"Trade-off Analysis for Distributed Dynamic Channel Allocation in Cellular Networks"

Dissertation
Submitted in the Partial Fulfillment of the Requirement for
the Award of Degree of

MASTER OF TECHNOLOGY
IN
COMPUTER SCIENCE & TECHNOLOGY

BY

HEMANT GAUR

UNDER THE SUPERVISION OF

PROF. P.C. SEXENA

SCHOOL OF COMPUTER & SYSTEMS SCIENCES
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI-110067

SCHOOL OF COMPUTER & SYSTEM SCIENCES
जवाहरलाल नेहरू विश्वविद्यालय
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI -110067 (INDIA)

CERTIFICATE

This is to certify that the dissertation titled *"Trade-off Analysis for Distributed Dynamic Channel Allocation in Cellular Networks"*, which is being submitted by **Mr. Hemant Gaur** to **School of Computer & Systems Sciences, Jawaharlal Nehru University, New Delhi** in partial fulfillment of the requirement for the award of the degree of **Master of Technology in Computer Science & Technology**, is a record of bonafied work carried out by him under the supervision of **Prof. P.C. Saxena** during the Monsoon Semester, 2003.

The matter results reported in this thesis have not been submitted in part or full to any other University or Institution for the award of any degree etc.

Prof. Karmeshu
Dean, SC&SS.
Jawaharlal Nehru University
New Delhi 110067

Prof. P.C. Saxena 5/1/04
Professor, SC&SS
Jawaharlal Nehru University
New Delhi 110067

Prof. P. C. Saxena
School of Computer & Systems Sciences
Jawaharlal Nehru University,
New Delhi-110067

ACKNOWLEDGEMENT

This dissertation would not have been possible without the guidance and the help of a large number of people. These people stood by my side and gave me constant encouragement and made me believe in myself.

My first thanks goes to my supervisor **Prof. P.C. Saxena** who not only served as my guide but also encouraged and challenged me throughout my academic program. He and the other faculty members especially **Prof. C.P. Katti** patiently guided me through the dissertation process, never accepting less than my best efforts.

I am grateful to them for providing me enough infrastructures through Data Communication and Distributed Communication Group (DCDCG) laboratory to carry out my work. It is great honor for me to be a member of this group.

I would also like to extend a special thanks to the **Prof. Karmeshu**, Dean SCSS and other faculty members of the SC & SS, JNU, especially **Prof.Parimala N.**, **Prof. GV Singh**, **Prof. KK Bhardwaj**, **Prof. S. Balasundaram**, **Dr. DK Lobiyal**, **Dr. Sona Minz**, and **Dr. RC Phoha** for their able guidance, which made me equipped to carry on this work.

I would like thanking my classmates and friends for their help on the discussion related to my work and providing constructive criticism,

resources and useful suggestions. My special mention in this regard goes to **Jitender**, **Mohit**, **Manish**, **Kulwant**, **Sudesh**, **Rakesh**, **Mintoo**, **Vishal and Biplab**.

I would also like to make a mention of the non-teaching staff of the school who has worked hard to provide us with a very smooth and comfortable environment for the research work.

Finally I will never find words enough to express the gratitude that I owe to my parents and loving sister, **Tamanna**. Throughout the period they were extremely co-operative and patient with my odd working hours and mood swings. Their tender love and affection has always been the cementing force for building the blocks of my academic career. The all round support rendered by them provided the much needed stimulant to sail through the phases of stress and strain.

Needless to add any errors and omissions found in the dissertation are entirely my responsibility and I will be extremely grateful if they were pointed out to me.

To all of you, thank you.

2ND JAN, 2004 HEMANT GAUR
New Delhi

Dedicated to my parents…

ABSTRACT

In recent years there has been tremendous increase in the demand of various forms of mobile communications. The integration of mobile communication and computing in fixed network introduces new issues and problems in distributed computing. With a limited frequency spectrum and increasing demand for mobile communication services, the problem of channel assignment becomes increasingly important. As the use of mobile communications systems grows, the need arises for new and more efficient channel allocation techniques. Distributed Dynamic channel assignment (DDCA) has been increasingly used to improve capacity in cellular telephone market because of the ability to reuse channel more efficiently by taking into account local propagation conditions and traffic loads. These channel allocation schemes have number of optimization criteria and a number of interesting tradeoffs in trying to meet them. Analyzing them is particularly important for applying frequency allocation algorithms in practice. However this can be very complex in distributed setting which implies a larger number of parameters to consider. This, besides analytical evaluation, necessitates the use of experimentation in tuning the algorithms' performance. We study and analyze these optimization criteria and associated tradeoffs. Also we analyze the tradeoff between increasing satisfiable requests and decreasing overhead per connection using frequency reservation for satisfying requests locally and propose dynamic tradeoff tuning strategies for optimizing the performance.

Contents

Chapter 1 Introduction

Chapter 2 Graph Coloring and Problem Modeling

Chapter 3 Distributed Dynamic Channel Assignment

Contents

Chapter 4 Distributed Long-Lived List Coloring Algorithms

Chapter 5 Performance Tradeoffs and Analysis

5.5 Results

CHAPTER 1
INTRODUCTION

1.1 Introduction

Technological advances and rapid development of handheld wireless terminals have facilitated the rapid growth of wireless communications and mobile computing. Taking ergonomic and economical factors into account, and considering the new trend in telecommunication industry to provide ubiquitous information access, the population of mobile users will continue to grow at a tremendous rate. Another important developing phenomenon is the shift of many applications to multimedia platforms in order to present information more effectively.

The tremendous growth of the wireless/mobile user population, coupled with the bandwidth requirements of multimedia applications, requires efficient reuse of the scarce radio spectrum allocated to wireless/mobile communications. The rapid development of the cellular telephone network has increased the need for good solution techniques for the *frequency assignment* problem for *cellular networks*. The main difference between radio television broadcasting and cellular networks is the need of individual connection for every customer. In radio or television networks thousands or even million customers receive the same signal transmitted by one single antenna. In cellular network, a signal is transmitted between one transmitter and one receiver. The service area of cellular network has to be covered with a substantially large number of antennae to satisfy the demand.

Efficient use of radio spectrum is also important from a cost of service point of view, where the number of base station required to service a given geographical area is an important factor. Reduction in the number of base stations, required to service a given geographical area is an important factor. A reduction in the number of base stations, and hence in the cost of service can be achieved by more efficient reuse of the radio spectrum. The basic prohibiting factor in radio spectrum reuse is interference caused by the environment or other mobile host. Interference can be reduced by developing efficient radio subsystems and by making use of channel assignment techniques.

In the radio and transmission subsystem, techniques such as deployment of *time and space diversity*, use of *low noise filters* and *efficient equalizers*, and deployment of efficient *modulation schemes* can be used to suppress interference and to extract the desired signal. However *co-channel interference* caused by *frequency reuse* is the most restraining factor on overall system capacity in the wireless network. The main idea behind channel assignment algorithm is to make use of radio propagation *path loss* characteristics in order to minimize the *carrier–to-interference* ratio and hence increase the radio spectrum reuse efficiency.

1.2 Cellular Communications

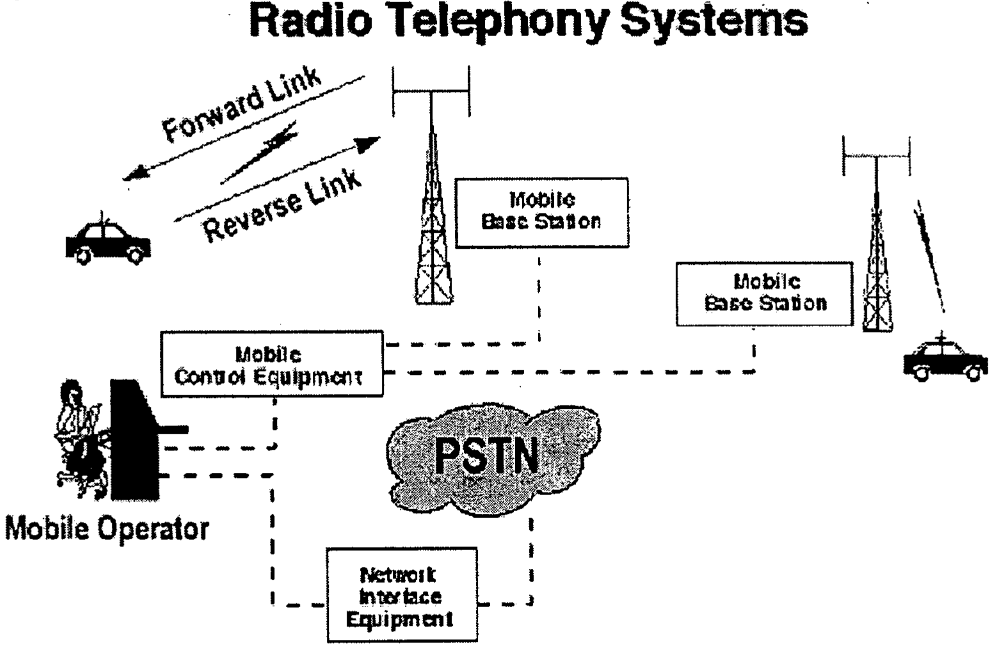
A cellular mobile communication system uses a large number of low power wireless transmitters to create cells, the basic geographical service area of a wireless communication system. Variable power levels allow cells to be sized according to the subscriber density and demand within particular region. As mobile user travel from cell to cell, their conversation are *"handed off"* between cells in order to maintain seamless service. Channels (frequencies) used in one cell can be reused in another cell some distance away. Cells can be added to accommodate growth, creating new cells in unserved area or overlaying cells in existing areas.

1.2.1 Mobile Communication Principles

Each mobile uses a separate, temporary radio channel to talk to the cell site. The cell site talks to many mobiles at once, using one channel per mobile.

Channels use a pair of frequencies for communications, one frequency, *the forward link*, for transmitting from the cell site, and one frequency, *the reverse link*, for the cell site to receive calls from the users. Radio energy dissipates over a distance, so mobiles must stay in the range of base station to maintain communications. The basic structure of mobile networks includes *telephone system* and *radio services*. Mobile radio service operates in a closed network and has no access to telephone system and mobile telephone service allows interconnection to the telephone network.

Figure 1: Basic Mobile Telephone Service Network



1.2.2 Mobile Telephone System Using Cellular Concept

Interference problems caused by mobile units using the same channel in adjacent areas provide that all channels could not be reused in every cell. Areas had to be skipped before the same channel could be reused. The interference effects were not due to the distance between the areas, but to the ratio of the distance between areas to the transmitter power (radius) of the areas. By reducing the radius of an area by fifty percent, service providers could increase the number of potential customers in the area fourfold. So large number of calls can be served by reducing the area to few hundred meters.

The cellular concept employs variable low-power levels, which allow cells to be sized according to the subscriber density and demand of given area. As the population grows, cells can be added to accommodate that growth. Frequencies used in one cell cluster can be reused in other cells. Conversations can be handed from cell to cell to maintain constant phone service as the user moves between cells.

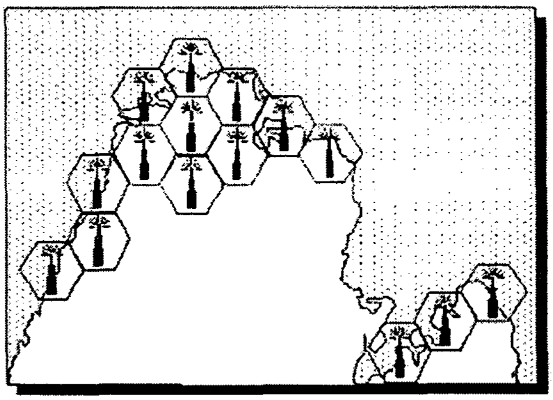


Fig 2 : Mobile Cellular System using Cellular Architectute

The cellular radio equipment can communicate with mobiles as long they are within range. Radio energy dissipates over distance, so the mobiles must be within operating range of the base station.

1.3 Cellular System Architecture

To improve the quality of service and to support more users in the system efficient use frequencies is needed for mobile cellular coverage. In modern cellular telephony, rural and urban regions are divided into areas according to specific provisioning guidelines. Development parameters, such as amount of *cell splitting* are considered by engineers experienced in cellular system architecture to determine cell sizes. Provisioning for each region is planned according to engineering plan that includes cell, clusters, frequency reuse and handovers.

1.3.1 Cells

A cell is basic geographical unit of cellular system. The term *cellular* comes from honeycomb shape of areas into which region is divided. Cells are base stations transmitting over small geographical areas that are represented as hexagons in general.

1.3.2 Clusters

A *cluster* is group of cells. No channel are reused within a cluster

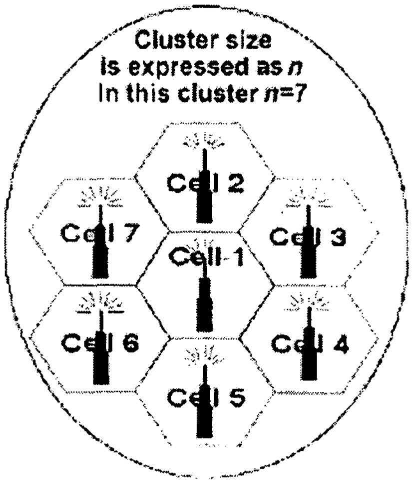


Fig 3 - A seven Cell Cluster

1.3.3 Frequency Reuse

Because only small number of radio channel frequencies were available for mobile systems, engineers had to find a way to reuse radio channels in order to carry more than one conversation at a time. The restructuring of the mobile telephone system architecture into the cellular concept provided a solution called *frequency reuse.*

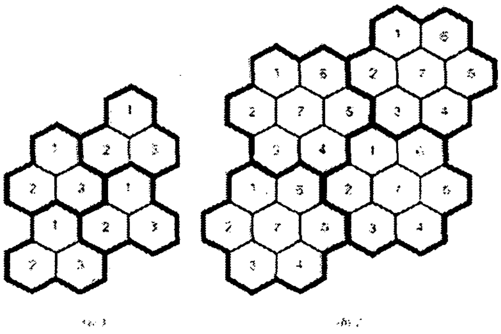


Fig 4:Cellular Frequency Reuse Plan for Two Cluster Numbers

The concept of frequency reuse is based on assigning to each cell a group of channels that is completely different from neighboring cells. The coverage area of the cells is called the *footprint.*

This footprint is limited by a boundary so that the same group of channels can be use in different cells that are far enough away from each other so that their frequencies do not interfere. As shown in fig 5, cells with the same number have same set of frequencies. Here because the number of available frequencies is 7, the *frequency reuse factor* is 1/7. That is each cell is using 1/7 of available cellular channels

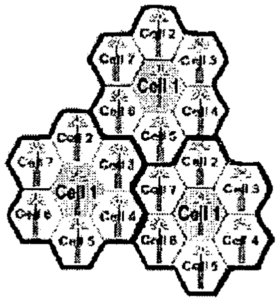


Fig 5: Frequency Reuse

1.3.4 Cell Splitting

Economic considerations made the concept of creating full systems with many small areas impractical. To overcome this difficulty, system operators developed the idea of *cell splitting*. As the service area becomes full of users, this approach is used to split a single area into smaller ones. In this way, urban centers can split into as many areas as necessary in order to provide acceptable service levels in heavy traffic regions, while larger, less expensive cells can be used to cover remote rural regions.

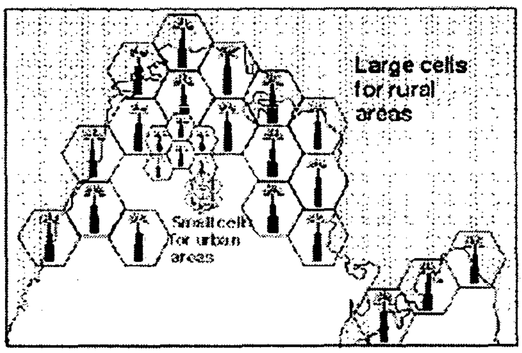


Fig 6: Cell Splitting

1.3.5 Handoff

The final obstacle in the development of the cellular network involved the problem created when the mobile subscriber traveled from one cell to another during a call. As adjacent areas do not use the same radio channels, a call must either be dropped or transferred from one radio channel to another when the user crosses the line between adjacent cells. Because dropping of calls is unacceptable, the process of *handoff* was created. *Handoff* occurs when the mobile telephone network automatically transfers a call from radio channel to radio channel as mobile crosses adjacent cells.

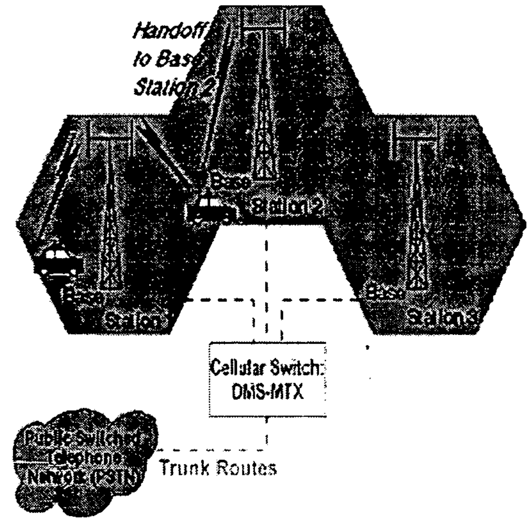


Fig:7 Handoff Between Adjacent Cells

During a call, two parties are on one voice channel. When a mobile unit moves out of coverage area of a given cell site, the reception becomes weak. At this point, the cell site in use requests *handoff.* The system switches the call to a stronger frequency channel in a new site without interrupting the call or alerting the user. The call continues as long as the user is talking and the user does not notice the hand off except for interruption for few milliseconds.

1.4 Cellular System Components

The cellular system offers mobile same service as provided by fixed stations to portable telephone stations over conventional wireless loops. It has capacity to serve tens of thousands of subscribers in a major metropolitan area. The cellular communication system consists of the following four major components that work together to provide mobile service to subscribers

1.) **Public Switched Telephone Network** (PSTN)

2.) **Mobile Telephone Switching Office** (MTSO)

3.) **Cell Site With Antenna System**

4.) **Mobile Subscriber Unit (MSU)**

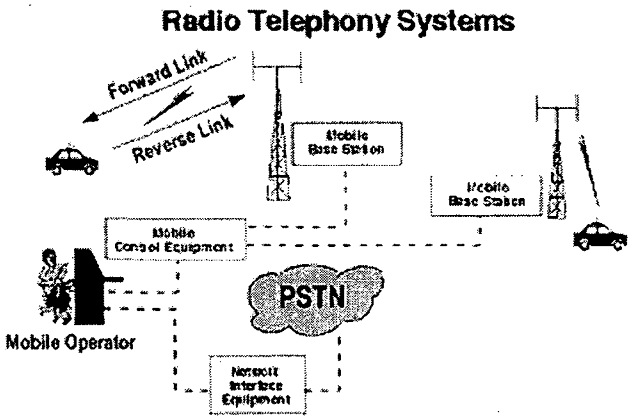


Fig 8: Cellular System Components

1.5 Channel Allocation Schemes

A given radio spectrum (or bandwidth) can be divided into a set of disjoint of non-interfering radio channels. All such channels can be used simultaneously while maintaining an acceptable radio signal. In order to divide a given radio spectrum into such channels many techniques such as *frequency division* (FD), *time division* (TD), or *code division* can be used. In FD, the spectrum is divided into disjoint frequency bands, whereas in TD the channel separation is achieved by dividing the usage of the channel into disjoint time periods called time slots. In CD, the channel separation is achieved by using different modulation codes. Furthermore, more elaborate techniques can be designed to divide a radio spectrum into a set of disjoint channels based on combining the above techniques.

Let $S_i(k)$ be denoted as the set *(i)* of wireless terminals that communicate with each other using the same channel k. By taking advantage of physical characteristics of the radio environment, the same channel k can be reused simultaneously by other set j if the members of the set i and j are sufficiently apart. All such sets, which use the same channel, are referred as *co-channel sets* or simply *co-channels*. The minimum distance at which co-channels can be reused with acceptable interference is called *the "co-channel reuse distance "*. This is possible because due to propagation path loss in the radio environment, the average power received from the transmitter at a distance d is proportional to $P_T\, d^{-\alpha}$ where α is the number in the range of 3-5 depending on the physical environment, and P_T is the average transmitter power.

As a consequence, an operator should carefully choose the frequencies on which each station transmits to avoid high interference levels. The selection of the frequencies is made in a way that interference is avoided, or second best, is minimized, is called *Frequency Assignment Problem* (FAP).

Depending on the application the conditions that should be satisfied by the frequency plan may vary.

1.6 Interference Constraints

There are three types of interference constraints i.e. Co-Channel, Adjacent, and Co-Site constraints that must be avoided in channel assignment within the same cell and from the different neighboring cell.

1.) **Co-Channel Constraint** (CCC), where same channel cannot be assigned to certain pairs of radio cells simultaneously.

2.) **Adjacent Channel Constraint** (ACC), where adjacent channels in the frequency domain cannot be assigned to adjacent cells simultaneously.

3.) **Co-Site Constraint** (CSC), where channels in the same cells must be separated by certain distance in the frequency domain.

Depending upon the maximum number of the possible simultaneous calls in a cell, the demand of channel for individual cells are set. The problem of channel allocation is to assign carriers to the cells so that the traffic demand for the cell is met.

1.7 Channel Allocation Strategies

Channel allocation schemes can be divided into number of different categories depending on the comparison basis. When channel assignment algorithms are compared based on the manner in which co-channels are separated, they can divided as follows

1.) **Fixed Channel Allocation** (FCA), in which the area is partitioned into a number of cells and number of channels are assigned to each cell according to some reuse pattern, depending upon the desired signal quality. Since allocation depends upon static demand of different cells, these algorithms though being simple, do not adapt to changing traffic conditions and user distribution.

2.) **Dynamic Channel Allocation** (DCA), in which all channels are placed in a pool and are assigned to new calls as needed. Demands at different cells change dynamically due to change in traffic patterns with time. At the cost of higher complexity, DCA schemes provide flexibility and traffic adaptability.

3.) **Hybrid Channel Allocation** (HCA), in which each cell is allocated a fixed set of permanent carriers and a number of channels are set aside to be dynamically allocated depending upon changing requests from the cells. This technique was deigned by combining FCA and DCA schemes because DCA strategies are less efficient than FCA under high load conditions.

Channel assignment schemes can be implemented in centralized and distributed fashion. In the centralized schemes a central controller assigns the channel, whereas in distributed schemes a channel is selected either by a local base station of the cell from which the call is initiated or selected autonomously by the mobile. In a system with cell based control, each base station keeps information about the current available channels in the vicinity, Here the channel availability information is updated by exchange of status information between base stations. Finally, in autonomously organized distributed schemes the mobile chooses a channel based on local CIR measurements without the involvement of a central call assignment entity.

Chapter 2

GRAPH COLORING AND PROBLEM MODELING

2.1 Graph Coloring

Graph coloring is a problem that came about from people wanting to color adjacent countries on a map. From the coloring problem many results in graph theory and related studies, such as *scheduling* and *resource allocation* have been obtained.

2.1.1 Vertex Coloring

Given a graph G=(V, E) and a coloring function **c: V → C**, we are to find an assignment of colors from C such that for all adjacent vertex (v, u) in V, we have **c (v) ! = c (u)**. The most interesting property of the set C of colors is its size. We are typically concerned with the smallest number of colors needed such that no adjacent vertex has the same color. Such a number is known as the ***chromatic number***, denoted by χ (G).

2.1.2 Edge Coloring

Given a graph G=(V, E) and a coloring function **c: V →C**, an assignment of colors such that for all adjacent edges, (e, j) in E, we have **c (e) ! = c (j)**. Similarly to vertex coloring problem, we are typically concerned with the smallest such number for the size of the set C where no adjacent edge has the same color. Such a number is known as the ***chromatic index***, denoted by **~χ(G)**. The result of coloring a graph is that we find k disjoint sets of vertices/edges associated with each color from {1,2...k}.

2.1.3 Generalized List Coloring

Each node v_i has an associated set of colors L_i, which models the set of free frequencies from which the frequencies to be allocated to v_i must be chosen. Finally, each node v_i has an associated number r_i of color (channel) requests; r_i may change in time. The requirement is to *properly* color each node v_i of G, i.e. that each v_i gets r_i colors from its list L_i in a way that none of the colors used for v_i is used for any of its neighbors in G. Thus it requires that each v_i is properly colored with *one* color from its list L_i.

2.1.4 List Chromatic Number $\chi_L(G)$

List Chromatic Number of G is the smallest number k for which, for *any* assignment of a list Li of size at least k to every vertex $vi \in V(G)$, it is possible to color G so that every vertex gets one color from its list. If $\chi_L(G) \leq k$ then G is said to be ***k*-choosable**. The list chromatic number $\chi_L(G)$ of a graph G is not necessarily the same as its chromatic number $\chi(G)$. Also $\chi(G) \leq \chi_L(G)$ as shown by the following figure.

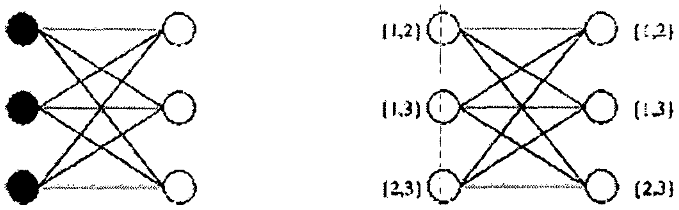


Fig 9: $K_{3,3}$ is 2 colorable but not 2-choosable.

2.2 Previous Work

The problem of allocating frequencies to mobile hosts by the base stations is viewed and analyzed as a generalized version of the *list-coloring* problem. The list-coloring problem was introduced in [14]. Sequential protocols for solving it can be found in [12]. For comparison, we briefly discuss the known sequential solutions to the problem, which can be used to provide centralized approaches for solving channel allocation.

Any protocol for the generalized list coloring problem that is *long-lived* (i.e. it can be invoked multiple times, with lists and requests that vary in time) can be used for solving the channel allocation problem in a dynamic manner. Moreover, the optimization criterion for minimizing the list sizes in the generalized list coloring is in order with the optimization criterion for maximizing *request satisfiability* in channel allocation.

2.3 Problem Modeling

An algorithm, which minimizes the list sizes, which are necessary to properly color, each node v_i with r_i colors from its list, also maximizes the *request satisfiability* in the system (or, similarly, minimizes the *blocking factor*). Hence, this provides a good way to measure the efficiency of a solution with respect to request satisfiability as the size of the list, which is sufficient to ensure that the node gets properly colored.

2.3.1 Signal Interference Graph

The cellular network is described with an interference graph $G = (V, E)$. Each node $v_i \in V(G)$ ($1 \leq i \leq n$), also called *process* or *station*, models a base station of the network. There is an edge between two nodes if there is possible signal interference when the same frequency is used concurrently in the cells of the respective stations. $N(v_i)$ denotes the set of neighbors of node v_i.

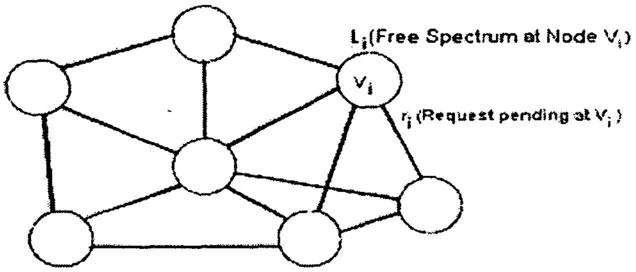


Fig 10: Signal Interference Graph

2.4 Sequential List Coloring

We present in brief what is known in the sequential case for list coloring. This is used in centralized channel allocation.

- ***General graphs,*** *lists of size Δ +1.*

  Given a graph of maximum degree Δ, one straightforward way of coloring the graph with Δ +1 colors is to consider vertices one at a time and assign them a color different from their neighbors. Clearly, we need at most Δ + 1 colors to ensure that for each vertex we can find one color different from its neighbors. Therefore, if each list is of size at least Δ + 1 we can always list-color the graph.

- ***Directed acyclic graphs,*** *lists of size equal to (number-of outgoing-neighbors + 1)*

  If we have an initial acyclic orientation of the edges of the graph then we could first color all vertices that are sinks by picking an arbitrary color from their lists. We then remove these sinks and color the new sinks created taking care to assign them a color different from their neighbors. It is now easy to see that a list coloring is always possible if each vertex has a list of size more than its out-degree in the initial orientation1. This is actually the best-known bound for list coloring [12].

- ***Planar graphs,*** *lists of size* 5.

  An elegant argument due to Thomassen [3] shows that every planar graph is 5-choosable. In particular, a planar graph is list colorable if every vertex in the graph not on the infinite face has a list of size 5 while vertices on the infinite face need only have lists of size 3. This is optimal since there are planar graphs known which are not 4-choosable.

These sequential solutions can naturally be used to provide centralized solutions to the dynamic channel allocation problem as follows. The channel manager maintains the set of frequencies that are in use at each base station. Every time the channel manager receives requests for frequencies from base stations it runs one of the sequential algorithms to allocate the frequencies.

2.5 Implied Bounds for Generalized List Coloring

Lemma 1. Lists of size (Δ +1) r are sufficient to properly list color each node of a graph whose maximum degree is Δ.

Lemma 2. In a directed acyclic graph G = (V, E) in order to properly list-color each node V_i with r_i colors, it suffices that

$$\left|L_i\right| = \sum_{(v_i, v_j) \in E} r_j + r_i$$

Where $(v_i, v_j) \in E$ implies that there is an edge connecting the two vertices and the direction towards v_j.

CHAPTER 3

DISTRIBUTED DYNAMIC CHANNEL ASSIGNMENT

3.1 Introduction

Recently, distributed channel acquisition algorithms have received considerable attention due to their high reliability and scalability. Due to short-term temporal and spatial variations of traffic in cellular systems, FCA schemes are not able to attain high channel efficiency. To over come this, DCA schemes were designed. In contrast to FCA, there is no fixed relationship between channels and cells in DCA. All channels are kept in a central pool and are assigned dynamically to radio cells as new calls arrive in the system. After the call is completed the channel is returned to the central pool. In DCA a channel is eligible for use in any cell provided that signal interference constraints are satisfied.

Distributed Dynamic Channel Assignment (DDCA) is being increasingly used to improve capacity in both the cellular and cordless telephone markets, because of the ability to reuse channels more efficiently by taking into account local propagation conditions and traffic loads. Common to many DDCA schemes is the incorporation of a coexistence etiquette used during channel selection, which aims to minimize interference caused to other users, thereby allowing different users/networks/companies to share a common bandwidth.

3.2 Distributed DCA Schemes

Micro-cellular systems have shown great potential for capacity improvements in high-density personal communications networks. However propagation characteristics will be less predictable and network control requirements more intense then the present systems. Several analysis and simulation results have shown that centralized DCA schemes can produce near optimal channel allocation, but at the expense of high centralization overhead. Distributed schemes are therefore more attractive, due to the simplicity of the assignment algorithm at each base station. The proposed distributed DCA schemes can be categorized as

3.2.1 Cell Based

A channel is allocated to a call by the base station at which the call is initiated. These use local information about the current available channels in the cell's vicinity. The channel pattern information is updated by exchanging status information between base stations. The cell-based scheme provides near-optimal channel allocation at the expense of messages exchanged between the base stations, especially under heavy traffic load.

3.2.2 Signal Strength Measurements

In these schemes a base station use only local information without the need to communicate with any other base station. Thus the system is self-organizing and channels can be placed or added anywhere, as needed to increase capacity or to improve radio coverage in a distributed fashion. These schemes allow fast real time processing and maximal channel packing at the expense of increased co-channel interference probability with respect to ongoing calls in adjacent cells, which may lead to undesirable effects such as interruption, deadlock, and instability. This scheme is not desirable, because of energy constraints.

Dynamic channel assignment has long been proposed as a method to increase capacity in a mobile cellular system. This is achieved through more efficient utilization of the available spectrum, while the minimum C/I (**CIR** or **carrier to interference ratio**) for each call in the system is preserved. Dynamic channel assignment schemes in general have higher complexity and require greater regional state information. They have larger system capacity due to the fact that each cell in DCA may be allocated with any channel according to the traffic loading and electromagnetic compatibility constraints.

Distributed dynamic channel allocation (DDCA) is a fundamental resource management problem in mobile cellular networks. It has a flavor of distributed mutual exclusion but is not exactly a mutual exclusion problem because a channel may be reused in different cells. Also the same resource here can be used by more than one cells which are sufficiently far apart.

3.3 Mobile Computing Requirements

3.3.1 Energy Conservation

Most of the communication between mobile computers is in the form of several short bursts of data transfer. The mobile hosts have a limited energy source, in the form of a battery pack. Wireless communication drains the energy of the mobile hosts. Hence, energy should be conserved at a mobile host by keeping its involvement in the channel allocation process to a minimum. This can be achieved by minimizing the number of messages it has to exchange with the mobile service station during channel selection.

3.3.2 Minimize Hand-offs

Voice communication can tolerate hand-offs as short breaks in communication go undetected by the human ear. However, such breaks can lead to complications in data transfer to/from a mobile host. So, a channel allocation algorithm should not induce any hand-offs, over and above those caused by the movement of the mobile hosts between cells.

3.3.3 Exploit Locality of Reference

Most computer applications exhibit high temporal and spatial locality of data reference. If the data items in great demand reside in a mobile host, they should be moved to a mobile service station from where they can be accessed over the fixed wire network. Until such a transfer takes place, or if such a transfer is not possible, the data references translate into frequent arrivals of requests at the mobile service station to establish communication sessions with the mobile host. A channel allocation strategy should be able to adapt to such traffic.

3.4 DDCA Advantages

- **Better Fault Tolerance**

  The algorithm does not need a central network switch and hence there is no single point of failure. The locality of the failure is limited only to the smaller part of network, called *failure locality*.

- **Scalability**

  The MSS only needs to exchange information with its neighbors within the co-channel interference range. Thus the load is shared by the base stations and not by the central switch. The symmetry of the channel allocation procedure across the entire network makes the system scalable.

- **Lesser Traffic**

  They can also exploit the temporal locality of load distribution to make quick decisions about channel allocation rather then all network traffic directed towards the central switch.

- **Faster Response Time**

  The distributed system has faster response time due to the local processing and lesser network traffic.

- **Efficient Use of Channels**

  Unlike the FCA algorithms, these algorithms can adapt to changing load distribution in the network.

- **Preserves Locality Inherit to Problem**

  The mobile service station of a cell makes all the decisions about channel allocation in that cell based on the information available locally.

- **No single point of failure**

  It is more robust than existing DCA algorithms as it does not depend on a central network switch whose failure can bring down the entire network.

- **Lower Cost**

Moreover, a fast and expensive mainframe acting as the MTSO can be replaced by a set of microprocessor based switches at the MSSs. These switches can collectively outperform the Mainframe and cost much less.

3.5 Algorithm Evaluation Criteria

- **Maximize *satisfiable requests*(or bandwidth utilization)**

  The optimization criterion for minimizing the list sizes in the generalized list coloring is in order with the optimization criterion for maximizing *request satisfiability* in channel allocation. This implies that it should adapt fast to temporal variations in channel demand in different cells.

- **Minimize *connection setup time***

  The time when a communication request is issued to the time the session is established. The connection set-up time is measured as the number of rounds of message-exchanges with the neighboring base stations and can be as high as the size of the spectrum. The size of messages too can be as high as the size of the frequency spectrum.

- **Minimize *communication complexity***

  It is the amount of information (number and size of messages) that needs to be exchanged per request. The size of messages can be as high as the size of the frequency spectrum.

- **Avoid *intra-cell handoffs***

Hand-off is a situation when some frequency used by a mobile host for a communication session has to change during the session, in which case the communication is interrupted until a new frequency is found.

- **Minimize *failure locality***

A fault-tolerant solution should guarantee that what happens in a specific neighborhood is not reflected in the whole network. Failure locality measures the size of the network affected by a faulty station in terms of the distance from that station. Arbitrary long process waiting chains, which implies a failure locality equal to the diameter of the network.

- **Minimize bound on *response time***

The time it takes for a request by a base station to be resolved is measured in units of maximum message delay in the network. It's the implied overhead for the connection setup time. Both the mean response time for queued requests and the mean for all requests are measured.

- **No *starvation* until process fails**

The protocol for allocating and releasing frequencies might lead to a situation in which a busy cell which has collected a large set of frequencies prevents its neighbors from satisfying their own requests, thus starving these base stations.

- **Minimize/avoid the involvement of the MH**

Since mobile hosts generally operate on a limited battery power supply, power-consuming operations such as network communications, CPU operations and disk accesses should be kept to a minimum. A mobile host's involvement in the session request process should therefore be minimal.

- **Avoid co-channel Interference**

  The same frequency cannot be used by cells within a certain broadcast range, as the broadcast ranges of two neighboring base stations may overlap, causing *co-channel interference*.

Clearly, there are also tradeoffs, i.e. not all these *optimization criteria* can be met concurrently. To study the effect of these tradeoffs on the performance of various DDCA algorithms is the main focus of this dissertation .

3.6 Mobility and Distributed Computations

New problems and issues, which are introduced due to the mobility of hosts, include: *bandwidth management* and *frequency allocation*, *naming* and *addressing* mobile hosts, *logical network structure maintenance* and *routing*, *locating the mobile hosts, file system organization, data management, consistency guarantees* of *multicasting* under host migration, *coping* with the less secure and less reliable environment of wireless and broadcast communication, *coping with resource (energy) constraints* of mobile hosts and, definition of appropriate *performance measures* for evaluating solutions.

Although some of the problems should be addressed at the application layer most of them should be solved at the network layer to provide transparency to the users of the network. A mobile host can connect to the network from different locations at different times; distributed algorithms therefore, cannot rely on the assumption that a participant maintains a fixed and universally known location in the network at all times.

3.6.1 System Model

It should be noted that due to mobility and to the heterogeneous nature of the system entities, mobile and fixed hosts must be modeled differently, in order to capture new concepts that appear in these networks [2]. Apart from the dynamic nature of the network architecture, namely that logical links among mobile hosts cannot be mapped to fixed sequences of physical links, the new concepts include

1. The *bandwidth* of a wireless connection is significantly less than that between fixed hosts and, therefore, that communication cost over the wireless and the wired portion of the network should be considered differently

2. *Data transmission and reception*, *CPU operations* and *disk accesses* consume power at mobile hosts, which, henceforth, may frequently operate in a *"dose"* mode or even disconnect from the network. It should also be noted that this system model views portable computing as a restrictive form of mobile computing, since a portable computer may connect to the network from different locations, but it essentially uses some wired connection, and therefore, cannot simultaneously move and maintain its connection to the network.

3.6.2 Impact of Mobility

A mobile host can connect to the network from different locations at different times. Distributed algorithms therefore cannot rely on the assumption that a participant maintains a fixed and universally known location in a network at all times.

1. ***Search:***

   With the introduction of mobile hosts, the cost of sending a message is not necessarily fixed for a given *<source, destination>* pair and may need to explicitly include a search component.

2. ***Wireless connection:***

   The wireless medium allows a MSS to communicate with all the MHs located in its cell with a single message transmission. The cost of such a message is independent of the number of recipients. However *multicasting* can cause problem as composition of the wireless network within a cell change dynamically as MHs enters and leave the cell.

3. ***Logical structures:***

   The main purpose of such a structure is to provide a certain degree of order and predictability to the communication amongst the participants. Messages exchanged within such structures follow only selected logical paths.

4. ***Disconnection:***

   A mobile host that disconnects in the midst of an algorithm execution may cause the execution to suspend till its reconnection. Disconnection in a mobile environment is distinct from failure. By anticipating possible disconnection, a mobile host can execute a disconnection protocol before it physically detaches from the network.

5. ***Doze mode:***

In this mode, the clock speed is reduced and no user computations are performed; instead, a mobile host simply waits passively to receive any message sent to it from the rest of the network; if such a message is received, it resumes its regular mode of operation. Like disconnection, this is a voluntary operation but here a mobile host is reachable from the rest of the system It aims at reducing power consumption.

CHAPTER 4

DISTRIBUTED LONG-LIVED LIST COLORING

4.1 Introduction

A brief description of the algorithms studied is given here; The algorithms are based on the modeling of the cellular network as an *interference graph*, where vertices denote cells (base stations), and an edge between two cells denotes the potential for co-channel interference.

4.2 Deterministic Distributed List Coloring

The DET_DLC solution, proposed by Garg et al. [13], models the problem as a *generalized list-coloring problem*. Each node (base station) of a graph has a list of colors (free frequencies) and is required to color itself with a number of colors (the number of pending requests) without any conflicts with its neighbors. The solution achieves the equivalent of the best known sequentially achievable upper bound for request satisfiability, implied by a theorem due to Alon and Tarsi [12]. The solution assumes an initial acyclic orientation of the graph which can be obtained by a vertex coloring and orienting each edge towards its lower-colored adjacent vertex and uses this graph for resolving priorities among competing base stations. It is based on a mutual exclusion approach - where only one base station in a neighborhood is able to select available frequencies at any given point in time. It employs a double doorway mechanism [10] to ensure synchronization among the neighbors with small delay and no starvation. The solution has a failure locality of 4, worst case response time of O(Δ) and a messaging complexity of

$O(\Delta^2)$, where Δ is the degree of the graph (the maximum degree among all vertices in the graph). This algorithm achieves the equivalent of the bound in *lemma 2* (Chapter 2). To achieve this, the solution needs some extra synchronization.

4.2.1. A preliminary solution: *Protocol* PREL-DET-DLC

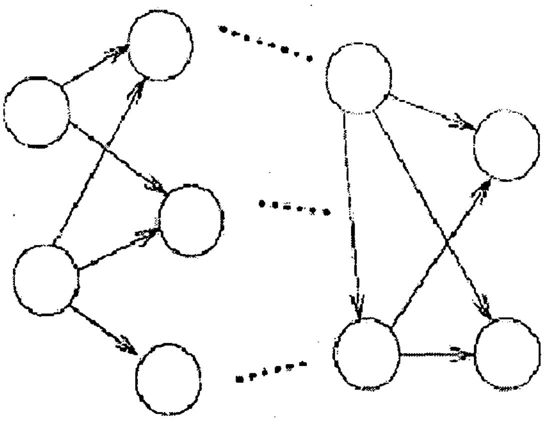
An acyclic orientation can be obtained by first executing a distributed vertex coloring protocol and then orienting each edge from the higher to the lower color. Arbitrary graphs can be colored in a distributed manner with $\Delta + 1$ colors [1], while for planar ones 6 is sufficient. To maintain the orientation, a privilege token pr_{ij} is associated with each edge (v_i, v_j). The node for which the edge is incoming holds the token (i.e. its privilege variable is true).

- Starting from such an acyclic orientation let the sinks concurrently and independently list-color themselves first. There is no conflict between them when they choose colors since no two sinks are adjacent.

- When a sink has chosen colors, it communicates to its neighbors the colors it chose and reverses the orientation of its incident edges (by sending to its neighbors the respective token for each edge).

- The resulting orientation is again acyclic, so the new sinks can list-color themselves next.

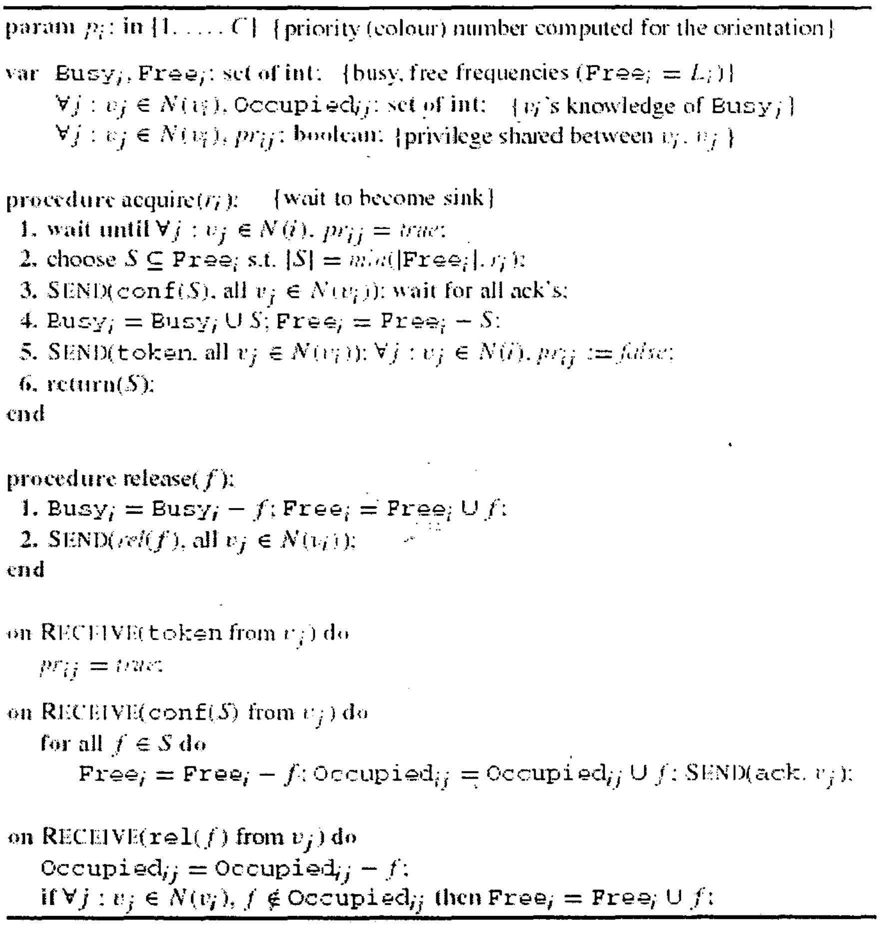
By repeating this procedure, in time proportional to the length of the longest ***waiting chain*** (directed path) in the initial orientation, all the nodes of the graph will be list-colored once, provided that for each node, the sum of the number of its requests and of those of its outgoing-edge neighbors in the initial orientation does not exceed its list size.



Maximun Waiting Chain Length = Number of Colors = Max Degree +1
(for planar graphs)

Fig 11: Solution using Orientation

The initial orientations chosen above, the longest chain length depends only on local measures; namely, it is at most Δ + 1 and 6 for general and planar graphs, respectively. Hence, this will work well as a *one-shot protocol* (i.e. if each node is interested in choosing colors once), since it also gives the best known guarantee for request satisfiability.



```
param $p_i$: in {1, ..., C}  {priority (colour) number computed for the orientation}

var  Busy_i, Free_i: set of int:  {busy, free frequencies ($Free_i = L_i$)}
     $\forall j : v_j \in N(v_i)$, Occupied_ij: set of int:  {$v_j$'s knowledge of Busy_j}
     $\forall j : v_j \in N(v_i)$, $pr_{ij}$: boolean:  {privilege shared between $v_i, v_j$}

procedure acquire($r_i$):    {wait to become sink}
  wait until $\forall j : v_j \in N(i)$, $pr_{ij} = true$;
  choose $S \subseteq Free_i$ s.t. $|S| = min(|Free_i|, r_i)$;
  SEND(conf(S), all $v_j \in N(v_i)$); wait for all ack's;
  $Busy_i = Busy_i \cup S$; $Free_i = Free_i - S$;
  SEND(token, all $v_j \in N(v_i)$); $\forall j : v_j \in N(i)$, $pr_{ij} := false$;
  return(S);
end

procedure release(f):
  $Busy_i = Busy_i - f$; $Free_i = Free_i \cup f$;
  SEND(rel(f), all $v_j \in N(v_i)$);
end

on RECEIVE(token from $v_j$) do
   $pr_{ij} = true$;

on RECEIVE(conf(S) from $v_j$) do
   for all $f \in S$ do
       $Free_i = Free_i - f$; $Occupied_{ij} = Occupied_{ij} \cup f$; SEND(ack, $v_j$);

on RECEIVE(rel(f) from $v_j$) do
   $Occupied_{ij} = Occupied_{ij} - f$;
   if $\forall j : v_j \in N(v_i)$, $f \notin Occupied_{ij}$ then $Free_i = Free_i \cup f$;
```

Fig 12: Protocol PREL_DET_DLC for node V_i

4.2.2 Theorem Protocol PREL_DET_DLC is a correct distributed solution for the generalized list coloring and the channel allocation problems. In the worst case, for a station v_i to get r_i frequencies (colors) from its $Free_i$ list at time t, it must be the case that

$$|Free_i| \geq r_i + \sum_{v_v \in N(v_i):\, pr_{ij} = false} r_j$$

at all time t. For this under continuous demand of frequencies at most 3 Δ messages are needed. The failure locality of the solution is Δ + 1 (6 if *G* is planar) for synchronous systems, or *n* for asynchronous systems. The worst-case response time is O(*Δ)* for synchronous systems and O*(n)* for asynchronous systems. The size of the messages is 2 bits, except from Δ messages, which are O(r) bits long.

4.2.3 *Protocol* DET_DLC

In a synchronous system the solution described above works fine. However, for an approach which is suitable for any system, the double doorway synchronization mechanism, Choy and Singh [10] is used.

An initial vertex coloring of the graph is assumed, in order to serve as a priority scheme. Assume *C* colors are used. Between neighboring nodes that compete for frequencies at the same time, the one with smaller color (p_i) chooses first. In order to avoid starvation of the nodes with low priority due to preemption by their higher priority neighbors, double doorway synchronization is used.

- ***The asynchronous doorway***

  It is equivalent to the process asking for permission to enter and waiting until it receives one (i.e. a synchronization message = $sync_1$) from each of its neighbors once.

  1. **Guarantees starvation freedom** – Provided that a process that has already crossed the doorway defers giving permission to its neighbors until after it has chosen frequencies.

2. **Exponential response time for worst case** (O (Δ^C)) – This is because after each process v_i has crossed the doorway, its neighbors with smaller color may also cross and they may do so one by one, in such a "*staggered*" timing that each one of them prevents *vi* from entering the *choosing phase* exactly at the time that v_i is ready to enter it.

- ***The synchronous doorway***

It is equivalent to the process waiting until it finds an instant when none of its neighbors is past this second doorway, i.e. at the same time the last *synchronization message* received by each one of its neighbors must be = sync_2

1. Prevent the above-described exponential scenario.

2. It can result in starvation.

However the *two doorways combined* cancel each others drawbacks and furthermore guarantee that each process that tries to enter the choosing phase can only possibly be taken over at most once by each of it's faster neighbors. After having crossed both doorways (sending sync_2 messages to all neighbors) a process v_i has to wait for its neighbors that are also past both doorways, to pass to it the respective privileges (figure 13, line 3 of procedure *acquire*). Initially this is meant to implement the edge-reversal procedure.

4.2.4 Privilege Release Mechanism

The use of a ***privilege-release mechanism*** guarantees good failure locality as it reduces the waiting chain length.

- *A competing node v_i, which has crossed both doorways, first requests the privileges from its higher priority neighbors and then (i.e. after having obtained those) from the lower priority ones.*

- *A node, which is non-competing or has not crossed both doorways, always gives a requested privilege.*

- *A competing node v_i, which has crossed both doorways, gives the privilege to a:*
    - *lower priority neighbor if it has not collected the privileges from all its higher priority neighbors; otherwise it defers answering until after it has finished choosing frequencies.*
    - *higher priority neighbor if it has not collected all the other privileges.*

```
var $\mathtt{Busy}_i, \mathtt{Free}_i$: set of int: {busy, free frequencies ($\mathtt{Free}_i = L_i$)}
    $\forall j : v_j \in N(v_i)$, $\mathtt{Occupied}_{ij}$: set of int: {$v_i$'s knowledge of $\mathtt{Busy}_j$}
    $\forall j : v_j \in N(v_i)$, $l_{ij}$: msgtype: {last synchronization message received from $v_j$}
    $\forall j : v_j \in N(v_i)$, $pr_{ij}$: boolean: {privilege shared between $v_i, v_j$;
                                                    one is true at a time}

param $p_i$: in $\{1, \ldots, C\}$ {priority (colour) number computed for the orientation}

procedure acquire($r_i$):
  1. $\forall j : v_j \in N(i)$, wait until $l_{ij} \neq \mathtt{sync}_1$;
                              {1st doorway: wait until true once for each $v_j$}
     SEND($\mathtt{sync}_1$, all $v_j \in N(v_i)$);
  2. wait until $\forall j : v_j \in N(i)$, $l_{ij} \neq \mathtt{sync}_2$;
                              {2nd doorway: wait until concurrently true for all $v_j$'s}
     SEND($\mathtt{sync}_2$, all $v_j \in N(v_i)$);
  3. < request and wait for all privileges $pr_{ij}$ >;
  4. choose $S \subseteq \mathtt{Free}_i$ s.t. $|S| = min(|\mathtt{Free}_i|, r_i)$;
  5. SEND($\mathtt{conf}(S)$, all $v_j \in N(v_i)$); wait for all ack's;
  6. $\mathtt{Busy}_i = \mathtt{Busy}_i \cup S$; $\mathtt{Free}_i = \mathtt{Free}_i - S$;
  7. SEND($\mathtt{sync}_3$, all $v_j \in N(v_i)$); {signal for doorways}
  8. return($S$);
end

procedure release($f$):
  1. $\mathtt{Busy}_i = \mathtt{Busy}_i - f$; $\mathtt{Free}_i = \mathtt{Free}_i \cup f$;
  2. SEND($rel(f)$, all $v_j \in N(v_i)$);
end

on RECEIVE($\mathtt{sync}_x$ from $v_j$) do
    $l_{ij} = \mathtt{sync}_x$;

on RECEIVE($\mathtt{conf}(S)$ from $v_j$) do
    for all $f \in S$ do
        $\mathtt{Free}_i = \mathtt{Free}_i - f$; $\mathtt{Occupied}_{ij} = \mathtt{Occupied}_{ij} \cup f$; SEND(ack, $v_j$);

on RECEIVE($\mathtt{rel}(f)$ from $v_j$) do
    $\mathtt{Occupied}_{ij} = \mathtt{Occupied}_{ij} - f$;
    if $\forall j : v_j \in N(v_i)$, $f \notin \mathtt{Occupied}_{ij}$ then $\mathtt{Free}_i = \mathtt{Free}_i \cup f$;
```

Fig 13: Protocol DET_DLC for node V_i

4.2.5 Theorem

Protocol DET-DLC is a correct distributed solution for the generalized list-coloring and channel allocation problems. In the worst case, after a process v_i has executed step 1 of procedure acquire, it is guaranteed to get r_i colors (frequencies) from its $Free_i$ list if

$$|Free_i| \geq r_i + \sum_{v_j \in N(v_i)} r_j$$

at that point in the execution. For this, the worst case response time and the number of messages needed are $O(\Delta^2)$ ($O(\Delta)$ if G is planar). The failure locality of the solution is 4. The size of the messages is 2 bits, except from Δ messages which are $O(r_i)$ bits long.

4.3 Randomized Distributed List Coloring

Previous solution for the *generalized distributed list coloring* problem relied on mutual exclusion. To determine what is achievable without any advanced synchronization. Randomized approach that assigns colors to a node such that they do not conflict with the neighbors; the number of colors assigned depends (*with high probability*) on the ratio of the size of the list and the degree of the node.

4.3.1 Randomized Distributed Approach

The RAND_DLC solution by does not rely on a *mutual exclusion* approach. Instead, each base station is able to select a random subset of the available frequencies, and then, through a negotiation phase with its neighbors, determine which of the frequencies can be used. The solution ensures that no frequency can be concurrently used by any two neighbors. The expected number of frequencies that a station can acquire is a fraction of 1/ 4 Δ of the free frequencies at the station. The message complexity of the solution is 3 Δ, and the failure locality is one.

```
var $Busy_i$, $Free_i$: set of int;   {busy, free frequencies ($Free_i = L_i$)}
    $\forall j : v_j \in N(v_i)$, $Occupied_{ij}$: set of int; {$v_i$'s knowledge of $Busy_j$}

procedure acquire( ):
  1. Pick, randomly, $S \subseteq Free_i$ s.t. $|S| = \epsilon|Free_i|$
     $Busy_i = Busy_i \cup S$; $Free_i = Free_i - S$;
  2. SEND(pick($S$), all $v_j \in N(v_i)$); wait for all replies;
  3. for all $f \in S$ do
         if some neighbour has replied 'No' for $f$ then
           $Busy_i = Busy_i - f$; $S = S - f$;
           if $\forall j, v_j \in N(v_i)$, $f \notin Occupied_{ij}$ then $Free_i = Free_i \cup f$;
  4. SEND(conf($S$), all $v_j \in N(v_i)$);  {$S$ is now the set of acquired frequencies}
  5. return($S$);
end

procedure release($f$):
  1. $Busy_i = Busy_i - f$; $Free_i = Free_i \cup f$; $S = S - f$;
  2. SEND(rel($f$), all $v_j \in N(v_i)$);
end

on RECEIVE(pick($S$)) do
   for all $f \in S$ do
       if $f \in Busy_i$ then REPLY('No', $f$) else REPLY('Yes', $f$);

on RECEIVE(conf($S$) from $v_j$) do
   for all $f \in S$ do
       $Free_i = Free_i - f$; $Occupied_{ij} = Occupied_{ij} \cup f$;

on RECEIVE(rel($f$) from $v_j$) do
   $Occupied_{ij} = Occupied_{ij} - f$;
   if $\forall j, v_j \in N(v_i)$, $f \notin Occupied_{ij}$ then $Free_i = Free_i \cup f$;
```

Fig 14: protocol RAND DLC for node V_i

Let node v_i have degree Δ and a list L_i of cardinality f. Each competing node, picks randomly, an ε fraction of the colors in L_i ; thus, each color in the list is picked with a probability ε. The node then checks with its neighbors to ensure that some color it picked is not picked by a neighbor; the color is dropped if such is the case. Thus a color picked by two neighbors could potentially be dropped by both of them. Node v_i, as in protocol DET-DLC, maintains two sets $Free_i$, the set of frequencies that are available to be picked and $Busy_i$, the set of frequencies that are in use; it also maintains for each neighbor v_j a set $Occupied_{ij}$, its view of the set $Busy_j$.

4.3.2 Theorem

Protocol RAND-DLC is a correct distributed solution for the generalized list-coloring and channel allocation problems. With the exchange of 3Δ messages of size $O(r)$ bits and in 3 rounds of communication a node v_i is expected to get at least

$$\mu = \frac{|\mathrm{Free}_i|}{4\Delta}$$

colors (frequencies) from its Free_i list, while with probability at least $1 - e^{-1}$ it gets at least $r = \mu - \sqrt{2\mu}$ colors (frequencies). The solution has failure locality 1.

4.4 Multiple Mutual Exclusion

The solution by Prakash et al. [15] (DIST_DCA) has the following properties

(i) It is distributed (based on a successful centralized technique).

(ii) It is deadlock free, starvation free and fair and has bounded latency.

(iii) It has earlier been shown experimentally to perform well under stable traffic.

DIST_DCA is based on a *multiple mutual exclusion* scheme: a base station attempts to acquire one frequency at a time, initially getting them from the local set of available frequencies, and, later, when that is exhausted, *borrowing* them from neighboring stations. The mutual exclusion mechanism is based on a priority scheme that uses timestamps based on *Lamport's logical clocks*, to order the base stations requests. This method of ordering however has a number of negative properties, one being the length of the base stations waiting chain, which can grow to the size of the network. This also implies a failure locality of the same size, i.e. the size of the network. However, in an environment, where the load is uniform and does not change frequently, and where there are no faults, the algorithm performs reasonably well and hence is used as a benchmark.

CHAPTER 5

Performance Tradeoffs and Analysis

5.1 Introduction

Frequency allocation in cellular networks has a number of *optimization criteria* and a number of interesting *trade-offs* involved in trying to meet them. Analyzing them is particularly important for applying frequency allocation algorithms in practice. Distributed setting implies a larger number of parameters to consider, but is more suitable for dynamic solutions. In [13], it was shown how to analytically measure and provide worst-case guarantees regarding *request satisfiability* and how to provide *fully dynamic* frequency assignment (i.e. *on demand*) with low communication and time overhead, which is particularly important for adapting to temporal changes in frequency demands in each cell. In applying these methods, the *trade-off* between connection set-up time and request satisfiability had to be considered. By designing appropriate *dynamic balancing* strategies, the trade-offs involved could be equilibrated in a way so that the gains are substantial, while the losses remain small.

5.2 Performance Goals and Objective

Following are the goals of any channel assignment protocol and hence act as an important factor for the design and analysis. As pointed out it is not possible to meet all of them concurrently so there are associated tradeoffs. Our aim is to find the optimal strategy so that the gains are maximized.

- Maximize ***satisfiable requests*** (or ***bandwidth utilization***)
- Minimize ***connection setup time***
- Minimize ***communication complexity***
- Avoid intra-cell ***handoffs***
- Minimize ***failure locality***
- ***No starvation*** until process fails
- Minimize bound on ***response time***

5.3 Performance Evaluation

Strategies for applying and fine-tuning the performance of the previous algorithms should not affect the worst-case guarantees. The amount of frequencies *allocated* at each base station follows the *offered load* at that station. We studied and analyze the performance of the original algorithms and their tunable versions. The set of experiments model a variety of situations, including *dynamic*, *non-uniform load* and base station *failures*, having the main variable parameter actually be the *temporal-change-of-traffic-distribution*; it measures an extensive range of performance metrics (including *utilization, frequency reuse, fault-tolerance*).

5.3.1 Performance Evaluation Matrices

Below we define the measures we are used to evaluate the algorithms. v_i denotes a base station, N(*i*) denotes it's neighborhood, i.e. itself and the nearby base stations with which the potential for co-channel interference can occur and $Used_i$ denotes the number of frequencies it actually uses each time in order to serve requests.

- **Response Time:** the implied overhead for the connection setup time. Both the mean response time for queued requests and the mean for all requests are measured.

- **Direct Connections:** the frequency requests that can be serviced locally, without the need for the base station to acquire more, thus implying a faster response time. This, combined with the response time gives a more complete picture of the connection setup time, as a direct connection will have a significantly faster response time than a queued request.

- **Dropped Calls:** the number of requests that could not be satisfied.

- **Total Messages:** the total number of messages (or *communication complexity*) over the whole network for the entire experiment execution.

- **Frequency Reuse:** the ratio of the number of calls served in a neighborhood over the number of frequencies that are used for them, i.e.

  $$\frac{\sum_{j\in N(i)}\left|Used_j\right|}{\left|Used\right|}, \text{ where } Used = \bigcup_{j\in N(i)} Used_j$$

  It is measured when a call is dropped.

- **Bandwidth Utilization:** the total number of frequencies in use, divided by the spectrum size, i.e.

  $$\frac{\left|Used\right|}{Spectrum}, \text{ where } Used = \bigcup_{j\in N(i)} Used_j$$

  A higher value indicates that the algorithm is effectively estimating the number of frequencies to acquire and retain. It is measured when a call is dropped.

- **Number of Affected Base Stations:** the number of base stations, which have to wait indefinitely due to the failure of others.

Having as our main aim to simulate, the traffic for a real cellular network, the key parameter we used for capturing this is the use of "*hot-cell configurations*" which is a distribution of the load across the network. In each configuration a cell may be *hot*, *normal*, or *cold*, corresponding to having high, moderate or low load, respectively. The configurations changed throughout the execution, reflecting non-uniform, dynamically changing load.

5.3.2 Simulation Parameters

- **Network size:** 49 cells (7x7 cellular grid; internal and border cells have 6 and 2-3 neighbors, respectively)
- **Spectrum size:** 500 frequencies
- **Types of cells:** hot, normal, and cold.
- **Changes in hot-cell configurations:** modeling dynamic and non-uniform load.
- **Network (message) delay:** 1 to 3 milliseconds.
- **Process (base-station) step time delay:** 1 to 2 microseconds.
- **Mean call duration:** exponentially distributed with a mean $1/\mu = 3$ minutes.
- **Mean call arrivals rate λ:** Poisson distribution, in hot-cells λ=85/min, normal cells, λ=45/min cold cells, λ =20/min.
- **Length of simulated execution:** 1,00,000 requests, each one requesting one channel (frequency)
- **Failures:** Up to three crash-failures occur at arbitrary stations at arbitrary points in the execution.

5.3.3 Randomized Backoff

In the case of RAND_DLC, when a frequency is required, then we attempt to acquire one by picking a subset from the set of free frequencies known to station. If none are available, we can either drop the call or retry. Retrying immediately is likely to produce the same result, so it is worthwhile to *backoff* and try again after a set time period. This time period must of course be determined by the time that a mobile user is prepared to wait for a connection. For this reason, one retry attempt was used, and the back-off time was chosen randomly between 100 and 500 milliseconds.

5.4 Maximizing Satisfiable Requests

5.4.1 Frequency Availability vs. Response Time

Aiming at maximizing the number of *satisfiable requests*, the best strategy is to assign frequencies to base stations on demand, leaving the unused frequencies to satisfy calls in different base stations. On the other hand, considering minimizing the overhead per connection, it is best for each base station to reserve some of the frequencies for satisfying as many requests as possible locally.

5.4.2 Frequency Acquisition and Retention

In the default behavior of DET_DLC, a base station i will select a subset S, of the free frequencies, such that $|S| = \min(|Free_i, r_i|)$, where r_i is the number of frequency requests for process/base station i. However, in certain situations a frequency acquisition strategy can take into account that some acquisition/ retention of extra frequencies can improve the performance. In the default behavior of the RAND_DLC, a fixed fraction of the free frequencies are chosen. The expected number of frequencies to be picked without conflicts with neighbors that are choosing concurrently is maximized using a $1/(\Delta+1)$ fraction. This value ensures that a process will select a set of frequencies inversely proportional to the number of neighbors it has. However, if the load is not evenly distributed this may be not the best choice; for example a *cold cell* (one which has a light load) may acquire frequencies which could better be used by one of its neighbors which has a higher load. It may be some time until these unused frequencies are made available again. We can have acquisition strategies that are able to respond to changes in the load of the network. In doing so we must maintain the *locality* of the solution. The second observation we can make of both DET_DLC and RAND_DLC, is that a base station v_i releasing a frequency after it has been used, makes it available for other neighbors, but it (v_i) may have queued requests that require the use of the frequency. The net effect of this is that the call is queued and v_i must acquire another frequency using the respective algorithm. The problem here is to develop a fair retains strategy *without starving* the low-load cells and hence the tradeoff.

5.4.3 Frequency Retention /Acquisition Strategies

- **Little's Law:**

  This method gave significant improvements in time and communication overhead per request, at a very small cost in dropped calls and bandwidth utilization compared with the default DET_DLC and RAND_DLC solutions studied.

- **Current Waiting Queue Size:**

  The r_i for a station v_i can be helpful through two parameters. The first parameter, called the *free ratio*, reduces the number of frequencies acquired (or retained), in an attempt to additionally ensure fairness in frequency acquisition and see the effects in the overall algorithm performance. The second parameter, *minimum ratio*, is used when the queue size at a station is very small. The minimum ratio is calculated as the minimum non-zero Little's-Law estimate encountered so far for the base station. These attempts to ensure that some frequencies are acquired or retained, so that cold cells are able to service their calls directly, without incurring additional communication cost. A base station uses these two parameters in order to compute the Queue Ratio and the estimate of frequencies to be acquired/retained.

Parameter	Description
Δ_i	degree (# of neighbours) of v_i)
$Busy_i$	set of frequencies held by v_i
λ_i	mean request-arrival rate at v_i
T	mean call duration
r_i	actual number of requests in waiting queue at v_i
$Little'sLaw$	mean number of requests estimated at v_i ($\lambda_i T$)
$free\_ratio$	fraction of frequencies to be left in neighbourhood for fairness
min_ratio_i	$min(\lambda_i T) \neq 0$ observed by v_i
LittlesLaw Strategy	$\lambda_i T$ - $\lvert Busy_i \rvert$
$QueueRatio_i$	$r_i(1+1/(\Delta+1))(1-free\_ratio)$
QueueRatio Strategy	$max(QueueRatio_i, min\_ratio_i)$

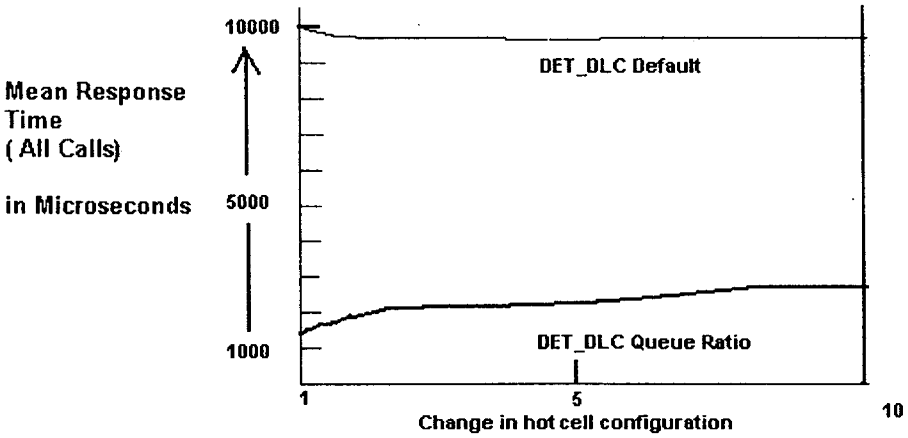
Table 1. Definitions Used in Tuning Strategies

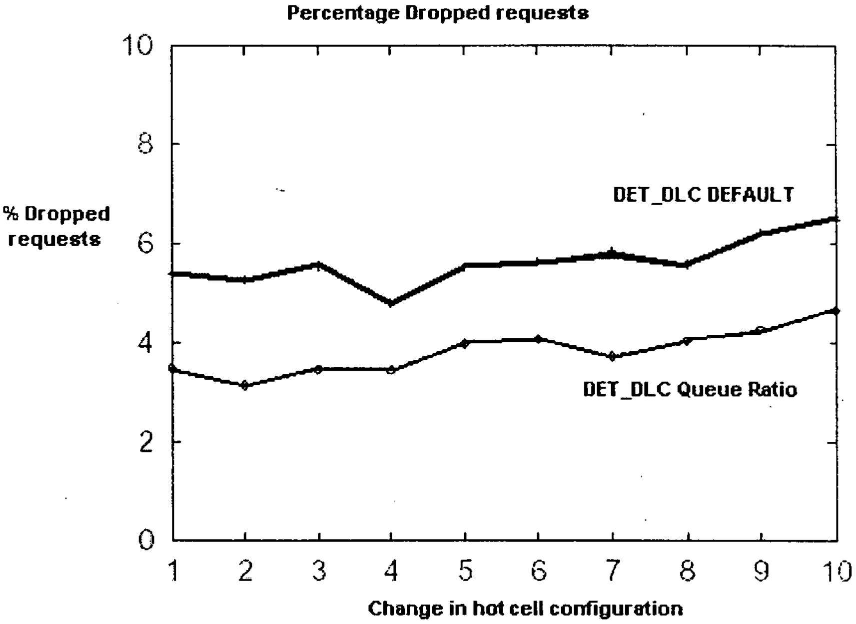
5.5 Results

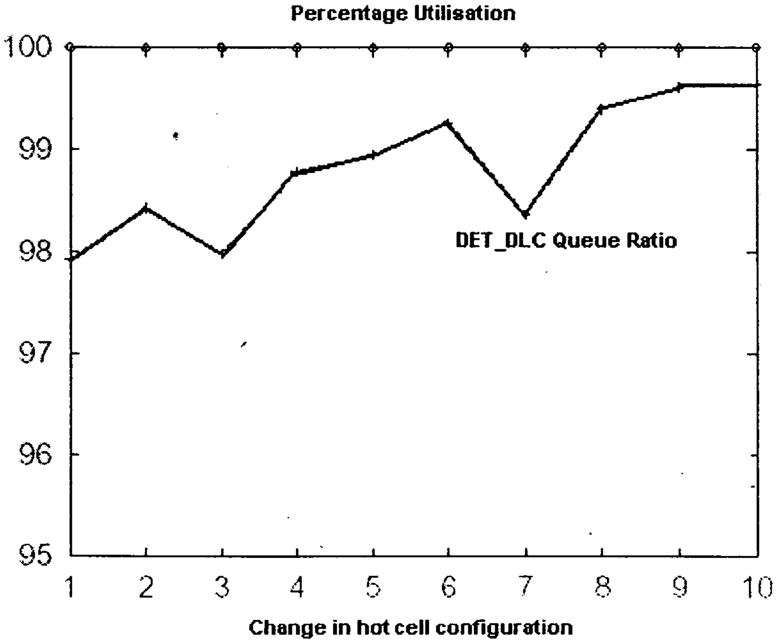
The results for all algorithms are summarized graphically in an attempt to give a complete overview of them. These were derived from above mentioned data values. The measures seem to converge after 10 changes of hot-cell configurations.

5.5.1 DET_DCA

The results for all algorithms are summarized graphically in an attempt to give a complete overview of them. These were derived from above mentioned data values. The measures seem to converge after 10 changes of hot-cell configurations. Figures shows some key results regarding improvements and trade-offs implied by the tuning strategies to DET_DLC.




Percentage Dropped requests
% Dropped requests
10
8
6
4
2
0
DET_DLC DEFAULT
DET_DLC Queue Ratio
1
2
3
4
5
6
7
8
9
10
Change in hot cell configuration


Percentage Utilisation
100
99
98
97
96
95
DET_DLC Queue Ratio
1
2
3
4
5
6
7
8
9
10
Change in hot cell configuration

Expected Trade-offs and Results

- If we take the perspective of trying to maximize the number of direct connection requests, we can see that for DET_DLC this can be improved significantly, by as much as 77%, as the default algorithm does not meet any requests directly.

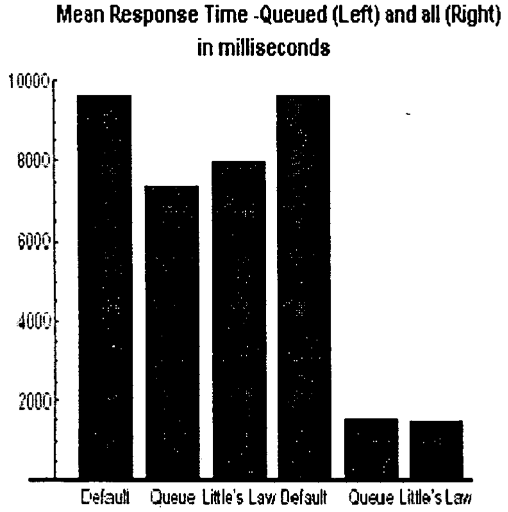
- There are trade-offs involved in acquiring extra frequencies and retaining frequencies, the most important being an increase in the number of dropped connection requests of 1.8% for the *QueueRatio* strategy and 1.2% for the *Little's Law* based strategy.

- There are some other benefits in trying to maximize the number of requests we can service directly, such as a great reduction in the number of messages.

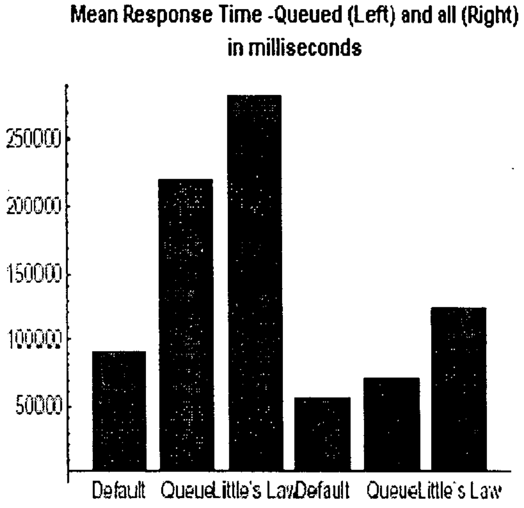
5.5.2 RAND_DLC

For RAND_DLC, the benefits are not so pronounced. This is because in the default algorithm the number of frequencies acquired can in many cases be larger than the number of requests, implying availability of frequencies to service requests immediately. Using the tuning strategies we can increase the number of direct connections by 17% for the *QueueRatio* strategy and 10% for *Little's Law*, with the trade-off being that less requests can be satisfied in total. As with DET_DLC, this reduction in satisfiabilities reflected in the results for the utilization and frequency reuse, with the *QueueRatio* strategy faring marginally better than the *Little's Law* based strategy. One negative result is the increase in response time, both for queued requests and overall. This is due to the action of the back-off strategy.

5.5.3 Mean Response Time

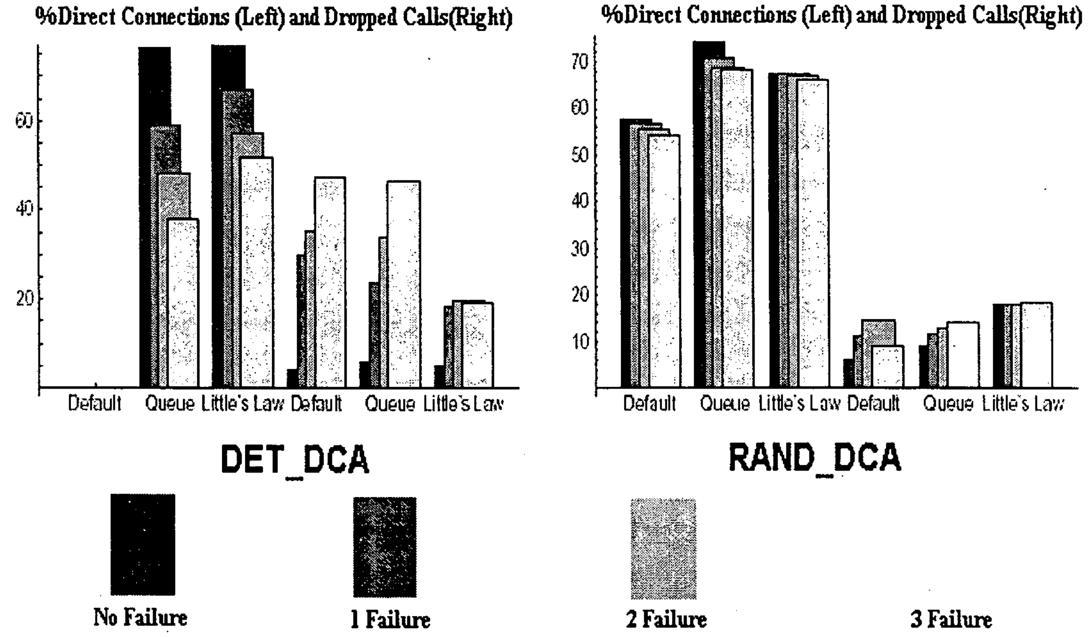


DET_DCA

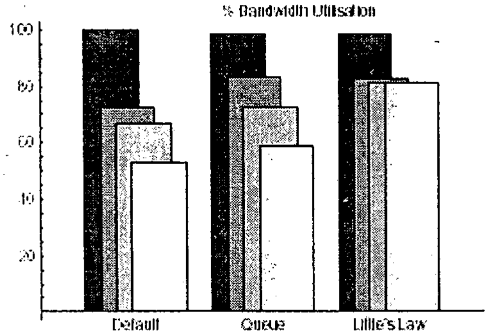


RAND_DCA

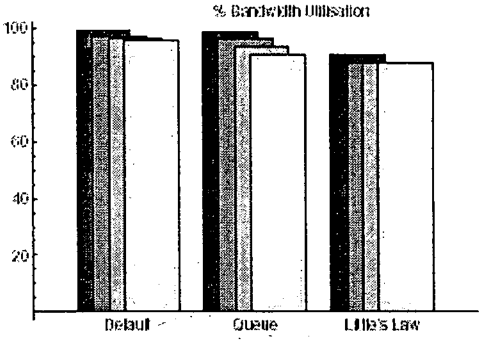
5.5.4 Direct Connections/Dropped Calls



5.5.5 Bandwidth Utilization

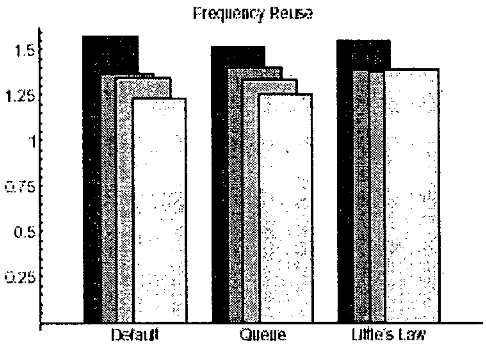


DET_DLC

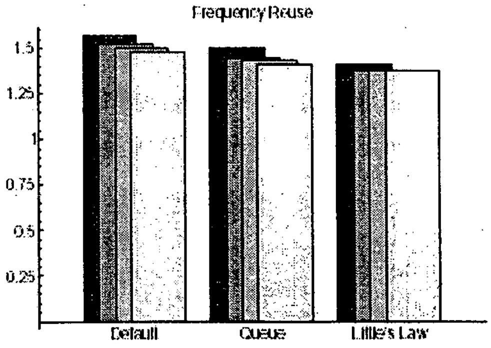


RAND_DLC

5.5.6 Frequency Reuse



DET_DLC



RAND_DLC

Conclusion

With rapidly growing interest in area of wireless communications in recent years, the wireless resources allocation problem has received tremendous attention. As a result, a vast amount of research has been done to extend the earlier work as well as introduce new techniques. Most of the recent work has been in the area of distributed, adaptive, measurement based, power control based, priority based, and overlay channel allocation schemes. In addition, a vast amount of results have been published which provide an insight into the performance, complexity, and stability of different channel allocation algorithms. With recent trends in the areas of *microcellular* networks and wireless access broadband networks where multimedia applications will be extended to the end users over wireless links, we are faced with new, interesting and important challenges to wireless resource allocation problem. These challenges have arisen as a result of emerging and new technologies, a result of recent advances in design of low power handheld wireless terminals, the design of advanced radio modems and antennas and finally recent development in he area of spread-spectrum systems. These emerging new areas will introduce a new set of constraints in the resource and channel allocation problems. The solution of these problems will play an important role in providing ubiquitous access to multimedia applications in personal communication network.

Frequency allocation in cellular networks has a number of *optimization criteria* and a number of interesting trade-offs involved in trying to meet them. Analyzing them is particularly important for applying frequency allocation algorithms in practice. Distributed setting implies a larger number of parameters to consider, but is more suitable for dynamic solutions. In this thesis we have measured the performance of some of the best known distributed channel assignment algorithms analytically and by simulation. The comparative results are studied for original algorithms and their calibrated versions by means of various performance matrices.

We have analyzed tradeoff between connection setup time and request satisfiability at variable load patterns. In particular the provision by each base station to require/retain extra frequencies for being able to accommodate calls locally results in ensuring faster response time. If we take the perspective of trying to maximize the number of direct connection requests, we see that the performance can be increased dramatically as the host can be granted the channel without running the algorithms to find the free channel.

Also, for the network to be able to handle rapidly changing load patterns, the tradeoff between the amount of network traffic generated and the response time is very significant especially in centralized setting, where each node has to communicate with a central node. By designing appropriate *dynamic balancing* strategies, the trade-offs involved could be equilibrated in a way so that the gains are substantial, while the losses remain small.

Further work in this direction can be the analysis and design of the *frequency allocation algorithms* that can select the frequency by using the "***frequency reuse information*** ", but the failure locality and the guarantees in the original algorithms are to be maintained. These schemes can take into account the temporal or spatial channel demands along with other factors.

References

[1] B.Awerbuch, A.Goldberg, M.Luby and S.Plotkin, *Network Decomposition and Locality in Distributed Computation*, Proceedings of IEEE 30th Symposium on Foundations of Computer Science, 1989, 364-369.

[2] B.R. Badrinath, A. Acharaya and T. Imielinski, *Impact of Mobility on Distributed Computations*, Operating Systems Review 27(2), April 1993,15-20.M

[3] C. Thomassen, Every planar graph is 5-choosable, Journal of Combinatorial Theory, Series B 62 (1994) 180–181.

[4] D.Bertsekas and R. Gallager, *Data Networks*, Prentice Hall International Edition, 1987.

[5] H.U. Simon, *The Analysis of Hybrid and Dynamic Channel Assignment*, Technical Report, University of Sarland 1988.

[6] J. Chuang, *Performance Issues and Algorithms for Dynamic Channel Allocation*, IEEE Journal on Selected Areas in Communications, Vol. 11,no.6, August 1993,955-63.

[7] Jerry D. Gibson, *The Mobile Communications Handbook* (second edition) CRC press 23.5-23.6.

[8] Lydian –An Educational Animation Environment for Distributed Algorithms and Protocols. http://www.cs.chalmers.se/~lydian/(2001)

[9] M.Biro, M.Hujter and Z.Tuza, *Precoloring Extensions*, Interval Graphs, Discrete Mathematics 100 (1992) 267-279.

[10] M.Choy and A. Singh, *Efficient Fault Tolerant Algorithms for Resource Allocation in Distributed Systems*, ACM Transactions in Programming Languages and Systems 17(3), May 1995, 535-559.

[11] Marina Papatriantafilou, David Rutter, Philip Tsigas, *Distributed Frequency Allocation for Cellular Networks: Tradeoffs and Tuning Stratagies*, Techinical Report, Chalmers Institute of Technology.

[12] N. Alon and M.Tarsi, *Colorings and Orientations of Graphs*, Combinatorica, 12(2), 1992,125-134.

[13] Naveen Garg, Marina Papatriantafilou and Philippas Tsigas, *Distributed long Lived List Coloring: How to Dynamically Allocate Frequencies in Cellular Networks,* Wireless Networks 8,2002, 49-60.

[14] P. Erdos, A.L. Rubin and H.Taylor, Choosability in graphs, *Proceedings of the West Coast Conference on Combinatorics*, Graph Theory and Computing (Congr. no.26),125-157

[15] R. Prakesh, N.G. Shivarti, M.Singhal, *Distributed Dynamic Fault Tolerant Channel Allocation for Mobile Computing*, IEEE Transactions on Vehicular Technology, Vol. 48(6) November 1999,IEEE New Jersy,1874-1888

[16] S.G. Hild, *A Brief History of Mobile Telephony*, Technical Report No.372, University of Cambridge, Computer Laboratory, January 1995.

[17] V.Barbosa and E. Gafni, *Concurrency in Heavily Loaded Neighbourhood Constrained Systems*, ACM Transactions on Programming Languages and Systems 11(4), October 1989, 562-584.