a call, even if it lies in a below threshold area. This model makes sense if the intermediary has sufficient battery power, which can be used to increase transmit power to reach the base station. "No replacement" means that the same pool of mobiles is kept throughout the simulation. The initial available set of idle to calling mobiles is considered as the idle to calling ratio.

2. "*No constraint and With Replacement*" –"With replacement" means that new idle mobiles are added as intermediaries are allocated.  So the idle to calling ratio is kept constant throughout the simulation.

3. "*With constraint and With Replacement*" – Only idle mobiles that do not lie in a BTA can be allocated as intermediaries. In addition new idle mobiles are added to maintain a constant idle/calling ratio.

4.  *"With constraint and No Replacement"* – This approach considers the above

    constraint and doesn't replace the idle mobiles allocated as intermediaries.

Figures 5.7, 5.8, 5.9 and 5.10 show the comparison of our theoretical model with

simulation for these three approaches. For Approaches 1 and 2, shown in Figures 5.7,

and 5.8, we see that the theoretical model underestimates the probability of a stretched

call whereas the model is fairly accurate for Approach 3 (Figure 5.9) and Approach 4

(Figure 5.10).
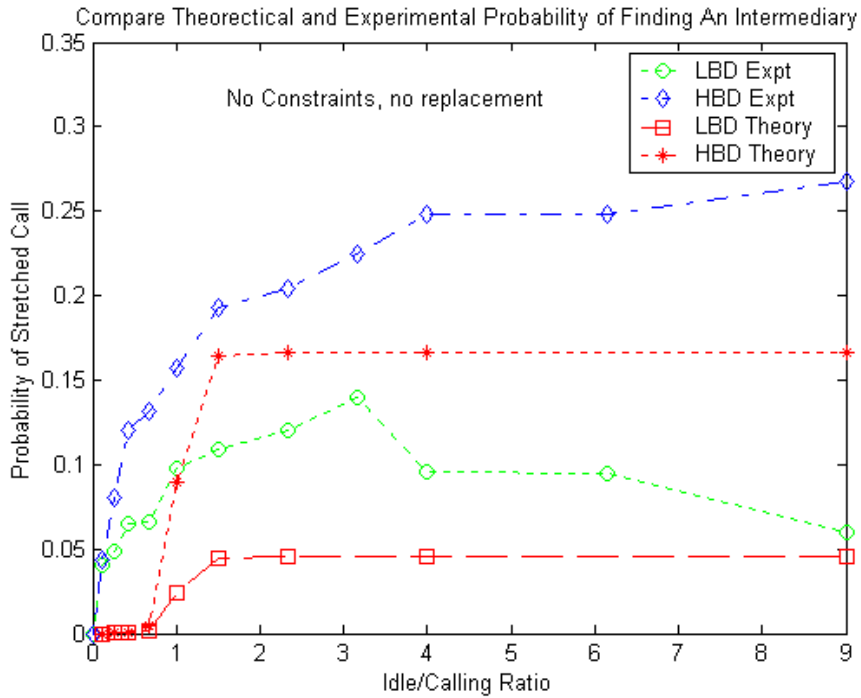


Figure 5. 7 Approach 1: No constraint, no replacement
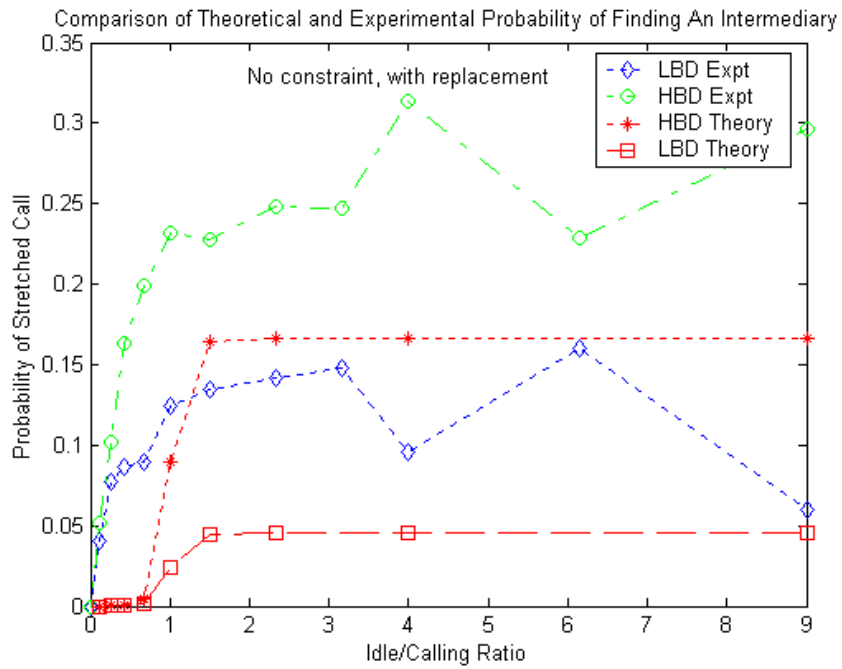
Figure 5. 8 Approach 2: No constraint, with replacement
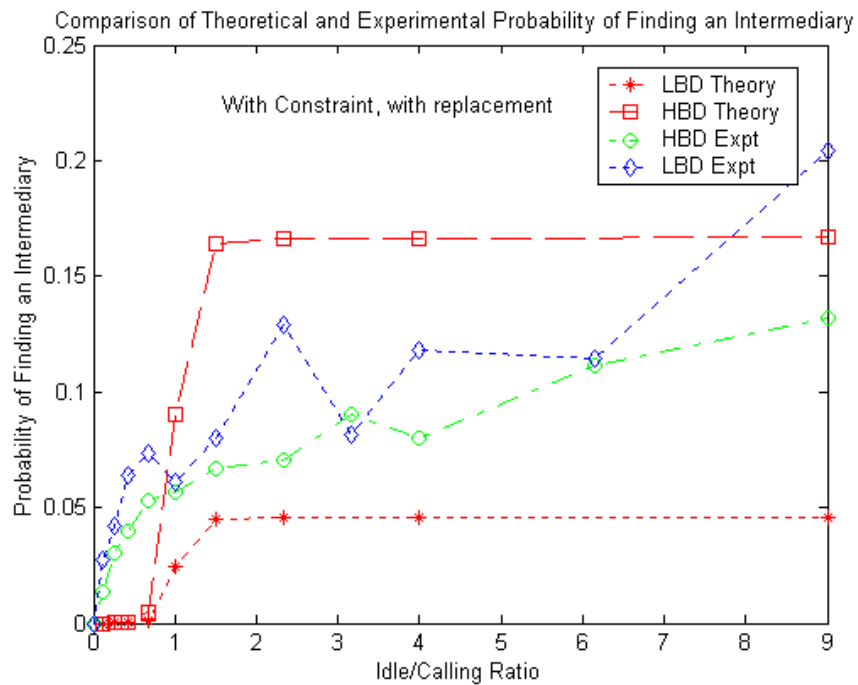


Figure 5. 9 Approach 3: With constraint, with replacement
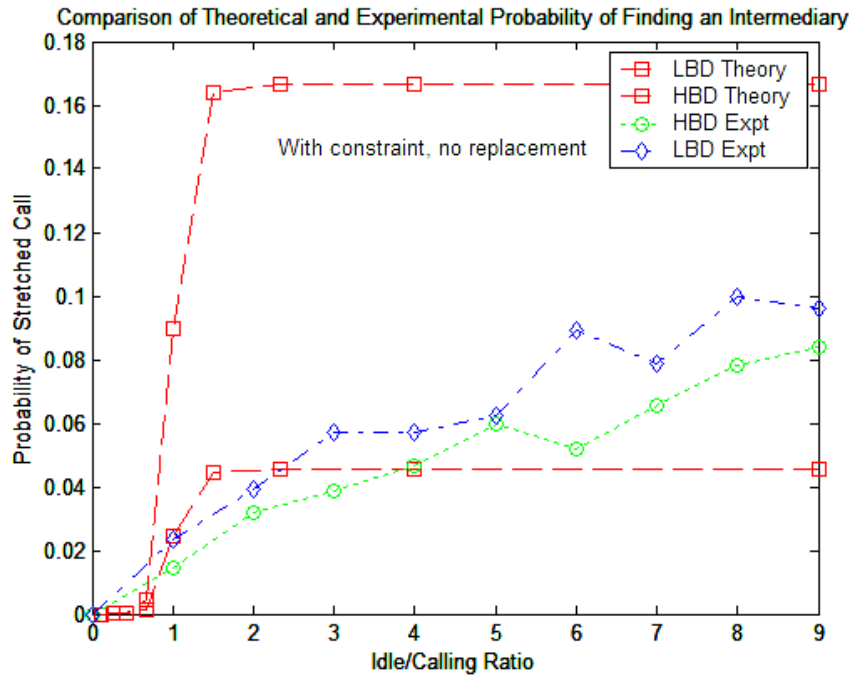
Figure 5. 10 Approach 4: With constraint, no replacement

## 5.5 Simulation Results

The following plots use the terminology

- Percentage of mobiles that lie in BTAs ("Below Threshold Areas")

- Percentage of mobiles that use intermediaries ("Found Intermediaries")

- A ratio of the above two ("Intermediaries/BTA ratio")

### 5.5.1 Coverage Results

In the figures below, first we have LBD cases and then the HBD cases.

Approach 1: No constraint, no replacement
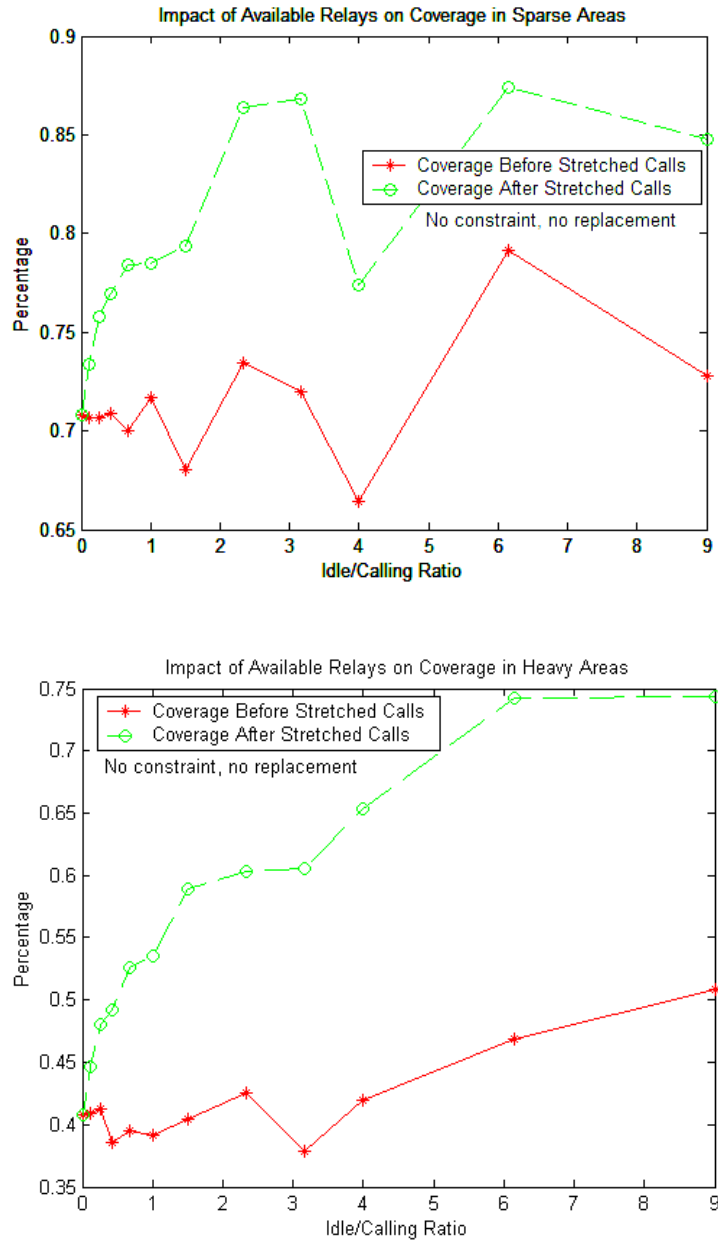


Figure 5. 11 No constraint, no replacement approach; Impact of available relays in LBD and HBD areas

Figure 5.11 shows the improvement in coverage for the "No constraint, no replacement" approach for LBD and HBD areas. Both LBD and HBD show improvements in coverage, though for HBD the improvement is quite high, of the order of 50%. For the LBD case, the improvement is smaller since there are fewer BTAs (as indicated by the solid line) to begin with. However, the improvement is over 15%.

Approach 2: No constraint, with replacement



Figure 5. 12  No constraint, with replacement approach; Impact of available relays in LBD and HBD areas

Figures 5.12 show the "No constraint, with replacement" approach for LBD and HBD areas. This approach improves over the previous no replacement approach for both LBD and HBD, since increasing the availability of intermediaries invariably helps in greater coverage. For an idle to calling ratio of one, there is an almost 20% improvement in coverage, and for idle to calling ratio of three, the coverage jumps from nearly 38% to approximately 69%. Higher values of idle to calling ratio don't give any more improvement in coverage. This means that having only a few additional relays nearby can greatly improve the possibility of having no dead spots.

Approach 3: With constraint, with replacement
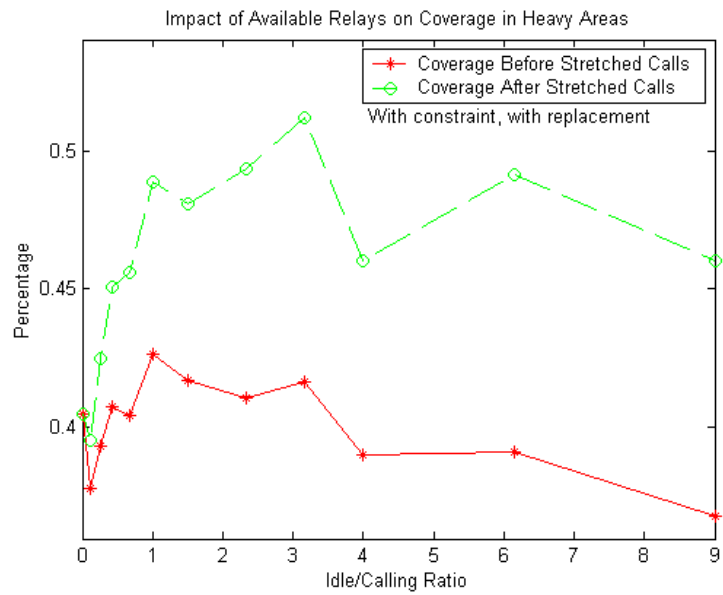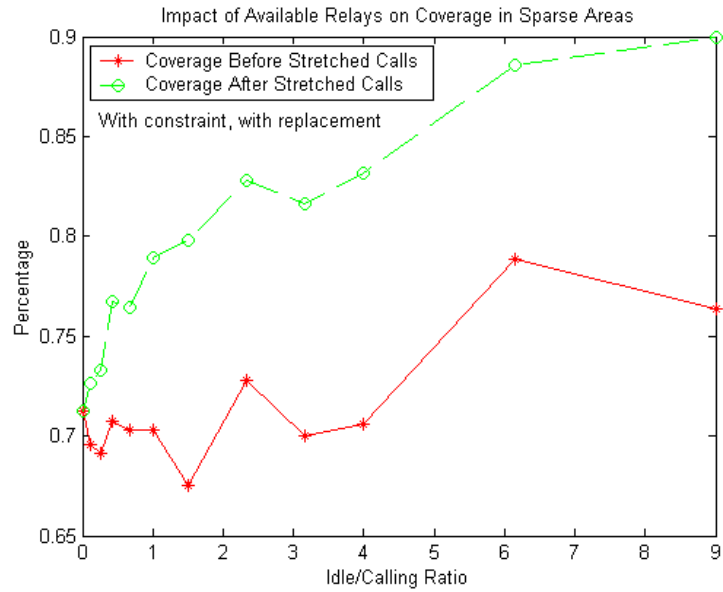




Figure 5. 13 With constraint, with replacement approach; Impact of available relays in LBD and HBD areas

In Figure 5.13 both the LBD and HBD plots show that the stretched mode curve follows the non-stretched mode curve. When comparing LBD with HBD, LBD shows greater improvement in coverage, going upto 90% for Idle to Calling ratio of 9. There is a steady improvement in coverage for LBD, which is similar to the trend seen in previous LBD figures. For HBD, the improvement in coverage flattens at approximately 50% even though there are more relays. This is because the relays also have higher chance of falling in BTAs.

Approach 4: With constraint, no replacement
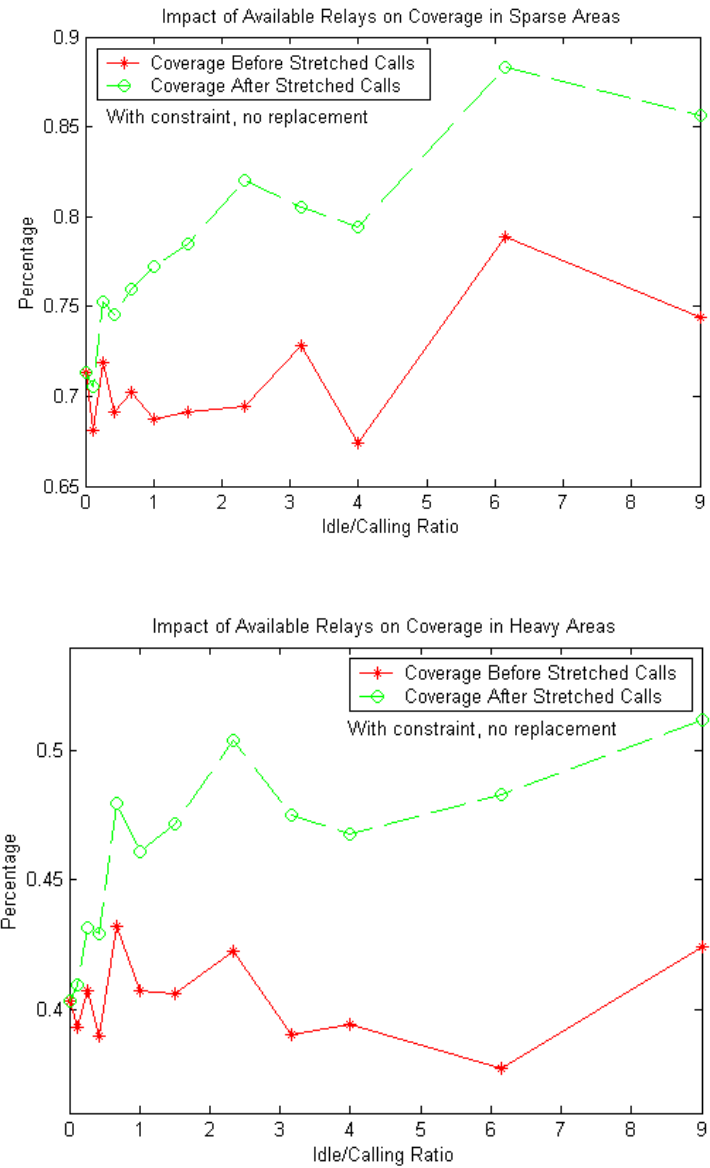


Figure 5. 14 With constraint, no replacement approach; Impact of available relays in LBD and HBD areas

Figure 5.14 shows the improvement in coverage for the "With constraint, no replacement" approach for LBD and HBD areas. For the LBD case the stretched mode provides similar improvement as the previous LBD figures. However, for HBD the coverage improvement doesn't go beyond 50%. With no replacement and with constraint, there is hardly any chance of getting many intermediaries who do not lie in BTAs.

In all the above figures both for LBD and HBD cases, the interesting region was where the stretched mode started to give improvement in coverage over the direct call mode. This corresponded to a idle/calling ratio of between 0 and 1. We explore this region in more detail in section 5.5.2.

5.5.2 Impact of Idle to Calling Ratio

As we increase the number of idle mobiles vs. total number of mobiles, we get an idea of possible relays. The amount of coverage is given by the percentage of the cell that is covered. We would like to explore how much this ratio affects the coverage. It is to be noted that LBD areas have lower number of areas that need a relay since most of the paths are line of sight. On the other hand, HBD areas have a higher number of areas that need a relay and many more paths are not line of sight paths. The first factor of "need of relay" is accounted by "Below Threshold Areas" in the following figures.

Another point to be noted is that though HBD areas have a higher need for relays, the process of finding relays is harder than in the LBD case since most of the possible candidates for relays might themselves be lying in a BTA area. When an

intermediary is found for a mobile it is denoted by "Found Intermediary" in the following figures. The "Intermediary vs. BTA Ratio" gives the ratio of "Found Intermediary" and "Below Threshold Areas".

Approach 1: No constraint, no replacement
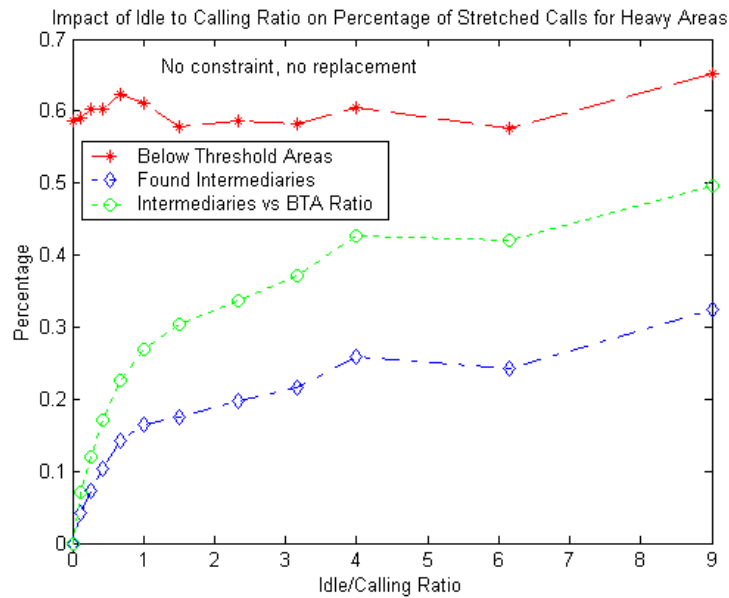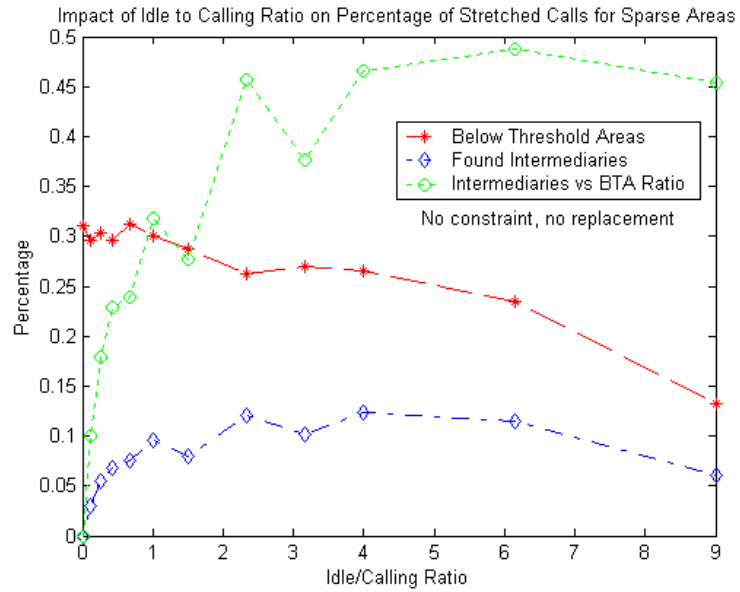


Figure 5. 15 No constraint, no replacement approach; Impact of available relays in LBD and HBD areas

For "No constraint, no replacement" approach, we can see in Figure 5.15 that the BTA were less than 30%. As the Idle/Calling ratio grew, we can see that more mobiles are able to find a relay. There is a sharp increase in the "Found Intermediaries" curve, as more relays are allocated to mobiles requesting a relay. Thus at Idle/Calling ratio of 1, each requesting mobile can get a relay. The amount of coverage increases steadily to about 30%. Increasing Idle/Calling ratio beyond that increases coverage to almost 48%. In case of HBD the BTA are nearly 60% for all cases of Idle/Calling ratios. Again as Idle/Calling ratio reaches 1, we get the highest possible increase in coverage.
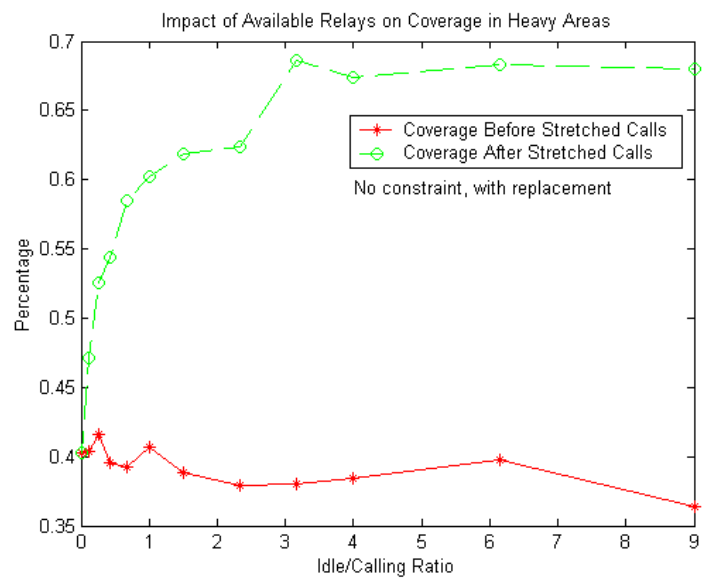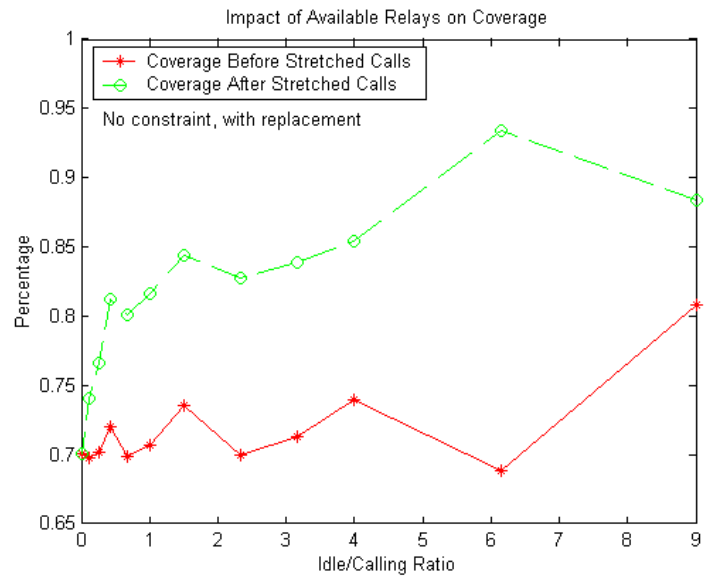
Approach 2: No constraint, with replacement
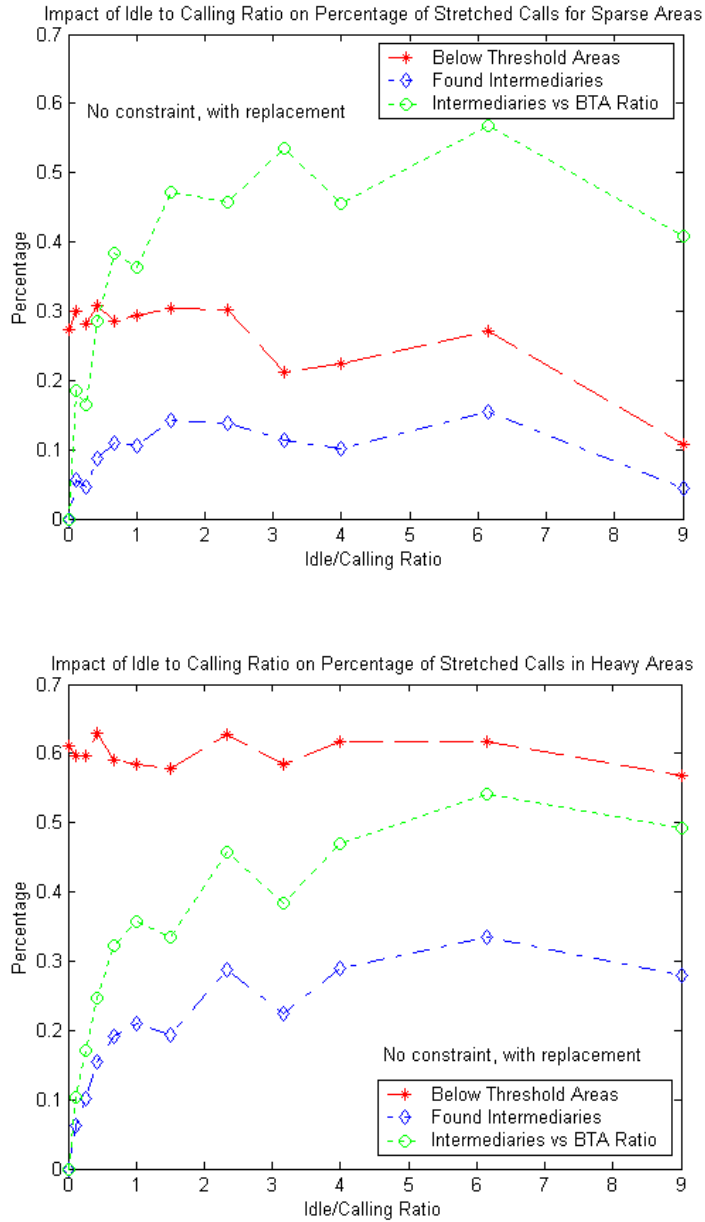


Figure 5. 16 No constraint, with replacement approach; Impact of available relays in LBD and HBD areas

For "No constraint, with replacement" approach in Figure 5.16, we expect

some improvement in the amount of coverage as the same sized pool of idle mobiles is

available for idle nodes. As can be seen in the LBD case, we do get upto 55% coverage. In the HBD case the improvement is nearly 51%. Again the coverage rises until the Idle to calling ratio is about 6 for both HBD and LBD cases.

Approach 3: With constraint, with replacement
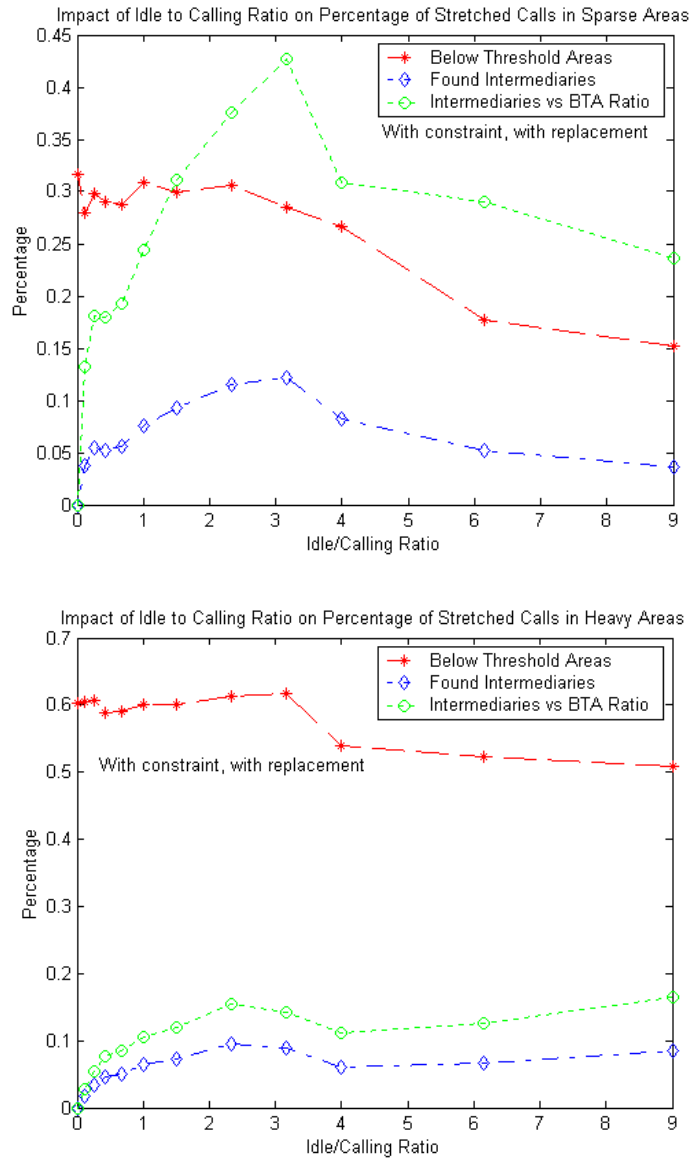




Figure 5. 17 With constraint, with replacement approach; Impact of available relays in LBD and HBD areas

For the "With constraint, with replacement" approach in Figure 5.17, we expect less coverage than in earlier cases as the constraint will limit the relays that are chosen. This is reflected in this case as the coverage drops to the maximum of 18%. In the LBD case, the coverage rises to nearly 40% till the Idle/calling ratio rises to 3 and then drops after that.

Approach 4: With constraint, no replacement
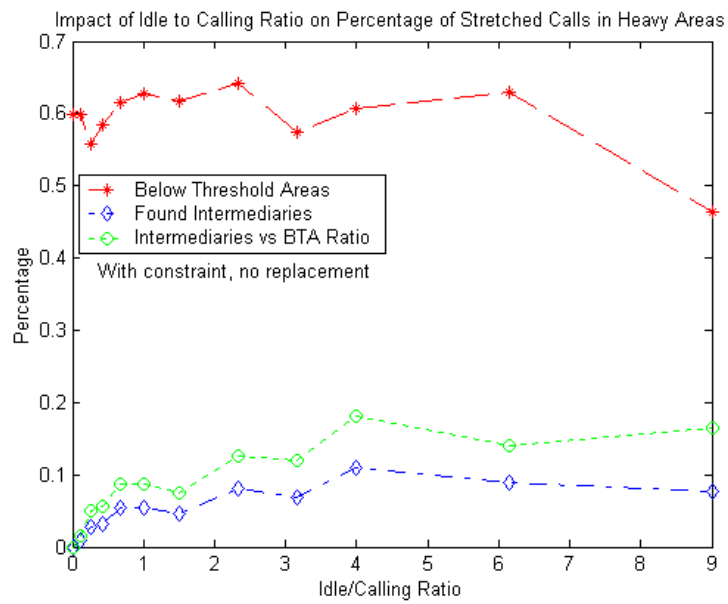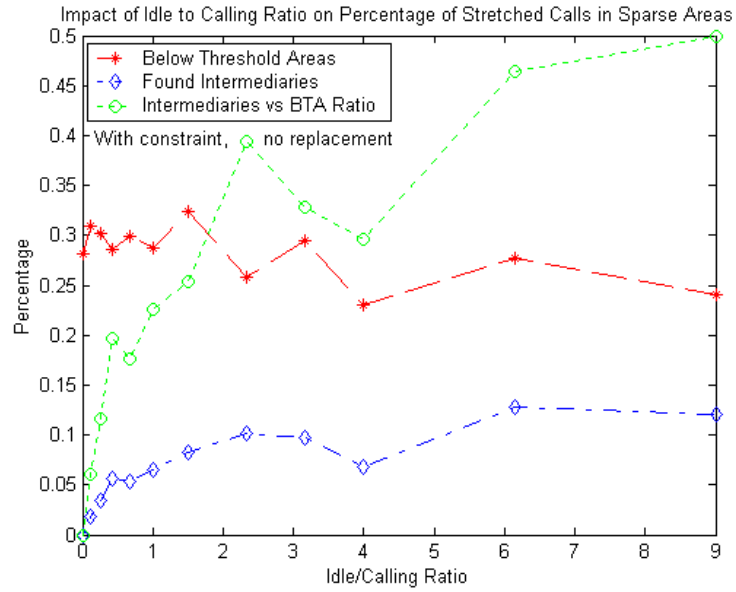


Figure 5. 18  With constraint, no replacement approach; Impact of available relays in LBD and HBD areas

For the "With constraint, no replacement" approach in Figure 5.18, we see that more intermediaries are found as the Idle to calling ratio keeps on increasing, reaching 50% for Idle/Calling ratio of 9. For the HBD case we do not see much improvement compared to the earlier case.

Here, before the ratio hits one, there is a steady increase in the number of intermediaries that were found. This also corresponds to a higher number of calling nodes. This holds lot of promise, as the more crowded a cell is, a little relaxing of call density can bring about greater probability of finding a carrier.

5.5.3 Impact of smaller values of Idle to Calling Ratio

We explore the steepest region of "Found Intermediaries" in the previous figures. In the following figures, all approaches in LBD show steady rise in coverage as the Idle/Calling ratio increases from 0 to 3.5.

Approach 1: No constraint, no replacement
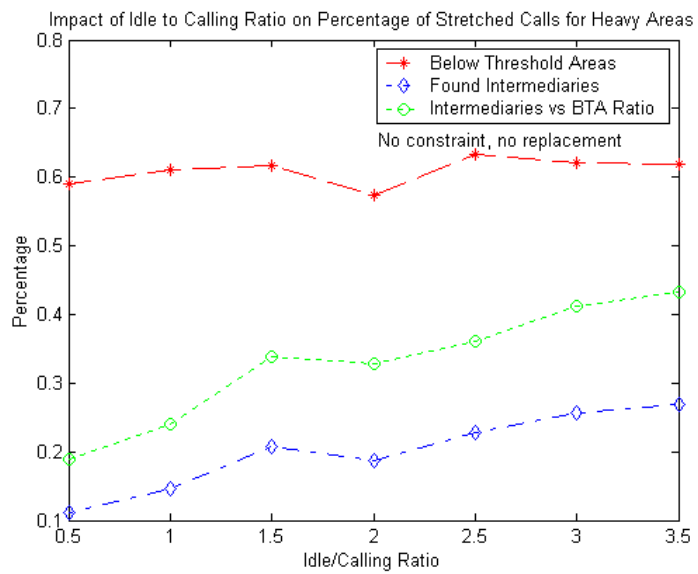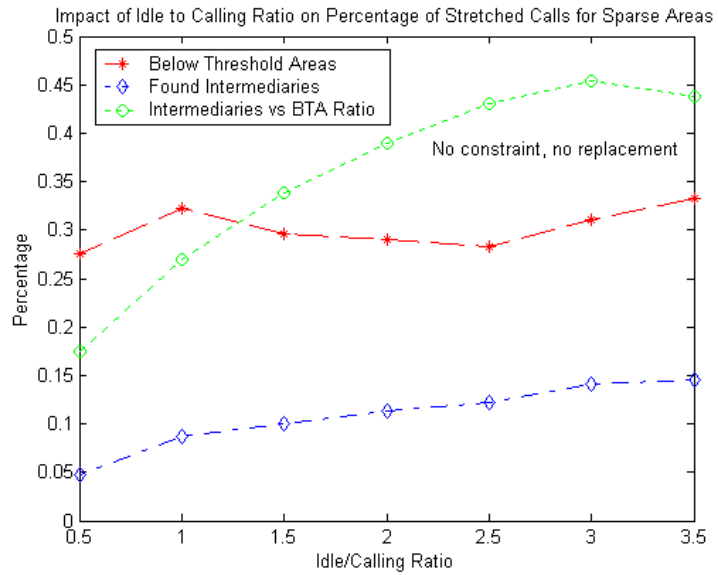


Figure 5. 19  No constraint, no replacement approach; Impact of available relays in LBD and HBD areas

Approach 2: No constraint, with replacement


Impact of Idle to Calling Ratio on Percentage of Stretched Calls in Sparse Areas


Impact of Idle to Calling Ratio on Percentage of Stretched Calls in Heavy Areas
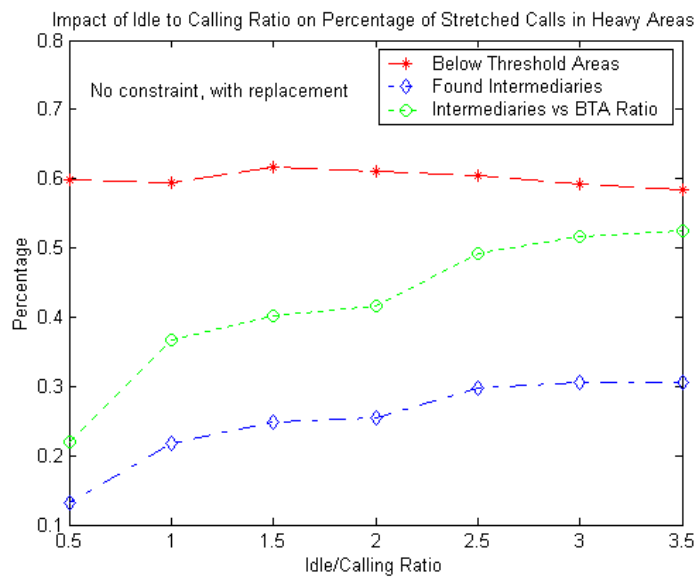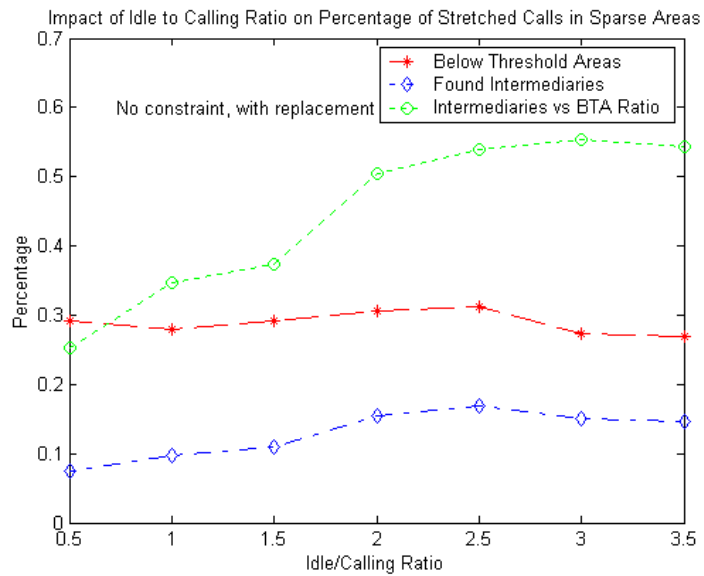
Figure 5. 20  No constraint, with replacement approach; Impact of available relays in LBD and HBD areas

In "No constraint" based approaches in Figures 5.19 and 5.20, we see a higher rise in coverage, with the curve almost leveling off after Idle/Calling ratio reaches 2.5. The coverage reaches almost 45% in the "No constraint, no replacement" approach, while it reaches almost 55% in the "No constraint, with replacement" approach.

Approach 3: With constraint, with replacement
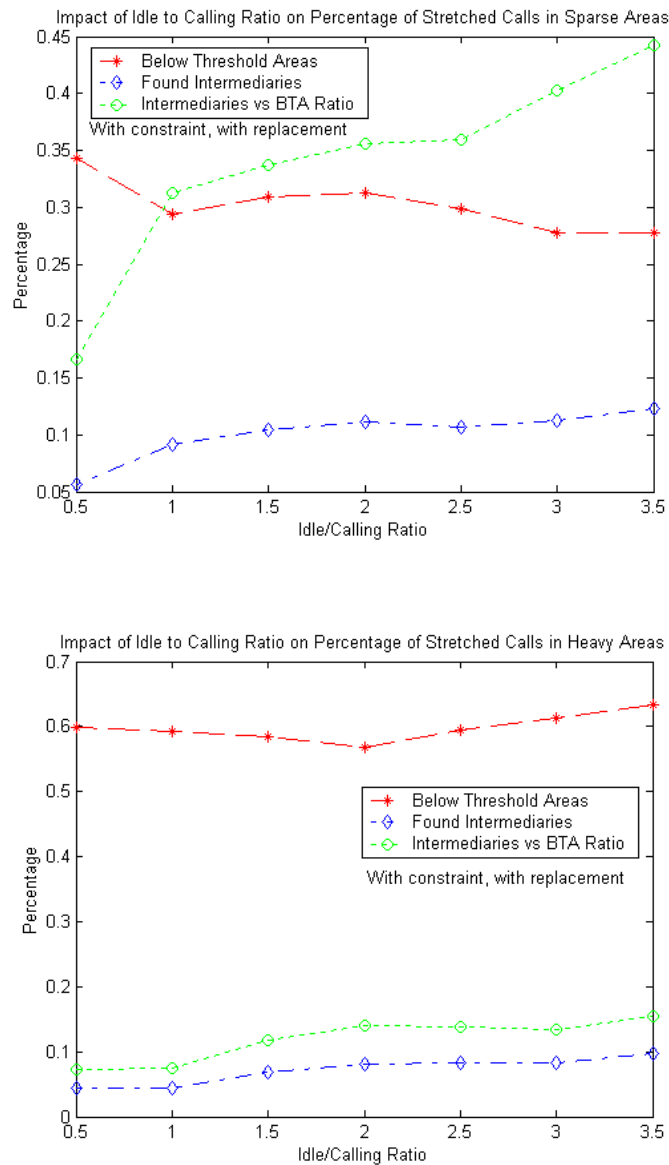




Figure 5. 21  With constraint, with replacement approach; Impact of available relays in LBD and HBD areas

Approach 4: With constraint, no replacement



Impact of Idle to Calling Ratio on Percentage of Stretched Calls in Sparse Areas



Impact of Idle to Calling Ratio on Percentage of Stretched Calls in Heavy Areas
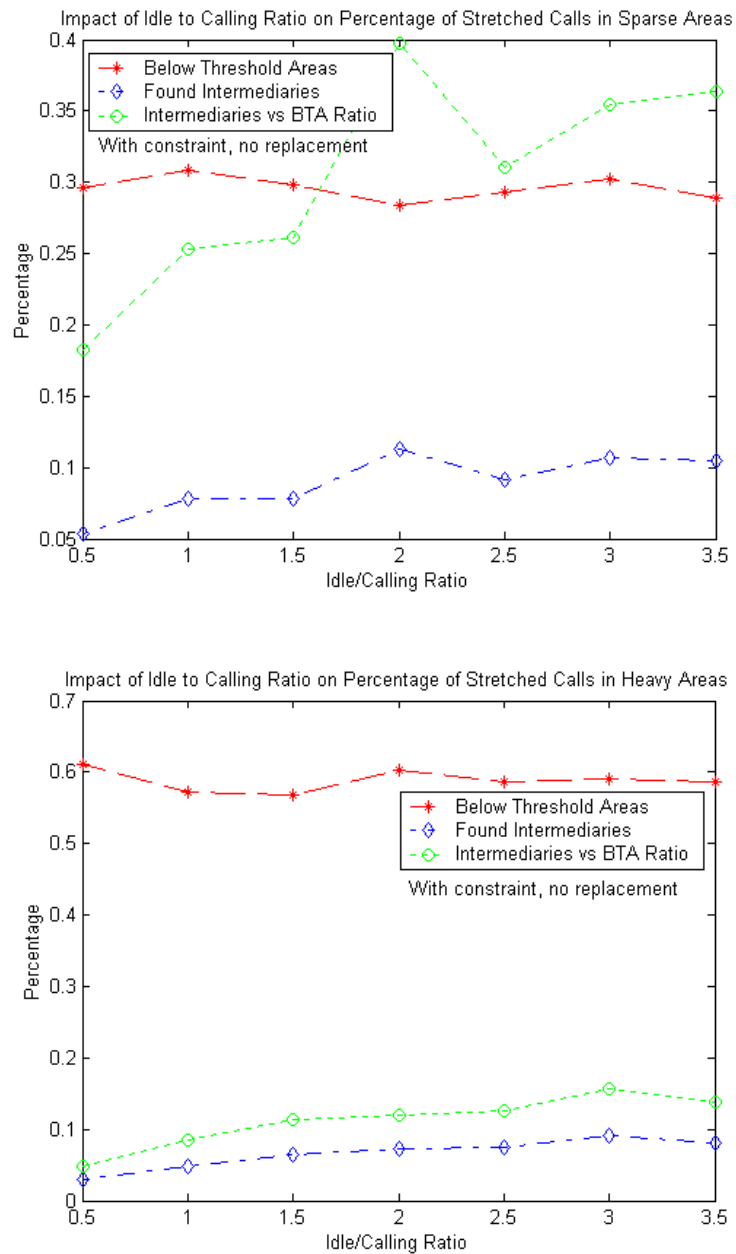
Figure 5. 22  With constraint, no replacement approach; Impact of available relays in LBD and HBD areas

In contrast, the "With constraint" based approaches, in Figures 5.21 and 5.2, show coverage reaching 45% in the "With constraint, with replacement" approach and

almost 35% in the "With constraint, no replacement" approach. The HBD case shows more increase in the "No constraint" based approaches and has highest values of nearly 40% in the "No constraint, no replacement" approach and highest values of almost 50% in the "No constraint, with replacement" approach. The "With constraint" based approaches severely limit the coverage in HBD cases, with coverage limited to a maximum of 15% in the "With constraint, with replacement approach" and a maximum of 12% in the "With constraint, no replacement" approach.

These figures show that the constraint is the primary hindrance in the effort to find intermediaries. When we use the constraint, even a "with replacement" policy can't give more than 10% intermediaries. Without the constraint, for higher ranges of Idle to Calling ratio, up to 50% of nodes falling in BTAs are able to find intermediaries, for both LBD and HBD.

We have shown a significant improvement in the reliability of coverage by using a stretched call, in which an intermediary is used to carry calls from mobiles located in BTAs. We considered heavy and sparse areas and computed the probability of finding an intermediary when the mobile is at a below-threshold area. Our results indicate that with an increase in idle to calling ratio, the coverage increases as does the percentage of stretched calls.

# 6. Interference in interference limited systems

A common measure of link quality is the Signal to Noise (SNR) or Signal to Noise plus Interference (S/[N+I]) ratio [Jones and Skellern, 1995]. In CDMA systems, the sources of interference lie both inside and outside the cell. As all cells use the same frequency, co-channel interference arises. In the forward link all mobiles have their Walsh code, which is orthogonal to all other codes in the cell, and thus no interference is expected. However, multipath contributes towards intra-cell interference [Lee and Miller, 1998]. In the reverse link, codes are not orthogonal and as a result give rise to same-cell interference. The signals generated in a cell permeate other cells giving rise to other-cell interference.

When relays are introduced in interference limited systems, it is imperative to know whether they themselves will contribute to the interference in the system. So we explore single cell interference in more detail. In order to analyze the interference in the cell we need to look at two main scenarios, based on whether or not the "constraint" is applied. The "constraint", as defined in the earlier chapter, refers to choosing a relay based on lowering the overall power consumption.

Scenario I: The system applies the constraint for choosing the relay.

This scenario is applicable when a mobile needs a relay to conserve its battery as well as overall system power.

Scenario II: The system does not apply the constraint for choosing a relay.

This scenario is applicable when a mobile is in a dead spot and needs to connect. The chosen relay may lie farther away from the base station.

We had earlier identified two types of implementations for relays – in the first we assume that the relay to mobile link is under power control and in the second we drop this assumption. We now evaluate system-wide power control by means of constrained optimization, and consider how the addition of relays can affect the optimization. We will see how the above two different implementations affect power control. We will then use the concept of outage contours to give us a realistic idea of relay based interference.

## 6.1 Power constrained optimization

The Signal to Noise Ratio (SNR) for node i is given as,

$$\left(\frac{E_b}{N_o}\right)_i = \frac{W}{R_i} \frac{h_i P_i}{\sum_{j \neq i} h_j P_j + \eta_o W}$$

6. 1

Where $\eta_o$ is the background additive white Gaussian noise with one sided power spectral density, $h$ is the channel gain, W is the bandwidth and R is the rate.

Let vector of $\dfrac{E_b}{N_0}$ for N nodes be denoted by

$$\Gamma = [\gamma_1, \gamma_2, ...., \gamma_N]$$

6. 2

Each user specifies a maximum power limit and a minimum rate requirement.

Power limits are represented by,

$$p = [p_1, p_2, ...., p_N]$$

6. 3

Rate limits are represented by,

$$\vec{r} = [r_1, r_2, ..., r_N]$$

6. 4

Channel gains for users are specified by the vector $\vec{h}$.

In a single cell CDMA system, power control is formulated as follows:

For, i = 1,2,…,N, we have

$$\frac{W}{R_i} \frac{h_i P_i}{\sum\limits_{j \neq i} h_j P_j + \eta_o W} \geq \gamma_i$$

6. 5

The power and rate constraints are,

$$0 < P_i \leq p_i$$
$$R_i \geq r_i$$

6. 6

The objective is to minimize,

$$\sum_{i=1}^{N} P_i$$

6. 7

Subject to above constraints for P and R. This leads to

$$\vec{A}\vec{P}^* = \eta_o W \vec{1}$$

6. 8

where

$$\vec{P}^* = \left[P_1^*, P_2^*, P_3^*, \ldots, P_N^*\right]$$

6. 9

is the optimal vector and $\vec{1}$ is a unity matrix

$$\vec{1} = \left[1,1,1,\ldots, 1\right]^T$$

6. 10

$$\vec{A} = \begin{pmatrix} \dfrac{Wh_1}{r_1'} & -h_2 & \ldots & -h_N \\ -h_1 & \dfrac{Wh_2}{r_2'} & \ldots & -h_N \\ \ldots & \ldots & \ldots & \ldots \\ -h_1 & -h_2 & \ldots & \dfrac{Wh_N}{r_N'} \end{pmatrix}$$

6. 11

To simplify,

$$W \frac{h_i P_i^*}{\sum\limits_{j \neq i} h_j P_j^* + \eta_o W} = \gamma_i r_i$$

6. 12

$$W h_i P_i^* = \gamma_i' \left( \sum\limits_{j \neq i} h_j P_j^* + \eta_o W \right)$$

6. 13

Where,
$$\gamma_i' = \gamma_i r_i$$

6. 14

$$\frac{W h_i P_i^*}{\gamma_i'} - \sum\limits_{j \neq i} h_j P_j^* = \eta_o W$$

6. 15

Expanding the equation,

$$\frac{W h_1 P_1^*}{\gamma_1'} - h_2 P_2^* - h_3 P_3^* - h_4 P_4^* - \ldots\ldots - h_N P_N^* = \eta_o W$$

$$- h_1 P_1^* - \frac{W h_2 P_2^*}{\gamma_2'} - h_3 P_3^* - h_4 P_4^* - \ldots\ldots - h_N P_N^* = \eta_o W$$

$$- h_1 P_1^* - h_2 P_2^* - \frac{W h_3 P_3^*}{\gamma_3'} - h_4 P_4^* - \ldots\ldots - h_N P_N^* = \eta_o W$$

$$- - - -$$

$$- h_1 P_1^* - h_2 P_2^* - h_3 P_3^* - h_4 P_4^* - \ldots\ldots - \frac{W h_N P_N^*}{\gamma_N'} = \eta_o W$$

6. 16

Subtracting the above equations from each other (second from first, third from second,

etc.) we get,

$$\left( \frac{W h_1}{\gamma_1'} - (-h_1) \right) P_1^* + \left( \frac{W h_2}{\gamma_2'} - (-h_2) \right) P_2^* = 0$$

$$\left( \frac{W h_2}{\gamma_2'} - (-h_2) \right) P_2^* + \left( \frac{W h_3}{\gamma_3'} - (-h_3) \right) P_3^* = 0$$

$$- - -$$

$$\left( \frac{W h_{N-1}}{\gamma_{N-1}'} - (-h_{N-1}) \right) P_{N-1}^* + \left( \frac{W h_N}{\gamma_N'} - (-h_N) \right) P_N^* = 0$$

6. 17

122

Simplifying,

$$\left(\frac{Wh_1}{\gamma'_1} + h_1\right)P_1^* = \left(\frac{Wh_2}{\gamma'_2} + h_2\right)P_2^*$$

$$\left(\frac{Wh_2}{\gamma'_2} + h_2\right)P_2^* = \left(\frac{Wh_3}{\gamma'_3} + h_3\right)P_3^*$$

$$- - -$$

$$\left(\frac{Wh_{N-1}}{\gamma'_{N-1}} + h_{N-1}\right)P_{N-1}^* = \left(\frac{Wh_N}{\gamma'_N} + h_N\right)P_N^*$$

6. 18

Substituting the values in the above equations, we get,

$$\frac{Wh_1}{\gamma'_1}P_1^* - h_2\frac{\left(\frac{Wh_1}{\gamma'_1} + h_1\right)}{\left(\frac{Wh_2}{\gamma'_2} + h_2\right)}P_1^* - h_3\frac{\left(\frac{Wh_1}{\gamma'_1} + h_1\right)}{\left(\frac{Wh_3}{\gamma'_3} + h_3\right)}P_1^* - \ldots - h_2\frac{\left(\frac{Wh_1}{\gamma'_1} + h_1\right)}{\left(\frac{Wh_2}{\gamma'_2} + h_2\right)}P_1^* = \eta_o W$$

$$\frac{Wh_1}{\gamma'_1}P_1^* + h_1 P_1^* - h_1 P_1^* - \left(\frac{Wh_1}{\gamma'_1} + h_1\right)P_1^*\left[\frac{1}{\left(1 + \frac{W}{\gamma'_2}\right)} + \frac{1}{\left(1 + \frac{W}{\gamma'_3}\right)} + \ldots + \frac{1}{\left(1 + \frac{W}{\gamma'_N}\right)}\right] = \eta_o W$$

$$\left(1 + \frac{W}{\gamma'_1}\right)P_1^* h_1\left[1 - \sum_{j=1}^{N}\frac{1}{\left(1 + \frac{W}{\gamma'_j}\right)}\right] = \eta_o W$$

6. 19

This means positivity of P* implies the following condition,

123

$$\left[ 1 - \sum_{j=1}^{N} \frac{1}{\left(1 + \dfrac{W}{\gamma_j'}\right)} \right] > 0$$

$$\sum_{j=1}^{N} \frac{1}{\left(1 + \dfrac{W}{\gamma_j'}\right)} < 1$$

6. 20

With relay, we have

$$\left(\frac{E_b}{N_o}\right)_i = \frac{W}{R_i} \frac{h_r P_r}{\sum_{j \neq i,r} h_j P_j + \eta_o W + I_{relay}}$$

6. 21

The first part of the interference expression is due to the signals from all mobiles except the relay and the current mobile. The last part of the expression, namely, $I_{relay}$, is due to the mobile to relay transmission.

Problem formulation with QoS constraints:

If a relay exists, 'r' denotes the relay,

$$\frac{W}{R_r} \frac{h_r P_r}{\sum_{j \neq i,r} h_j P_j + \eta_o W + I_{relay}} \geq \gamma_r$$

6. 22

If a relay does not exist, for all i= 1, …, N

$$\frac{W}{R_i} \frac{h_i P_i}{\sum_{j \neq i} h_j P_j + \eta_o W} \geq \gamma_i$$

6. 23

All other mobiles then see the relay as just another mobile.

Power and rate constraints are as follows with the last constraint only applicable for Scenario I.

124

$$0 < P_i \leq p_i$$
$$R_i \geq r_i$$
$$p_{i-r} + p_r < p_i \qquad\qquad 6.\,24$$

where, $p_{i-r}$ denotes the mobile to relay connection. The last constraint, as given by

Scenario I, conveys that the relay is only chosen when sum of $p_{i-r}$ and relay power,

$p_r$, is less than the original mobile to base station power, $p_i$. Thus, we need to,

Minimize

$$\sum_{i=1}^{N} P_i \qquad\qquad 6.\,25$$

Subject to the above constraints.
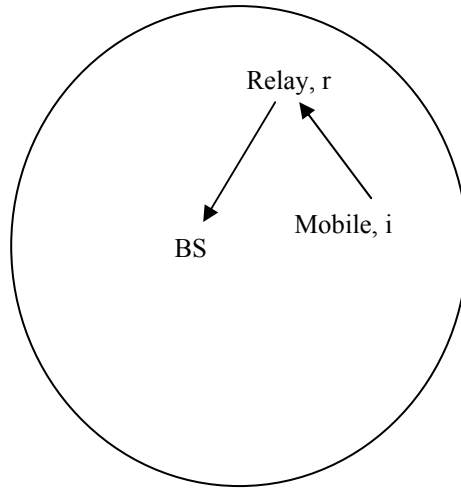


Figure 6. 1  Choice of relay for mobile

Solve the linear equations,

$$\frac{W}{R_i} \frac{h_i P_i^*}{\sum\limits_{j \neq i,r} h_j P_j^* + \eta_o W + k I_{relay}} = \gamma_i \qquad\qquad 6.\,26$$

Where k= 0, if a relay is not involved, k=1, if a relay is involved in the transmission,

and r is included if such a relay exists.

$$W \frac{h_i P_i^*}{\displaystyle\sum_{j \neq i,r} h_j P_j^* + \eta_o W + kI_{relay}} = \gamma_i R_i$$

$$Wh_i P_i^* = \gamma_i' \left( \sum_{j \neq i,r} h_j P_j^* + \eta_o W + kI_{relay} \right)$$

$$\frac{Wh_i P_i^*}{\gamma_i'} - \sum_{j \neq i,r} h_j P_j^* - kI_{relay} = \eta_o W$$

6. 27

Let first mobile's call be relayed by mobile number two (the relay), and third

mobile's call be relayed by the fourth mobile. Then we have the equations:

$$\frac{Wh_1}{\gamma_1'} P_{relay1}^* - h_2 P_2^* - h_3 P_{relay3}^* - \ldots\ldots - h_N P_N^* = \eta_o W$$

$$( \;) + \frac{Wh_2}{\gamma_2'} P_2^* - h_3 P_{relay3}^* - \ldots\ldots - h_N P_N^* - \underline{\underline{h_1 P_{relay1}^*}} = \eta_o W$$

$$- h_1 P_{relay1}^* - h_2 P_2^* + \frac{Wh_3}{\gamma_3'} P_3^* - \ldots\ldots - h_N P_N^* = \eta_o W$$

$$- h_1 P_{relay1}^* - h_2 P_2^* - ( \;) + \frac{Wh_4}{\gamma_4'} P_4^* - \ldots\ldots - h_N P_N^* - \underline{\underline{h_3 P_{relay3}^*}} = \eta_o W$$

$$- - - - -$$

$$- h_1 P_{relay1}^* - h_2 P_2^* - h_3 P_{relay3}^* - \ldots\ldots + \frac{Wh_N}{\gamma_N'} P_N^* = \eta_o W \qquad \text{6. 28}$$

The double underlined items are the interference due to relays even though a

shorter distance gives a higher path gain resulting in smaller power requirements. But

here we are replacing the smaller power with the still smaller power of the mobile-to-

relay connection. Therefore, for simplicity's sake we have not changed the path gain

of the mobile to relay connection.

Earlier we had,

$$P^* = \begin{pmatrix} P_1^* \\ P_2^* \\ P_3^* \\ . \\ . \\ . \\ P_N^* \end{pmatrix}$$

6. 29

Now we have,

$$P_{new}^* = \begin{pmatrix} P_{relay1}^* \\ P_2^* \\ P_{relay3}^* \\ . \\ . \\ . \\ P_N^* \end{pmatrix}$$

6.    30

Scenario I

Since,  $\quad P_{relay1}^* < P_1^*$  and  $P_{relay3}^* < P_3^*$

So,  $\quad\quad \sum P_{new}^* < \sum P^*$

6. 31

Also same the condition for the equations exist, i.e., positivity of $P_{new}^*$ implies the following condition,

$$\left[ 1 - \sum_{j=1}^{N} \frac{1}{\left(1 + \dfrac{W}{\gamma_j^{'}}\right)} \right] > 0$$

$$\sum_{j=1}^{N} \frac{1}{\left(1 + \dfrac{W}{\gamma_j^{'}}\right)} < 1$$

6. 32

Whether $P_{relay}^{*}$ is under power control or not, it satisfies the power constraint and so its value will always be less than the original mobile's power. Being under power control will only mean that the base station can optimize its value. This means that we have not changed power control equations but the overall system power is less. This leads to less inter and intra cell interference.

Scenario II

Since the relay's power may or may not be less than the mobile's power,

$$\sum P_{new}^{*} \leq \sum P^{*}$$

or

$$\sum P_{new}^{*} \geq \sum P^{*}$$

6. 33

Here also, equation 6.32 applies based on positivity of $P_{new}^{*}$. So even though power control is maintained, this implies a smaller value for mobile and relay power. Let us solve the set of equations numerically to see their behavior. Let there be only two mobiles in the cell and a relay is introduced. We need to see the power values of the mobiles before and after introduction of the relay. Since these power equations are non-linear, a slack variable is introduced to make them linear.

```
% COMMENT
% Set of equations
a =
   4.8000   2.2000   1.0000        0
   2.7000   6.6000        0   1.0000
b =
   10   10   10   10
c = a/b
c =
   0.2000
   0.2575

% add a relay
a1 =
   4.8000   2.2000   1.4000   1.0000        0        0
   2.7000   6.6000   0.6000        0   1.0000        0
   1.2000   3.5000   4.7000        0        0   1.0000
b1 =
   10   10   10   10   10   10
c1 = a1/b1
c1 =
   0.1567
   0.1817
   0.1733
```

This example in Matlab shows solving linear equations to get the power values given by arrays c and c1. The target SIR is taken to be 10 (not exact units). Mobile 1's power is reduced from 0.2 to 0.1567 and mobile 2's power is reduced from 0.2575 to 0.1817. So the mobiles reduce their power when the relay is introduced to maintain the same target SIR at the base station. The mobiles adjust their power level whenever the relay satisfies either of the scenarios.

This brings us to the necessity of power control for the relay-mobile connection. If the mobile continues to send its signal with the same power, then we don't have any savings and it will add to the interference. The new signal power for Scenario I should satisfy,

$$P_R + P_{MR} < P_M$$
$$P_{MR} < P_M - P_R$$

6. 34

So, open loop power control for the relay should start with the $P_{MR}$ value. The base station informs the relay to transmit at this power. Closed loop power control can be achieved if the relay and mobile are able to coordinate through Power Control Group (PCG) slots, reducing their power even further.

We also distinguish between the range of the cell, i.e., coverage and useful service area of coverage.  If we use relays to increase coverage, we may add to the total interference. However, we are more concerned with the useful service area, and any additional power has to be under power control. Increase of power at a certain radius may mean that those nearer to the base station have to alter their power level, but it all balances out as shown by the above linear equations,  since we have seen that addition of a relay is just like adding another mobile in the cell.
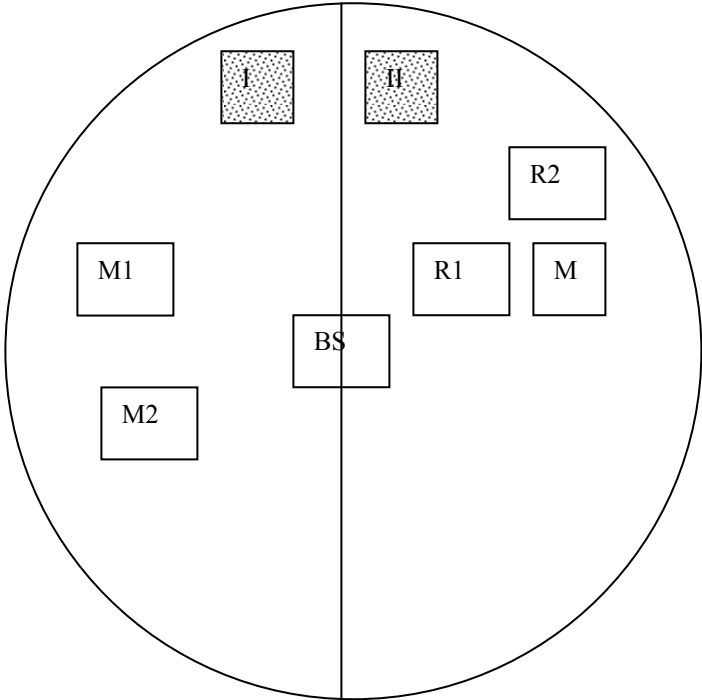


Figure 6. 2  Spatial effect for choice of relay

Now let us explore the impact of different locations of the relay as in Figure 6.2. There are two halves in the cell given by I and II. Mobile M needs a relay. So according to Scenario I, it will choose relay R1, but according to Scenario II, it will choose relay R2. Now it does seem that R1 saves power, and so there should be less interference in the system. On the other hand, relay R2 lies further away from the mobiles in the Ist half of the cell, M1 and M2. So, R2 will give less interference to mobiles M1 and M2. This shows that not only the constraint but spatial locations of mobiles affect the amount of interference in the cellular system. We also note that if the range of the relay is constrained to one half of the cell, only then will we really see less total system power consumption. Note that the power control equations are global (i.e., cell-wide), but the choice of the relay range can supplement the minimization of system power. So we explore this in terms of outage contours to see the effect of a mobile on total cell interference.

## 6.2 SIR and Outage Contours

In interference-limited environments interference exceeds the amount of noise. Due to interference, when a link cannot be maintained at a desired quality, outage occurs [Jones and Skellern, 1995]. It may be noted here that [Cook, 1987] and [Jones and Skellern, 1995] analysis is for free space path propagation where the path loss exponent is two, resulting in nice circular contours. To incorporate higher path loss components makes the derivation of link equations for more than one interferer

complex (see Appendix B). So we try to analyze the effect of more relays via a computer simulation.

As shown by the Table 6.1, Universal Mobile Telecommunication System (UMTS) spectrum allocation shows that Frequency Division Duplexing (FDD) uplink and downlink are separated and as such do not interfere with each other [Laiho, Wacker and Novosad, 2002]. Time Division Duplexing (TDD) and FDD uplink are adjacent. The intra-cell interference in FDD – CDMA cell is less as both uplink and downlink use separate frequency bands. This scenario changes with TDD-CDMA as the same frequency region is being utilized in downlink and uplink. The relay acts as a mini base station in the cell. We plot SIR contours to know how much and how widespread relay induced interference will be. The contours show at each point exactly how much each relay can interfere. This enables us to properly design a CDMA cell. This also helps in locating regions where outage will most likely occur.

The Signal to Interference Ratio (SIR) is defined as desired signal divided by signal received from all interferers [Rappaport, 1996],

$$SIR = \frac{S}{\sum_{int\,erferers} I} \qquad\qquad 6.\,35$$

where S is the signal power required to make connection to the base station and I is the signal interference power due to all the relays. Noise is not considered here, as it is an interference limited system. We have used the Hata based path loss equations and the reverse link power of the mobile is taken as defined in [Lee and Miller, 1998] (see Appendix A).

The region where SIR goes below a threshold of 10 dB can be considered an outage region as the mobile intending to place a call is unable to do so. We plot SIR contours and outage contours. We explore the effect of few selectively placed relays and also study randomly placed relays to get an optimum number of relays. Introduction of one relay creates a local minimum in the cell, as also suggested in the AMPS based study [Jones and Skellern, 1995]. As we add more relays, we can see how they can create local minima. In the simulation the mobiles are randomly distributed and we can see whether the relays increase the interference. This is important if relays are to find the senders or if finding the relays has to be in the jurisdiction of the base station. If the relay's presence increases the interference then it is better for them to be listening devices rather than broadcasting their presence. We want to know if there are any regions where relays should be avoided, and when we reach the optimum number of relays.

| Freq-uency | 1710 - 1785 | | 1805 - 1880 | 1880 - 1900 | 1900 - 1920 | 1920 - 1980 | 2010 - 2025 | | 2110 - 2170 |
|---|---|---|---|---|---|---|---|---|---|
| | GSM 1800 uplink | | GSM 1800 downlink | DECT RX/TX | UTRA TDD RX/TX | UTRA FDD uplink | UTRA TDD RX/TX | | UTRA FDD downlink |

Table 6. 1 UMTS spectrum allocation, Frequency in MHz [3GPP, 1998]
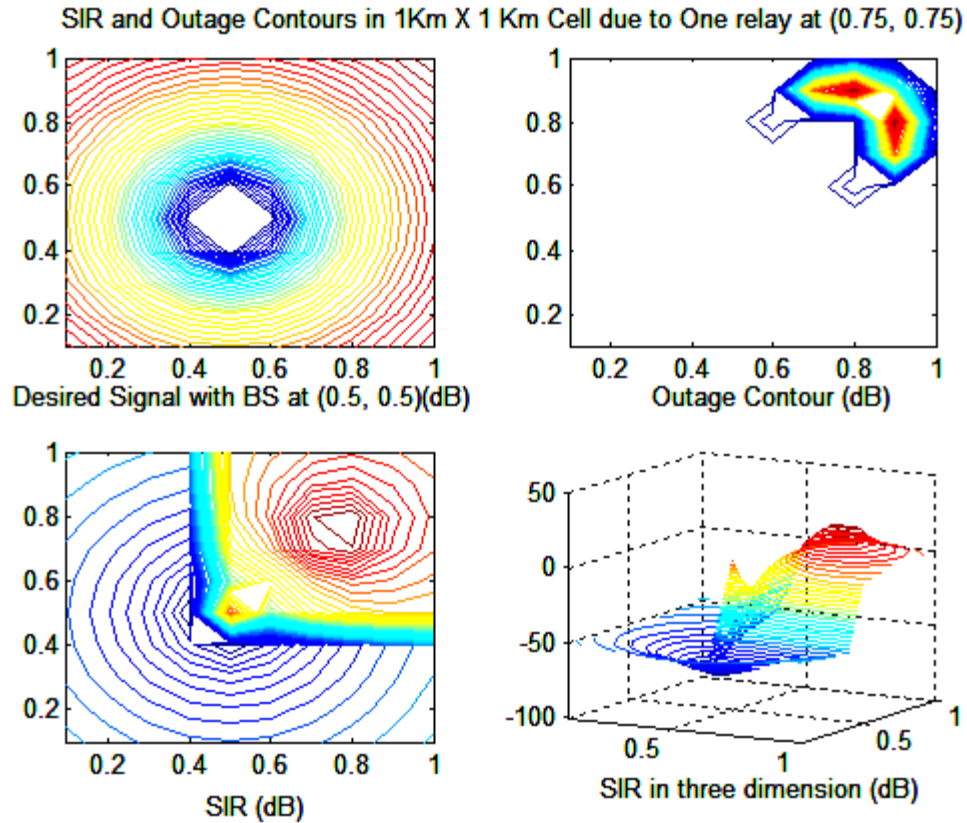
Figure 6. 3  Signal, SIR and Outage contours due to one relay

The cell is 1Km x 1Km in dimensions. The outage is defined when SIR is less than 10dB. The relays are chosen according to Scenario I, but the range of relays is dictated by Scenario II. It is also to be noted that if relays are under power control, it does not matter if the constraint is a factor in the choice of relays. As we saw in the last section, the overall system power is realigned so that the target SIR is maintained. Thus the assumptions for the simulation are:

- FDD CDMA is used as it has separate frequency bands for uplink and downlink.

- Power control at relays is assumed – so relay centric contours show minima at relays.

- Relays here have mobile type maximum sending power, 23dBm, if car type higher power relays are used, they have to limit their transmitting power so as not to create extra interference.

- Relays have the range equal to quarter of the cell. It means that a relay at the periphery can not interfere with a mobile situated at its diametrically opposite end.

The outage region is measured by means of ImageJ [ImageJ, 2001], an image analysis software. This software measures the concerned region pixel-wise and gives the percentage area. The simulation gives a snapshot of SIR and outage over the cell. The densely populated contours are the regions where the SIR or the outage is high. This tells a potential mobile of the amount of interference it might see.
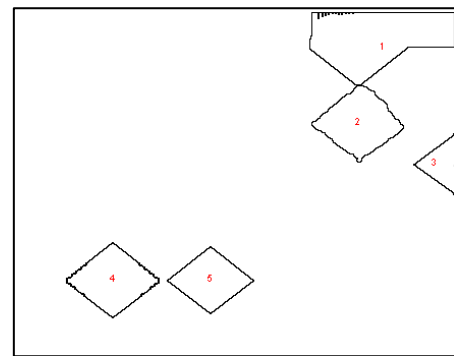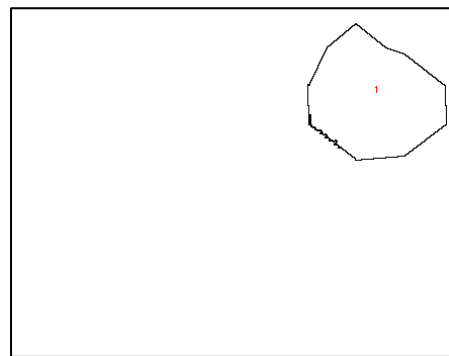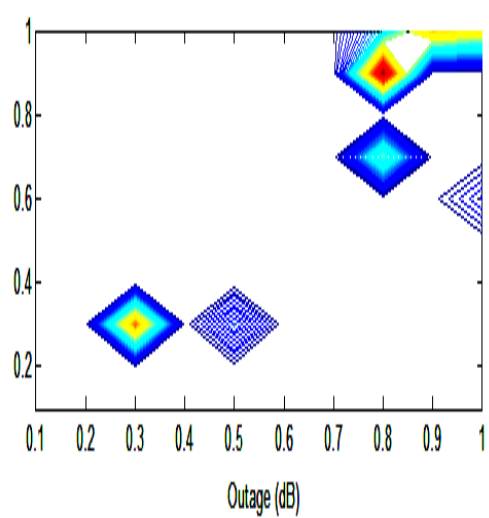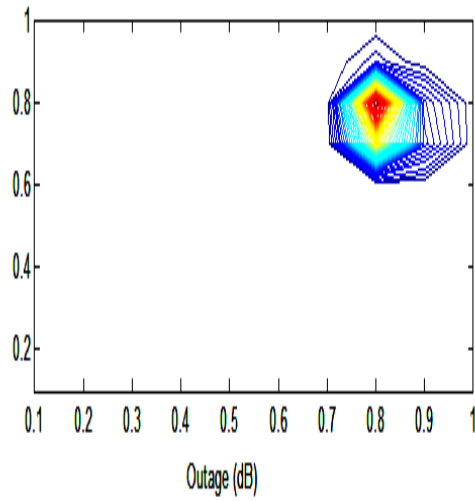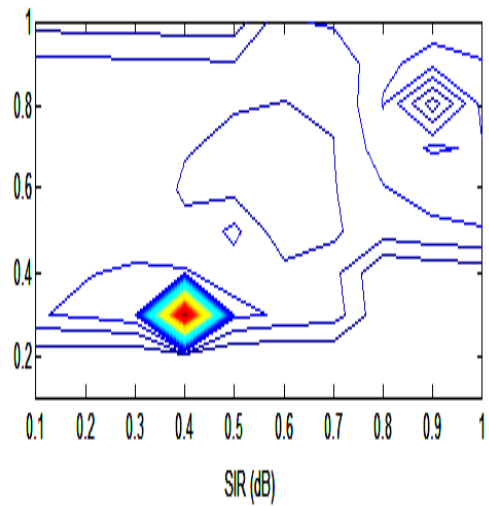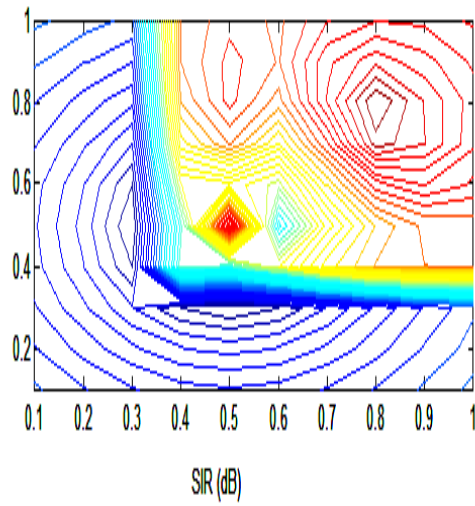
Left Column:

Figure 6. 4  SIR and outage contour due to two relays in same quadrant

Right Column:

Figure 6. 5  SIR and outage contour due to two relays in opposite quadrants
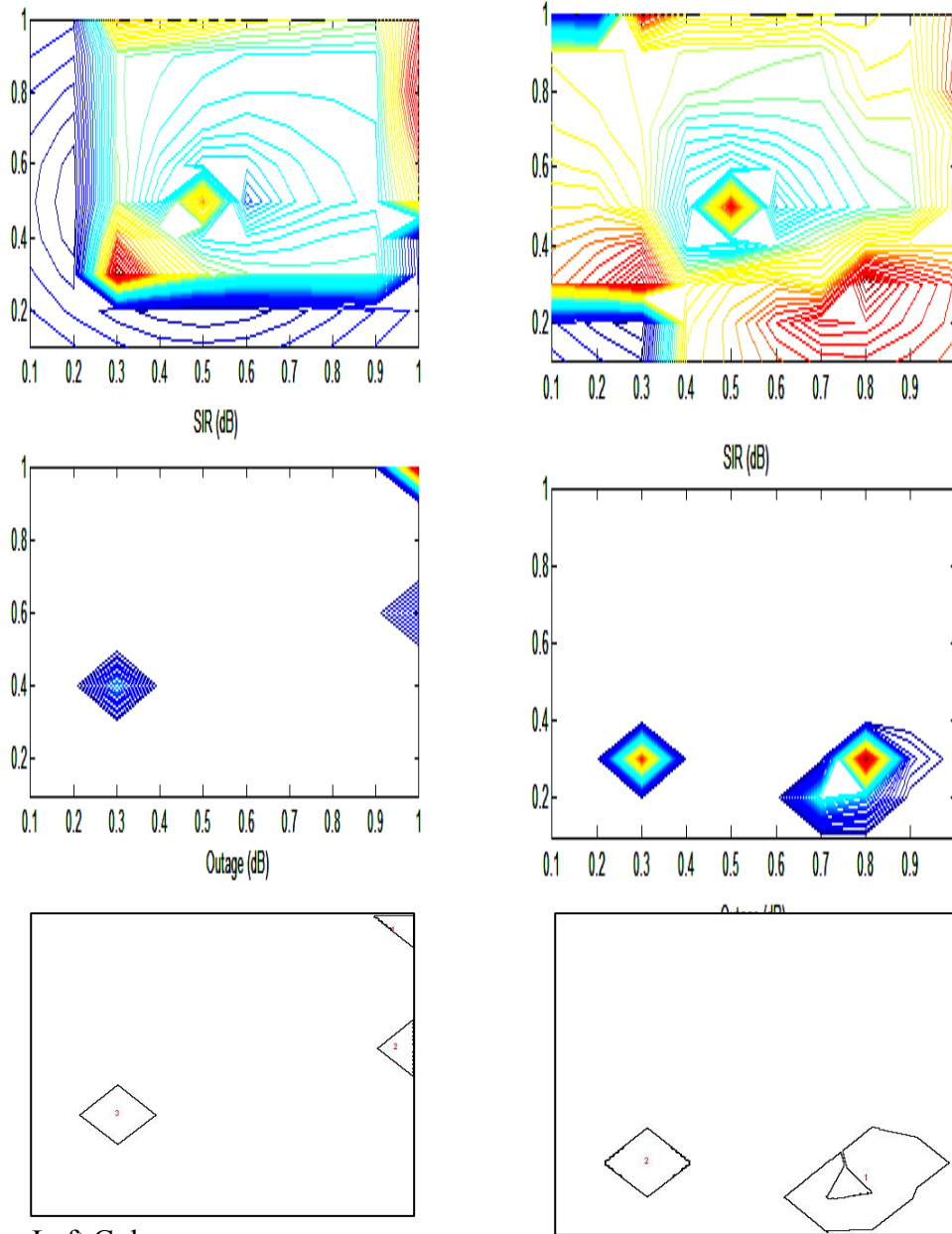
Left Column:
Figure 6. 6  SIR, outage contour and outage region due to three relays in three opposite quadrants

Right Column:
Figure 6. 7  SIR, outage contour and outage region due to four relays in four opposite quadrants
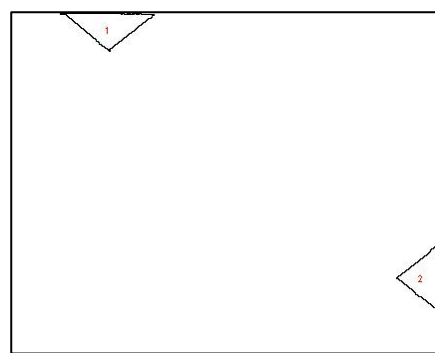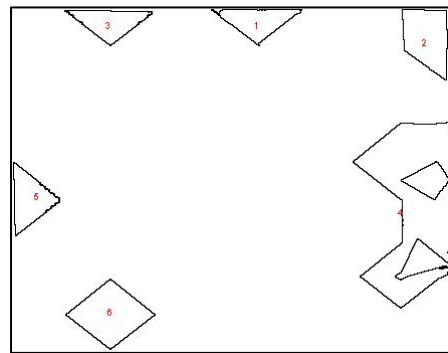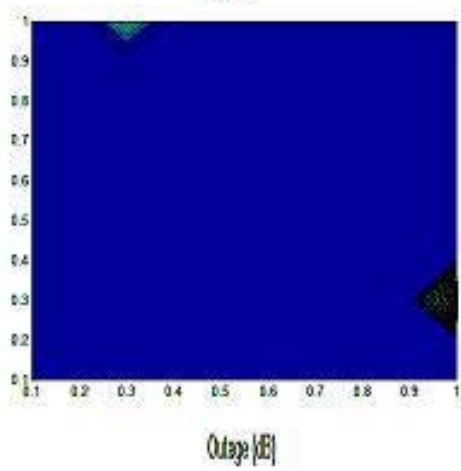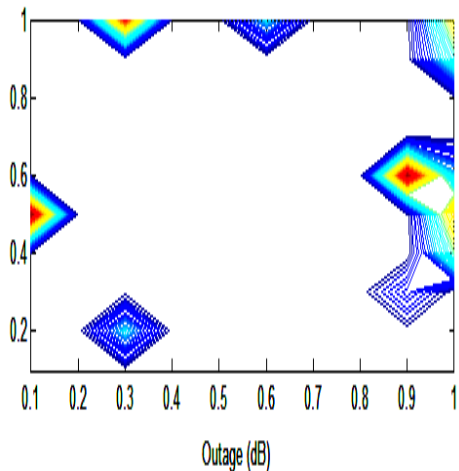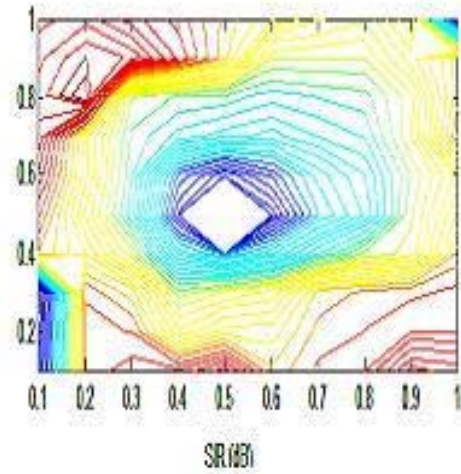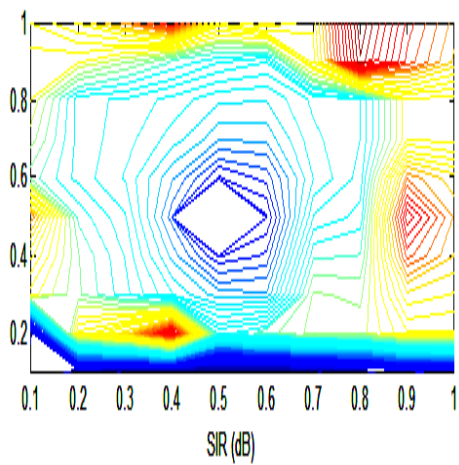
Left Column:

Figure 6. 8  SIR, outage contour and outage region due to five randomly placed relays

Right Column:

Figure 6. 9  SIR, outage contour and outage region due to ten randomly placed relays

SIR (dB)
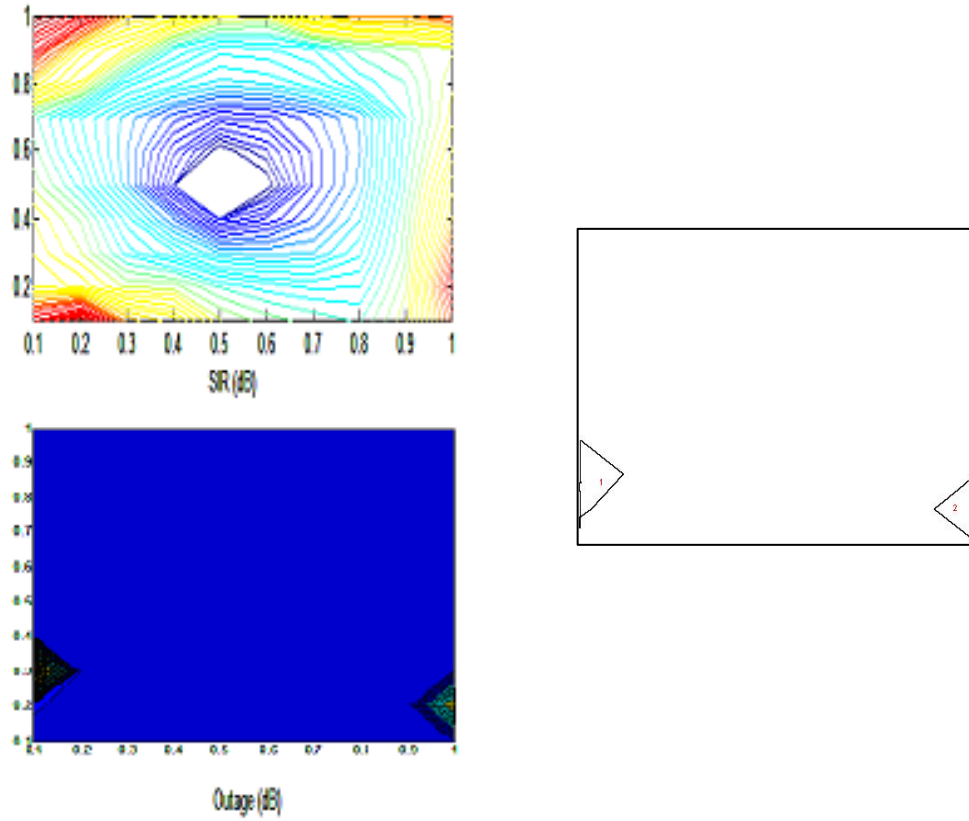


Outage (dB)

Figure 6. 10 SIR, outage contour and outage region due to twenty randomly placed relays

| Fixed Relays | Percentage area affected by outage |
|---|---|
| Two in same quadrants | 8.5% |
| Two in opposite quadrants | 12% |
| Three in opposite quadrants | 3.4% |
| Four in opposite quadrants | 9.2% |

Table 6. 2  Area affected by outage for fixed relays

| Randomly placed relays | Percentage area affected by outage |
|---|---|
| Five relays | 2.8% |
| Ten relays | 1.13% |
| Twenty relays | 0.8% |

Table 6. 3  Area affected by outage for random relays

## 6.3 Results and Discussion

The Signal to Interference (SIR) contours show the distortion in signal level. Figure 6.3 shows the desired signal contour, SIR contour and outage contour due to interference introduced by one relay at the top right quadrant. The relay creates a local minimum at its location. In Jones and Skellern [1995], the number of interferers was one and even that could produce major disturbance in the smooth circular curves.

The Figures 6.4-6.10 show SIR and outage plots. We see that the quadrant where the relay is placed suffers the most from outage. Figure 6.4 shows the effect of placing two relays in the same top right quadrant. Figures 6.5-6.6 have relays selectively chosen so as to be distributed in each quadrant. Figure 6.5 shows the effect of placing two relays in opposite quadrants. This creates disturbance in both the quadrants, resulting in more outage area. Figure 6.6 shows the effect of placing three relays in three different quadrants. Figure 6.7 shows the effect of placing four relays in four different quadrants. The outage area due to two relays in opposite quadrants and

four relays in four quadrants is similar. However the outage area for three relays case is very small.

Figures 6.8-6.10 have randomly placed relays. Figures 6.9 and 6.10 outages are shown with filled contour, as it is easy for ImageJ analysis. If we add more relays in the cell, more local minima are created, resulting in outage regions. As the number of relays increases, the area of outage moves towards the periphery of the cell. If the number of relays is less than 10, say 2-3 in each quadrant, then the outage region is more prominent. For a larger number of relays, as shown by randomly placing 20 relays, the cell size gets distorted. It is because interferers are so distributed that each mobile is able to see a similar disturbance. Only the peripheral region of the cell gets uneven amount of interference, resulting in decreasing of cell size. [Fujiwara, Takeda, Yoshino and Otsu, 2002]  also recognize the amount of interference possible. Since their simulation is in the downlink, they do not allocate relays if they are nearer to base station, since the base station overwhelms the signals. Only the mobiles further from the base station are able to relay for others.

Tables 6.2 and 6.3 give the amount of outage area created by relays for fixed relays and randomly placed relays respectively. It is to be noted that the amount of outage keeps on decreasing as the number of relays increases. The more relays or mobiles there are, the less their power will be in order to satisfy the power control equations. Thus the relays in the system have less individual power as well as shorter range. As noted earlier, the outage for higher number of relays is more towards the periphery of the cell, signifying a smaller range of the cell due to relays.

The bound for number of terminals, N, is given by,

$$N \leq 1 + \left( \frac{W}{R} \right) / \gamma^*$$
<div align="right">6. 36</div>

where W is the bandwidth, R is rate and $\gamma^*$ is target SIR. For W= 1.2288 MHz and R = 9600 bits/sec and $\gamma^* = 10$, we get N to be less than 13.8. If the target SIR is lowered to 9, then N is less than 15.2. Thus we see that if power is saved with choice of relays then the cell capacity increases. If only useful service area was increased then the total interference in the system may or may not decrease and the capacity of the system may not be affected.

We deduce the following steps for the selection of relays:

- Create an imaginary line from the major roads crossing the cell. In the roads the path is Line Of Sight (LOS) and the signal travels far, like an urban canyon. We also don't want this LOS region to be the place for outage.

- The circular cell now has four quadrants based on road intersections.

- If there are already 2-3 relays in the quadrant, try another quadrant, so that outage region is not concentrated in one quadrant.

The introduction of relays creates regions of outage in the cell. For a smaller number of relays, these regions can be thought of as mini-cells. So the relays should be so chosen as to avoid the LOS regions to minimize the interference. For a higher number of relays, the interference distorts the cell. To overcome the effects of interference, [Zadeh and Zabbari, 2002] suggest use of more efficient links. A solution can be the use of Multiple-In and Multiple-Out (MIMO) radio channels, where there are more channels between the transmitter and receiver pair by having more antennas

at each site. In the reverse link in the CDMA cell, there is more of a chance of interference as the signals are not orthogonal. [Zadeh and Zabbari, 2002] suggest Multi-user Detection (MUD), where more signals can be detected. Thus the advantages of relaying can be kept while keeping interference at the minimum.

# 7. Utility

In the introduction to this research, we had stated that one way to encourage users to act as relays is to provide them with some financial benefit. In this chapter we study this idea in some more detail. In order to evaluate how users may be motivated to act as relays, we use the economics terminology of 'utility'. Utility is generally defined as the amount of satisfaction that is attained by an individual by receiving some item or service. Thus, if an individual gets rewarded with some amount of money for relaying, that he/she feels more than compensates for the loss of battery power, then that person has a positive utility and will be willing to act as a relay.

The question of how to define utility in a meaningful way is non-trivial because, when we optimize the system based on these utility functions, we would like to ensure that our primary goals of reducing overall power consumption and increasing coverage are met. There have been some papers in the literature that have defined utility functions for the cellular environment; their goals, however, are generally different. Power is one of the parameters to be considered in defining the utility function as is also done in other papers [Gunturi and Paganini, 2003]. [Zhou, Honig and Jordan, 2005] [Lu, Goodman, Wang, Erkip, 2004] use signal to interference ratio and data rate to define a utility function and consider maximizing the total utility for all the users in the system. However, the target of the paper is to arrive at a utility based power control for CDMA networks. [Goodman and Mandayam, 2000] define utility as based on the amount of error free bits transmitted per unit of energy.

Unlike all the previous papers where utility was computed by the base station, in our system we allow each relay to operate independently and selfishly. By selfishness, we mean that a relay can relay as many packets as required to maximize its utility. The reason behind this is that we want to encourage relaying while simultaneously promoting decentralized decision making (i.e., the base station is not the one who unilaterally makes relay allocation decisions). We note, however, that this selfish behavior could very easily cause interference-based outages in other parts of the cell. Thus, we assume that each relay is able to get power control information to minimize the effect of relay-induced interference.

We have until now only considered voice services since most previous cellular technologies were mainly focused on voice services. For maintaining Quality of Service (QoS) in this environment, delay avoidance is the most important parameter for voice services. So for voice the utility will be minimum delay along with SIR and battery power constraints. In such a case relays introduce additional delay and at most one relay can be introduced without exceeding delay requirements. Third-generation cellular services include both voice and data. In this chapter, we focus primarily on data services and we show how relays can enhance utility for data services. For data services, error free transmission is the most important QoS parameter. Keeping this in mind, we evaluate how inclusion of relays can affect the behavior of the system. In data related services, the benefit aspect of utility should include the concept of error free transmission.

$$U \propto f(\Pr(NoError)) \qquad \text{7. 1}$$

where, 'f' is a function of Probability of no error.

If we take into account energy spent for transmission, we can use the utility as defined in [Goodman and Mandayam, 2000]. Utility is based on the number of error free bits transmitted per unit of energy, given here in bits per joule.

$$U = \frac{T(p)}{p}$$

$$\text{7. 2}$$

where T is the throughput giving the number of error free bits transmitted, and p is the power spent for the transmission. If there are M bits in a frame, out of which L is the number of information bits, then there are (M-L) bits for error correction. Throughput for mobile $i$ is defined as

$$\text{7. 3}$$

$$T(p_i) = \frac{LRf(\gamma_i)}{M}$$

where R is the rate of transmission (bits/sec) and $f(\gamma_i)$ denotes the probability of no error for all M bits,

$$f(\gamma_i) = (1 - \Pr ob(error))^M \qquad \text{7. 4}$$

where the probability of error is a function of $\gamma_i$, the signal to interference ratio for mobile i, and the coding used.

## 7.1 Relay based transmission – centralized decision making

Before looking at our selfish model for relay assignment, let us formulate the utility model for the centralized case. This is the model studied by previous researchers where the base station makes relay assignments based on maximizing some global utility function. If a mobile, $i$, is in direct contact with the base station, then its utility is given as,

$$U_i = \frac{LRf(\gamma_i)}{Mp_i} \frac{bits}{Joule}$$

7. 5

Now if relay $j$ is used, it is chosen so as to save power. If $p_{ij}$ is the power required for transmission from the mobile to the relay, then the combined power of $p_j$ and $p_{ij}$ should be less than $p_i$. This indicates that the new utility for the mobile i and relay j is

$$U_i(j) = \frac{LRf(\gamma_i)}{Mp_{ij}} \frac{bits}{Joule}$$
$$U_j = \frac{LRf(\gamma_j)}{Mp_j} \frac{bits}{Joule}$$

7. 6

where $U_i(j)$ indicates that this is the utility when mobile goes through a relay.

*If we compare the utilities of the mobile, we find that new utility when using a relay is larger (since less power is used by it), indicating more satisfaction for the mobile.* Since relays are used for saving power, the total power in the relay-assisted system is less than the total power in a non-relay system, where $U_i(j) + U_j$ is less than $U_i(BS)$, where the mobile goes directly to the base station. *This indicates less interference at other mobiles and thus a lowering of the Signal to interference ratio.*

**7.2 Cost vs. Benefit – decentralized decision making**

The model that we envision for relay assisted communication is that each mobile tries to maximize its utility. Likewise, each relay also does a cost benefit comparison when it agrees to carry a call. Our formulation is very different from that of previous studies in that we consider a purely market based model with relays and mobiles being independent agents. Thus, the cost and benefit formulation we use is more like a standard economic model. Therefore, the utility is then defined as,

$$Utility = Benefit - Price \qquad\qquad 7.\,7$$

In our model we consider relays as being selfish. Thus, we need to reformulate the utility expressions taking into consideration the benefit obtained and a cost.

1.  Benefit and cost for mobile

The benefit a mobile derives by being able to communicate some L bits of information is based on how important that communication is to the mobile and is independent of whether a relay is used or not. In other words, the benefit has nothing to do with the mode of communication but just with the act of communicating itself. Thus, we write the benefit for mobile $i$ of communicating L bits of information as,

$$B_i = FL(Joules) \qquad\qquad 7.\,8$$

where F (Joules/bit) denotes the amount of energy the mobile is willing to spend for communicating those L bits. Note that F is not the actual energy expended but rather the maximum energy the mobile is willing to sacrifice to send those L bits. To understand this consider the following situation. Say a user has a low battery and is

on a road trip which may last some hours without hope of recharging. Now, the user can make phone calls to friends to chat or save the battery in case there is an emergency. Thus, a rational user will allocate a low F value for calling friends whereas the value for calling in case of an emergency would be high. As the trip comes to an end, these F values may well change with the priority for calling a friend now higher (say they need to get together for lunch).

The cost of communicating depends on how the communication occurs and how much energy is actually expended. Let $j$ be a relay (in case of a direct to base station link, the relay is trivially the base station itself). Then, we write the cost of communication for $i$ as,

$$c_i(j) = \frac{p_{ij}(Watts)}{R_{ij}(bits/\sec)} M_{ij}(bits) + z_j(i)(Joules) \qquad 7.9$$

where $p_{ij}$ is the power used by mobile $i$ to transmit to relay $j$, $R_{ij}$ is the data rate used, $M_{ij}$ is the actual amount of bits sent, and $z_j(i)$ is the price charge by the relay. The formulation is based on the observation that energy is the primary commodity in cellular communications.

2. Benefit and cost for the relay

Unlike a mobile, the relay derives a benefit only in terms of the price the mobile pays. Thus,

$$B_j(i) = z_j(i) = Kc_j(i)(Joules) \qquad 7.10$$

where K >=1 and $c_j(i)$ is the cost to the relay defined as follows,

$$c_j(i) = \frac{p_j(Watts)}{R_j(bits/\sec)} M_j(bits) \qquad\qquad 7.11$$

where we assume that the code rate $L_j/M$ and data rate $R_j$ used by the relay can be different from that of the mobile. $p_j$ is the power used for transmission to the base station.

Combining the equations from above, we can now write the utility for mobile $i$ and relay $j$ as follows,

$$U_i(j) = FL - \frac{p_{ij}}{R_{ij}} M_{ij} - K \frac{p_j}{R_j} M_j \qquad\qquad 7.12$$

$$U_j(i) = (K-1)\frac{p_j}{R_j} M_j \qquad\qquad 7.13$$

## 7.3 Results and Discussion

We use M =80 bits, L=64, and a constant bit rate = 9600 bits/sec. As an example, let us consider a mobile that is stuck at a dead spot. It can not connect to the base station, so we take its power to be infinite. Let F be 0.001.

| Direct or Relay | Value of K | Power (watt) | Utility of Mobile |
|---|---|---|---|
| Direct to BS | K=0 | $p_i = \infty$ | $U_i(BS) = -\infty$ |
| With Relay A | K=2 | $p_{ij} = 0.5$, $p_j = 0.8$ | $U_i(j) = 0.0465$ |
| With Relay B | K=3 | $p_{ij} = 0.2$, $p_j = 1.2$ | $U_i(j) = 0.0323$ |

Table 7. 1 Example of utility of mobile

We tabulate the utility of the mobile in Table 7.1. The highest value of utility for the mobile is when it chooses relay A. We see that relays charge differently and this is reflected in varying K values. We need to know the strategy that both the relay(s) and mobile need to maximize their individual utilities. The strategy that gives higher utility will be the optimal solution. Let us see in Figure 7.1, how this can be done. We start with the same scenario of a mobile stuck at a dead spot. It chooses a relay with higher utility. Now as it moves further, its own direct link to base station improves. It again checks the utility of using relays versus connecting directly to the base station and determines that it has a higher utility when connecting directly. So it does not choose a relay.
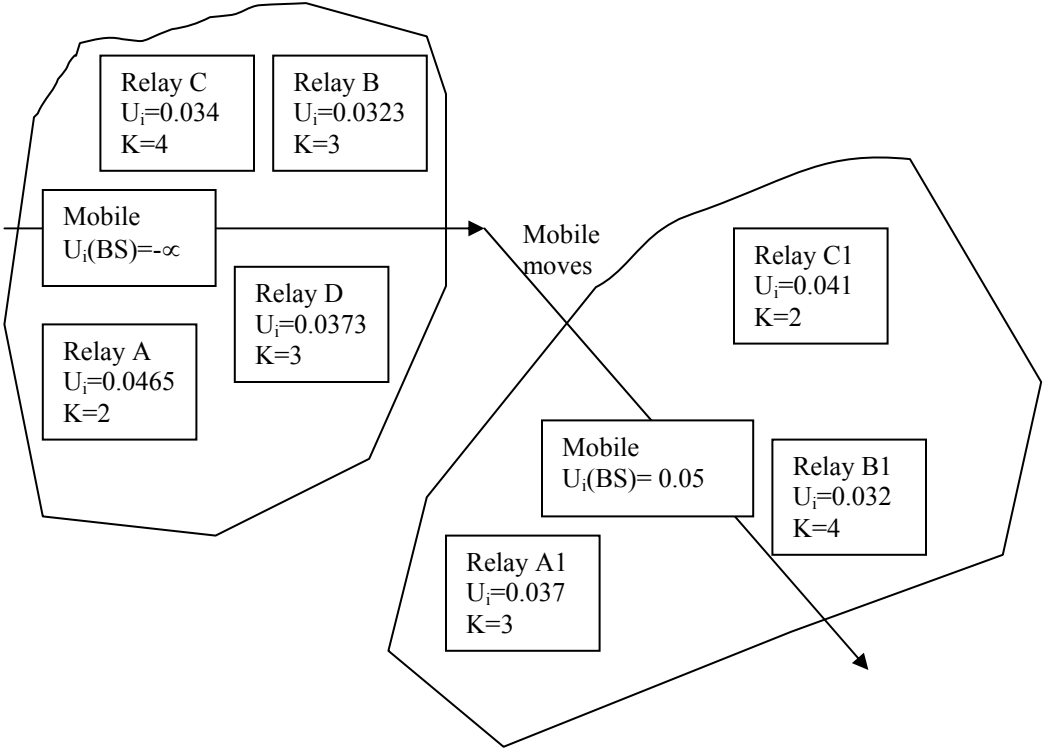


Figure 7. 1 Utility of mobile as it moves

We have until now seen how the utility of a mobile varies but it is also important for a relay to maximize its utility. The game that relays play is adjusting their K values since if a charges a high value of K, it has a chance of being left out as a relay whereas a low value is probably not worthwhile. Note that the position of a relay relative to the mobile and base station has a great impact on the range of K values that it can select. Thus, if a relay is near the base station and has less power requirements, it can experiment with a range of K. So how does a relay stay competitive and maximize its utility? It is easy to see that this situation can be best modeled as a non-collaborative game where each relay competes. It is a game of imperfect information as each relay acts independently.

| Relay A | Relay B | Relay C | Relay D |
|---|---|---|---|
| K=2 | K=3 | K=4 | K=3 |
| $p_{ij}$=0.5, $p_j$=0.8 | $p_{ij}$=0.2, $p_j$=1.2 | $p_{ij}$=0.4, $p_j$=0.8 | $p_{ij}$=0.5, $p_j$=0.9 |
| $U_i$=0.0465 | $U_i$=0.0323 | $U_i$=0.034 | $U_i$=0.0373 |
| $U_j$=0.0066 | $U_j$=0.02 | $U_j$=0.02 | $U_j$=0.015 |

W   L   L   L

Ist Iteration

| Relay A | Relay B | Relay C | Relay D |
|---|---|---|---|
| K=2 | K=2 | K=4 | K=2 |
| $p_{ij}$=0.5, $p_j$=0.8 | $p_{ij}$=0.2, $p_j$=1.2 | $p_{ij}$=0.4, $p_j$=0.8 | $p_{ij}$=0.5, $p_j$=0.9 |
| $U_i$=0.0465 | $U_i$=0.0323 | $U_i$=0.034 | $U_j$=0.049 |
| $U_j$=0.0066 | $U_j$=0.02 | $U_j$=0.02 | $U_j$=0.0075 |

L   L   L   W

IInd Iteration
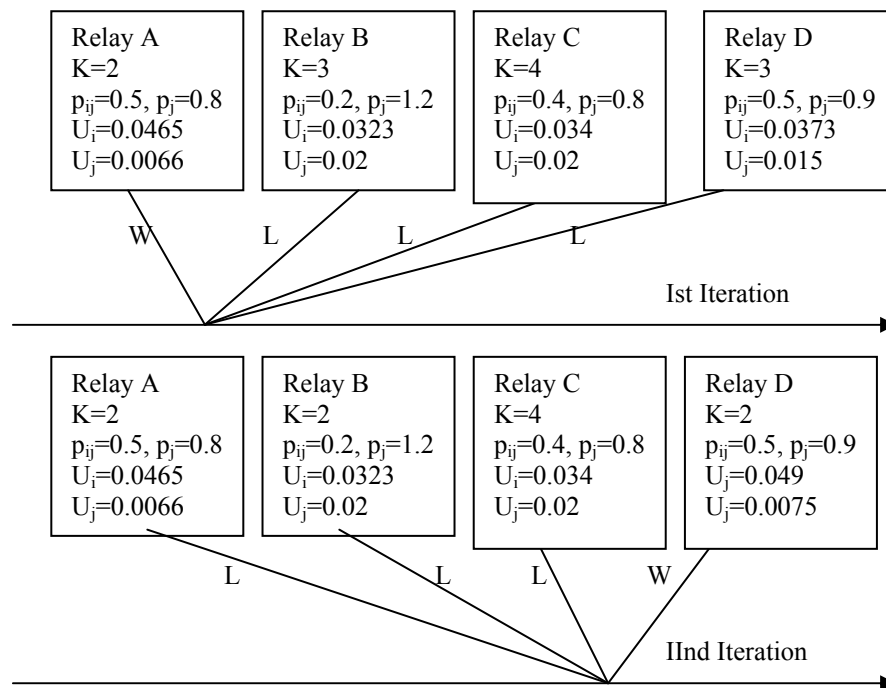
Figure 7. 2  Variation of relays parameters

Let us look into the strategy that relays A, B, C and D employ assuming a multi-round game. Let us take a pair $(K, p_j)$, and assume that the data rate and code rate is constant (even though this varies in the real world). In Figure 7.2, the first iteration chooses relay A denoted as winner *W*, while other relays are denoted as losers *L*. So what should other relays have done to come away with a win? Consider the next iteration. Assuming that the spatial geometry has not changed, the relay cannot change $p_j$ so we change the value of K instead. We see that in the IInd iteration, relay D wins as the mobile has the highest utility when choosing relay D (since D reduced its K), even though relay D's utility reduces.

It is to be noted that once can be relays while having other relays carry their own calls! Thus the values of K can't be kept static. They have to be dynamic depending upon the state of the mobile. Also values for F are dynamic and depend upon the specific data the mobile wants to send at any given time. Depending upon the urgency of the mobile more iterations of the game can be played to choose a better relay. Here the base station can also play a very important role by arbitrating between mobiles and relays as well as providing information about relay locations to all relays (changing the way the game evolves).

The question remains as to who will be able to relay and when. Even though there are large gaps in transmission, one may be hesitant to spend its own battery power. In case of urban areas, people arriving in offices can plug their phone when not in use and allow their mobiles to relay. Even when unplugged many users may want to benefit by favorable pricing for relays.

We need to point out one of the advantages of the relaying mechanism as described. We can achieve spatial diversity by building hierarchical networks, with smaller picocells inside larger cells where the relays act as a nucleus of a smaller picocell, inside the cell. So instead of spending a lot of money to place base stations for picocells we achieve spatial diversity without a high cost. Also, our system is "on-demand" as it can be created as link conditions deteriorate. As to the question of whether relaying increases capacity, it should be pointed out that there is only a limited amount of real estate, i.e., channels that will be shared by mobiles as well as relays in our system. Thus using a relay does not add to the total number of users in the cell in our system, unless all mobiles use relays to reduce the target SIR or there is some other framework, like another channel or an ad hoc network is in place.

## 8. Conclusion

In this thesis we analyzed the benefits and constraints of using relays in cellular environments. Based on what we have learned, we can now conclude by answering the questions we had raised in Chapter 2 and Chapter 3 and explored in later chapters.

The number of relays for a call in cellular networks depends upon whether the call is voice or data based services. For voice calls, there are strict latency constraints, while for data the constraint is reliability. Therefore we can add more relays for data service, but for voice service only one relay is possible.

For a cellular system, there are two ways in which relays can be allocated. One is a centralized base station based approach and another is a decentralized approach where relays advertise themselves. A benefit of the centralized approach is that the relay allocation can be very efficient the base station is aware of global needs in the. The decentralized scenario is only possible if relays have more receiving capability and can distinguish between incoming signals. This Intelligent relay can then determine if it can contribute to the cellular system as a relay. It is to be noted that the hardware and software complexity for a relay depends upon whether we are using centralized or decentralized approaches. In the centralized approach most of the software complexity is at the base station but the mobiles need enough hardware to receive and send the signals to both mobiles and the base station. However, in decentralized case the software and hardware have to be included in the relay so that intelligent decisions can be made. Here, both modulation and demodulation, coding

and decoding have to be in duplicate with enough buffering. In any case, if power control information is to be sent to mobile-relay link then the hardware has to be in duplicate.

We found that both the Greedy approach and Intelligent relay approach saved almost four times the energy when compared with direct to base station calls for all cases studied. The optimum approach saved six times when compared with direct to base station call. However, both optimum and Intelligent relay approach have higher computational complexity. Using a detailed discrete event simulator based on the Greedy approach we found that the total amount of power savings varied from 1.76 dB to 8.45 dB and with the base station at the corner of the cell case showing higher savings. The number of relay-relay handoffs increases with the number of mobiles and call rate in the Greedy approach. Finally, a higher cell gain reduces energy consumption for direct as well as relayed calls.

We explored the impact of relays in increasing useful coverage area. We saw that the use of relays can increase the useful service area by about 10% with real life scenarios. In heavy building density areas there is more of a need of relays as compared to low building density areas. However, the chance of finding relays is greater in low building density areas. It is seen that having more available idle nodes helps in choosing relays, so we conclude that the base station should admit more mobiles even if they are not calling. If the relays are not under power control, then they add to the interference in the cellular system. In interference limited environment the relays have to be under power control to maintain balance. If the relays broadcast themselves in the cell, then they create more dead spots. Thus either the base station

should be responsible for allocating relays or the relays should be provided with enough intelligence to perform power control on the relay to mobile link. Finally, we observed that as a guideline, it is a good idea to avoid LOS regions such as wide roads so that additional outage is not created. Moreover, the relays should be distributed all over the cell.

If the objective of the system is not to increase the range of the cell, then the capacity of the system does not change when using relays. Furthermore, if the target Signal to Interference ratio is reduced by power reduction in the cell (using relays), then the capacity of the cell can increase. This is again a policy decision: for example, all mobiles at the periphery of the cell could be forced to use relays and thus minimize both inter and intra-cell interference.

There are several interesting extensions of the work presented here. For one, for purely data services, using multiple relays is possible but with the caveat that there may be more interference in the system. This problem also needs to be studied in the context of 4g wireless systems where data services are predominant. Next, the utility game can be extended to consider a market-based economic model. Thus, users may participate simply to make profit with numerous interesting implications (e.g., individuals may start offering wireless internet connections from their homes for a price and so forth).

# References

[3GPP, 1998] www.3GPP.org

[3GPP2, 2000] www.3GPP2.org

[Aggelou and Tafazolli, 2001] George N. Aggelou and Rahim Tafazolli, "On the Relaying Capacity of Next Generation GSM Cellular Networks", IEEE Personal Comm., pp. 40-47, Feb 2001.

[Banerjee and Misra, 2002] Suman Banerjee and Archan Misra, "Minimum energy paths for reliable communication in multi-hop wireless networks", MOBIHOC-02, 2002.

[Berg, 1995] J.E. Berg, "A recursive method for street microcell path loss calculations", Personal, Indoor and Mobile Radio Communications (PIMRC 95), Vol.1, p. 140-143, Sept. 1995.

[Bertoni, 2000] Henry L. Bertoni, "Radio propagation for modern wireless systems", Prentice Hall PTR, 2000.

[Borlin, 1999] Niclas Borlin, "Assignment problem using Hungarian method in Matlab", Sweden.

[Bourg and Seeman, 2003] David M. Bourg and Glenn Seeman, "AI for game developers", O'Reilly, 2003.

[Bronzel, Rave, Herhold and Fettweis, 2001] M. Bronzel, W. Rave, P. Herhold, G. Fettweis, "Interference Reduction in Single Hop Relay Networks", 11th Virginia Tech Symposium on Wireless Personal Communications, 2001.

[Carpaneto and Toth, 1980], "Algorithm 548: Solution of the assignment problem [H]", ACM Transactions on Mathematical Software, 6(1):104-111, 1980.

[Coinchon, Salovaara, and Wagen, 2001] Mathias Coinchon, Ari-Pekka Salovaara, Jean-Frederic Wagen, "The Impact of Radio Propagation Predictions on Urban UMTS Planning", International Zurich Seminar on Broadband Communications, 2001.

[Consul, 1989] P.C. Consul, "Generalized Poisson Distributions, Properties and Distributions", Marcel Dekker, Inc., New York, 1989.

[Diasemi, 2003] www.diasemi.com

[Fujiwara, Takeda, Yoshino and Otsu, 2002] Atsushi Fujiwara, Shinji Takeda, Hitoshi Yoshino and Toru Otsu, "Area coverage and capacity enhancement by multihop connection of CDMA cellular network", Proceedings of IEEE Vehicular Technology Conference (VTC), Vol.4, p. 2371-2374, Fall 2002.

[Garg, 2002] Vijay K. Garg, "Wireless Network Evolution, 2G to 3G", Prentice Hall Communications Engineering and Emerging Technologies Series, 2002.

[Gonzalez, Gordon and Horowitz, 1997] R. Gonzalez, B. Gordon, M. Horowitz, "Supply and Threshold Voltage Scaling for Low Power CMOS", IEEE J. Solid-State Circuits, vol. 32, no. 8, Aug. 1997.

[Goodman and Mandayam, 2000] David Goodman and Narayan Mandayam, "Power control for wireless data", IEEE Personal Communications, vol. 7, No. 2, pp. 48-54, April 2000.

[Gunturi and Paganini, 2003] Sarma Gunturi and Fernando Paganini, "Game theoretic approach to power control in cellular CDMA", VTC 2003.

159

[Hadjiyiannis, Chandrakasan and Devdas, 1998] George Hadjiyiannis, Anantha Chandrakasan and Srinivas Devadas, "A low power, low bandwidth protocol for remote wireless," Wireless Networks 4, pp. 3-15, 1998.

[Harrold and Nix, 2002] T.J. Harrold and A.R. Nix, "Performance analysis of intelligent relaying in UTRA-TDD", Proceedings of IEEE Vehicular Technology Conference (VTC), p.1374-1378, 2002.

[Hata, 1980] M. Hata, "Empirical formula for propagation loss in land mobile services, IEEE Transactions on Vehicular Technology, Vol. 29, p. 317-325, Sept 1980.

[Herhold, Rave, Fattweis] Patrick Herhold, Wolfgang Rave and Gerhard Fettweis, "Relaying in CDMA networks: Pathloss reduction and transmit power savings".

[Ho and Stuber, 1995] Ming-Ju Ho and Gordon L. Stuber, "Capacity and power control for CDMA microcells", Wireless Networks, Vol. 1, Issue 3, pp. 355-363, 1995.

[Hsieh and Sivakumar, 2001] H.Y. Hsieh, R. Sivakumar, "Performance Comparison of Cellular and Multi-hop Wireless Networks: A Quantitative Study", ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS), June 2001.

[Hsieh and Sivakumar, 2001] H.Y. Hsieh, R. Sivakumar, "Performance Comparison of Cellular and Multi-hop Wireless Networks: A Quantitative Study", ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS), June 2001.

[Hsieh and Sivakumar, 2004] Hung-Yun Hsieh and Raghupathy Sivakumar, "On using peer-to-peer communication in cellular wireless data networks", IEEE Transactions on Mobile Computing, Vol. 3, No.1, p. 57-72, Jan- Mar 2004.

[ImageJ, 2001] http://rsb.info.nih.gov/ij/

[Jain, 1991] Raj Jain, "The art of computer systems performance analysis – Techniques for experimental design, measurement, simulation, and modeling", John Wiley & Sons, Inc, New York, 1991.

[Jamalipour and Yabusaki, 2003] Abbas Jamalipour and Masami Yabusaki, "3G Mobile Network Technologies and Experiences", IEEE Wireless Communications, Vol. 10, No.1, pp. 4-6, Feb 2003.

[Jones and Skellern, 1995] Brendan C. Jones and David J. Skellern, "Designing the Ubiquitous Microcellular Network", Workshop on Applications of Radio Science, 1995.

[Lakkavalli and Singh, 2003] Shashidhar Lakkavalli, Suresh Singh, "Stretched Call Model for Next Generation Cellular Networks", IEEE Vehicular Technology Conference, 2003.

 [Lakkavalli, Negi and Singh, 2003] S. Lakkavalli, A. Negi and S. Singh, "Stretchable Architectures for Next Generation Cellular Networks", Proceedings of the International Symposium on Advanced Radio Technologies, p. 59-65, 2003.

[Lam, Cox and Widom, 1997] Derek Lam, Donald C. Cox and Jennifer Widom, "Teletraffic Modeling for Personal Communication Services", IEEE Communications Magazine, p. 79-87, Feb 1997.

[Lee and Miller, 1998] Jhong Sam Lee and Leonard E. Miller, "CDMA Systems Engineering Handbook", Artech House Publications, Boston, 1998.

[Lin and Hsu, 2000] Ying-Dar Lin and Yu-Ching Hsu, "Multihop cellular: A new architecture for wireless communications", Proceedings of INFOCOM , p. 1273-1282, Vol.3, 2000.

[Lin, Hsu, Oyang, Tsai and Yang, 2000] Ying-Dar Lin, Yu-Ching Hsu, Kuan-Wen Oyang, Dong-Su Yang, Tzu-Chieh Tsai, "Multihop Wireless IEEE 802.11 LANs: A Prototype Implementation", Journal of Communications and Networks, vol.2, no.4, Dec. 2000.

[Lu, Goodman, Wang, Erkip, 2004] Xiaoan Lu, David Goodman, Yao Wang, Elza Erkip, "Complexity bounded power control in video transmission over a CDMA wireless network", Globecom 2004.

[Macgill, 1999] http://www.vision.ee.ethz.ch/~buc/astar/AStar.html

[Maxim, 2001] www.maxim-ic.com

[Neskovic, Neskovic and Paunovic, 2002] A. Neskovic, N. Neskovic and D. Paunovic, "Macrocell Electric Field Strength Prediction Based Upon Artificial Neural Networks", IEEE Journal on Selected Areas in Communications, Vol. 20, No.6, pp1170-1177, 2002.

[ODMA, 1999] 3rd Generation Partnership Project, "Opportunity Driven Multiple Access", 3G TR 25.924.

[Okumara, 1968]  Yoshihisa Okumara, "Field strength and its variability in VHF and UHF land mobile service", Review of the Electrical Communications Laboratory, Vol. 16, No. 9-10, p. 825-873, Sept-Oct 1968.

[Papadakis, Kanata and Constantinou, 1998] N. Papadakis, Athnasios G. Kanata, Philip Constantinou, "Microcellular Propagation Measurements and Simulation at 1.8 GHz in Urban Radio Environments", IEEE Trans. Veh. Technol., vol. 47, pp. 1012-1026, 1998.

[Pinos, Aydin and Hamdy] Jordi de los Pinos, Levent Aydin and Walid Hamdy, "Performance comparison of packet data bearer service modes in CDMA2000", Proceedings of PIMRC 2003.

[RadioFrame Networks, 2003] www.radioframenetworks.com

[Rappaport, 1996] Theodore S. Rappaport, "Wireless Communications", Prentice Hall PTR, 1996.

[Rouse, Band, McLaughlin, 2002] T. Rouse, I. Band and S. McLaughlin, "Capacity and power investigation of Opportunity Driven Multiple Access in TDD-CDMA based systems", IEEE International Conference on Communications, p.3202-3206, 2002.

[Simunic, 2002] Tajana Simunic, "Dynamic Management of Power Consumption", Power Aware Computing, ed. R. Graybill, R. Melhem, Kluwer Academic Publishers, 2002.

[Singh and Raghavendra, 1998] S. Singh and C. Raghavendra, "Pwer-aware routing in mobile ad-hoc networks", Proceedings of Mobicom, 1998.

[Sreng, Yanikomeroglu and Falconer, 2003] V. Sreng, H. Yanikomeroglu, D.D. Falconer, "Relayer Selection Strategies in Peer-to-Peer Relaying in Cellular Radio Networks", IEEE Vehicular Technology Conference Fall, Orlando, Florida, 2003.

[Sreng, Yanikomeroglu and Falconer, 2003] V. Sreng, H. Yanikomeroglu, D.D. Falconer, "Relayer Selection Strategies in Peer-to-Peer Relaying in Cellular Radio Networks", IEEE Vehicular Technology Conference Fall, Orlando, Florida, 2003.

[Sunsted, 1995], Todd Sundsted. "How-To Java: Observer and Observable". JavaWorld, 1995.

[Tameh, Nix and Molina, 2003] E.K. Tameh, A. Nix, A. Molina, "The Use of Intelligently Deployed Fixed Relays to Improve the Performance of a UTRA-TDD System", IEEE Vehicular Technology Conference, 2003.

[Veeravalli and Sendonaris, 1999] V.V. Veeravalli and A. Sendonaris, "The coverage – capacity tradeoff in cellular CDMA systems", IEEE Transactions on Vehicular Technology, Vol. 48 (5), p. 1443-1451, Sept. 1999.

[Viswanathan and Mukherjee, 2003] Harish Viswanathan, Sayandev Mukherjee, "Performance of Cellular Networks with Relays and Centralized Scheduling", IEEE Vehicular Technology Conference, 2003.

[Wu, Qiao and Tonguz, 2001] Hongyi Wu, Chunming Qiao, Ozan Tonguz, "Performance Analysis of iCAR (Integrated Cellular and Ad-hoc Relay System)", ICC, Vol. 2, pp. 450-455, 2001.

[Wu, Qiao and Tonguz, 2001] Hongyi Wu, Chunming Qiao, Ozan Tonguz, "Performance Analysis of iCAR (Integrated Cellular and Ad-hoc Relay System)", ICC, Vol. 2, pp. 450-455, 2001.

[Wu, Qiao, De and Tonguz, 2001] Hongyi Wu, Chunming Qiao, Swades De and Ozan Tonguz, "Integrated cellular and ad hoc relaying systems", IEEE Journal on Selected Areas in Communications, Vol. 19, No. 10, p. 2105- 2115, Oct 2001.

[Zadeh and Jabbari, 2001] Ali N. Zadeh and Bijan Jabbari, "Performance analysis of multihop packet CDMA cellular networks", Proceedings IEEE GLOBECOM, 2001.

[Zadeh and Jabbari, 2002] Ali ,N. Zadeh and Bijan Jabbari, "Self- organizing packet radio ad hoc networks with overlay (SOPRANO)", IEEE Communications Magazine, p. 149- 157, June 2002.

[Zhou and Yang, 2002] Jie Zhou, Yang Richard Yang, "PARCelS: Pervasive Ad-hoc Relaying for Cellular Systems", Proceedings of Med-Hoc-Net, Sardegna, Italy, Sept 2002.

[Zhou and Yang, 2002] Jie Zhou, Yang Richard Yang, "PARCelS: Pervasive Ad-hoc Relaying for Cellular Systems", Proceedings of Med-Hoc-Net, Sardegna, Italy, Sept 2002.

[Zhou, Honig and Jordan, 2005] Chi Zhou, Michael L. Honig and Scott Jordan, "Utility-based power control for a two cell CDMA data network", TWC 2005.

## Appendix A: Reverse Link Power Budget

The Signal to Noise Ratio (SNR) at the base station [Lee and Miller, 1998] is given by,

$$ SNR = \frac{P_m G_c G_m (1/L_{med})}{(N_0 W)_c + \left(\dfrac{M}{F_e} - 1\right)\alpha_r P_m G_c G_m (1/L_{med})} \qquad \text{A. 1} $$

where,

$P_m$ = mobile's power amplifier output
$G_c$ = cell antenna gain, including cable losses
$G_m$ = mobile antenna gain, including cable losses
$L_{med}$ = median reverse link path loss
$\alpha_r$ = average voice activity factor
M = capacity of cell
$F_e$ = frequency reuse efficiency
$(N_0 W)_c$ = thermal noise at the cell receiver

The denominator in equation A.1 is total of noise spectral power density, $\sigma^2$ and $I_T$, the total average interference power on reverse link as,

$$ \sigma^2 \equiv (N_0 W)_c $$
$$ I_T \equiv \left(\frac{M}{F_e} - 1\right)\frac{\alpha_r P_m G_c G_m}{L_{med}} \qquad \text{A. 2} $$

The cell loading, X, is defined as,

$$ X = \frac{I_T}{\sigma^2 + I_T} \qquad \text{A. 3} $$

Thus,

$$\sigma^2 + I_T = \frac{I_T}{X} = \frac{\sigma^2}{1-X} \qquad\qquad\qquad A. 4$$

Thus, the total noise and interference at the base station receiver is,

$$\left(N_0 W\right)_c + \left(\frac{M}{F_e} - 1\right) \alpha_r P_m G_c G_m \left(1/L_{med}\right) = \frac{\left(N_0 W\right)_c}{1-X} \qquad A. 5$$

Substituting this in equation A.1, we get the required SNR as,

$$SNR_{req} = \frac{P_m G_c G_m / L_{max}}{\left(N_0 W\right)_c}\left(1-X\right) \qquad\qquad\qquad A. 6$$

$$L_{max} = \frac{P_m G_c G_m}{\left(N_0 W\right)_c}\frac{\left(1-X\right)}{SNR_{req}} \qquad\qquad\qquad A. 7$$

Also, SNR$_{req}$ can be written in terms of received bit energy-to-noise density ratio, E$_b$/N$_0$, as,

$$SNR_{req} = \frac{E_b / N_0}{W/R} = \frac{E_b / N_0}{PG} \qquad\qquad\qquad A. 8$$

where, PG is the processing gain, W is the bandwidth and R is the rate.

Thus, equation A.7 becomes,

$$L_{max} = \frac{P_m G_c G_m}{\left(N_0 W\right)_c}\frac{\left(1-X\right)}{\left(E_b / N_0\right)/PG} \qquad\qquad\qquad A. 9$$

In dB units, equation A.8 becomes,

$$L_{\max}(dB) = P_m(dBm) + G_c(dB) + G_m(dB) - SNR_{req}(dB)$$

$$- (N_0 W)_c(dBm) + 10\log_{10}(1 - X) \qquad\qquad \text{A. 10}$$

The noise spectral power density at the base station receiver may be written as noise figure and is taken to be 5dB [3GPP, 1998]. $E_b/N_0$ is taken to be 7 dB for urban area. Bandwidth is 1.2288 MHz, rate R is 9600 bits/sec and cell loading factor, X is 0.5. Cell and mobile antenna gain are taken as 10 dBi or as required.

## Appendix B. SIR Contours

[Cook,1980], [Cook, 1987] analyses the effect an interferer can have on a line of sight link. For a single interferer case a generalized link equation for omnidirectional antennas is given as,

$$R_L^n = R_J^n K \qquad\qquad \text{B. 1}$$

where n, the path loss exponent is greater than or equal to 2. The geometry is shown in Figure A.1. $R_L$ is the distance from the transmitter to the receiver, i.e., mobile and base station in our case of relays. $R_J$ is the distance from the receiver to the interferer, i.e., mobile to relay. Tx and Rx are the transmitter and receiver respectively. J is the interferer. For n=2, and $D = K^{2/n}$, we have contours of constant, D and K, [Cook, 1987] derived below.

Radius, a, and offset $R_0$, can be given as

$$a = \frac{SD^{1/2}}{1-D} = \frac{SK^{1/n}}{1-K^{2/n}}$$
$$R_0 = \frac{SD}{1-D} = \frac{SK^{2/n}}{1-K^{2/n}} \qquad\qquad \text{B. 2}$$

The parameter K, describes the relative effectiveness of the link performance with respect to the interference in an interference limited environment. K is defined as x [Cook, 1980], [Cook, 1987], [Jones and Skellern, 1995],

$$K = \frac{P_t G_r W_j L_s}{Z P_j W_b} \qquad\qquad \text{B. 3}$$

where, $G_r$ is the net antenna gain between interferer and receiver, $L_s$ is the system loss factor, $P_j$ is the interference power and $W_j$ and $W_b$ are the bandwidths of the interfering and wanted signals respectively.
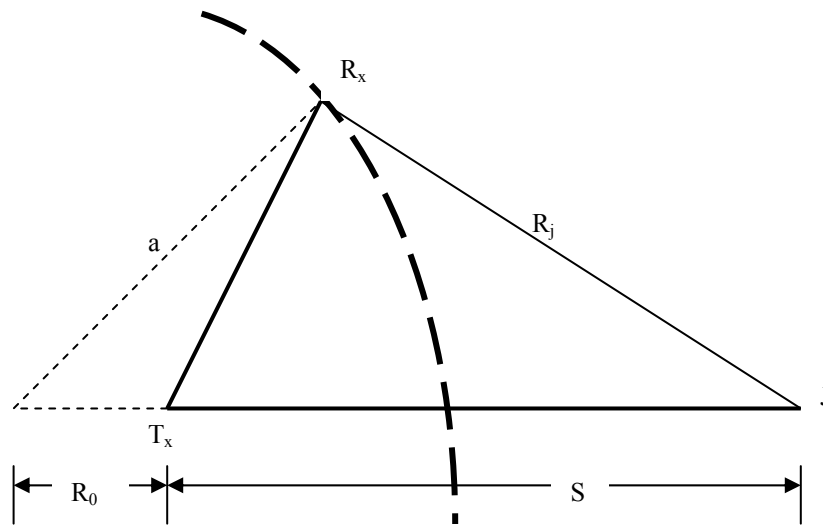


Figure A.1: Geometry for K-contour derivation [Cook, 1987]

[Jones and Skellern, 1995] also show outage contours as defined by the locus where S/I=Z. The outage contours at the mobile end is a family of circles with centers (a, b=0) and radii $R_0$. [Cook, 1987] actually calls them signal to interference ratio (SIR) contours, which is a more apt choice. We follow this definition of SIR contours and define outage contours as where the signal falls below a threshold.

## Appendix C. Comparison of Related Work

| Paper, Single Cell / Multi Cell | Relay Mechanism | Variables | Simulation | Results |
|---|---|---|---|---|
| iCAR, Multi cell = 19 [Wu, Qiao and Tonguz, 2001] | Ad hoc, Hybrid Network | Call blocking probability | Formula derived | Better load balancing |
| TDMA cellular mulithop, Single cell [Sreng, Yanikomeroglu and Falconer, 2003] | Not specified | System coverage, path loss, distance | 400 by 400 meters cell, 2.5 GHz | Increases coverage, better scheme based on path loss rather than distance |
| Multihop ad hoc networks [Zadeh and Jabbari, 2002 ] | Hybrid network | NA | NA | NA |
| Multihop Cellular Network (MCN) [Lin and Hsu, 2000] | Hybrid network | Throughput | NA | Higher throughput for MCN |
| Peer-to-peer network [Hsieh and Sivakumar, 2001] | Hybrid network | Throughput, power | 1500 by 1500 meter, 100 mobiles | Change peer-to-peer model to cellular for higher speed |
| PARCelS, multi cell =7 [Zhou and Yang, 2002] | Mobile as Relay in Hybrid network | Traffic load, congestion, | Seven cells with radius as one, 100 data channels | Balance traffic load, reduce congestion |
| Cellular network performance [Viswanathan and Mukherjee, 2003] | Hybrid network | Throughput | Sectored cell, radius 2km | Throughput gain achieved |
| Stretched Network [Lakkavalli and Singh, 2003] | Mobile as relay in cellular network | Frame sharing structure in uplink and downlink | TDD, FDD frame sharing proposed | Compares TDD and FDD frame mechanism for relays |
| Intelligent, fixed relays, UTRA-TDD [Tameh and Molina, 2003] | Repeater as relay in cellular network | Capacity, power | Seven cells, 21 relay nodes | Improvement in capacity, reduction in power |

Table A.1. Comparison of related work

[Zadeh and Jabbari, 2002] identify the advantages of combining both cellular and

ad hoc networks as:

- Increasing robustness and scalability of the system.

- Supporting dynamic topology through packet ratio

- Load balancing in the network due to routing opportunity

- Extending the cell coverage area

- Providing a broad connectivity based on multihop technology

- Exploiting spatial diversity through adaptive routing

- Reducing total consumed power

- Increasing the network capacity

We discuss various papers that have touched upon either one or many such aspects.

[Hsieh and Sivakumar, 2004] have discussed joint cellular and peer-to-peer communication in ad hoc networks, where the cellular network model and the peer-to-peer network model are defined using 802.11. In the cellular wireless network model, a cellular path is just the one hop path between host and base station. In the peer-to-peer network model, a multihop path between peer hosts allows for communication to exist without a base station. To study their scheme they use the ns-2 network simulator. The IEEE 802.11 protocol is used to define peer-to-peer networks using the Distributed Coordination Function (DCF) of 802.11. CSMA/CA is used as the MAC protocol and Dynamic Source Routing (DSR) as the routing protocol. The cellular model is defined by the Point Coordination Function (PCF) of the 802.11 protocol, where access points serve as the base stations. The physical and MAC layer for the PCF mode is assumed to serve as a typical cellular model. Even though cellular systems such as CDMA use Direct Sequence Spread Spectrum (DSSS) as in 802.11, the model is inherently different since CDMA is basically a multi user model as against 802.11, where only a single user is granted access at a time.

The target in the paper [Hsieh and Sivakumar, 2004] is to see the effect of internet data access in hybrid peer-to-peer/cellular networks. Since internet access

172

primarily uses TCP as the transport layer, 90% of the simulated traffic is TCP flows. They compare end to end throughput for multihop and single hop flows. The average end-to-end throughput per flow is lower in the single hop case as against the multihop case. Since all the traffic has to move via the base station they identify base station as the bottleneck. It may be pointed out here that in a real cellular network base station is endowed with lot more capabilities than an access point and as such can afford more channels and power. Thus their conclusion that the BS is a bottleneck may not be valid. They also note that mobility in peer to peer network models introduces route failures and network partitions.

The paper on multihop ad hoc networks [Zadeh and Jabbari, 2002] use the term "self organizing" to signify the dynamic changing conditions of the wireless channel and the need to regroup the network. The paper gives cell splitting as an analog of a multihop cellular model. Cell splitting splits the cell into smaller cells and thus lowers the transmission distance. The new hierarchical cellular architecture requires such smaller pico cells overlaid on the existing cell to achieve load balancing. However this necessitates the establishment of new base stations which is a costly affair. Similar shortening of links can be achieved by multihop forwarding where wireless routers route the traffic. In this paper, a router node is a unique entity as it not a source or destination of the traffic.

[Wu, Qiao, De and Tonguz, 2001] identify the need of balancing traffic in heavily congested areas called hot spots. The congestion is time dependent and may achieve peak values at rush hours. This means that there are few data channels (DCH) available at those times resulting in blocked calls. The system is called Integrated

173

Cellular and Ad hoc Relaying (iCAR) network. They use ad hoc relaying stations (ARS) positioned at strategic locations to dynamically balance the load. The ARSs are under direct control of MSC making it at the same level as base station logically. Unlike other papers [Hsieh and Sivakumar, 2004] and [Zadeh and Jabbari, 2002], this paper provides the required mixing of the cellular and ad hoc models needed to realize a realistic system. Each ARS is provided with two interfaces, the C and R interface. C interface operates around 1.9 GHz (PCS) and R interface uses unlicensed band at 2.4 GHz (ISM band). The simulations are carried in the GloMoSim simulator.

This paper identifies three types of relaying: primary, secondary and cascaded relaying. We have to remember that since the main purpose of introducing relaying in this paper is to shift a call to another cell, we require ARS in each cell to carry the call for mobiles and then to forward it to another cell. So the communication in primary relaying is from mobile in heavily congested cell A, to ARS in cell A, to ARS in lightly loaded cell B onwards to base station in cell B. Secondary relaying means that the mobile has to use many hops, i.e., many ARSs are used as intra cell hopping till an ARS finally is able to do inter cell relaying. Cascading relaying means relaying through more than two cells to transfer the call. The metrics in this paper are call blocking probability, throughput and additional signaling overhead used by relaying. The simulations show that primary relaying reduces the call blocking probability to an acceptable level of 2%. Similarly secondary relaying further reduces the call blocking probability. It is to be noted that even though system wide call blocking probability decreases, the less congested areas taking extra call load report increase in call blocking probability. For throughput purposes, throughput ratio is used, which is

taken as the ratio of received data over the data to be transmitted. For low traffic loads, i.e., for average holding times less than 140 second, 99% throughput ratio can be obtained by relaying. For high traffic loads, limited capacity lowers the throughput ratio. The amount of signaling overhead for primary relaying is very small for BTS and only 1% for MSC. For secondary and cascaded relaying BTS has high signaling overhead, while MSC has at most 20%. It is to be noted here that even though ARS is logically under the MSC, the BTS also has to bear signaling overhead.

[Zhou and Yang, 2002] introduce the PARCelS system that is very similar to the iCAR system. The main difference lies in the fact that relays are not special devices as in iCAR. Any mobile such as a laptop can be used as a relay. They also report load balancing and avoiding traffic congestion. Their solution can balance load even before congestion happens.

[Viswanathan and Mukherjee, 2003] study the throughput gains achieved by relays in a hybrid network. Their simulations show that up to 70% throughput gain is possible with four relays. The effect of total power distributed in the cell via base station and relays is shown. They show that throughput increases when more relays share the power.

[Lin and Hsu, 2000] also combine features of ad hoc and Single - hop Cellular Network (SCN) and come up with the "Multihop Cellular Network (MCN)" architecture. In a multihop cellular network, mobiles help to relay packets. Two possible architectures of MCN are defined in the paper, MCN-b and MCN-p. The difference lies in the fact that either the number of base stations is reduced by relaying or the transmission ranges of both mobile and base station is reduced. In MCN-b, the

175

number of base stations is reduced and the distance between two neighboring base stations increases to $k_b$ times the distance in SCN. In MCN-p, the transmission ranges of base stations and mobiles is reduced by $1/k_p$ of the original distance in SCN. The routing path is from mobile to base station in cell A, to another base station and then to mobile in cell B. The mobile to base station path may be multihop. In an implementation of multihop cellular, mobile stations run a bridging protocol [Lin, Hsu, Oyang, Tsai and Yang, 2000]. It is to be noted that even though base stations in MCN-b are further apart, no multihop path exists between base stations as in iCAR [Wu, Qiao, De and Tonguz, 2001] where ARSs relay between base stations. The paper does not address the possible increase of interference due to higher power required to communicate between base stations. Both SCN and MCN use the 802.11 protocol and is implemented in PARSEC, a discrete event simulation language. They report that throughput of MCN is higher than that of SCN. Also the throughput of MCN decreases as the range of transmission decreases. This decrease of range increases the number of hop counts.

Until now many architectures have used 802.11 based ad hoc networks to provide relaying or specially equipped relays such as ARSs. For CDMA based cellular networks many approaches are possible as described in papers given below. It is to be noted however, that many of the papers are contemporaries of our research.

[Fujiwara, Takeda, Yoshino and Otsu, 2002] describe a multihop CDMA cellular system. They point out that data transmission is proportional to its bit rate. If the transmit power is kept constant, the transmission bit rate increases and practical coverage area decreases. In fact if bit rate is increased from several hundred kilobits

per second to several tens of megabits, the practical area shrinks to 7% of the original area. They also recognize the interference arising out of multihops, calling it multihop interference. The paper recognizes the distance nearer to the base station as the area where high bit rate transmission is possible in a single hop, while the area further away from the base station is for low bit rate transmission using multiple hops. The Monte Carlo simulations use a multicell downlink model where each mobile is capable of relaying. Routing is done by selecting the minimum total transmit power route. The coverage area in a single hop was found to be 58% of the cell, while for multihop, the coverage area is approximately 90% of the cell. The system capacity improvement in multihop is slight due to inter hop interference.

[Zadeh and Jabbari, 2001] in their analysis of multihop CDMA cellular network identify two packet forwarding strategies. One is minimum path loss strategy, where relay with minimum link propagation path loss is chosen. Another is minimum path loss with forward progress, where forward progress signifies that the transmission direction towards the base station is chosen along with minimum path loss.

[Zadeh and Jabbari, 2002] discuss the architecture of a self organizing multihop network using CDMA in TDD mode. The steps in self organization involve Neighbor discovery, connection setup, channel assignment, mode selection and mobility management and topology updating.

In the cellular model, there has been a proposal of Opportunity Driven Multiple Access (ODMA) [ODMA, 1999] for cellular multihop. ODMA is an enhancement of UMTS-TDD mode and is standardized by 3GPP [3GPP, 1998]. In ODMA mobiles route amongst themselves by keeping an account of their neighbors.

[Lakkavalli and Singh, 2003]  studies the amount of delay possible in multiple hops in ODMA based cellular model. The main finding is that the delay varies with the number of hops in between the mobile and base station. This makes the ODMA model unsuitable for real time services such as voice. Instead, a two hop cellular model is suggested to keep delays minimum.

[Lakkavalli and Singh, 2003] report on a CDMA based solution for relay. They provide various mechanisms by which relays can be incorporated within the CDMA network using mobile handsets as relays. [Tameh, Nix and Molina, 2003] present repeater like fixed relays in cellular systems, that show significant improvement in capacity and reduction in mean mobile transmit power with relays.

We see that relays can relay calls not only for intra cell calls but inter cell calls too. The difference in design is that if load balancing between the cells is the goal then inter cell relaying is used and the relays are under the jurisdiction of MSCs. On the other hand for other goals such as power reduction and coverage, intra cell relaying is employed. If relays use the underlying cellular technology, then they fall under the jurisdiction of BSCs, making the design process simpler. Thus in this thesis, we assume that any mobile or fixed station can relay calls for other mobiles. By not using ad hoc technology for relays, we also avoid inherent problems of ad hoc networks.