Bounding the Performance of Multi-hop Cellular Systems for Real Time Traffic

Master's Thesis Report

Department: Electrical and Computer Engineering

Advisor: Prof. Ozan Tonguz

Student Name: Xindian Long

Submitted to the

Department of Electrical and Computer Engineering

Carnegie Mellon University

for Partial Fulfillment of the Degree of Master of Science

October 25, 2005

Contents

1 Introduction

In recent years, the proliferation of Internet and wireless users has exposed the serious limitations of legacy wireless networks. Compared with wired networks, wireless networks have significantly lower data rates. The lack of sufficient bandwidth has long been the biggest obstacle in the development of wireless systems. Studies on wireless networks targeting throughput enhancement, extending coverage area, load balancing, and power conservation have recently gained increasing attention in the research community. These studies have resulted in several hybrid wireless network architectures that combine multi-hop radio relaying and existing infrastructures of wireless networks to provide high-capacity wireless networks [1] [2] [3] [4] [5].

Present-day wide area wireless networks are based on the cellular model. The major difference between wireless networks and wired networks is that wireless signals at the same channel frequency interfere with each other in space. Interference amongst wireless singals using the same carrier frequency channels is known as co-channel interference. This makes bandwidth a precious resource in the operation of wireless networks. It is well known that signal strength attenuates when the seperation between a transmitter and a receiver is increased. When interfering signals are seperated from one another by a sufficient distance, co-channel interference can be reduced to a tolerable level. Frequency reuse refers to the concept of using wireless channels on the same carrier frequency to cover different areas seperated by a sufficient distance [6].

Instead of covering an entire local area with one base station using high power, the wireless operators divide the coverage area into multiple cells and establish one base station in each cell. Each base station transmits at a power level sufficient to cover its own cell and does not cause high interference to a remote cell. Therefore, cells sufficiently far apart can use the the same set of channels. Cells can be of any shape in reality, while they are typically modeled as hexagons for analytical purposes. The set of cells using different channels form a cluster, and the size of clusters is an integer obtained from the following equation [6]: $N = i^2 + j^2 + ij$, where i, j are nonzero integers.

Co-channel reuse ratio is defined as [6]: $\frac{D}{R} = \sqrt{3N}$, where D is the distance between the centers of nearest co-channel cells, and R is the cell radius. As the cluster number increases, the distance between co-channel cells increases, the interference decreases, and consequently better signal to interference ratio can be achieved. In a practical system, co-channel interference considerations govern the choice of cluster size.

Traditionally, mobile hosts (MHs) in one cell reach the base station (BS) via one wireless link; i.e., in one hop. Current cellular networks are single hop cellular networks (SCN). Electromagnetic spectrum is reused only between cells. In SCN, communication occurs via channels assigned or allocated to each cell. Cellular networks have infrastructure support for real time traffic. Base stations are responsible for tracking the location of MHs, allocating channels, setting up, clearing, and handing calls over.

Multi-hop relaying was initially introduced in ad-hoc networks. In ad-hoc networks, bandwidth is reused at each hop. There is no central controller or infrastructure support for real time traffic. Communication occurs in an ad hoc and contentious manner. Consequently, there is a high protocol overhead and it is a challenge to provide QoS guarantees in wireless ad hoc networks. However, reusing channels in each hop instead of only between cells can enhance the capacity of the system.

Figure 1: Reuse of channels between cells and within a cell

Several hybrid systems have been proposed to combine the advantages of both systems. Multi-hop cellular networks (MCN) [1] [2] [7] [8] [9] [10] [11] [12] [13] represent hybrid cellular architectures in which a mobile host can reach the base station in multiple hops. In MCN, the transmission power of MHs and BSs on the data channels is reduced to a fraction (1/k) of the cell radius, where k is a design parameter that operators choose based on several factors that we will discuss later in this thesis. A transmission range less than the cell radius makes it possible for two communication hops in the same cell to use the same channel while satifying the SIR requirements (as long as they are seperated by a sufficient distance). Figure 1 shows the MHs and their transmission ranges with different filling patterns representing different channels. As one can observe, MHs in different cells as well as MHs within the same cell can use the same channel.

In addition to a higher level of frequency reuse and enhanced capacity, MCN can offer several other benefits [1] [14] [15] [16]. First, the coverage of the network can be increased without adding more base stations. This allows MCN to provide service in 'dead spots' of a cell as well. Second, since the signal only needs to cover a small distance, the energy consumption of the nodes can be reduced. Finally, MCN also handle 'hot spots' better. It is fair to say that MCN are more tolerant to base station failures, have higher scalability, and cause less contention and collision than ad-hoc networks.

In MCN, a mobile station that has no direct connection with a base station can use other mobile stations as relays. The MCN architecture relies on the assumption that other mobile nodes will cooperate. A few incentive mechanisms have been proposed by several researchers to encourage this kind of cooperation [15] [17] [18] [19] [20].

The integrated cellular and Adhoc relay (iCAR) system uses a number of ad hoc relaying stations (ARSs) to divert the traffic from a highly loaded cell to lightly loaded neighbouring cells [3]. In [21], a detailed performance analysis for iCAR in terms of maximum possible capacity gain and blocking probability is provided. The impact of limited number of ISM-band Ad hoc relay channels are studied in [22] [23] for dynamic load balancing and load sharing.

In MADF [24], mobile units and forwarding channels are used to forward data from a host cell to surrounding cells without going through the base stations. Aggelou and Tafazolli proposed the adhoc GSM(A-GSM) architecture as an extension to the GSM cellular architecture for providing extended service coverage to dead spots [11]. In [5], B. S. Manoj et al. proposed a multihop throughput enhanced wireless architecture(TWiLL). TWiLL is designed for limited mobility systems such as wireless local loops. Simulation results show that TWiLL introduces tremendous performance improvements over the conventional Wireless Local Loop. Directional throughput-enhanced wireless in local loop (DWiLL) [25] is similiar to TWiLL except that directional antennas are used, which can reduce interference signicantly and, hence, improve throughput.

A performance comparison based on simulation, between MCN and iCAR is presented in [26]. It shows that both iCAR and MCN substantially enhance the network throughput. However, MCN suffers from problems such as network partitionings, high control overhead, and high dropping probability when hosts are highly mobile. In [1], throughput for the packet transmission is analyzed based on the RTS/CTS method. Protocols for real-time and best-effort traffic are proposed in [8] [10] for MCN.

To the best of our knowledge, there is no analysis reported in the open literature on the performance of MCN architecture for real time traffic, such as voice communication. To gain additional insight into the MCN network operation, this thesis analyzes the performance of multi-hop cellular network for real time traffic in terms of capacity and blocking probability under the constraints of relay channel availability, spatial spectrum limitation, cell radius R, and co-channel reuse level. Protocol overhead and its relationship with co-channel reuse ratio and mobility is also examined and quantified. Numerical results are given which show the performance bounds and best operating region of MCN for different parameters.

2 MCN Architecture and Protocol with Real Time Traffic Support

In this thesis report, the architecture and protocol with real time support proposed in [9] is assumed. The available bandwidth is split into one control channel and many data channels. Unlike the SCN in which data channels are assigned to cells, in MCN, data channels can be used throughout the system area provided that on-going transmissions using the same channel are seperated by an adequate distance and do not interfere with each other. While the transmission range on the data channels is reduced, the transmission range on the control channel is still equal to the cell radius. Control channels are used for mobile station registration, neighbour discovery, routing, call set up and clear, handoff, etc.

When a node makes a call by sending a CallRequest over the control channel to the base station, a path is computed by the base station. The base station also chooses the channels to be used along each wireless hop in the route. The path and allocated channel information is then broadcast to the intermediate relay nodes over the control channel. Upon receiving of this information, the sender, relay nodes, and the destination become aware of the call setup. If the call cannot be established, the caller is informed by the base station over the control channel. To allocate a channel to each hop, channels are checked in a predetermined order and the first available channel that satisfies the constraints is used. In the duration of the call, a node may move away and thus be no longer reachable; interference may increase when transmitters and receivers belong to different communication hops but using the same channel move close, and thus render one or more channels unusable. These all make the original wireless path break. MH detecting this will send a RouteError request to the base station, which will compute a new path and broadcast the information to all MHs in the cell.

The MCN architecture tries to assign as much responsibility as possible to the base stations. The base stations host the location database and the topology information database. Computation of routes and assignment of channels are done at the BSs.

The BS requires information about the topology of the nodes in its cell to compute paths to MHs. Beacon and neighbour table update mechanisms are used to provide topology information. Every MH in the system is required to participate in the neighbour discovery process. All nodes periodically (every T seconds) transmit beacons with a transmission range of R/2 using TDMA. Each node maintains a neighbour table based on the hello beacons it receives. Each node sends incremental updates to the BS over the control channel. Thus, the BS has up-to-date information about all links in the cell. The BS will maintain the topology graph of nodes in its service area in a suitable data structure. It computes the route according to the topology graph when receiving a request from a MH.

3 Problem Statement

The basic principle of operation of a multi-hop cellular network is to reduce the data transmission power of MHs and BSs. A MH reaches the BS or another MH in multiple hops. This facilitates the spectrum reuse inside the cell between different hops in addition to the reuse between cells. We assume the cell radius is R_c. The transmission range is the range that the receiver can receive the data effectively. In MCN, the transmission range R_t is reduced to a fraction of the cell radius: $R_t = \frac{R_c}{k}$. The system only establishes connection links between mobile stations when they are within a distance of R_t.

We define the following terms in the context of SCN with basic fixed channel allocation schemes. Real time traffic such as voice communication is bidirectional in nature, so each connection will include an uplink and a downlink channel. We consider the bidirectional connection as one channel when calculating N_c, N_e, and other related channel resources unless stated otherwise.

N_f : Frequency reuse factor (cluster size), $N_f = 3, 7, 11...$, etc.

N_b : Number of base stations in the coverage area,

N_c : Number of channels available in the entire spectrum over the coverage area,

N_e : Available channels for base stations in each cell in $N_e = N_c/N_f$.

Co-channel interference prohibits nearby mobile stations from using the same channel. Therefore, a parameter β is introduced to control the interference. Only when the potential interfering transmissions are at a safe distance apart (βR_t), the corresponding channel can be reused. In reality, this can be realized by making MHs listen to other MHs transmitting with the same channel. β is thus the co-channel reuse ratio in the MCN.

For practical deployment considerations, we assume that only some of the MHs in the network have the capability to relay traffic for other MHs. When a mobile station is relaying traffic for another mobile station, it is acting as a relay station (RS). Relay stations can communicate with the base station and mutiple mobile stations at the same time on different channels. They use the regular cellular system frequency bands. We further assume that each relay station can relay (and initiate) at most K_r number of connections at the same time due to hardware constraints of the physical devices.

Our ultimate goal in this study is to quantify the performance of MCN in terms of throughput and blocking probability for different parameters, compare it with SCN, and discover the maximum throughput improvement possible with the MCN architecture. We also try to quantify the overhead introduced by MCN and its relationship with active calls in the system and the speed of mobile users.

4 Performance Analysis

4.1 Per-hop Resource Usage

Figure 2: Nodes communicating using the same channels and the interference area

In cellular systems, different transmissions can use the same channel simultaneously only when the interference is small enough. Therefore, the covered geographical area of these channels is a critical resource. When a particular channel is used for the transmission between two nodes, the resource consumed is the channel and the area affected by this transmission.

As shown in Figure 2, Nodes A_1 and A_2 are communicating using channels C_1 and C_2, where one is the uplink and the other is the downlink. At the same time, nodes B_1 and B_2 are also communicating with these two channels C_1 and C_2. We define these kind of hops using the same pair of channels as co-channel hops. For each channel, the sender is transmitting at a power with tranmission range of R_t, therefore the other node sending in the same channel has to be at a distance of at least βR_t away, as described in Eq.(1). To maintain the minimum required signal to interference ratio (SIR), one must satisfy the following constraints when assigning channels to MHs initiating or relaying calls:

$$\begin{aligned} d(A_1, B_2) &\geq \beta R_t, \\ d(A_2, B_1) &\geq \beta R_t. \end{aligned} \tag{1}$$

Observe that we do not have explicit requirements for the distance between two senders using the same channel: A_1 and B_1, A_2 and B_2. From Eq.(1), one can prove that (see the detailed proof in section 8.2).

$$d(O_1, O_2) \geq \sqrt{\beta^2 - 1} R_t \tag{2}$$

where, O_1, O_2 is the midpoint of line A_1A_2, B_1B_2 respectively. The condition in Eq. (2) is equivalent to defining a circular area with a radius of $d_{if} = \frac{\sqrt{\beta^2-1}R_t}{2}$ for each hop, and ensuring that these areas for co-channel hops do not overlap. We define this area as the interference area and d_{if} as the interference distance. To maintain the required SIR in the system, it is neccessary that the interference areas of co-channel hops do not overlap with each other.

The co-channel reuse ratio β controls the average SIR in the system. In single hop cellular networks, mobile stations reach the base station in one hop. Only one side of the connection, the mobile station is moving and handoff occurs only when the mobile station moves across the cell boundary. However, in the MCN, both ends of a hop can be moving. When MHs using the same channels move close, interference increases and handoff may be required when their interference area overlap with each other. In section 5, we will discuss the relationship between β and the number of handoffs.

4.2 Available Channels in Each Cell

In MCN, the transmission power is reduced since the objective is to reach only a fraction of the cell radius, $\frac{R_c}{k}$, where R_c is the cell radius. However, this does not neccessarily render all the channels (N_c) usable at the base station in each cell. For example, when the distance between two neighbouring base stations is less than βR_t, these two base stations cannot use the same channel; otherwise, the system can not ensure the minimum SIR required. For relay stations, we can allocate the same channel when two MHs are seperated by a sufficient distance, but the base station positions are fixed. When neighbouring base stations are not seperated by a distance sufficient to satisfy the SIR requirements, clustering may still be needed. The actual configuration depends on the parameter k and the SIR requirements. The SIR requirements can be translated into a requirement of co-channel resue ratio β. For base stations in the MCN, it is also required that $\frac{D}{R_t} > \beta$, where D is the closest distance of co-channel cells and R_t is the tranmission range. One can observe that the distance between two neighbouring cells is $\sqrt{3}R_c$; hence, if $\frac{D}{R_t} = \frac{\sqrt{3}R_c}{\frac{R_c}{k}} < \beta$, then proper clustering is still needed in the system.

It is well known that the distance between co-channel cells is $\sqrt{3N_f}R_c$ [6], where N_f is the cluster size. Therefore, the ratio of co-channel cell distance and transmission range in MCN is: $\frac{\sqrt{3N_f}R_c}{\frac{R_c}{t}}$. Given the requirement $\frac{D}{R_t} > \beta$, we need to choose the minimum cluster size such that:

$$\begin{aligned} &N_f > \frac{1}{3}\left(\frac{\beta}{k}\right)^2 \\ &N_f = i^2 + j^2, i, j = 1, 2, 3, 4, \cdots \end{aligned} \tag{3}$$

The number of channels available to base stations in each cell is then $N_e = \frac{N_c}{N_f(k,\beta)}$. It is also the maximum MH-BS connnections that can be set up in each cell.

4.3 Bounding the Average Active Calls in the System

4.3.1 Interference Constraints

Considering all the MHs communicating with the same pair of channels, the sum of the interference areas of all these links, defined as the channel's cumulative interference area, is bounded by the entire coverage area (A) of the cellular system. This will give an upper bound on the number of co-channel hops that a certain area can support simultaneously. Therefore, for all the N_e channels, the sum of all channels' cumulative interference area is bounded by the spatial spectrum product: $A \cdot N_e$.

In reality, a call from a MH in one cell, in most cases, has a destination outside of that cell, which means that the MH needs to go through the base station rather than reaching the destination at another MH in the same cell. Additionally, calls received by a MH in a cell are mostly from a source outside that cell. We ignore the case of same cell calls and consider only MH-BS connections in this thesis. The main objective is to obtain a bound on the average active MH-BS connections in the system. It is worth noting that, in this thesis, when we use the words "active calls" in a cell, we refer to the active MH-BS connections in that cell.

Assume that the current active MH-BS connection number in the system is M, and let S_l denote the sum of interference areas over each hop on MH-BS connection l. Then the interference area for all the M connections is given by:

$$\begin{aligned} &\textstyle\sum_{l=1}^{M} S_l = \sum_{l=1}^{M}\left(\sum_{j=1}^{n_l} \mathcal{A}(h_{lj})\right) \\ &\mathcal{A}(h_{lj}) = \pi d_{if}^2 \end{aligned} , \tag{4}$$

where $\mathcal{A}(h_{lj})$ is the interference area for hop h_{lj}, and n_l is the number of hops traversed in connection l.

In Eq. 4, rearranging all the $\mathcal{A}(h_{lj})$ terms according to the channels they are using, one obtains:

$$\sum_{l=1}^{M} S_l = \sum_{i=1}^{N_e} \Bigg(\sum_{\{\forall j,l|\mathcal{C}(h_{lj})=c_i\}} \mathcal{A}(h_{lj}) \Bigg), \tag{5}$$

where c_i is channel pair i and $\mathcal{C}(h_{lj})$ represents the channel pair used on hop h_{lj}.

Taking expectations on both sides of Eq. (4), one has:

$$\begin{array}{l} E\Big(\sum_{l=1}^{M} S_l \Big) = E(n) \cdot \pi d_{if}^2 \cdot E(M) \\ E\Big(\sum_{l=1}^{M} S_l \Big) \geq \frac{E(d)}{R_t} \cdot \pi d_{if}^2 \cdot E(M) \end{array}, \tag{6}$$

where, $E(M)$ is the average number of active calls, $E(n)$ is the average number of hops in a connection, and $E(d)$ is the average distance from MH to BS.

From Eq. (5) we have:

$$\Big(\sum_{l=1}^{M} S_l \Big) < A \cdot N_e \tag{7}$$

Combining Eq. (6) and Eq. (7), the average number of active MH-to-BS connections can be bounded by the following inequality:

$$E(M) < \frac{A \cdot N_e}{\frac{E(d)}{R_t} \cdot \pi d_{if}^2} \tag{8}$$

We define the normalized average active calls ($\overline{E(M)} \triangleq \frac{E(M)}{N_c \cdot N_b}$). It actually represents the average active calls per cell normalized by the entire available channel number N_c.

Since $N_b = A / \Big(\frac{3\sqrt{3}R_c^2}{2} \Big)$, $\frac{E(M)}{N_c \cdot N_b}$ can be bounded by the following formula:

$$\overline{E(M)} = \frac{E(M)}{N_c \cdot N_b} < \frac{6 \cdot \sqrt{3} \cdot k}{\eta \cdot \pi \Big(\beta^2 - 1 \Big)} \cdot \frac{1}{N_f}, \tag{9}$$

where $\eta = \frac{E(d)}{R_c}$.

Assuming randomly distributed MHs in the coverage area, it is easy to show that (see section 8.1 for details):

$$\eta = \frac{6\sqrt{3} - \pi}{12} \approx 0.60$$

4.3.2 Relay Node and Channel Constraints

The achievable thoughput in terms of average active connections in the system is also limited by the number of relay nodes and channels available in the coverage area and their capabilities.

In each MH-BS connection, if the number of hops involved is n_l, then the relay channels used in the connection is: $2 * (n_l - 1)$. Summing over all connections, one has:

$$\sum_{l=1}^{M} 2 * (n_l - 1) < N_r \cdot K_r,$$

where N_r is the number of MHs in the system and K_r is the average relay channels of each MH.

Taking expectations on both sides, one gets:

$$E(M) \cdot (E(n_l) - 1) < \frac{N_r \cdot K_r}{2},$$

where $E(n_l) = \frac{E(d)}{R_t} = \eta \cdot k$. This can be further simplified to:

$$E(M) < \frac{N_r K_r}{2\Big(k\eta - 1\Big)}$$

$$\overline{E(M)} = \frac{E(M)}{N_c \cdot N_b} < \frac{N_r}{N_b} \cdot \frac{K_r}{N_c} \cdot \frac{1}{2\Big(k \cdot \eta - 1\Big)} \tag{10}$$

This is the performance bound of expected MH-to-BS connections in terms of relay nodes available.

4.4 Blocking probability Analysis

Equations (9) and (10) give the performance bounds on MH-to-BS connections. Combining these two equations and the result in section 4.2, one obtains:

$$\overline{E(M)} < \begin{cases} \frac{1}{N_f(k,\beta)} \\ \frac{6\cdot\sqrt{3}\cdot k}{\eta\cdot\pi\left(\beta^2-1\right)} \cdot \frac{1}{N_f} \\ \frac{N_r}{N_b} \cdot \frac{K_r}{N_c} \cdot \frac{1}{2(k\cdot\eta-1)} \end{cases} \tag{11}$$

where $N_f(k, \beta)$ is the cluster size determined from Eq. (3).

By defining EM_{max} as:

$$EM_{max} = min\Big(\frac{1}{N_f(k,\beta)}, \frac{6\cdot\sqrt{3}\cdot k}{\eta\cdot\pi(\beta^2-1)}\cdot\frac{1}{N_f}, \frac{N_r}{N_b}\cdot\frac{K_r}{N_c}\cdot\frac{1}{2(k\cdot\eta-1)}\Big)$$

one gets:

$$\overline{E(M)} < EM_{max} \tag{12}$$

The expression in (12) gives some insight into the blocking probabilty. From Little's Law, we know that $E(M) = \frac{\lambda}{\mu}\cdot(1-P_B)$, where λ is the call arrival rate in the entire cellular system, and μ is the average call holding time. If we define the load as $\rho = \frac{\lambda/N_b}{N_c\mu}$, then:

$$P_B > 1 - \frac{EM_{max}}{\rho} \tag{13}$$

In addition, we still need to consider the reachability of relay stations in the cell. Assuming random and independent distribution of MH locations, the number of MHs in a certain area πr^2 follows a Poisson distribution with parameter ξ, where $\xi = \frac{N_r}{N_b\cdot\frac{3\sqrt{3}}{2}R_c^2}$ represents the density of MHs in the area. Then, the probability that a MH will find another MH within the radius of R_t is $1-e^{-\xi\pi R_t^2}$. Let $P_c(h)$ represent the probability of sucessfully setting up a call with h hops. Its value is determined by the probability that the MH can find a relay station in each hop; hence,

$$P_c(h) < (1-e^{-\xi\pi R_t^2})^h. \tag{14}$$

For a connection with h hops, the probability, $P_{br}(h)$, that the call will be blocked because there are no available relay stations is given by:

$$P_{br}(h) > 1-(1-e^{-\frac{2\sqrt{3}\pi}{9k^2}\frac{Nr}{N_b}})^h.$$

By reducing the transmission distance of MHs and BSs to a fraction $(1/k)$ of the cell radius, the cell is divided into k layers with radius difference of R_t. One can observe that the probability of a connection to the base station having h hops is greater than the probability that the initial MH located in layer h of the cell, $P_l(h) = \pi\frac{2h-1}{k^2}$. Hence:

$$P_B > \begin{cases} 1-\frac{EM_{max}}{\rho} \\ \sum_h \Big(P_l(h)\cdot P_{br}(h)\Big) \end{cases} \tag{15}$$

Equation (15) gives a lower bound for the blocking probability of the system. However, the analysis above does not capture the fact that all calls need to connect to the base station, which makes the area near the base station a hot spot. Intuitively, this will result in more blocked calls than that has been illustrated above.

4.4.1 Markov Chain Modeling of the system

Exact blocking probability analysis is very complex because of the dynamic nature of relaying nodes and non-uniform resource distribution. However, one can derive a lower bound assuming ideal conditions. Since the transmission range is reduced to $1/k$ of the cell radius, we can divide the cell into k layers, as shown in Figure 3.

Figure 3: The cell is divided into k layers. MHs in layer i need to go through at least i hops to reach the base station

Mobile nodes in layer i would need to go through at least i hops to reach the base station. To simplify the analysis, we assume that there is a sufficient number of relay nodes available so that mobile stations can always find a proper relay station in the direction of the shortest path to the base station. The call originating mobile stations are assumed to be optimally located to pack as many calls as possible in the same area. We further assume that one can adjust the allocated channels of existing calls on the fly in order to allow maximum reusage.

Under the aforementioned assumptions, one can model the multi-hop connections to the base station as a multi-link circuit-switching network as shown in Figure 4. Each link corresponds to a hop from one MH in one layer to another MH in the next layer in Figure 3. The capacity of each link ($C_k, C_{k-1}, \ldots, C_1$, etc.) is determined by the freqency and spatial resources available in each layer. The arrival rate at each node λ_i is the call arrival rate in each layer.

If we assume that calls arrive according to Poisson processes and the length of call holding time is exponentially distributed, we can model the system as a multi-dimensional Markov Chain. The state is defined as $(m_1, m_2, \ldots, m_{k-1}, m_k)$, in which $m_i (i = 1, 2, \ldots, k)$ represents the number of on-going calls originating from layer i.

Figure 4: Modeling the system as a circuit switching network with the links representing hops beween neighbouring layers

Figure 5: State transition diagram of the Markov Chain. The state representing on-going calls in all layers of the cell.

Figure 5 shows the state transition diagram, in which $(m_1, m_2, \ldots, m_{k-1}, m_k)$ is a valid state. The state $(m_1, m_2, \ldots, m_{k-1}, m_k)$ is a valid state only when it satisfies the channel and space constraints in the cellular systems.

Demand matrix describes the link requirements of different types of calls; in our case, calls originating from different layers of the cell. Let the demand matrix be defined as $\mathcal{M} \triangleq \{a_{ij}\}$. A call originating from layer j uses a_{ij} circuits on link i. It is easy to show that $\vec{m} = \{m_1, m_2, \ldots, m_k\}$ has a unique stationary distribution given by:

$$\begin{array}{c} \pi(\vec{m}) = G(\vec{C})^{-1} \cdot \prod_{(l=1,\ldots,k)} \left((\frac{\lambda_l}{\mu})^{m_l} \cdot \frac{1}{m_l!} \right) \qquad \vec{m} \in \mathcal{L}(\vec{C}) \\ \mathcal{L}(\vec{C}) = \{\forall \vec{m} | \mathcal{M}\vec{m} \leq \vec{C}\} \\ G(\vec{C}) = \left(\sum_{\vec{m} \in \mathcal{L}(\vec{C})} \prod_{(l=1,\ldots,k)} \left((\frac{\lambda_l}{\mu})^{m_l} \cdot \frac{1}{m_l!} \right) \right) \end{array} , \qquad (16)$$

where $\vec{C}$ is the capacity vector $(C_1, C_2, \ldots, C_k)^T$, $G(\vec{C})$ is the normalizing constant, and $\mathcal{L}(\vec{C})$ represents the collection of all valid states in the system.

The probability for calls originating from layer r to find frequency resource to reach the base station is then given by

$$P_f(h) = G(\vec{C})^{-1} G(C - \mathcal{M}\vec{e}_r), \qquad (17)$$

where $\vec{e}_r$ is a unit vector with all its elements equal to zero except the rth one, and $\mathcal{M}\vec{e}_r$ is the link usage vector when there is only one active call in the entire cell, which originates from layer r.

If the total call arrival rate in one cell is λ_c, and the arriving calls are uniformly distributed in the cell, then the call arriving rate in layer i is

$$\lambda_i = \lambda_c \cdot \frac{2i - 1}{k^2}.$$

Since Eq. 14 shows the probability of sucessfully finding relay nodes for a path of h hops, the overall blocking probability in the cell is given as

$$P_B = 1 - \sum_{i=1,\ldots,k} \left(\lambda_c \cdot \frac{2i-1}{k} \cdot \left((1 - e^{-\xi \pi R_t^2})^h \cdot G(\vec{C})^{-1} G(\vec{C} - \mathcal{M}\vec{e_i}) \right) \right). \tag{18}$$

4.4.2 The State Space

Figure 6: For co-channel hops (A_1A_2 and B_1B_2) in the same layer, the interference distance between them determines the minimum value of angle θ.

To solve the equations in the previous section, we need to determine the state space $\mathcal{M}(\vec{C})$. The available channels and geographic area limit the number of active links on each layer. Let the total number of channels available to the cell be N_e. Considering two nearby links communicating simultaneously with the same channel pair in Figure 6, we have:

$$\theta \geq F(r_1, r_2, d_{if}) = cos^{-1}(\frac{r_1^2 + r_2^2 + 4d_{if}^2}{2r_1r_2}). \tag{19}$$

When these two links are in the same layer i, the range of r_1 and r_2 are determined by i, then:

$$\begin{gathered} \theta \geq \mathcal{F}(i, d_{if}) = Min(F(r_{1m}, r_{2m}, d_{if})), \\ r_{1m} \in \left\{ \ (i-1)R_t, \ i \cdot R_t \ \right\}, \\ r_{2m} \in \left\{ \ (i-1)R_t, \ i \cdot R_t \ \right\}. \end{gathered} \tag{20}$$

For each layer, the maximum number of links communicating with the same channel pair is

$$\mathcal{N}(i) = \frac{2\pi}{\mathcal{F}(i, d_{if})}. \tag{21}$$

The link capacity vector is

$$\vec{C_a} = \{C_i\}, \qquad C_i = N_e \cdot \mathcal{N}(i). \tag{22}$$

Figure 7: A circle with radius R_d and a communication hop at a distance of R_a. α is the portion of the circle that the interference area of the hop covers.

The matrix describing the link demands of calls from each layer is

$$\mathcal{M}_a = \left\{ \begin{array}{ccccc} 1 & 1 & 1 & \dots & 1 \\ 0 & 1 & 1 & \dots & 1 \\ 0 & 0 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{array} \right\}. \qquad (23)$$

Up to this point, the state space of the system is determined by the inequality

$$\mathcal{M}_a \cdot \vec{m} < \vec{C}_a.$$

However, links in adjacent layers interfere with each other as well. From the analysis in section 4.1, we know that each co-channel hop occupies an interference area with radius d_{if}. Interference areas of co-channel hops cannot overlap with each other.

Considering any circle centered at the BS with radius R_d, and a communication hop from layer i to layer $i-1$, Figure 7 shows the portion of the circle covered by the interference area of this hop. The radian value of the covered circle is α. It can be shown that,

$$\alpha(i, R_d) = 2 \cdot cos^{-1} \frac{R_a^2 + R_d^2 - d_{if}^2}{2R_a R_d}$$

The value of R_a can be obtained from the fact that it is a hop from layer i to layer $i-1$.

Communication hops in each layer occupies a portion of the curve. Considering hops at every layer of the cell, we define:

$$\vec{d}(R_d) \triangleq \frac{1}{2\pi}\{\alpha(1, R_d), \dots, \alpha(i, R_d), \dots, \alpha(k, R_d)\}$$

There are infinite number of choices for the curve and its radius R_d, hence there can be infinite number of vectors $\vec{d}(R_d)$. In the calculation, we will sample a few number of curves with different R_d and choose those that give tight constraints.

Let

$$D = \left\{ \begin{array}{l} \vec{d}_1(R_{d1}) \\ \vec{d}_2(R_{d2}) \\ \dots \\ \vec{d}_i(R_{d_i}) \\ \dots \\ \vec{d}_w(R_{d_w}) \end{array} \right\} \tag{24}$$

and $\vec{n} = \{n_1, n_2, \dots, n_k\}^T$, in which n_i represents the number of communication hops on layer i. One obtains:

$$\begin{aligned} & D \cdot \vec{n} < \vec{C}_b \\ & \vec{C}_b = \{\underbrace{1, \dots, 1, \dots, 1}_{w}\}^T \end{aligned} \tag{25}$$

Since $\vec{n} = \mathcal{M}_a \cdot \vec{m}$

$$D \cdot \mathcal{M}_a \cdot \vec{m} < \vec{C}_b.$$

Let:

$$\vec{C}^* \triangleq \left\{ \begin{array}{c} \vec{C}_a \\ \vec{C}_b \end{array} \right\}, \quad M^* \triangleq \left\{ \begin{array}{l} \mathcal{M}_a \\ D \cdot \mathcal{M}_a \end{array} \right\}, \tag{26}$$

One thus obtains:

$$M^* \cdot \vec{m} < \vec{C}^* \tag{27}$$

Eq. 27 is the final constraint that determines the state space. $\mathcal{M}^*$ and $\vec{C}^*$ denote the demand matrix and capacity vector, respectively. An example of matrix $\mathcal{M}^*$ and $\vec{C}^*$ when $k = 7$ and $\beta = 3$ is given below:

$$\mathcal{M}^* = \left\{ \begin{array}{ccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 2 & 2 & 2 & 2 & 2 & 2 \end{array} \right\} \tag{28}$$

$$\vec{C}^* = N_e \cdot (1 \quad 1 \quad 1 \quad 7 \quad 9 \quad 12 \quad 14 \quad 1)^T$$

Given the demand matrix $\mathcal{M} = \mathcal{M}^*$ and link capacity vector $\vec{C} = C^*$, the stationary distribution and blocking probability can be computed from Eq. (16) and Eq. (18).

4.4.3 Computing the Blocking Probability

The analysis in sections 4.4.1 and 4.4.2 gives an explicit framework for the lower bound of the blocking probability. However, this is far from providing the complete solution. For some cases it is impractical to compute G and the blocking probability directly. Note that the state space increases rapidly when the link capacity C_i increases. The computational complexity grows exponentially when the number of layers increases. It is easy to show that the computational complexity is $\mathcal{O}(k \cdot N_c^k)$. On a Windows XP Pentium III system, computing the blocking probability directly from Eq. (16), and (18), the program takes around 3 seconds to finish when $N_c = 100$, and $k = 3$. This implies that it would take more than 800 hours to compute the blocking probability when $N_e = 100$, and $k = 6$.

To reduce the computational cost, decompostion methods proposed in [28] [29] [30] break down the system into subnetworks and can reduce the computation time by several orders of magnitude. In [29], several cut sets are identified from the system and analysis is reduced to the dimension of the subsystems. Cut sets are collections of links and routes that the routes in one set do not use links in the other set. Li extends the decomposition method in [30] so that the noninterfering property of cut sets can be relaxed.

By observing the demand matrix in the previous section, since we find no obvious cut sets in our system, we cannot directly apply the aforementioned method. In order to compute the blocking probability efficiently, we extend the decomposition method so that no explicit cut set is required to break down the computation.

From Eq. (16), we know that

$$G(\vec{C}) = \left(\sum_{\vec{m} \in \mathcal{L}(\vec{C})} \prod_{(l=1,\ldots,k)} \frac{\rho_l^{m_l}}{m_l!} \right) \qquad \rho_l = \frac{\lambda_l}{\mu} \ . \tag{29}$$

Now we divide the type of calls (i.e., calls from different layers) into two sets, R_1 and R_2. Using the scenario described by Eq. (28) as an example, R_1 is the collection of call types originating from layers $4, 5, 6$, and 7, and R_2 is the collection of call types originating from layers $1, 2$, and 3. Let $\mathcal{L}_i(\vec{X})$ be the collection of valid call status when only call of type R_i is permitted given link capacity vector $\vec{X}$; i.e.,

$$\mathcal{L}_i(\vec{X}) \triangleq \{\forall \vec{m} | \mathcal{M}\vec{m} \leq \vec{X} : m_j = 0, \forall j \notin R_i\}$$

Let $\vec{U_1}$ be the vector describing the link demands of call set R_1 in the following way:

$$\vec{U_1} = \{u_1(i)\} \quad u_1(i) = \begin{cases} 0, \text{ if } \forall c_j \in R_1, a_{ij} = 0 \\ 1, \text{ otherwise} \end{cases} \tag{30}$$

where c_j represents call type j.

We define a projection operator

$$\mathcal{P}_1(\vec{n}) = \{x_j\} \quad x_j = n_j \cdot u_1(j).$$

Eq. (29) can now be transformed into:

$$\begin{aligned} G(\vec{C}) &= \textstyle\sum_{\forall \vec{\xi}} \left(G_1(\vec{C} - \vec{\xi}) \cdot Q_2(\vec{\xi}, \vec{C}) \right) \\ \vec{\xi} \in \Omega &= \left\{ \forall \vec{\xi} | \mathcal{P}_1(\mathcal{M}\vec{m}) = \vec{\xi}, \vec{m} \in \mathcal{L}_2(\vec{C}) \right\} \\ Q_2(\vec{\xi}, \vec{X}) &= \textstyle\sum_{\forall \vec{m} | \mathcal{P}_1(\mathcal{M}\vec{m}) = \vec{\xi}} \left(q_2(\vec{m}, \vec{X}) \right) \\ q_r(\vec{m}, \vec{X}) &= \textstyle\sum_{\vec{m} \in \mathcal{L}_r(X)} \left(\prod_j (\frac{\rho_j^{m_j}}{m_j!}) \right) \end{aligned} \tag{31}$$

$Q_2(\vec{\xi}, \vec{C})$ is the unnormalized probability of the link status whose projection on U_1 is $\vec{\xi}$ assuming that only calls in set R_2 are permitted. $G_1(\vec{X})$ is the normalized constant on the reduced calling set $\mathcal{L}_1(\vec{X})$.

$$G_i(\vec{X}) = \left(\sum_{\vec{m} \in \mathcal{L}_i(X)} \prod_{(l=1,\ldots,k)} \frac{\rho_l^{m_l}}{m_l!} \right) \tag{32}$$

The blocking probability of call type r is given by

$$\begin{aligned} P_{B_r} &= 1 - G(\vec{C})^{-1} G(\vec{C} - \mathcal{M}\vec{e}_r) \\ G(C - \mathcal{M}\vec{e}_r) &= \textstyle\sum_{\forall \vec{\xi}} \left(G_1(\vec{C} - \mathcal{M}\vec{e}_r - \vec{\xi}) \cdot Q_2(\vec{\xi}, \vec{C} - \mathcal{M}\vec{e}_r) \right) \\ \vec{\xi} \in \Omega_r &= \left\{ \forall \vec{\xi} | \mathcal{P}_1(\mathcal{M}\vec{m}) = \vec{\xi}, \vec{m} \in \mathcal{L}_2(\vec{C} - \mathcal{M}\vec{e}_r) \right\} \end{aligned} \tag{33}$$

The complexity for computing $Q_2(\vec{X})$ is $\mathcal{O}(N_e^{|R_2|})$, and the complexity for computing $G_1(\vec{X})$ is $\mathcal{O}(N_e^{|R_1|})$. Let $|\Omega|$ be the maximum size of Ω_r, then the computational complexity for this algorithm is $\mathcal{O}\Big(|\Omega| \cdot k \cdot (N_e^{|R_1|} + N_e^{|R_2|}) \Big)$. The value $|\Omega|$ depends on the way we divide R_1 and R_2, the exact value of the demand matrix and the link capacity vector. It quantifies the number of sets of call status whose calls are from call type collection R_2, and each set of the call status results in link usages that have the same constraint on calls from collection R_1. The more zero elements in vector $\vec{U_1}$, the lower the value of $|\Omega|$ will be. When there are more uniformity and zero elements in the demand matrix, it is more likely that we will have the same $\vec{\xi}$ (in Eq. (33)) for different call status $\vec{m}$, hence a lower value of $|\Omega|$.

Given the demand matrix in (27), if R_1 represents call types originating from layers $p+1$ to k, and R_2 includes call types originating from layers 1 to p, we can observe that for matrix $\mathcal{M} = \{a_{ij}\}$, there is $a_{ij} = 1 \ \forall\ 1 \le i \le p\ \&\ 1 \le j \le p$, which means ξ_i always equal to each other for all $i \le p$. Other than that, for any $1 \le j \le p$, and $u_i \neq 0$, there are non-zero elements only at line $i = k+1$. Therefore, the value of $|\Omega|$ is less than N_e^2. The computational cost is reduced when we have a large k.

From the above analysis, we know that although there are no obvious cut sets in the network, there are some zero elements and uniformity in the demand matrix. This makes the sizes of Ω_r much less than the size of $|\mathcal{L}_2(\vec{C})|$. Thus, the algorithm reduces the computational complexity by exploiting the sparsity and uniformity of the demand matrix.

Figure 8: The speed that two MHs moving into each other can be obtained from their relative velocity projecting on the axis connecting these two MHs

5 Protocol Overhead Estimation

The MCN architecture and protocol with real time support described in section 2 includes lots of overhead on topology updating and routing. The base stations keep the location and topology information and compute the route. We assume that the base station is able to complete the necessary computations and ignore the *computing overhead*. We believe this is a reasonable assumption given the current trends in available processing power. Given this assumption, we proceed to analyze the *communication overhead* enforced by the protocol.

The current method for topology discovery is based on the use of hello beacons transmitted periodically. Each base station is assigned a special small updating channel to transmit beacon signals so that the protocol is contention and collision free. Each of these updating channels has enough bandwidth for one mobile station to transmit hello beacons at a frequency of $1/T$ seconds. We assume that the neighbourhood discovery channel uses different frequency bands than other data and control channels, and each MH's updating channels are in seperate time slots.

The base stations assign the time slots to MHs during the registration. Assuming that a hello beacon consisting of a node identifier (six bytes should be enough), one header and one checksum, one beacon message is $l = 8$ bytes. Now, if the the number of MHs in the cell is $N_m = 1000$ and updating period is $T = 5$ *second*, the total bandwidth required is $\frac{N_m * l}{T}$ $byte/sec = 12.8$ $kbps$.

After discovering neighbours from the received beacon signals, the MHs must also transmit the neighbour information to the base station. Because this information is transmitted incrementally, the MHs only inform the BS about new MHs that have moved into its transmission range and MHs that have moved out of its transmission range. The communication overhead needed is obviously related to the moving speed of mobile stations. We assume that the mobiles are moving towards each other at a speed of $v(m/s)$, which can be obtained from the relative velocity projecting on the axis connecting these two MHs, as shown in Figure 8. In section 8.3, we will show how one can obtain the average relative speed between two MHs ($\overline{v}$) from the distribution of MH moving velocity.

Given the transmission range R_t, the current MH will inform the BS about the neighbouring MHs crossing the boundary of radius R_t during time peroid T . When two MHs are moving close, the reported MH is in the area between two circles of radius $R_t + vT$ and R_t, as shown in Figure 9-a. When they are moving away, the reported MH is in the area between two circles of radius R_t and $R_t - vt$, as shown in Figure 9-b. The node density is given by $\xi = \frac{N_m}{\frac{3\sqrt{3}}{2}R_c^2}$. Given a randomly distributed moving speed and direction, it can be shown that any two MHs have equal probability of moving close and moving away (see section 8.3 for proof). Let $\overline{v}$ represent the average relative moving speed between any two MHs ($\overline{v}$ is a scalar and it is always true that $\overline{v} > 0$), the average number of nodes the current MH needs to report is: $\xi \cdot 2\pi R_t \overline{v} T$. The total bandwidth needed is then given by:

$$N_m^2 \frac{4\pi\overline{v}l}{3\sqrt{3}kR_c}, \tag{34}$$

where l is the message length to report one node.

Figure 9: MHs crossing the boundary at radius R_t centering around the current MH are those need to be reported to the BS

Assume that there are 1000 nodes per cell, the cell radius is $R_c = 500\ m$, $k = 3$, and the average speed is $\overline{v} = 4\ m/s$, the average bandwidth requirement is 412 $kbps$. Comparing with the toal GSM bandwidth in a typical cell, $60 \cdot 200\ kbps$, this corresponds to an overhead of less than 4%.

It has been shown via simulations that in MCN, there is an inceasing overhead due to handoffs [26]. We will estimate the communication overhead resulting from handoff calls. Assume that a call needs handoff when for any node in its route to the base station, there is an interfering channel at a distance of βR_t away. Given M number of active calls in a cell, for any one of the active calls (call i, with $h(i)$ hops), in a short time peroid of Δt, the probability that a handoff is needed is given by:

$$h(i) \cdot \frac{1}{N_{ae}} \cdot \xi_a \cdot \frac{1}{2}\pi((\beta R_t + v\Delta t)^2 - (\beta R_t)^2), \tag{35}$$

where ξ_a is the density of nodes that are engaged in active calls, and N_e is the number of total channels used in the cell. The average number of handoffs needed in time t_r is:

$$N_h = \overline{h} \cdot \frac{1}{N_e} \cdot \xi_a \cdot \pi\beta R_t \overline{v} t_r, \tag{36}$$

where $\overline{h}$ is the average hop number, and $\overline{v}$ is the average speed. Observing that $\xi_a = \frac{\sum_{j=1}^{M} h(j)}{A_c}$, and $\sum_{j=1}^{M} h(j) \cdot \pi\frac{(\beta^2-1)}{4} R_t^2 < A_c \cdot N_{ae}$, from Eq. (36), one can get

$$N_h < \frac{4\overline{h} \cdot \beta \cdot \overline{v} \cdot t_r}{(\beta^2 - 1)R_t}. \tag{37}$$

If we have $\overline{h} = k = 3, R_c = 500\ m, \beta = 3, \overline{v} = 4\ m/s$ (corresponding to a random distribution of moving speed between $0 - 10m/s$), and one handoff on average takes 128 bytes of message transmitted, the average bandwidth requirement is about 110 *bps* per active call. Considering the voice channel of 13 *kbps* in GSM, the overhead is less than 1%. One can observe from Eq. (37) that the handoff overhead will increase as the speed of MHs increases.

6 Results and Discussion

6.1 Results on capacity bounds

Figure 10: Capacity bound imposed by spatial freqency production

Figure 11: Capacity bound imposed by spatial freqency production

In this thesis, we first studied the bounds on average active calls with different parameters. In Figures 10 and 11 , the curves show the capacity bounds in terms of average active calls in each cell normalized by the total available channels in the system. Solid curves show the MCN performance bounds imposed by spatial frequency factors.

Dashed curves show the fraction of channels available for each cell in SCN networks and hence the maximum active calls in the cell for different cluster sizes (N_f) and cochannel reuse ratios (D/R). In Figure 10, different curves are plotted as a function of transmission factors k for different values of the interference parameter β, and in Figure 11, different curves are plotted for different k values as a function of β. From these results, one can see that the bounds on active calls get looser when k increases, and get tighter as β increases. They are direct results from Eq. 9. The discontinuity in both plots result from the change of cluster size in MCN.

When the cochannel reuse reatio (D/R) in SCN equals the parameter β in MCN, they have the same average SIR, hence a given user experiences the same quality of service in both networks. To compare other performance metrics under the same SIR, one can look at the performance curve of MCN when $\beta = 4.5$ and the curve of SCN with cochannel reuse reatio $D/R = 4.6$. One can observe that, in most regions, MCN provides a large performance improvement over SCN, but when k is small, the improvement over the SCN networks is relatively small.

Figure 12 shows the bounds from both the spatial freqency factor and relay node/channel availability. Dashed lines are bounds imposed by different node densities for different transmission range factor k values, while the solid lines are bounds imposed by spatial freqency factor, and dotted lines indicate the performance of SCN. Node density is defined as $\frac{N_r}{N_d}$, i.e., the number of relay nodes per cell. Relay node capability $\frac{K_r}{N_c}$ is set to 0.05 in all plots. One can see that as we increase k, the bound imposed by spatial factor gets looser, whereas the bound imposed by relay node density gets tighter. This is because as k increases, transmitting power is reduced, so there is more frequency reuse in the same space; however, the call needs to go through more relay nodes to reach the base station. For a designated cochannel reuse ratio β and node density $\frac{N_r}{N_d}$, the corresponding solid line and dashed line together determine the bound on average active calls. Given the required value of β and node density, one needs to choose the tranmission range factor k wisely. For example, when $\beta = 6.0$, $\frac{N_r}{N_d} = 40$ and $k = 4$ would be good choices for maximum performance improvement.

Figure 12: Capacity bound imposed by spatial freqency production and relay node/channel availability, $\frac{K_r}{N_c} = 0.05$

6.2 Results on blocking probability

In this section, we present the analysis and simulation results on blocking probability in the MCN system given that the number of usable channels in the cell remains the same.

Blocking probability bounds derived from Markov Chain analysis of section 4.4 are ploted in Figure 13 for MCN networks, comparing them with the bound for SCN networks with the same cochannel reuse ratio ($D/R = \beta$). In both cases, MCN networks have lower bounds that are much lower than the blocking probability of SCN networks.

Figure 13: Blocking probability compared with SCN

Blocking probabilities for different k values are plotted in Figure 14 as a function of β; blocking probabilities for different β values are plotted in Figure 15 as a function of k. For both of the plots, we assume that the number of available channels for the cell is the same ($N_e = 40$), and utilization $\rho = 0.6$.

One can observe from Figure 14 that the bound for blocking probability gets looser rapidly with β when β is small, and the curve becomes flat when β is large. This is easy to understand because although a larger β value indicates a larger average SIR in the system, it also means that each hop will occupy more area so the channel reuse ratio is decreased, and hence the system has a higher blocking probability. However, when β is large enough such that for almost each communication hop in the cell, the interference area of the hop covers the entire cell, then increasing the β value does not have a significant effect on the bounds obtained from the analysis. The increase in the value of β would affect the performance of the other cells but our analysis does not account for this factor yet.

Figure 15 shows that the blocking probability bounds gets looser slowly as k increases. This is because when β is larger than a certain value, i.e. large enough for a communication hop at the *ith* layer of the cell to interfere with the communication close

Figure 14: Blocking probability as a function of β for different k values

Figure 15: Blocking probability as a function of k for different β values

to the base station, increasing k actually increases the number of communication channels that will interfere with the communication near the base station, which is called the "hot spot" effect. This seems contradictory to the analysis we have given in previous sections, but recall that curves are plotted under the assumption that the usable channel in the cell N_e is the same. Actually a larger k value will result in a higher N_e value in reality.

To investigate the tightness of the bounds derived for blocking probability, we compared our analysis with simulation results. Simulation is written in C#. C# is a language developed by Microsoft which tries to combine the benefits of C++ and Java. We choose this language because of its similarity to C++ language and its strong capability to handle data structures. Shortest path algorithm and the simple "first available channel allocation" schemes are used. Table-1 shows the parameters used in the simulation and the analysis. MH locations are uniformly distributed in the cell. The random way point model is used to model the movement of MHs and the maximum speed is set to 10 m/s.

Table-1 Parameters used in the simulation and analysis

Description	Value
BTS Number	1
MH Number	600
Cell radius	500m
Average inter-arrival time of calls for each MH	3600 second
Cluster Size	1

Figure 16: Blocking probabilty from analysis and simulation, Number of Channels = 40

Figure 17: Blocking probabilty from analysis and simulation, Number of Channels = 40

Figure 18: Communication overhead due to topology discovery

Figures 16 and Figure 17 compare our analysis results with the simulation results. In Figure 16, $k = 2$, $\beta = 3$ and in Figure 17, $k = 3$, $\beta = 5$. One can observe that the blocking probabilities from simulation results are all above the analysis results, as expected, and they are within 20% range in most situations. It is important to emphasize that the blocking probability bounds derived in our analysis and shown in Figures 16 and Figure 17 represent the best-case performance. The simulation results in both figures, on the other hand, represent results that were obtained with reasonable system parameters (see Table-1), naive channel allocation schemes and routing methods.

6.3 Results on protocol overhead

In this section, we present the results on protocol overhead estimation. Figure 18 shows the communication overhead required for topology discovery. Figure 18(a) shows the overhead versus the number of nodes in each cell. Four different curves are for four different values of average relative speed $\overline{v}$. Figure 18(b) shows the overhead versus average relative speed. Five different plots are corresponding to five different values of node number in the cell. The topology discovery overhead is related to the total registered number of mobile stations in the cell, and not related to the accepted number of calls.

Figure 19 shows the additional communication overhead due to handoffs. Figure 19(a) shows the per call handoff overhead versus the co-channel reuse ratio in the MCN. Four different curves are for four different values of average moving speed. Overhead is specified per active call because the more active calls in the system, the more handoffs needed. The larger the co-channel reuse ratio, the lower the handoff overhead.

Figure 19: Communication overhead (per active call) due to handoffs

Figure 19(b) shows the overhead vs. average relative speed between two MHs. Observe that the overhead increases linearly with the relative speed. Four different plots correspond to four different values of co-channel reuse ratio β.

7 Conclusions and future work

Because of the complexity of analyzing the performance of real time traffic in multi-hop cellular networks, no closed form expression for blocking probability exists in the open literature. Previous simulation results show that multi-hop cellular networks can provide a performance improvement over single-hop cellular networks, but no one has shown how much improvement multi-hop cellular networks can ultimately achieve. People also choose parameter k and β based on experience and intuition. In this study, we have attempted to quantify how much performance improvement multi-hop cellular networks could yield and have also tried to estimate the protocol overhead. The study also provides a guide for choosing the operating system parameters, such as k and β.

In this study, we have analyzed the performance of multi-hop cellular networks by finding the maximum average active calls in the system, and by modeling a single cell system as a Markov process. Both methods yeild a lower bound on blocking probability. Maximum mean active call analysis leads to a tighter bound when β is large compared to k, since the Markov Chain analysis does not account for the interference between different cells. When k is larger than β, Markov Chain analysis gives a tighter bound because it considers more detailed interaction between wireless links in different layers. We have compared our analysis with simulation results as well and have validated the prediction of our analysis. The analysis yields a reasonable tight lower bound for the blocking probability.

Our results show that although multi-hop cellular networks can provide performance improvements in most cases, one still needs to choose operating parameters carefully to allow for maximum performance enhancement. This work presents *a bound on the maximum performance improvement*. In other words, the bounds derived for blocking probability should be interpreted as the best-case performance that multi-hop cellular networks can achieve. How much improvement a multi-hop cellular networks can actually achieve in practice depends on the implementation, including the channel allocation scheme and the routing algorithm used. Mobility will also reduce the attainable performance improvement.

Mobility is a factor that will introduce extra overhead in multi-hop cellular networks and this has been shown by simulations in other studies [26] [27]. This thesis presents an estimation of protocol overhead and overhead introduced because of increased handoffs. The relationship between the overhead and the number of MHs in the system and average moving speed is presented. We have shown that the overhead is reasonable and not excessive in a wide range of circumstances.

To achieve maximum capacity gain in multi-hop cellular networks, it appears that one should pack as many nodes transmitting with the same channel as possible in the coverage area. On the other hand, channels are reused more frequently and interference occurs more frequently due to the mobility of nodes. To circumvent this either mobility prediction or more spatial distance between nodes transmitting via the same channel is needed. Further research is needed to determine whether sophisticated channel allocation schemes can be used to meet the aforementioned two contradicting requirements.

The Markov Chain analysis conducted in this thesis does not capture the interference between cells. Although the area near the base station is the "hot spot" that causes most of the call blocking, and the interference at the cell boundary is not significant, further research is needed to quantify the impact of interference between different cells.

8 Appendices

8.1 Derivation of average MH-BS distance

Assuming that the cell shape is a hexagon, the coverage area of a cell is $\frac{3\sqrt{3}}{2}R_c^2$. Also assuming that the MH location in the area follows a uniform random distribution, the probability density function (pdf) of the location is a constant given by:

$$f = \frac{2}{3\sqrt{3}R_c^2} \tag{38}$$

Figure 20: Average MH-BS distance

In Figure 20, R_c is the cell radius. $R_a = \frac{\sqrt{3}}{2}R_c$, is the distance from the center to any border of the hexagon, r is the distance from center to any point on the cell boundary, and $\alpha(r)$ is the angle between the line r and R_c.

The area of the hexagon is also be given by:

$$\mathcal{A} = \pi R_a^2 + \int_{R_a}^{R_c} r\theta(r)dr, \tag{39}$$

where $\theta(r) = 12\alpha(r)$.

Since $\mathcal{A} = \frac{3\sqrt{3}}{2}R_c^2$, from Eq. (39), one obtains the following formula:

$$\int_{R_a}^{R_c} r\theta(r)dr = \frac{3\sqrt{3}}{2}R_c^2 - \pi R_a^2 \tag{40}$$

The expected distance from a mobile station to the base station is given by:

$$\begin{aligned} E(d) &= f \cdot \left(\int_0^{R_a} 2\pi r^2 dr + \int_{R_a}^{R_c} r^2\theta(r)dr \right) \\ E(d) &\geq f \cdot \left(\frac{2\pi}{3}R_a^3 + R_a \cdot \int_{R_a}^{R_c} r\theta(r)dr \right) \end{aligned} \tag{41}$$

Substituting Eq. (38) and Eq. (40) into Eq. (41), the expected distance of a mobile station to the base station can be simplified to:

$$E(d) \geq \frac{6\sqrt{3} - \pi}{12}R_c. \tag{42}$$

Figure 21: Derivation of Interference Distance

8.2 Derivation of Interference Distance

In Figure 21, A_1A_2, B_1B_2 are two sets of nodes comunicating using the same set of channels. Let E be the midpoint of line A_1B_1, we have $EO_1//B_1A_2$, $EO_2//A_1B_2$, and $|EO_1| = \frac{|A_2B_1|}{2}$, $|EO_2| = \frac{|A_1B_2|}{2}$. So, $\alpha = \pi - \theta$.

Since $\alpha = \alpha_1 + \alpha_2$, we have $\theta = \pi - (\alpha_1 + \alpha_2)$.

Since it is always true that $|A_1B_2| > |B_1B_2|$, α_1 is an acute angle. Let:

$$y = cos(\alpha_1) = \frac{|A_1B_2|^2 + |A_1B_1|^2 - |B_1B_2|^2}{2 \cdot |A_1B_2| \cdot |A_1B_1|}$$

Since y is minimized when $\frac{\partial y}{\partial(|A_1B_1|)} = 0$, one can determine that the value of y is minimized (α_1 is maximized) when $|A_1B_2|^2 = |A_1B_1|^2 + |A_1A_2|^2$. So:

$$\alpha_1 \leq sin^{-1}(\frac{|B_1B_2|}{|A_1B_2|}) \tag{43}$$

Similiarly,

$$\alpha_2 \leq sin^{-1}(\frac{|A_1A_2|}{|A_2B_1|}). \tag{44}$$

By defining $D \triangleq |O_1O_2|$, from Figure 21, we know that:

$$D^2 = (\frac{|A_1B_2|}{2})^2 + (\frac{|A_2B_1|}{2})^2 - 2\frac{|A_1B_2|}{2}\frac{|A_2B_1|}{2}cos(\theta).$$

$$-cos(\theta) = cos(\alpha_1 + \alpha_2) = cos(\alpha_1)cos(\alpha 2) - sin(\alpha_1)sin(\alpha 2).$$

Combining this with equations (43) and (44), we get:

$$cos(\alpha_1 + \alpha_2) \geq (\frac{\sqrt{|A_1B_2|^2 - |B_1B_2|^2}\sqrt{|A_2B_1|^2 - |A_1A_2|^2} - |A_1A_2| \cdot |B_1B_2|}{|A_1B_2| \cdot |A_2B_1|},$$

and

$$D^2 = \frac{1}{4}\cdot(|A_1B_2|^2 + |A_2B_1|^2 + 2\cdot(\sqrt{|A_1B_2|^2 - |B_1B_2|^2}\cdot\sqrt{|A_2B_1|^2 - |A_1A_2|^2} - |A_1A_2|\cdot|B_1B_2|)).$$

Since $|A_1A_2| \leq R_t$, $|B_1B_2| \leq R_t$, $|A_1B_2| > \beta R_t$, $|A_2B_1| > \beta\alpha R_t$, we have:

$$D \geq \sqrt{\beta^2 - 1}R_t, \tag{45}$$

which establishes Eq. (2) of the thesis report.

8.3 Derivation of average relative speed of any two MHs

From Figure 8, one can see that the relative speed of two MHs moving close is described by: $v = v_1 cos(\alpha_1) - v_2 cos(\alpha_2)$. Let $X = v_1 cos(\alpha_1)$, $Y = v_2 cos(\alpha_1)$. Assume v_1, v_2 are random variables uniformly distributed between $[0, v_m]$, and α_1, α_2 are random variables uniformly distributed between $[0, 2\pi]$, then X and Y are statistically identical random variables. We will try to find the distribution function of X and Y.

The cumulative distribution function (cdf) of X is:

$$\begin{aligned} C_X(x) &\triangleq P(X \leq x) \\ &= \int_t \left(f_X(t) \cdot P(cos(\theta) \leq \tfrac{x}{t}) \cdot dt \right). \end{aligned} \tag{46}$$

When $x \geq 0$,

$$\begin{aligned} C_X(x) &= \int_0^x \left(f_X(t) \cdot P(cos(\theta) \leq \tfrac{x}{t}) \cdot dt \right) + \int_x^{v_m} \left(f_X(t) \cdot P(cos(\theta) \leq \tfrac{x}{t}) \cdot dt \right) \\ &= \int_0^x f_X(t)dt + \int_x^{v_m} \left(f_X(t) \cdot (1 - \tfrac{arccos(\frac{x}{t})}{\pi})dt \right) \\ &= 1 - \frac{arccos(\frac{x}{v_m}) - \frac{x}{v_m} \cdot log\left(\frac{v_m}{x} + \sqrt{\left(\frac{v_m}{x}\right)^2 - 1} \right)}{\pi}. \end{aligned} \tag{47}$$

When $x < 0$,

$$\begin{aligned} C_X(x) &= \int_0^{-x} \left(f_X(t) \cdot P(cos(\theta) \leq \tfrac{x}{t}) \cdot dt \right) + \int_{-x}^{v_m} \left(f_X(t) \cdot P(cos(\theta) \leq \tfrac{x}{t}) \cdot dt \right) \\ &= \int_{-x}^{v_m} \left(f_X(t) \cdot (1 - \tfrac{arccos(\frac{x}{t})}{\pi})dt \right) \\ &= \int_{|x|}^{v_m} \left(f_X(t) \cdot \tfrac{arccos(\frac{|x|}{t})}{\pi})dt \right) \\ &= \frac{arccos(\frac{|x|}{v_m}) - \frac{|x|}{v_m} \cdot log\left(\frac{v_m}{|x|} + \sqrt{\left(\frac{v_m}{|x|}\right)^2 - 1} \right)}{\pi}. \end{aligned} \tag{48}$$

From Eq. 47 and Eq. 48, one can obtain the probability density functions of X and Y, and consequently the distribution of v. The closed form of the probability density function $f_v(v)$ is difficult to get, but it can be calculated numerically. Figure 22 shows one example of the pdf of $f(v)$ given $v_m = 10$.

Figure 22: Distribution density function of MHs' relative speed with MH's maximum moving speed being 10 m/s.

The average relative speed of two MHs and their relationship with the maximum moving speed(v_m) of MHs is shown in Figure 23.

Figure 23: The relationship between average relative MH speed and maximum MH moving speed

In addition, from Eq. 47 and 48, one can observe that $F_X(-x) = 1 - F_X(x)$; therefore, $f_X(-x) = f_X(x)$. We will prove that $P(Y - X > 0) = P(Y - X < 0)$. For the two statistically identical random variables, X and Y, there is:

$$\begin{aligned} P(Y > X) &= \textstyle\int_u \Big(f_X(u) \cdot P(Y > X | X = u) \cdot du \Big) \\ &= \textstyle\int_u \Big(f_X(u) du \cdot (\int_u^\infty f_Y(v) dv) \Big) \\ &= \textstyle\int_u f_X(u) du - \int_u \Big(f_X(u) du \cdot (\int_{-\infty}^u f_Y(v) dv) \Big) \\ &= 1 - \textstyle\int_u \Big(f_X(u) du \cdot (\int_{-\infty}^u f_Y(v) dv) \Big). \end{aligned} \tag{49}$$

Since $f_X(-u) = f_X(u), f_Y(-v) = f_Y(v)$, it is easy to show that

$$\int_u \Big(f_X(u) du \cdot (\int_u^\infty f_y(v) dv) \Big) = \int_u \Big(f_x(u) du \cdot (\int_{-\infty}^u f_y(v) dv) \Big). \tag{50}$$

From Eq. 49 and 50, one obtains:

$$\int_u \Big(f_X(u) du \cdot (\int_u^\infty f_Y(v) dv) \Big) = \frac{1}{2}.$$

Therefore,

$$P(Y > X) = \frac{1}{2},$$

and

$$P(Y - X > 0) = P(Y - X < 0).$$

References

[1] Y. D. Lin and Y. C. Hsu, "Multihop Cellular: A New Architecture for Wireless Communication", in *Proceedings of IEEE Conference on Computer Communications (INFOCOM'2000)*, TelAviv, Israel, March 2000, pp. 1273-1282.

[2] H. Luo, R. Ramjee, P. Sinha, L. E. Li and S. Lu, "UCAN: A unified cellular and ad-hoc network architecture", in *Proc. of the 9th Annual International Conference on Mobile Computing and Networking (MobiCom'2003)*, pp. 353-367, San Diego, CA, USA, 2003.

[3] Hongyi Wu, Chunming Qiao, Swades De and Ozan Tonguz, "An integrated Cellular and Ad hoc Relaying system: iCAR", in *IEEE Journal on Selected Areas in Communications (JSAC), Issue on Mobility and Resource Management in Next Generation Wireless Systems*, vol. 19, No. 10, pp. 2105-2115, Oct. 2001.

[4] A. N. Zadeh, B. Jabbari, R. Pickholtz, and B. Vojcic, "Self-Organizing Packet Radio Ad Hoc Networks with Overlay (SOPRANO)", in *IEEE Communications Magazine*, 40, no.6, pp. 149-157, June 2002.

[5] B. S. Manoj, Christo Frank D, C. Siva Ram Murthy, "Throughput Enhanced Wireless in Local Loop (TWiLL)-The architecture, rotocols, and pricing schemes", in *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 7, issue 1, pp. 95-116, Jan. 2003

[6] V. H. Mac Donald, "The cellular concept", in *The Bell System Technical Journal*, 58(1), pp. 15-41, Jan. 1979.

[7] Ali Nabi Zadeh, Bijan Jabbari, "A high capacity multihop packet CDMA wireless network", in *Wireless Networks*, vol. 9, issue 4, pp. 365-372, July, 2003.

[8] R. Ananthapadmanabha, B. S. Manoj, and C. Siva Ram Murthy, "Multi-hop Cellular Networks: Architecture and Routing Protocols", in *Proc. 12th IEEE International Symposium on Personal, Indoor, and Mobile Radio Communication*, pp. 78-82, San Diego, USA, Sept. 30-Oct. 3, 2001.

[9] B. S. Manoj, Ananthapadmanabha, and C. Siva Ram Murthy, “Multihop Cellular Networks: Architecture and Protocols for Best-effort and Real-time Traffic Support”, in *Journal of Parallel and Distributed Computing*, vol. 65, issue 6, pp. 767-791, June 2005.

[10] K. Jayanth Kumar, B. S. Manoj, and C. Siva Ram Murthy, “RT-MuPAC: Multi-Power Architecture for Voice Cellular Networks”, in *Proc. 9th International Conference on High Performance Computing*, pp. 377-387, Bangalore, December 18-21, 2002.

[11] G. N. Aggelou and R. Tafazolli, “On the Relaying Capacity of Next-Generation GSM Cellular Networks”, *IEEE Personal Communications*, pp. 40-47, February 2001.

[12] Y. Bejerano, “Efficient Integration of Multi-Hop Wireless and Wired Networks with QoS Constraints”, in *Proceedings of the 8th Annual International Conference on Mobile Computing and Networking (MobiCom’2002)*, pp. 215-226, Atlanta, Georgia, USA, Sept. 23-28, 2002.

[13] H. Holma(Editor) and A. Toskala(Editor), “WCDMA for UMTS: Radio Access for Third Generation Mobile Communications”, Wiley, 2002.

[14] H. -Y. Hsieh and R. Sivakumar, “Towards a Hybrid Network Model for Wireless Packet Data Networks”, in *IEEE Symposium on Computers and Communications (ISCC)*, pp. 264-271, Taormina, Italy, July 2002.

[15] M. Kubisch, S. Mengesha, D. Hollos, H. Karl, and A. Wolisz, “Applying Ad-hoc Relaying to Improve Capacity, Energy Efficiency, and Immission in Infrastructure-based WLANs”, in *Proceedings of Kommunikation in Verteilten Systemen (KiVS 2003)*, pp. 195-206, Leipzig, Germany, February 2003.

[16] O. C. Mantel, N. Scully, and A. Mawira, “Radio Aspects of Hybrid Wireless Ad Hoc Networks”, in *IEEE VTS 53rd, Vehicular Technology Conference, 2001 (VTC 2001 Spring).* vol. 2, pp. 1139-1143, Rhodes, May 06-09, 2001.

[17] Naouel Ben Salem, Levente Buttyan, Jean-Pierre Hubaux, Markus Jakobsson, "Co-operation in Multi-hop Cellular Networks", *The School of Computer and Communication Sciences, EPFL*, Technical Report Nr. IC/2003/55.

[18] N. Ben Salem, L. Buttyan, J. -P. Hubaux, and M. Jakobsson, "A Charging and Rewarding Scheme for Packet Forwarding in Multi-hop Cellular Networks", in *Proceedings of the 4th ACM Interational Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc 2003*, pp. 13-24, Annapolis, Maryland, USA, June 1-3, 2003.

[19] Attila Weyland and Torsten Braun, "CASHnet - Cooperation and Accounting Strategy for Hybrid Networks", in *2nd Workshop on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt 2004)*, Cambridge, UK, March 24-26, 2004.

[20] M. Jakobsson, J. P. Hubaux, and L. Buttyan, "A Micro-Payment Scheme Encouraging Collaboration in Multi-Hop Cellular Networks", in *Proceedings of Financial Cryptography 2003*, Guadeloupe, Jan. 26-30, 2003.

[21] S. De, O. Tonguz, H. Wu, C. Qiao, "Integrated cellular and ad hoc relay (iCAR) systems: pushing the performance limits of conventional wireless networks", in Proceedings of the 35th Annual Hawaii International Conference on System Sciences, pp. 3899-3906, Jan. 7-10, 2002.

[22] O. K. Tonguz and E. Yanmaz, "On the Performance of iCAR Systems with Limited Number of ISM-band Ad hoc Relay Channels", in *Proc. IEEE Global Telecom. Conf. (GLOBECOM'02)*, vol. 2, pp. 1835 -1839, Nov. 2002.

[23] E. Yanmaz, S. Mishra, O. K. Tonguz, H. Wu, and C. Qiao, "Efficient Dynamic Load Balancing Algorithms using iCAR Systems: A Generalized Framework", in *Proc. IEEE Vehic. Tech. Conf. (VTC Fall 2002)*, vol. 1, pp. 586-590, Sept. 2002.

[24] X. Wu, S. H. G. Chan, and B. Mukherjee, "MADF: A Novel Approach to Add an Ad Hoc Overlay on a Fixed Cellular Infrastructure", in *Wireless Communications and Networking Conference, 2000. WCNC. 2000 IEEE*, vol. 2, pp. 549-554, Sept. 2000.

[25] B. S. Manoj and C. Siva Ram Murthy, "A High-Performance Wireless Local Loop Architecture Utilizing Directional Multi-Hop Relaying", Technical Report, *Department of Computer Science and Engineering, Indian Institute of Technology, Madras, India*, June 2002.

[26] B. S. Manoj, Christo Frank D, and C. Siva Ram Murthy, "Performance evaluation of throughput enhancement architectures for next generation wireless systems", in *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 6, no. 4, pp. 77-90, Oct. 2000

[27] Hung-Yun Hsieh, Raghupathy Sivakumar, "Performance comparison of cellular and multi-hop wireless networks: a quantitative study", in *Proceedings of the 2001 ACM SIGMETRICS international conference on Measurement and modeling of computer systems table of contents, SIGMETRICS: ACM Special Interest*, Cambridge, Massachusetts, United States pp. 113-122, 2001.

[28] Z. Dziong, James W. Roberts, "Congestion probabilities in a circuit-switched integrated services network", *Performance Evaluation*, vol. 7, issue 4, pp. 267-284, Nov. 1987.

[29] A. E. Conway, E. Pinsky, S. Tridandapani, "Efficient decomposition methods for the analysis of multi-facility blocking models", *Journal Associ. Comput. Mach. (JACM)*, vol. 41, no. 4, pp. 648-675, 1994.

[30] Li, C.Y., Wai, P.K.A. Li, V.O.K., "The decomposition of a blocking model for connection-oriented networks", in *IEEE/ACM Transactions on Networking (TON)*, vol 12, issue 3, pp. 549-558, June 2004.