TECHNISCHE UNIVERSITÄT WIEN

DISSERTATION

# A "Pricing" and Game Theory Based Approach to Radio Resource Management

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der technischen Wissenschaften

eingereicht an der
Technischen Universität Wien
Fakultät für Elektrotechnik und Informationstechnik

von

Dipl.-Ing. Mugdim Bublin
Matrikelnummer 9326133
A-1150 Wien, Anschützgasse 8/3

Wien, im November 2004 ..................................

First supervisor: Prof. Dipl.-Ing. Dr. Richard Eier
Institut für Computertechnik
Technische Universität Wien
Fakultät für Elektrotechnik und Informationstechnik

Second supervisor: Prof. Dipl.-Ing. Dr. Ernst Bonek
Institut für Nachrichtentechnik und Hochfrequenztechnik
Technische Universität Wien
Fakultät für Elektrotechnik und Informationstechnik

## *Zusammenfassung*

Radio Ressource Management (RRM) Algorithmen spielen eine wesentliche Rolle bei der effizienten Verwendung der knappen Funkressourcen in Mobilfunknetzen. Die Klasse der verteilten, selbstadaptiven RRM Algorithmen gewinnt wegen ihrer geringen Komplexität und ihres geringen Signalisierungsaufwands immer mehr an Bedeutung. In dieser Arbeit verwenden wir "pricing" und den spieltheoretischen Zugang zur Entwicklung von verteilten, selbstadaptiven RRM Algorithmen. In unserem Model maximiert ein Netzbetreiber seinen Gewinn, indem er die „Preise" für die Funkressourcen gemäß dem Netzzustand festlegt. Die mobilen Teilnehmer wählen dann die Ressourcen aus, die ihre Nutzfunktionen unter Verwendung der „Ressourcepreise" als Parameter maximieren. Wegen der speziellen Form unserer Nutzfunktion können manche bekannte RRM Algorithmen als zustandabhängige „Regeln" durch die Preise „erzwingen" werden.

Wir untersuchen weiter, welche „Regeln" (Algorithmen) „nicht „ausbeutbar" („evolutionär" stabil) in Netzen ohne „Preisen" wie in „ad-hoc" Netzen sind. Wir zeigen auch, durch die Verwendung der Spieltheorie, wie eine Kooperation unter „selbstsüchtigen" Nutzern in solchen Netzen entstehen kann. Weiter, wir setzten voraus, daß unsere „Spieler" „beschränkt rational" sind und beschränkte Informationen wie lokale Messungen zur Verfügung haben. Unsere Algorithmen suchen auch nicht nach absolut optimalen Lösungen, sondern nach „genug guten" Lösungen. Diese Eigenschaften ermöglichen praktische Implementierungen unserer Algorithmen mit niedrigen Berechnungs- und Signalisierungsaufwand.

Wir untersuchen Kapazitätsgewinne und Kompromisse von RRM Algorithmen, wie Power Control, Scheduling, Dynamic Channel Allocation oder Handover, durch Systemsimulationen in verschiedenen Umgebungen (städtisch und ländlich) für unterschiedliche Lasten und Dienste (Sprache und Paketdaten), mit und ohne Intelligente Antennen.

Die Ergebnisse dieser Arbeit sind sowohl für Mobilfunkgerätehersteller als auch für Netzbetreiber von Nutzen. Die Hersteller können die Ergebnisse verwenden, um effiziente, verteilte, adaptive RRM Algorithmen mit niedrigem Signalisierungsaufwand zu entwickeln. Betreiber können die Ergebnisse dazu verwenden, um optimale RRM Algorithmen für jeden Netzzustand durch Preisparameter ohne Bedarf an Softwareänderungen zu „aktivieren", oder um Teilnehmer optimal zu vergebühren, oder auch als Entscheidungsunterstützung dafür bereitzustellen, wann und wo welche RRM Algorithmen und Technologien wie Intelligente Antennen eingesetzt werden sollen.

### *Abstract*

Radio Resource Management (RRM) algorithms play a main role in the efficient usage of the scarce radio resources in wireless networks. A class of distributed, self-adaptive RRM algorithms gain increasingly on importance due to their low complexity and signalization overhead. In this work, we apply a “pricing” and game theory based approach to the design of distributed, self-adaptive RRM algorithms. In our model a network provider maximizes its gain by setting “prices” for the radio resources according to the network’s state. Wireless users then choose those resources that maximize their utility functions using “prices” of the resources as parameters. Due to special form of our utility function we can “enforce” by “prices” some well known RRM algorithms as state dependent “rules”.

We investigate further what are the non-exploitable (“evolutionary” stable) “rules” (algorithms) in the networks without prices like ad-hoc networks. We show also, using game theory, arguments how cooperation among myopic users might also arise in such networks. Further, we assume that our “players” are “bounded rational” and have limited information available like local measurements. Our algorithms also do not necessary search for absolutely optimal solutions but for “good enough” i.e. “satisfactory” solutions. These properties enable practical applications of our algorithms with low computation and signalization overhead.

We estimated capacity gains and trade-offs of some RRM algorithms by system-level simulations in different environments (urban and rural), for different loads and services (speech and packet data), with or without smart antennas.

Both wireless equipment manufacturer and network providers can benefit from the results of this work. Manufacturers can use the results to design efficient, decentralized, adaptive RRM algorithms with low signalization overhead. Providers can also use the results of this work to “activate” optimal RRM algorithms for each network state simply by “price” (parameter) settings without the need for software changes as well as to optimally charge users, or to support decisions which and when RRM algorithms and technologies like smart antennas should be employed.

## *Acknowledgment*

*I would like to thank all the people who contributed to my personal and professional life during the curse of my PhD work. In particular, I would like to express deep gratitude to my first supervisor Prof. Richard Eier for his support during my work. His suggestions and comments improved the content and the clarity of this work significantly. I would also like to express my gratitude to my second supervisor Prof. Ernst Bonek for agreeing to act as my second supervisor and to Univ. Ass. Elmar Trojer for carefully reading my dissertation and providing useful comments.*

*My great appreciation goes to my friends and colleagues from Siemens AG Austria, PSE department, from whom I learned very much. I owe special thanks to Mr. Peter Slanina for sharing with me a part of his great experience in a fascinating filed of mobile communication. A very special thanks to Michael Bartl, Georg Diernhofer, Mario Konegger, Dr. Gerald Ostermayer, Toplica Pacic, Thomas Stadler and Juraj Weisz and other colleagues from PSE, who have developed with me our mobile simulator during last five years, for their friendly help and cooperation all this time. I am deeply grateful to Kenan Slavic and Plamen Bratanov for helping me to develop "filling" for real network planning problems.*

*Finally, I would like to express my deepest gratefulness to my lovely wife Merima, who has been accompanying me each and every day during the mission, for her love, persistence and support. I am also deeply grateful for love and support of my parents, father Mehmed, who has encouraged and motivated me all the time, and mother Azra, who has always selflessly taken a care of her son.*

*“For since the fabric of the universe is most perfect and the work of a most wise Creator, nothing at all takes place in the universe in which some rule of maximum or minimum does not appear."*

**Leonhard Euler**

*“Life is a constrained optimisation problem.”*

**Rich Frost**

# Contents

# 1 Motivation and Overview

*"Today, mathematics is returning to the 19th century, to concrete computations, after seventy years of very abstract mathematics. The latest fashion is 19th-century mathematics. Some of the best work in mathematical physics is based on constructive approach: Instead of proving abstract existence theorem, you* ***produce an algorithm that delivers the solution****. It is a powerful methodology.*"

**David H. Sharp, Interview in Science 2004**

## 1.1 Motivation

Nowadays we are witnesses to a **tremendous increase in the number of users and services in modern wireless networks**. This trend is expected to be continued and even to be extended in the future. In order to enable communication with a wireless user, two basic resources are needed: Bandwidth (channels) and power. These **resources are limited** and **expensive**: Only a relatively small part of the spectrum is dedicated to commercial wireless communication and “prices” for the frequencies within this spectrum have achieved enormous values at some auctions (for example UMTS license prices in UK and Germany). By the employment of denser infrastructure, the same channels can be re-used more often in the cellular network but system costs would increase. Furthermore, by managing scarce radio resources, important **social trade-offs** should be made, since allocating more resources to one user means leaving less resources (or decreasing the resource quality) to all other users in the system. It is **the task of Radio Resource Management (RRM) algorithms to enable efficient usage of wireless network resources with reasonable infrastructure costs and signaling complexity**. RRM algorithms make decisions like: To **which base station(s)** to connect the user (Handover), **which channel** should be allocated to a user (Channel Allocation) and with **which power** (Power control) user should transmit. Taking right decisions at the right time is crucial for system efficiency i.e. **maximizing the number of users with sufficient signal quality** (satisfied users) or **maximizing total system throughput (data rate)**. It is task of network providers to find a trade off between these two conflicting objectives: maximizing number of satisfied users on the one hand, and maximizing system throughput on the other hand.

On the other side wireless users faces a trade-off between maximizing own signal quality or data rate on the one hand, and minimizing power consumption on the other hand. Also, the users have conflicting objectives too. Furthermore, the users’ objectives are oft in conflict with provider objectives.

RRM decisions can be taken **centrally** by signaling all relevant information to one decision-maker or in a **decentralized** manner letting each base station or mobile station make decisions concerning the base station or the mobile station according to locally available information like their own measurements. **Taking into account size and complexity of a wireless network, as well as rapidly changing propagation conditions, it is clear that a centralized resource management would be very costly, if possible, at a reasonable price, at all.** A whole network for the signaling of the size of the network for the users' data or even larger would be needed to provide a centralized RRM.

RRM algorithms should also be **scalable and flexible to adapt to changes** in traffic and propagation conditions: For example, an increase in the number of users applying possible new services, changes in an environment like new buildings, changes in infrastructure like new cells, and employment of new technologies like smart antennas. It is **much cheaper to let flexible RRM algorithms adapt themselves to the changed conditions** than to do new network planning with possible employment of new base stations or make changes in base stations or mobile hardware and/or software each time something in traffic and propagation conditions changes.

Further, searching for a global optimum in a permanently changing wireless environment would require an enormous signalization and computation effort even for the networks of moderate size. That is why,**(sub-) optimal**, **distributed** RRM decisions are needed, based on **local available information and measurements**. These eventual sub-optimal solutions should provide the approximately same performance as a globally optimal solution with far less signalization and computation overhead.

The above reasons give rise to interest on **decentralized, self-adaptive RRM algorithms**, which enable **(sub-) optimal**, **cost efficient solutions** with relatively **low complexity and signaling overhead**. Development and improvements of such algorithms have been subject to extensive research in almost all modern wireless systems (GSM, UMTS, WLAN) and even will gain in importance in future systems. Fourth generation (4G) wireless systems are expected to be adaptive, self-organizing radio networks supporting a wide variety of new applications and services, where RRM algorithms should play a key role in providing higher spectrum efficiency and better service quality.

## 1.2 State of the Art and our Model

In the past a lot of decentralized, measurement based RRM algorithms were proposed and analyzed [16], [34], [45], [46], [109], [12], [68], [110]. These algorithms are mainly based on

**heuristics rules** derived from the goal to maintain or maximize signal quality (signal-to-interference ratio (SIR)), as long as possible for as many users as possible. Distributed and self-adaptive RRM algorithms like channel segregation [28], [1] and iterative algorithms like distributed power control [108], [109], [25] are of great practical importance.

These algorithms were designed rather heuristically, for example: users should increase their power if their SIR lies below certain threshold (power control) [25], [108], [109], take the channels with the lowest interference (channel allocation) [12], take the base station to which the channel gain is highest (handover) [74]. The algorithms also **do not make clear distinctions between users' and networks' objectives**. Furthermore, different algorithms are "optimal" under different conditions. To the best of our knowledge, a little research is done about **combining the different algorithms in order to optimize system performance under all conditions**.

In recent times it is more and more recognized **that radio resource management is, like each resource management, an economic problem** i.e. allocation of scarce resources among competing agents [80], [62]. Each decision for one user influences the other users in the network. For example, the more power a user transmits, the better the signal for the user is, but higher interference to other users is generated; the better channel (lower interference) is allocated to a user the less "good" channels are left to the other users that possibly arrives later etc. On the one hand, it would be desirable to let each user make its decisions alone in order to reduce signalization and computation overhead. Furthermore, searching for a global optimum in a centralized manner would be in practice almost a formidable task due to a large number of users and a fast changing environment. On the other hand, in order to prevent users from disturbing other users, **we need a mechanism to direct their decisions to socially desirable behavior**. These are similar problems of resource allocation among competing agents in a decentralized manner that we also encounter in a democratic society and in a market economy. In the economy there are already well-developed tools and methods to analyze such problems like Decision and Game theory [100], [70], [26], [51]. A modern economy is market based i.e. resource allocation is performed in a distributed manner: **Each economic agent makes decisions (choose resources), which maximize only its utility under its budget constraints. "Prices" serve as mediators between consumers and producers and as signals of relatively resource scarcity**. Agents take "prices" as parameters of their utility function. Similarly, in a society the role of government or social institutions is to "direct" the agents (citizens) to socially desirable behavior by punishments and rewards, which serve as "prices" for social decisions. In a similar way **a network**

**provider can optimize resource usage in a decentralized manner by setting “prices” for radio resources (cells, channels, power) and letting users choose resources which maximize their utility with “prices” as parameters**.

An “economic” approach for efficient resource allocation based on “pricing” and game theory has already been proposed for fixed networks like the Internet and ATM [63], [65], [17], [95], [101] and for mobile networks [80], [62], [79], [92], [91], [94], [102], [105]. But these works do not take into account the “traditional” RRM algorithms like minimum interference DCA or SIR-based PC and do not investigate state dependent decisions. The previous works also do not emphasize that in wireless networks, like in real world, the “players” are “**bounded rational**” [88] i.e. have limited information available and do not make optimal but **“satisfactory” decisions** according to the “**rules of thumb**”. **The real networks are (like a real world) too complex to apply “classical” assumption of the game theory that players are fully rational i.e. they take always optimal decisions which maximize there utility function.** In order to make the optimal decisions players should have a full information about what is the state of the environment, what the other players know, what the other players know that we know etc, which would require enormous signalization overhead to obtain all relevant information. Furthermore, in mobile networks each few milliseconds the relevant information like channel gain and interference changes due to fast fading. As stated by Nobel laureate Herbert Simon [89]: **“The decision maker has a choice between optimal decisions for an imaginary simplified world or decisions that are "good enough," that satisfice, for a world approximating the complex real one more closely”**. For example a manager set the price for a product according to the “rule of thumb”: production cost + some margin (estimated by experience). Time and effort required to optimize the prices to the last decimal point would require much more costs than the potential gain of the optimal price would be.

In our work **we wanted to build “a bridge” between “old” heuristic-based algorithms and a “new” economic approach.** We tried to combine the advantages of both approaches: The wealth of practical and simulation experience available for the “old”, heuristically algorithms and more flexibility and firm theoretical fundaments from economics for the new approach. To this purpose **we applied “an economic” model for decentralized optimization based on “pricing” and game theory and tried to obtain well-known RRM algorithms as special cases**. We assumed **“bounded rational” players, which make “satisfactory decisions”** (provides “good enough” performance under given load) using **local available measurements**.

We differentiated between users' utility function, which comprises users needs like data rate or signal quality, and network utility function, which comprises provider needs like maximization of the number of satisfied users or total data throughput. As a mediator between these two utilities serves the resource "prices" set by network (cells) and used as parameter of the user's utility functions. In comparison to previous works we used a **different model and a form of the users' and network's utility function** and a (possible) **fictive "prices"** (parameters) instead of real (monetary) prices. **Due to appropriate choice of the utility function, some of the well known heuristically algorithms can be obtained as special cases of our model.**

In our model the users took those resources (channels, cells and powers), which **maximize their expected utility functions, maintaining their service constraints** without explicitly taking into account the utility of the other users (myopic users). Overall **network performance could be then optimized in a distributed manner by the appropriate setting of "prices" by network (cells)**. Cells (network) set "prices" according to the state (load, interference) of the cells (network). The aim is to maximize the cell (network) utility like the number of satisfied users or total system throughput. Some of the well-known "heuristically" algorithms, which are "good enough" for a given state, can be "activated" just by appropriate parameter ("price") setting, without the need for implementing each of the algorithms separately in software or hardware.

**Optimal algorithms' parameters ("prices") depend on the state** (load, interference) of a cell or network. We used rather rough state classification according to the load and interference (low, medium and high states), because our simulation results showed that differences among algorithms are significant for only these few states. These states usually last in practice for a few hours (like "rush hour") and when the state changes relevant cells can broadcast the new "prices", defining the "optimal" algorithms for the new state. This is an advantage in comparison to a traditional approach where only one algorithm is used for all states. A provider can use simulation results or network statistics to find the "best“ algorithms ("prices") for each state in the network (cell). **We defined the users' utility function so that we can by state-dependent "price" setting "enforce" some heuristic rules** like: "Use minimum interference DCA, when load is low" or "send with power just enough to achieve the minimum signal quality but not better in the case of a high load".

Furthermore, **we investigate which of these "heuristic rules" would be a "rational" choice in the case without "price"-settings controller(s) like in ad-hoc networks or among networks with different providers**. We show using (evolutionary) game theory methods that under certain circumstances cooperation might also emerge among myopic users in such networks. We provided also examples of the design of efficient, stabile and robust RRM algorithms using the game theory results.

## 1.3 Relation to Game Theory

The aim of this work was not to develop new mathematical theories but to **provide a framework for RRM algorithms design and give some examples of the algorithms designed according to the framework**. That is why we do not follow a classical mathematical "Definition-Theorem-Proof" approach. We believe that **optimization of wireless networks with large number of users and stochastic, fast changing environment is too complex to be accessible in praxis for thorough mathematical analysis with reasonable signalization effort** ("bounded rationality" assumption). Instead, **we use the mathematical results, heuristically reasoning and simulations as guidelines to design practical and efficient ("good enough") RRM algorithms.**

Please note that Game theory **does not provide us with any algorithm ready to be used for RRM**. Instead, Game theory, like any other mathematical theory, provides concepts, definitions, theorems and proofs, which might serve as **inspiration** and a "firm mathematical background" for the art of the algorithms design in complex, distributed, competitive environments.

Furthermore, we make use of the modern economic and Game theory results, based on **"bounded rationality", "satisfactory solution" and use of "rules of thumb"** in an evolutionary context, which are also subject of an intensive research in the current economic theory [51], [90].

We hope that this work represents a step further (however small) on the way to establishing the role of Game theory in RRM research, like, for example, linear algebra plays in signal processing research.

## 1.4 Thesis Overview

The sequel of this work is organized as follows: In **chapter 2 we introduce RRM problems and give an overview of existing RRM algorithms.** We emphasize distributed, measurement based algorithms, which require a low signalization and computation overhead which make them very interesting for practical implementation.

In **chapter 3 we take "an economic view" on the RRM problems and set up our model based** on a maximization of users' and networks' utilities with "prices" as mediators between these utilities. In chapter 3 we also **analyze how cooperation might emerge among myopic users** in networks without centralized controllers (like ad-hoc networks) using the results of game theory.

**In chapter 4, we give examples of RRM algorithms based on our "pricing" and Game theory framework and present some simulation results** for the RRM algorithms for different loads and in different environments (urban, sub-urban and rural). We also investigated gains of technologies like smart antennas and how smart antennas influence choice of RRM algorithms.

Finally, we conclude in chapter 5 and propose topics for further research.

In short, chapter 2 and appendixes give an overview of the existing results. In chapter 3 our theoretical model and in chapter 4 practical algorithms and simulation results are presented. The text can be read in predefined or in any other order, for example, one can first read chapter 2, than chapter 4 and finally chapter 3, which would be approximately the order in which the text evolved.

# *2 Radio Resource Management Problems and Algorithms*

*"If I have seen farther than others, it is because I was standing on the shoulders of giants."*

***Isaac Newton***

In this chapter we describe radio resource management (RRM) problems and tasks of RRM algorithms. We also give examples of some well-known RRM algorithms for power control, channel allocation, handover and admission control. We emphasize decentralized, measurement based RRM algorithms since they are of high practical importance because they have low signaling overhead, flexibility and implementation simplicity. In chapter 3 we show how these algorithms can be established as special cases of a more general "pricing" and game theory framework.

## 2.1 Tasks of RRM Algorithms

We consider the following scenario (see Figure 2-1): A mobile station arrives in an area covered by several base stations (access ports). The area where the signal level from a base station is above a certain level is called a cell. In the following we use the terms base station and cell as synonyms. The purpose of the base stations is to provide to mobile stations an access to the fixed network over a radio link. In order to establish communication over a radio link certain resources should be provided - at least one base station, one channel and a certain amount of power should be allocated to the mobile.

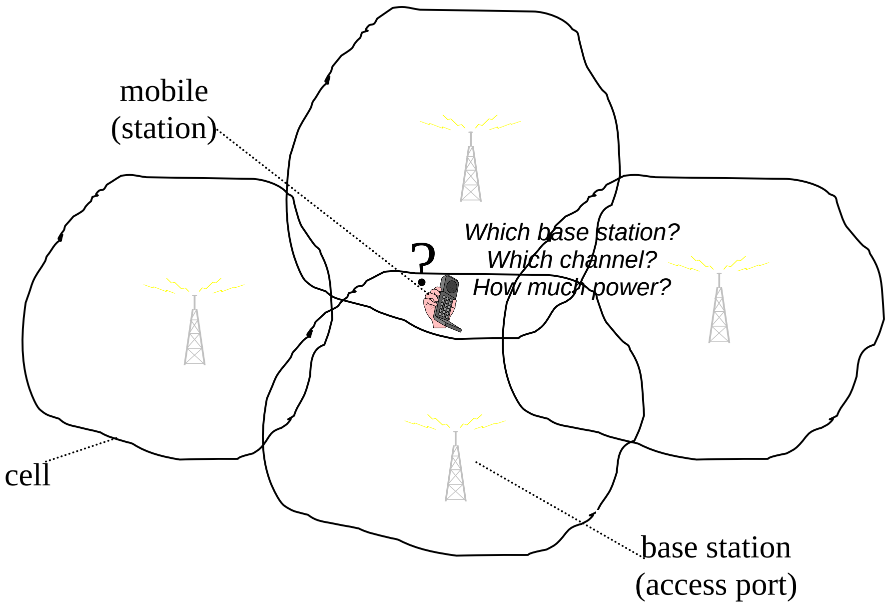



**Figure 2-1: Resource management problems in cellular networks**

The tasks of RRM algorithms are to decide which base station(s), channel(s) and how much power should be allocated to the mobile in each instant of time. These tasks are divided among several RRM algorithms in order to reduce implementation complexity. The RRM algorithms are:

- **Handover (HO):** HO decides to which base station (cell) mobile is connected and initiates change of a base station (intercell HO) or a channel within the same cell (intracell HO).
- **Admission Control (AC):** AC controls an access of mobiles to cells. AC can deny access to a cell if, for example, the load of the cell is too high.
- **Channel Allocation (CA):** CA allocates channel(s) for communication between users and base stations.
- **Power Control (PC):** PC allocates power to the users for communication with base stations.
- **Scheduler (SH):** Scheduler decides which users and when (among the users allocated to the base stations) should transmit and how much of their data.
- **Load Control (LC):** LC regulates data rates (for example, by selection of source and channel coder and/or modulation schema) of the users according to the load in the network (cell).
- **Congestion (CC):** CC reduces the data rate and/or drops users in the case of too high load (congestion) in the network (cell).

In the following we describe a typical scenario of activation of different RRM algorithms during a call. At the call beginning, Handover (HO) algorithm decides to which base station(s) the mobile should be connected. Only the base stations, which allow access to the cell are considered for HO: For example, if the base station is overloaded the access to the cell of the base station may be denied for all new mobiles. It is the task of the **Admission Control (AC)** algorithm of a cell to accept or refuse new users in the cell. Further, the mobile users should be provided with a sufficient signal quality from the base station i.e. the users' position has to be within the coverage area of the base station (cell). From all base stations, which provide sufficient signal level and allow access to their cells, HO might select the one with the highest signal level (highest gain, lowest path-loss see Appendix A). This would be an example of a so-called path-loss based handover. Other handover algorithms are also possible: For example, handovers which take into account load of the cell, expected interference in the cell etc.

When a base station is selected, **Channel Allocation (CA)** algorithm decides which channel is selected for the communication between the base station and the mobile. A channel denotes a bandwidth in time or space sufficient for communication between the base station and mobile. Channel definition is system dependent: For example, in a Time Division Multiple Access (TDMA) systems channels are time slots. In a Frequency Division Multiple Access (FDMA) systems channel is defined by frequency of the carrier. In a FDMA/TDMA systems like GSM channel is a (frequency, timeslot)-pair. In Code Division Multiple Access (CDMA) systems the channels are chip sequences called codes. In Space Division Multiple Access (SDMA) systems a wireless channel is defined by a space dimension where the mobile is located. In “hybrid” systems like UMTS TDD mode (a FDMA/TDMA/CDMA system) channel is defined by a (frequency, timeslot, code)-triple. UMTS TDD modus has 15 timeslots, 16 codes and usually only one frequency (“high cheap rate” TDD). Bandwidth around the carrier is usually assumed implicitly: For example, 200 kHz in GSM, 5 MHz in UMTS. In case of frequency hopping, a channel is represented by a hopping sequence – a set of (frequencies, time slot)-pairs. Note that FDMA, TDMA and CDMA use orthogonal channels: Users using different channels (almost) do not disturb each other i.e. interference generated in the channels other than the one used for communication is relatively low.

It is important to distinguish between two possible communication directions from base station to mobile, denoted as downlink (DL) and from mobile to base station, denoted as uplink (UL). In general, in both directions different channels are allocated. For example, in GSM or UMTS FDD, UL and DL use different frequencies from the symmetrical bandwidth. In UMTS TDD different timeslots (but same frequency) are used in UL and DL. It is also possible in TDD to change the number of timeslots devoted to UL and DL (in each cell) by using a variable switching point.

A simple channel allocation algorithm would be to select a channel randomly from the set of free channels in the cell (Random CA). Other CA algorithms are also possible such as CA which takes into account interference from other users on channels and/or channel gain from the users, service type of the users etc.

After a base station and a channel are selected, power should be chosen with which the user should communicate with the base station on the channel. Power Control (PC) algorithms should select the amount of power used for the communication in such a manner that enough signal quality is provided to the user itself on the one hand, and on the other hand, the other users in the system are disturbed as little as possible. The more power PC allocates to the user, the better the signal quality (higher data rate) of the user, but the higher interference is generated to the other users (on the same channel) in the system. A good PC algorithm should find a trade-off between satisfying the needs of the user and generating interference to the other users. A simple power control would be to always use constant (maximal allowed) power. More sophisticated PC algorithms also take into account channel gain of the users, interference on the channel, required signal quality (data rate) etc.

In contrast to real-time (RT) services (like speech), non-real-time (NRT) services (like packet data) have relatively uncritical delay requirements. That is why a channel and power do not have to be allocated immediately to a NRT user. The data of NRT users could be put in the base station queue and the Scheduler (SH) algorithms decides when the data are scheduled for the transmission. A simple scheduler algorithm would be First-In-First-Out (FIFO) scheduler, which schedules data according to the users’ arrival time in the queue. The more sophisticated scheduler might also take into account the channel quality of the users, data rate achieved by the users so far, users’ data rate requirements, etc. Note that a scheduler could be regarded as a special case of power control, where power is set to zero for the not scheduled users.

During the call, signal quality of the mobile users could degrade due to users’ movement, for example, interference on the channel might become high or users might leave the original cell. It would then be the task of Handover (HO) to initiate and perform channel change (intracell handover) or cell and channel change (intercell handover). The load of the network can also change and the users might be enforced by a Load Control (LC) algorithm to increase/decrease their data rates, for example, by switching to another source or channel codec or modulation schemas. If the cell or network becomes overloaded Congestion Control (CC) algorithm must then select some users, which should decrease their data rate or, in the case of a severe overload, drop some users.

The RRM algorithms work on a different time scale. Whereas HO decisions should be done each few seconds (when mobiles cross cell borders), power control is done each few milliseconds, in order to adapt quickly to changes in propagation conditions (see Appendix A). Further, RRM algorithms run in mobile and/or base stations (usually in software). Sometimes, some part of the algorithms is executed in mobile stations and some in base stations. For example, in DECT, mobile stations select channels for allocation and in GSM or UMTS the base stations (or other network element like RNC) select the channels using the measurements of the mobile (mobile assisted HO). In Table 2-1, an overview of RRM algorithms, their decisions and time and “logical” (MS or BS specific) place of their execution is given.

**Table 2-1**: **RRM Algorithms, their decisions, "logical place" of execution and time scale**

| Algorithms | Decision | Place | Interval |
|---|---|---|---|
| Handover (HO) | Change base station(s)/channel(s) | MS and BS | seconds |
| Admission Control (AC) | Admit users in a cell/system | BS | seconds |
| Channel Allocation (CA) | Which channel | MS and BS | seconds |
| Power Control (PC) | Allocate power | MS (UL), BS (DL) | microseconds – milliseconds |
| Scheduler (SH) | Which users transmit and how much data | BS | milliseconds - seconds |
| Load Control (LC) | Change data rata (codec, modulation scheme) | MS and BS | milliseconds - seconds |
| Congestion Control (CC) | Drop users / change data rata | BS | seconds |

Distributive measurement based RRM algorithms which are performed in the affected base stations or mobiles according to locally available measurements like channel gain and interference are of particular importance. In this way, decisions can be taken relatively quickly at the place where they are required. This is especially important for “fast” algorithms like power control, which has, for example, a period of several hundreds of μs in UMTS (“fast power control”) [21]. Furthermore, an excessive signalization is avoided for communication with a centralized RRM controller, which could be an almost formidable task in a complex and fast changing mobile environment, where information become obsolete very quickly. In the following, we describe some existing solutions to the distributive measured based RRM algorithms.

## 2.2 Power Control

### 2.2.1 System Model

The aim of power control is to allocate power to users in order to provide sufficient signal quality or data rate to the users but, at the same time, reduce interference to the other users in the system as much as possible. In order to determine a power control algorithm, we observe received signal quality and interference at a mobile at a certain moment of time after cell and channel allocation is already performed for the mobile (see Figure 2-2). We assume a wireless system with orthogonal channels with N transmitter in the system transmitting the power to N receiver at the same channel. We use the channel model as described in Appendix A.

Useful (carrier) signal power $C_i$ at the mobile i is defined as a product of the link gain $g_{ii}$ between mobile i and its base station i (see equation (A.2)), which sends a carrier signal to the mobile with the power $p_i$:

$$C_i = g_{ii} p_i \quad i = 1 \ldots N \tag{2.1}$$

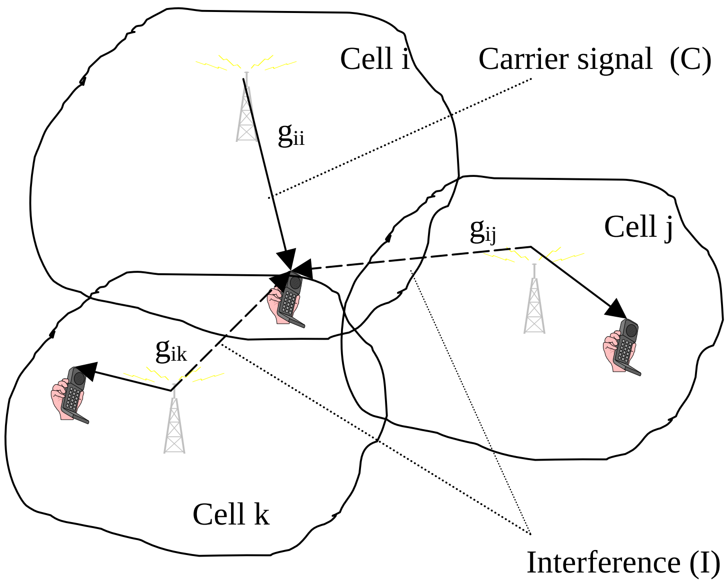



**Figure 2-2: Carrier signal and interference in a mobile network**

We denote with $g_{ij}$ the link gain between user *i* and the interfering base station *j*, which sends at the same channel with the power $p_j$. Interference $I_i'$ at the mobile *i* is the sum of the received powers of all other N-1 base stations (except his own base station i), which transmit on the same channel plus the noise power (denoted with $\eta_i$):

$$I_i^{'} = \sum_{j=1, j\neq i}^{N} g_{ij} p_j + \eta_i \quad i = 1...N \tag{2.2}$$

Modern wireless systems are interference limited systems since interference from other users is much higher than noise power [110]. Note that channel gains and interference are random variables and vary over time. Channel gain changes according to distance, slow fading and fast fading (see Appendix A). Interference changes with varying channel gain between receiver and interferers as well as with change in number of interferers due to arrivals and departures of (co-channel) users in the systems.

As a measure of signal quality, **Carrier-to-Interference Ratio (CIR)** is used. CIR at mobile i can be expressed then as [77] ratio between useful signal $C_i$ and interference $I_i^{'}$ (see equations (2.1) and (2.2)):

$$CIR_i = \frac{C_i}{I_i^{'}} = \frac{g_{ii} p_i}{\sum_{j=1, j\neq i}^{N} g_{ij} p_j + \eta_i} \quad i = 1...N \tag{2.3}$$

CIR is an important measure of the signal quality since possible data rate is an increasing function of CIR and consequently **Bit Error Rate** (**BER**) is a decreasing function of CIR for all modulation schemes. Data rate and BER are an ultimate measure of users' satisfaction. For example, NRT data users are regarded as unsatisfied if their data rates during the session were lower than 10% of their nominal data rate, and speech users are regarded as unsatisfied if their BER were above $10^{-3}$ for more than 5% of call duration [19].

PC algorithms can influence data rates and BER by controlling the power, which in turn, determines CIR and data rate or BER for a given coding and modulation scheme. In the following, we describe some iterative PC algorithms, which are of great practical importance.

### 2.2.2 Iterative PC Algorithm

In this section, we describe some iterative power control algorithms according to [21] and [110]. They belong to a class of measurement based PC algorithms, since they update users' powers according to the measured values of channel gain and interference. In order to set the transmitter power $p(n+1)$ at the next time period $n+1$, the PC algorithms need information about channel gain $g(n)$ and possibly interference $I^{'}(n)$ obtained from the measurements at the receiver and signalized back to the transmitter at the nearest previous time period $n$. Also note that transmitter power is limited in practice and cannot exceed some predefined level $P_{max}$.

Below we give examples of some PC algorithms:

**Fixed (Maximal Power) PC.** Simplest PC algorithms always set the power p to a fixed value, usually maximum power $P_{max}$ i.e. $p(n) = P_{max}$ for each time instant $n$.

The advantage of the Fixed PC algorithms is its simplicity – no changes of power and no feedback from the receiver about channel gain or interference is needed.

The disadvantages of the Fixed PC is high power consumption and high interference generating to the other users in the system. Even if the channel gain is high and interference at the receiver is low, the fixed PC always transmits with maximum power, although sufficient signal quality (CIR) at the receiver could be achieved with (far) less power.

Fixed PC was used in systems where interference from/to other users was not an issue due to great distance between the cells using the same channels (co-channel distance) and low users' density as in the first wireless systems and in the early phase of cellular systems [61]. Modern wireless systems like GSM/GPRS and UMTS are **interference limited** due to the high number of users and low co-channel distance used to provide better spectrum usage. That is why other PC algorithms like C-based and CIR-based PC are more often used in modern wireless systems.

**C-based PC.** The purpose of C-based PC is to keep for each mobile the received signal strength $C$ at a fixed predefined value $C_{thr}$ i.e.:

$$C = C_{thr} \overset{(2.1)}{\Rightarrow} p(n+1) = \frac{C_{thr}}{g(n)} \tag{2.4}$$

As can be seen from equation (2.4), some measurement and signaling overhead is needed since information of channel gain should be measured by the receiver and sent back to the transmitter in order to set transmit power properly.

C-based PC brings in general reduction of power consumption and generating interference in comparison with Fixed PC because less than $P_{max}$ is transmitted whenever $C_{thr}/g < P_{max}$.

C-based PC can be useful if all users should have the same power at the receiver, for example, for easy detection in UL of some CDMA systems like IS95 [29], [55].

Since C-based PC does not take into account interference at the receiver, poor signal quality (CIR) and thus high error probability can occur in the case of high interference at the receiver and conversely, in the case of low interference at the receiver, C-based PC transmits with unnecessary high power. That is why CIR-based PC is used more and more in practice (all 3G systems like UMTS FDD and TDD, cdma 2000 etc.) [21].

**CIR-based PC.** The purpose of CIR-based PC is to keep received signal quality CIR of each mobile at a fixed predefined value $CIR_{thr}$:

$$CIR = CIR_{thr} \overset{(2.3)}{\Rightarrow} p(n+1) = \frac{CIR_{thr} I^{'}(n)}{g(n)} = \frac{CIR_{thr}\, p(n)}{CIR(n)} \tag{2.5}$$

CIR-based PC tries to maintain the required signal quality of the users, and, at the same time, minimize power consumption and interference to the other users in the system. CIR based PC is used, for example, in UMTS [21], where each PC period the power is increased (decreased) for a step (1-2 dB) if CIR was below (above) $CIR_{thr}$ in the previous period.

Note that for CIR-based PC information about channel gain and interference or actual CIR should be measured at the receiver and sent back to the transmitter. By evaluating different PC algorithms, power costs like energy consumption of the transmitters (especially important for mobiles due to battery supply) and social costs due to interference to other users in the system should be compared with measurement and signaling costs needed to perform the PC algorithm. CIR-based PC needs more measurements than constant power or C-based PC, but consumes less power and generates less interference than constant power or C-based PC.

An important question is whether and when it is possible to achieve for all users, at the same channel, sufficient signal quality $CIR_{thr}$ and if an iterative PC algorithm like CIR-based PC given in (2.5) converges to a stabile (equilibrium) powers for all users on the same channel**.** We give the answers to these questions in the next subsection.

### 2.2.3 Convergence of CIR-based PC Algorithms

As can be seen from equation (2.3), CIR of one user i can be temporarily improved by increasing the power to the user $p_i$. In that case all other users would suffer from increased interference e.g. the denominator of equation (2.3) will also be increased for the users other than user i. The other users would then have to increase their powers and so on, until all users have CIR equal to their predefined threshold or at least one user has power equal to the maximum power $P_{max}$ but CIR lower than required CIR ($CIR_{thr}$). The question that naturally arises is: Under what conditions iterative PC according to (2.3) converges to some stable state (equilibrium) where all users achieve their required $CIR_{thr}$ and do not have to change their powers? The answers to these questions are given mostly in [108], [109] and [25] (for a survey see also [110], chapter 6). In the following, we summarize the convergence results for iterative CIR-based PC.

At first, we try to represent iterative CIR-based PC in a matrix form. To this aim, we start from equation (2.5) and rearrange it as follows:

$$p_i(n+1) \;=\frac{CIR_{thr} I_i^{'}(n)}{g_{ii}} \overset{(2.2)}{\Rightarrow}$$
$$p_i(n+1) = CIR_{thr} \sum_{j=1, j\neq i}^{N} \frac{g_{ij}}{g_{ii}} p_j(n) + \frac{\eta_i}{g_{ii}} \quad i = 1...N \tag{2.6}$$

We assume that during the convergence time of PC only powers of the users (and thus interference) change and the channel gains do not change. Denoting with $\boldsymbol{\eta}$ the vector with the element $\eta_i/g_{ii}$ at the i-th position and with **H** "normalized link gain matrix" or channel matrix (because it represents cross-correlations of users on the same channel) such that:

$$h_{ij} = \begin{cases} CIR_{thr} \dfrac{g_{ij}}{g_{ii}} & if\ i \neq j \\ 0 & if\ i = j \end{cases} \tag{2.7}$$

Using the matrix **H** with elements defined in (2.7) we obtain the following representation from (2.6) matrix form:

$$\mathbf{p}(\mathrm{n}+1) = \mathbf{H}\mathbf{p}(\mathrm{n}) + \boldsymbol{\eta} \tag{2.8}$$

The question is when the iterative PC according to (2.8) converges i.e. if there is a limit $\mathbf{p}^*$ - an equilibrium vector of users' powers, such that $\mathbf{p}$(n) ->$\mathbf{p}^*$ as n->∞ and $\mathbf{p}^*$satisfies the following equation:

$$\mathbf{p}^* = \mathbf{H}\mathbf{p}^* + \boldsymbol{\eta} \tag{2.9}$$

In order to estimate the existence of the equilibrium power vector $\mathbf{p}^*$ we denote with ε(n) an error vector at the iterations-step n i.e. ε(n) is the difference between vectors **p**(n) and $\mathbf{p}^*$: ε(n) = **p**(n) **-** $\mathbf{p}^*$**.** From (2.8) and (2.9) we obtain the following equation for error vector **ε** as a function of the iterations-step n and initial error vector **ε** (0):

$$\begin{aligned} \boldsymbol{\varepsilon}(\mathrm{n}) &= \mathbf{p}(\mathrm{n}) - \mathbf{p}^* = \mathbf{H}\left(\mathbf{p}(\mathrm{n}-1) - \mathbf{p}^*\right) = \mathbf{H}^2\left(\mathbf{p}(\mathrm{n}-2) - \mathbf{p}^*\right) \\ &= \mathbf{H}^{\mathbf{n}}\left(\mathbf{p}(0) - \mathbf{p}^*\right) = \mathbf{H}^{\mathbf{n}}\boldsymbol{\varepsilon}(0) \end{aligned} \tag{2.10}$$

From (2.10) can be seen that the error vector converges to **0** (vector with all entries 0) for any initial vector of users' powers **p**(0)**,** if and only if the matrix $\mathbf{H}^n$ converges to matrix with all entries zero (zero matrix **O**).

In order to estimate convergence of the matrix $\mathbf{H}^n$ when *n->∞* we make use of Jordan normal form (eigenvalue decomposition) of the matrix **H**. According to [98] every square matrix **H** can be represented in following (**Jordan normal or canonical**) form: $\mathbf{H} = \mathbf{M}\,\Lambda\,\mathbf{M}^{-1}$, where

$\Lambda$ is upper-triangular matrix with eigenvalues of the matrix **H** at the main diagonal. If **H** has only distinct eigenvalues, $\Lambda$ is a diagonal matrix with eigenvalues of **H** at the main diagonal.

Since $\Lambda^n$ -> **O** as n -> $\infty$ if and only if the largest eigenvalue of the matrix **H** (called spectral radius of **H** and denoted $\rho(\mathbf{H})$) is less than one [57], $\mathbf{H}^n = \mathbf{M}\,\Lambda^n\mathbf{M}^{-1}$ -> **O** as n -> $\infty$ if and only if the largest eigenvalue of the matrix **H** is less than one. Consequently, from (2.10) follows that error vector converges to **0** vector if and only if the largest eigenvalue of the matrix **H**, is less than one i.e. if and only if:

$$\rho(\mathbf{H}) < 1 \tag{2.11}$$

If (2.11) is satisfied **iterative PC (2.8) is contraction mapping** [31] in space of users powers and there is a limit $\mathbf{p}^*$ (fixed point or an equilibrium vector of users' powers), such that **p**(n) ->$\mathbf{p}^*$ as n->$\infty$ and $\mathbf{p}^*$satisfies (2.9). This is also true if the users adapt their power asynchronously [31].

From (2.9) follows than an equilibrium vector of users' powers $\mathbf{p}^*$ satisfies the following equation (where **I** is identity matrix):

$$\mathbf{p}^* = \mathbf{H}\mathbf{p}^* + \boldsymbol{\eta} \Leftrightarrow (\mathbf{I}-\mathbf{H})\mathbf{p}^* = \boldsymbol{\eta} \Leftrightarrow \mathbf{p}^* = (\mathbf{I}-\mathbf{H})^{-1}\boldsymbol{\eta} \tag{2.12}$$

Since **H** is a nonnegative matrix (all elements are greater or equal to 0) and $\rho(\mathbf{H}) < 1$, it follows that $(\mathbf{I}\text{-}\mathbf{H})^{-1}$ is nonnegative too [98]. Note that according to Perron-Frobenius theorem [98] the largest eigenvalue of nonnegative matrixes (like **H**) is positive and real. In that case $\mathbf{p}^*$ has all positive components (since $\eta$ is also positive). This means that all users would achieve their CIR target if their powers are unlimited. Since in practice users have limited maximal power, all users can achieve their CIR requirements only when the algorithms converges to the power vector $\mathbf{p}^*$ which satisfy condition $p_i^* \leq P_i^{max}$, i=1..N, where $p_i^*$ denotes the equilibrium power of the user i and $P_i^{max}$ is the maximal possible transmit power of the user i.

An important issue is also the **convergence speed** of CIR-based PC i.e. how many iterative steps does a CIR-based PC need to achieve equilibrium powers. The faster the iterative PC (2.8) converges, the faster a stable state in the system arises. Furthermore, the faster convergence of PC, the less signalization overhead (for signaling channel gains and interference) is needed and the lower probability that the matrix **H** changes due to fading. The largest eigenvalue of **H** has again the crucial role in convergence and convergence speed of CIR-based PC: According to equation (2.10), the lower $\rho(\mathbf{H})$ the faster error vector $\boldsymbol{\varepsilon}$ converges to **0**.

Summarizing [110]: **If the largest eigenvalue of H,** ρ(**H**), **is lower than one, iterative CIR-based PC** (2.8)) **converges, and the lower ρ(H) the higher the convergence speed of the PC**.

#### *2.2.4 Outer loop PC*

Since an ultimate measure of the users satisfaction is Bit Error Rate (BER) and BER is a function of CIR, the required signal quality $CIR_{thr}$ for the CIR-based PC (see (2.5)) should be changed during the time to adapt to changes in an environment and/or mobile speed. This task is done by the outer loop PC [21]. The task of the Outer loop PC is to keep $CIR_{thr}$ at or above the level required to keep the BER below some service specific limit ($BER_{thr}$, say $10^{-3}$ for speech). The curve *BER = f(CIR)* depends on the modulation and coding used in the system as well as on the mobile speed and environment (urban, rural). Outer loop tracks user BER (or some other soft-information like raw BER after the decoder, decoder soft-decisions values, actual CIR etc.) and increases the target CIR ($CIR_{thr}$) for CIR-based PC (2.5) if the BER is above a certain threshold or otherwise decreases $CIR_{thr}$ (see Figure 2-3).

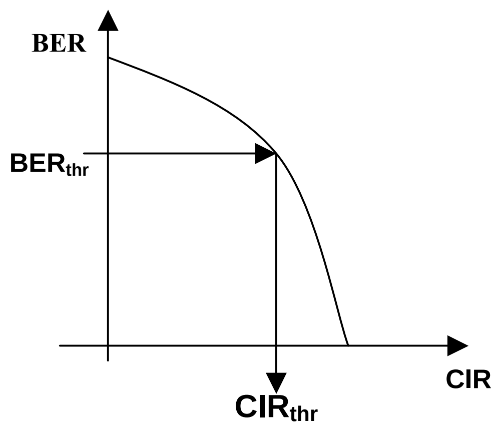



**Figure 2-3: Determination of $CIR_{thr}$ by Outer loop PC**

The outer loop PC is an example of an adaptive, decentralized, measurement based RRM algorithm, which changes the parameter ($CIR_{thr}$) for the other RRM algorithms (inner loop PC) according to the actual state of the user. In this case the state of a user is defined according to the actual user's *BER = f(CIR)* curve.

## 2.3 Water Filling PC

It is interesting that CIR-based PC although widely used in practice (UMTS, see [21]) is in general not optimal in the sense of the maximizing data rate from the information theoretical point. That is why we also describe water-filling PC, which is the optimal PC in terms of information theory i.e. water-filling PC is the power control strategy that maximizes data rate under power constraints.

In the following, we describe single user and multi-users case separately. In both of them water-filling PC is optimal or near optimal PC strategy under power constraints.

### 2.3.1 *Single User Case*

We have the following PC optimization problem in the single user case: Maximize expected (over distribution of channel gain $g$ and interference $I^{'}$) capacity:

$$C = E\left\{\log\left[1+\frac{p(g,I^{'})g}{I^{'}}\right]\right\} \tag{2.13}$$

Under average power constraint:

$$E\left\{p(g,I^{'})\right\}\le P_{\max} \tag{2.14}$$

Where $p(g,\ I^{'})$ is the power of the user when the channel gain is $g$ and interference $I^{'}$. In (2.13), Shannon's formula for channel capacity is used with $I^{'}$ denoting the total noise power experienced by the user [15].

Maximum of the expected capacity (2.13) under constraint (2.14) over a relatively long time period is called **ergodic capacity**. Some important assumption must be met in order to derive optimal power allocation strategy in the case of ergodic capacity [8], [71]:

- Both receiver and transmitter **have to know the channel state perfectly** (perfect CSI). Input symbols are selected in order to maximize mutual information between transmitted and received symbols conditioned on the channel gain $g$. The effect of imperfect channel information is hard to analyze. Extreme sensitivity of the well-known Gaussian coding scheme to the channel state information is shown in [54].

- In order that expectation of the channel gain $g$ can be translated to a time average, the **codewords have to be long enough** to capture the ergodicity of the channel fading process.

Taking into account the above assumptions, the maximum of (2.13) under constraint (2.14) can be obtained by using Lagrange multiplier technique: Multiplying the equation (2.14) with a constant $\lambda$ (Lagrange multiplier) and subtracting it from the equation (2.13) we obtain the following equation:

$$E\left\{\log\left[1+\frac{p(g,I^{'})g}{I^{'}}\right]\right\} \quad - \quad \lambda\left[E\left\{p(g,I^{'})\right\}-P_{\max}\right] \tag{2.15}$$

The maximum of (2.13) under constraint (2.14) is obtained setting the first derivate (by *p*) of (2.15) to zero (since logarithms is a concave function) i.e. the power should be set according to the following equation:

$$p(g, I') = \left( \frac{1}{\lambda} - \frac{I'}{g} \right)^{+} \tag{2.16}$$

We assumed that interference $I'$ does not depend on power *p*, for example in the case when there are a lot of users on the same channel influence of each single users on total interference is negligible. Lagrange multiplier λ is chosen in such a manner that equation (2.14) is satisfied, when power is set according to the equation (2.16). This strategy is called (time) "**water-filling**" [30] (see Figure 2-4) and is similar to water-filling over frequencies as originally described by Shannon in [85].

Water-filling PC increases power in the case of a "good" channel i.e. if the ratio between interference and channel gain ($I'/g$) is relatively low and decreases power when channel is "bad" ($I'/g$ high). This is just the opposite to what CIR-based PC does, which in order to maintain signal quality at a certain level (CIR-threshold)) increases power when channel is "bad" and decreases power when the channel is good (see Figure 2-4).

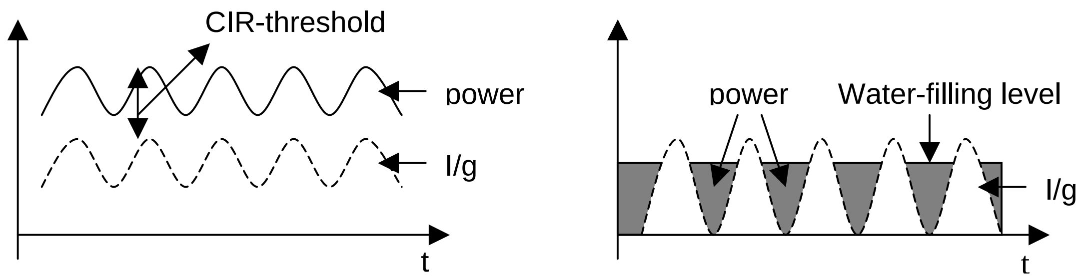



**Figure 2-4: Comparison of CIR-based PC (lefts) Water-filling PC (rights)**

Water-filling PC outperforms constant power PC in the case of low channel gain, but does not show much improvement over constant PC [8], [30].

### 2.3.2 Multiuser Case

In the multiuser case we have the following optimization problem: Maximize expected sum capacity (over joint-channel gains and interference distribution):

$$E\left\{\sum_{i=1}^{N}\mu_i \log\left(1+\frac{p_i(g_i,I_i^{'})g_i}{I_i^{'}}\right)\right\} \tag{2.17}$$

where $\mu_i$ is the priority, $p_i$ transmit power, $I_i^{'}$ interference variance and $g_i$ channel gain of user $i$. Under (average) power constraints (2.14) for each user alone and, in the broadcast channel case (DL in cellular networks), we also have constraints on total sum-power of the base station ($P_{max}$ - maximal broadcast power in DL):

$$E\left\{\sum_{i=1}^{N} p_i(g_i,I_i^{'})\right\} \leq P_{max} \tag{2.18}$$

It can been shown using Lagrange multiplier technique like in the single user case that also in the multiuser case, when Channel State Information (CSI) is available for both transmitter and receiver, (sub-) optimal power allocation over users for both multi-access [50], [97] and broadcast [58] channel also has a water-filling interpretation:

$$p_i(g_i,I_i^{'}) = \left(\frac{\mu_i}{\lambda} - \frac{I_i^{'}}{g_i}\right)^{+} \tag{2.19}$$

We assume that users use orthogonal channels like orthogonal codes as in CDMA in downlink $1/\lambda$ is the water-filling level determined in the manner that power constraints are satisfied. **If all users have the same priority, power is allocated to the user with highest channel gain (and lowest interference)** (see page 17 in [58]). In contrast to the single user case, in the multi-users case improvement of optimal power control over constant power control is significant [8], [50] because there is, with high probability, at least one user with a very good channel.

In practice following problems with water-filling PC and the ergodic capacity might arise:

- Both transmitter and receiver has to **know the channel state** (channel gain and interference) perfectly. This would cause high signalization overhead in fast changing fading environments.
- Sometimes users **might not even need maximal data rates** but just certain satisfactory data rates according to services needs and amount of data actualy available.

- One has to use **relatively large code length to capture the ergodicity of the channel fading process**. In practice, the code lengths and interleaving periods often have to be short due to the **delay requirements** of real-time services like speech. Such services require that certain **minimal data rate is met with a low probability of outage**.

To account to such problems the notion of **outage capacity** [36], [59] and some **sub-optimal** power allocation methods are used.

We say that outage has occurred [71] if in a given time segment of length $T$ the channel fading is such that the minimal required data rate $R_{min}$ cannot be supported because mutual information between input and output conditioned on channel gain is lower than $R_{min}$ i.e.:

$$\frac{1}{T} I(X^T; Y^T \mid g^T) < R_{\min} \tag{2.20}$$

Usually, quality of service requirements defines an outage probability $P_{out}$ i.e.:

$$P_{out} = \min \left\{ P_r \left[ \frac{1}{T} I(X^T; Y^T \mid g^T) < R_{\min} \right] \right\} \tag{2.21}$$

where $P_r(X)$ defines the probability that $X$ occurs. We can also say that the outage capacity at outage probability $P_{out}$ is $R_{min}$.

For the services with stringent delay constraints like speech and other real-time services, outage capacity plays an important role. It was shown in [36], [60] and [71] that optimal power allocation for services with stringent delay limitation in the case of block fading depends on the channel gain. A simple sub-optimal power allocation strategy is either threshold or constant (maximal) power strategy for both single and multi-users case and depends on channel gain:

*Single User Case*

**Low channel gain - Threshold policy** [60]: All power is allocated to a given block if the gain of that block exceeds the threshold. The threshold decreases as the delay deadline approaches. The threshold is zero for the last block.

**High channel gain - Constant power transmission** [30], [60]: In the case of high channel gains, user should transmit with maximal allowable power. This explains why optimal PC has diminishing improvement over fixed PC in the case of high CIR– in both cases user transmits with maximal power.

*Multiuser Case (Broadcast Channel)*.

**Low channel gain - Threshold policy** [8], [60]: All power is allocated to the user for which the product of priority and channel gain of that block is maximal and exceeds the threshold. The threshold depends on the remaining number of blocks and power. Also optimal PC policy is also a threshold policy as in the single user case.

**High channel gain - Constant power transmission** [60]. The user with highest priority transmits all the time with constant (maximal) power.

## 2.4 Scheduling

The task of the scheduling algorithms is to decide which users (data) from the queue are selected for the transmission and for how long (see [48], [49] for scheduling mechanism in general and [38] for packet data scheduling in UMTS). In Figure 2-5 a model of the scheduler and its queue is represented.

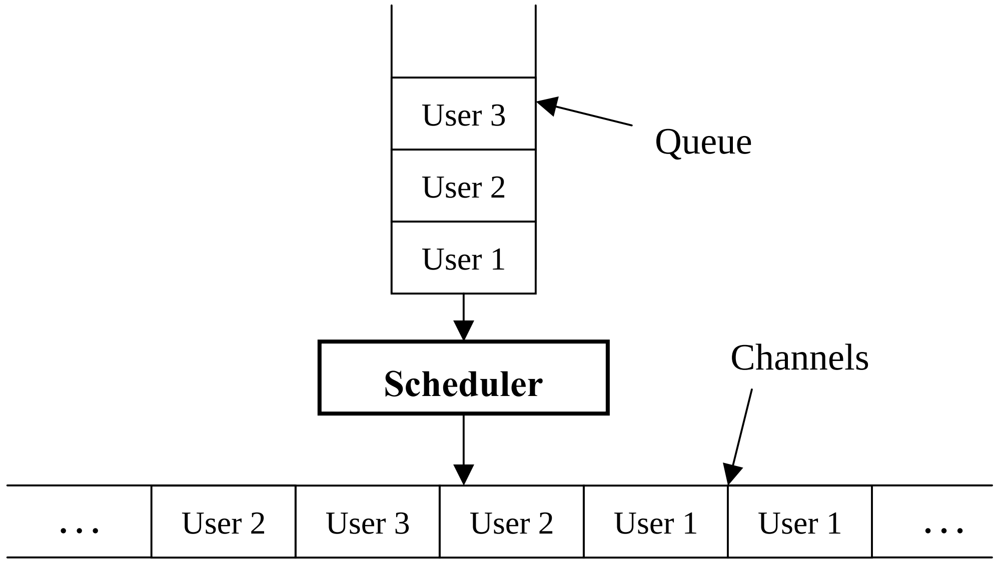



**Figure 2-5: A Model of the Scheduler**

Some possible scheduling mechanisms are [38]:

**“Fair” scheduling:** This strategy tries to allocate in average the same data rates to all users in the cell. The users who obtained the lowest data rates so far are scheduled first

**“Greedy” scheduling:** The users with better channel conditions i.e. with higher channel gains (for example, users near to BS) and lower interferences are scheduled more often in order to maximize total data rate in the cell.

**First-In-First-Out ("FIFO") scheduling:** The users are scheduled to the transmission according to their arrival times, the users who arrived first in the queue are also scheduled first.

**"Priority-based" scheduling:** The users with the highest priority are scheduled first. The priority can be set according to the users' service type i.e. the more "important" services could get higher priorities.

Scheduling can be also regarded as a special case of Power Control, for example, only the users who obtain power greater than 0 are actually scheduled for transmissions.

## 2.5 Channel Allocation Algorithms

### *2.5.1 Cell Concept and Problem of Channel Allocation*

The first mobile systems [61] usually used an antenna radiating with high power in order to cover a relatively large area. The users were differentiated between each other by assigning them different channels (frequencies) from a set of available channels. The problem was that such systems had very limited capacity e.g. the maximal number of users that can be served at any moment was equal to the number of channels in the set. The number of channels was relatively low because only a limited amount of spectrum could be used for personal mobile communication. Other potential frequencies were reserved for the military, radio, TV etc.

In order to achieve higher subscriber capacity and more efficient spectrum usage, AT&T introduced in 1971 a **cellular concept** [61] (see Figure 2-6).

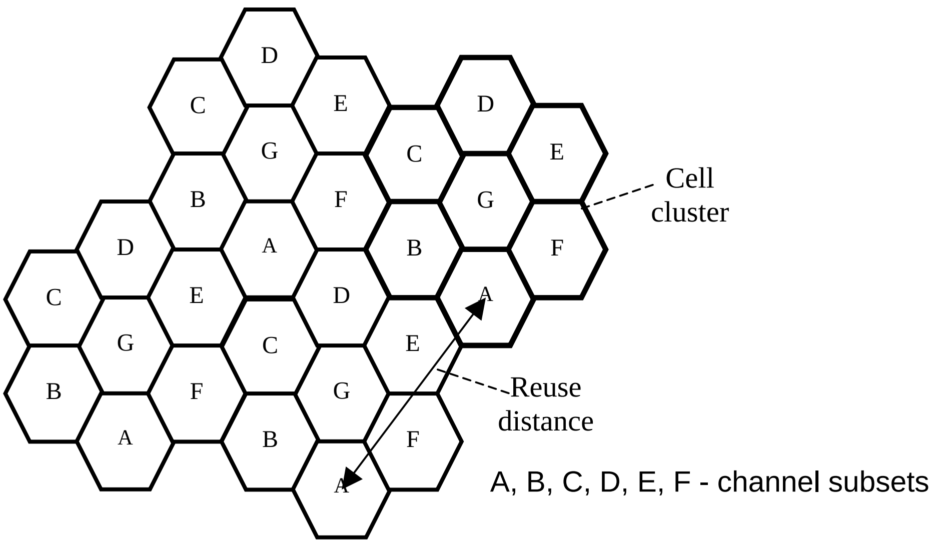



**Figure 2-6: Cellular area coverage (re-use factor 7)**

According to the cellular concept, the desired area is divided into zones called cells. Each cell uses a subset (A, B, C, D, E and F in Figure 2-6) of the available channels. The union of A, B, C, D, E, F and G is the set of all available channels in the system. A minimal set of concatenated cells using all available channels is called a cell **cluster**. As depicted in Figure 2-6, cells using disjunctive subsets A, B, C, D, E, F and G make a cluster. In this case we say that the re-use factor in the system is 7, since clusters consists of 7 cells and each cell use approximately 1/7 of total available channels. If the distance between cells is large enough, the cells can use the same channel subsets (**channel-reuse concept**), since interference from the users in the cells using the same channel can be kept low due to path-loss low: The power $P_D$ on distance D from the radiating antenna, decrease exponentially with the distance D (see Appendix A and [41]).

$$P_D \sim D^{-n} \tag{2.22}$$

*n* is path-loss coefficient and depends of the environment (usually between 3 and 5) [77]. Since each base station (BS) covers a much smaller area (usually 100 - 1000 m) than in original non-cellular systems (several kilometers), its power can be kept relatively low (usually several Watts). According to (2.22), this power falls exponentially with the distance between transmitter and the receiver. It means that the interference between two cells using the same subsets of channels (co-channel cells) can be kept arbitrarily low, if the cells lie at high enough distances.

In this way available channels could be **re-used** many times in the desired area. Consequently, the number of users which can be served by the system with the sufficient signal quality using a predefined set of frequencies, increases manifold in comparison to the systems without channel reuse and the same set of frequencies.

In addition to system capacity increase by channel reuse, the cellular concept has other important advantages: Traffic adaptability and reduction of required power.

**Traffic adaptability** means that if the traffic in a region increases, the capacity of the system could be enlarged by further dividing of the existing cells in that region into the new, smaller cells in the similar way that the larger area was originally divided into cells. This process is called **cell splitting**.

The second advantage of the cellular concept is **power saving due to shorter distances between mobiles and base stations**. This is especially important for mobile stations (MSs) due to limited battery power and the low cost and low size requirements.

However, with the cell concept a new problem arises: How should channels be assigned to cells and users? Various channel allocation algorithms give answers to this question.

The first obvious solution is to assign **in advance,** during the network planning process, channels to cells. Care should be taken to assign different channel sets to the neighbor cells and same channel sets to the cells in a large enough distance i.e. a certain channel allocation pattern should be used to maintain a desired re-use distance as represented in Figure 2-6. A cell is allowed to use only channels from its channel set. This kind of channel allocation is called Fixed Channel Allocation (FCA). FCA has been widely used in the past and even at the present time (GSM, TETRA).

The disadvantages of FCA are [1], [44], [68]:

- In the case of irregular traffic e.g. many mobile stations (MSs) in one cell and only a few MSs in the neighbor cell, **some cells can be overloaded and some cells can be under-loaded**. Overloaded cells do not have enough free channels to adapt required traffic, although there are enough unused channels in the neighborhood.
- Intensive **channel reuse planning is needed**, because cells in practice are not of a regular shape as represented in Figure 2-6.
- Each time when **new cells** are added to the systems or the **environment changes** (for example due to new buildings) **channel planning should be done again**.
- Channel **reuse planning is made for the worst-case scenario**. Therefore, a large amount of system capacity is often wasted.
- Channel reuse planning is **very difficult for micro-cellular systems** because the zone shapes are highly deformed due to irregularity of radio wave propagation in micro cell.

An alternative to FCA is **Dynamic Channel Allocation (DCA)** [68]. In DCA based systems **assignment of channels to cells is done dynamically, during the system operation**.

It is important to stress that most of the previous studies [4], [16], [34], [45], [28], [46], [1], [12] investigated DCA for "old" systems where the re-use factor was greater than 1 i.e. not all cells could use all channels available in the system. In modern wireless systems like UMTS coding gain, due to powerful signal processing techniques, is high (more than 10 dB) and thus requirements on signal quality at the receiver antenna (CIR) are lower. That is why re-use factor 1 is possible where all cells could use all channels and no channel planning is needed. Distributed, measurement-based DCA algorithms can then be applied for these modern systems in order to [68]:

- Maximize the number of satisfied users and/or total system throughput,
- Reduce signalization overhead (distributed algorithms),

- Avoid channel planning and
- Autonomously adapt to environment changes (dynamical algorithms).

In the following, we concentrate ourselves on the class of distributed, measurement based DCA algorithms, which enables decentralized and adaptive channel allocation with low signaling overhead and are likely to be used in 3G and 4G wireless systems.

### 2.5.2 *Distributed Dynamic Channel Allocation Algorithms*

Due to the complexity of modern wireless systems and rapid changes in signal propagation conditions, a class of **decentralized, measurement based dynamic channel allocation algorithms** (DCA) is of special importance. Decentralized measurement based DCA algorithms **reduce drastically the signaling amount in comparison with centralized algorithms** due to the use of locally (in one cell) available information (like interference and channel gain), and **enable flexible, self-adaptive solutions** to the different environment and propagation conditions due to easy of change of local parameters.

Examples of decentralized measurement based DCA algorithms are: Priority-based (channel segregation), Minimum Interference DCA and Autonomous reuse Partitioning DCA. The simplest DCA algorithm is Random DCA, which often serves as a benchmark to other DCA algorithms and does not need any measurement information. In the following, we give a short description of these algorithms.

#### *2.5.2.1 Random DCA*

The simplest DCA algorithm is Random DCA, which just **allocates channels to the users randomly from the set of free channels**. Random DCA algorithms have also been used in combination with some interference threshold comparison i.e. channels with interference lower than some thresholds are allocated randomly to the users [12].

In systems with fixed channel allocations to cells, channels are selected randomly from the set of available channels in the cell, but only a small subset of total available channels could be used in each cell (1/7, 1/12 of total number of channels) in order to keep re-use distance large and thus interference low. In modern wireless systems like UMTS, due to powerful signal processing (coding gain) and/or use of smart antennas, almost all channels have relatively low interference and all cells can use all channels (re-use factor 1). A simple Random DCA algorithm can then be applied to allocate channels to users. Note that well-known frequency hopping schemas (like in GSM) and spread spectrum systems (like in IS95 or in UMTS) are also based on the idea of some kind of channel “randomization” (in time and frequency), where one channel consists of many timeslots and frequencies in order to provide high diversity and good average channel quality.

#### *2.5.2.2 Channel Segregation*

Channel segregation is a good example of a distributed, self-adaptive, measurement based DCA algorithm [28], [1]. The algorithm is based on the **prioritization of channels** in each cell according to previous experience (measured interference) on the channels in the cell. Each cell has its own priority table, where it maintains the priority values for each channel. By service request, the interference $I'$ of the free channel with the highest priority is compared with interference threshold $I_{thr}$ (see Figure 2-7). If interference on the channel is lower than the threshold, the channel is allocated to the user and the channel’s priority is increased. Otherwise, the channel’s priority is decreased and a free channel with the next highest priority is compared with the interference threshold. In this way neighbor cells ”learn” to use different channels e.g. a certain kind of "Channel Segregation" between cells is established, like territorial segregation between some animals in nature.

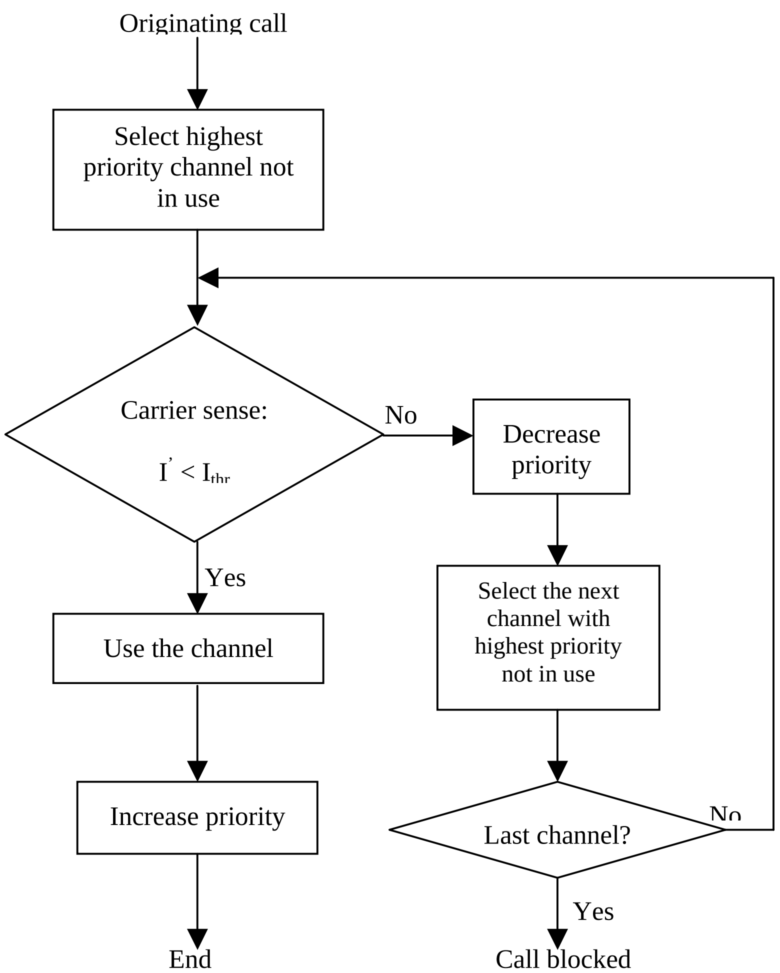



**Figure 2-7: Channel Segregation algorithm according to [28] and [1]**

According to [68] the advantages of channel segregation algorithms are:

- **No channel planning is needed.**
- **Autonomous and adaptive to traffic changes,** i.e. segregation is established independently from the number of users in the system and environment.
- **Decrease the number of intracell handovers**, since channels with low interference are selected for allocation.
- **Decrease load to switching system**, since each cell makes channel allocation decisions for itself.
- **Reduce blocking probability**, because channels used successfully in the past (those with high probability) are used as often as possible.
- **Quickly reaches a (sub-) optimal allocation**.

#### *2.5.2.3 Minimum Interference based DCA Algorithms with and without Threshold*

Priority-based DCA belongs to the class of **Interference based DCA algorithms**. The interference based DCA algorithms choose channels according to measured and averaged interference from other cells and possibly compare it with certain threshold. Some other interference based DCA algorithms are [12], [68]:

- **Minimum Interference DCA** selects the channel with lowest interference. This algorithm tries to minimize the overall interference in the system.
- **Lowest interference below a threshold DCA** selects the channel with lowest interference below a certain threshold. This algorithm tries to minimize interference in the system, while maintaining, at the same time, quality of each call above some minimum level.
- **Highest Interference below a threshold DCA** selects the channel with highest interference below a certain threshold. This algorithm tries to utilize spectrum more compactly as long as quality of each call is acceptable.
- Some **combination of above algorithms DCA** like marginal interference algorithms, which selects the most interfered channels if the interference is below a certain threshold, otherwise, selects the least interfered channel. This algorithm tries to balance signal quality and spectrum efficiency.

In [12] a comparison between above algorithms was done and it was shown that in the case of speech service the simple minimum interference algorithm, which needs no parameters (thresholds), oft outperforms other interference based DCA algorithms.

#### *2.5.2.4 Autonomous Reuse Partitioning (ARP)*

Intuitively, one would expect that a “good” DCA allocates channels with lower interference more often than the channels with higher interference, since the expected value of the CIR and thus signal quality is large on the channels with lower interference. However, the problem with the priority-based and minimum interference channel allocation is that the efficiency of the algorithms might depend on the order of the arrivals of users. For example, if the users close to the base station arrive before the users close to the cell border, the users closer to the base station get the better channels (with lower interference) than the users at the cell border. In that case, the users closer to the base station can achieve signal quality (CIR) much better than they actually need, because they have higher channel gain and lower interference (see equation (2.3)). On the other hand, some of the users at the cell borders might not achieve their minimal required CIR, because they have lower channel gain and higher interference. If we had assigned the channels with lower interference to the users at the cell border, and the channels with higher interference to the users close to the base station, we could possibly achieve for all (the most) users at least minimal required CIR. This idea is exploited by Autonomous Reuse Partitioning (ARP) DCA algorithms [34], [46], [68] (see Figure 2-8):

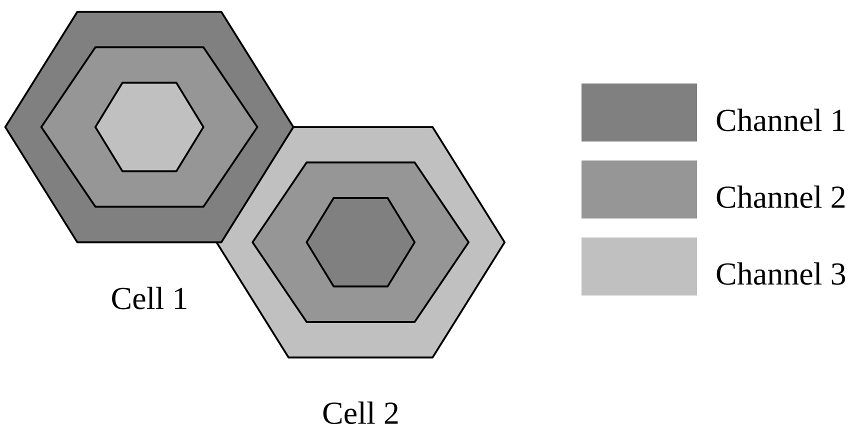



**Figure 2-8: Autonomous Reuse Partitioning (ARP) DCA**

**The “better” channels, with lower interference, are assigned to the users at the cell border and the worse channels, with the higher interference, are assigned to the users near to the base station.** The neighbor base stations should “cooperate” and allocate channels so that the users on the cell border of one cell use the channels which are allocated to the users near the base stations in neighbor cells and vice versa as can be (see Figure 2-8). As represented in Figure 2-8 channel 1 is used at the cell border in Cell 1 and near the base station in Cell 2. The goal is to minimize intercell interference on as many channels as possible by maximizing “channel packing” in the system. In this way, the number of users with sufficient signal quality in the system can also be maximized.

## 2.6 Admission Control and Handover

The task of HO is to decide **which cell (base station) a mobile user should be connected to**. The purpose of Admission Control (AC) is to decide **how "desirable" the admission of the users in the cell (system) is**. In an extremely case, AC can refuse the admission of users in the cell.

HO algorithms should at first ask the AC algorithms of the candidate cells if the users would be admitted in the cell or not and how desirable the admission of the users is. AC is a more cell oriented algorithm: Admit/not admit the users in a cell and indicate how “desirable” admission is. HO is more mobile users oriented - it tries to find the “best” cell for the mobile. In the following, we describe some basic HO and AC algorithms and criteria governing HO and AC decisions.

### *2.6.1 Handover*

When new users (new calls) arrive in a mobile system, decisions should be made to which base stations to connect the users. For example, the base stations to which the users have the highest channel gain (lowest path-loss) could be chosen. During the call the users move within a cellular system and propagation conditions change (see Appendix A) over time. That is why cell or channel within the cell should be changed. In UMTS FDD system a mobile can be also connected to several cells at the same time (Soft Hondover) [21].

Summarizing, the tasks of HO algorithms are [74], [20] (see Figure 2-9):

- Transfer a call from one cell to another (**intercell HO**)
- Choose the cell where the new call is started (**initial or new call HO**)
- Initiate channel change within the same cell (**intracell HO**)

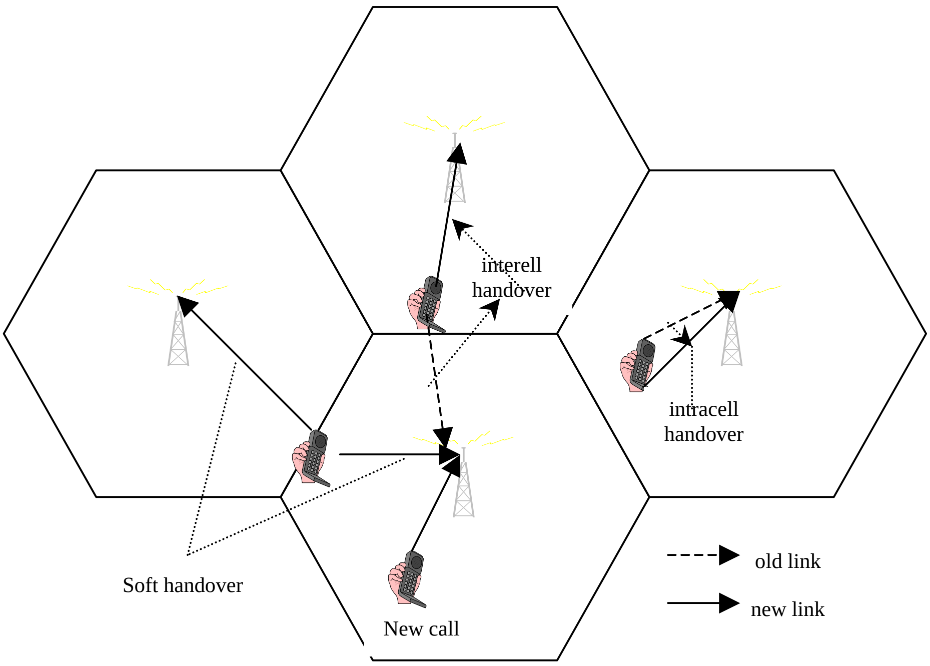



**Figure 2-9: Different kinds of HO**

Decisions about cell change are usually made [74] **according to signal quality.** As a "signal quality" the signal strength $C$ or CIR (see equation (2.3)) of some predefined (beacon) channel can be used. If the signal quality of cell B (target or candidate cell) becomes (for a HO-margin) better than the signal quality of cell A (home or active cell), HO is performed from cell A to cell B, as represented in Figure 2-10.

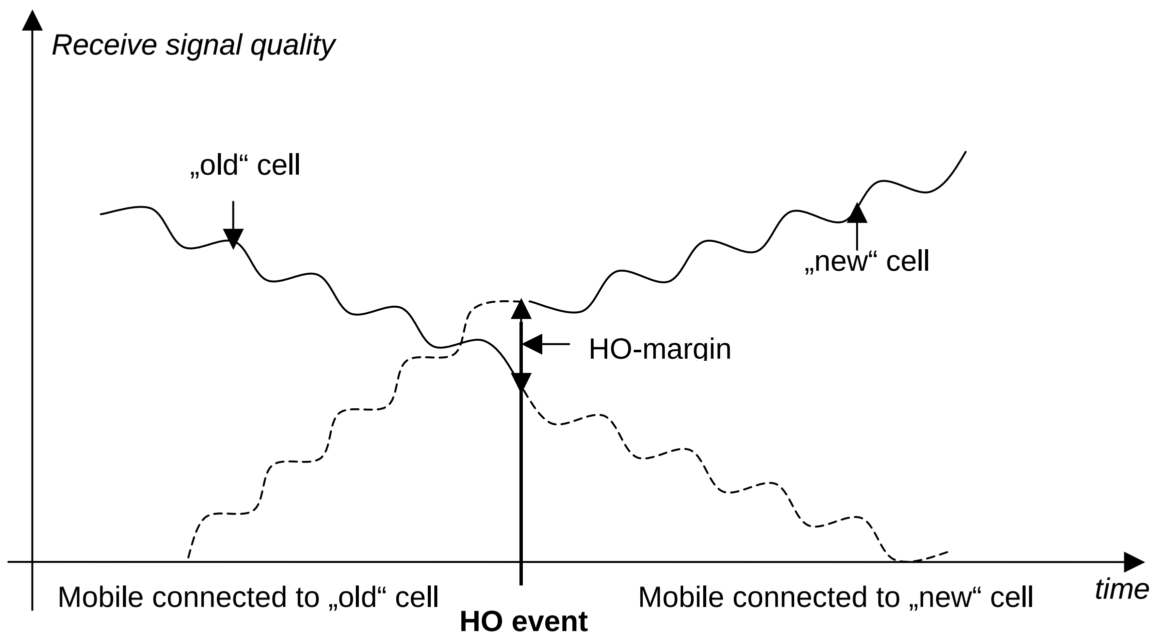



**Figure 2-10: Signal quality based HO decision**

**HO-margin** is used in order to avoid too many HOs, for example, at the cell border due to rapid fluctuation of signal quality difference ("ping-pong effect").

In order to reduce the number of HOs, it is also possible not to perform HO if the signal level of the active cell is satisfactory i.e. lies above a certain threshold, even if candidate cell provides a higher signal level.

HO decision can be made according to channel gain (the path-loss-based HO) and channel gain and interference (CIR-based HO), see Figure 2-11. In both cases the signal should be averaged over some time window. Further important issue which should be taken into account by HO decisions is admission control: If admission control of the candidate cells allows the admission of the new users in the cell or not.

```
Do each HO evaluation-period
    For each candidate cell in the list
        estimate Signal Quality (CIR, path-loss) of the  candidate cells
        If Signal Quality (the active cell) – Signal Quality (candidate cell) > HO_Margin
            If AC of the candidate cell allows access to the cell
                Perform HO to the candidate cell
            else
                Take the next candidate cell
            end
Until the end of the link
```

**Figure 2-11: An example of a Signal Quality based HO algorithm**

Each HO algorithms makes a trade-off between users' satisfactions, which decreases if HO is done too late and signaling load, which increases with the number of handovers.

#### *2.6.2 Admission Control*

**Admission Control (AC) decides to admit or not new users in the cell and how "desirable" this admission is**. The users should only be admitted in the system if their expected quality of service like minimum signal quality or data rate could be satisfied. For example, [19] defines the satisfied speech user as the user, whose Bit Error Rate (BER) during the call was for 95% of the time below a certain threshold. If AC in a cell expects, for example according to the cell statistic, that the probability that BER is above the threshold for more than 5% of the time, the speech users (at least the new one in the system) should not be admitted in the cell. AC can also make a "soft decision" expressing how desirable the admission of the users in the cell is. For example, if AC makes decisions between 0 and 1 (0 and 1 inclusive), 0 could mean users' rejection and 1 full acceptance, the values between 0 and 1 could express "desirability" of the admission - the closer the AC decision is to 1, the more desirable the users' admission in the cell is. These decisions can then be taken into account by HO algorithms together with other HO criteria like CIR or path-loss, in order to determine if, and to which cell, HO should be performed.

Furthermore, AC has to make a trade-off between new users and the users already in the system especially in the case of high loads. The users already in the system should have higher priority for the admission than new users, since call blocking is regarded as a less severe network failure than call dropping.

AC like some other RRM algorithms considers the following two issues for its decisions:

- **Traffic load**: The higher the cell load is the less desirable is the admission of the new and handover users in the cell. The new (in system) users should be accepted in the cell only if there are enough channels left free for the users arriving from other cells due to intercell HO.
- **Interference**: The higher the average interference in the cell the less desirable the admission of the new and handover users in the cell is. New users in the cell should be accepted only if increased interference level due to new users (in system) stays under a certain predefined interference threshold.

Since traffic load and interference are relevant not only for AC but also for other RRM algorithms like DCA, Scheduler, Load and Congestion Control, we describe these two issues in the following in more details.

#### *2.6.2.1 Traffic Models*

From the traffic modeling point of view, **a cell in a mobile network can be regarded as a queue with channels as servers**. The users arrive in the cell (queue) from the other cells (handover users) or start their calls in the cell (new call users). The users also end their calls in the cells or leave the cell due to handover to the other cells (see Figure 2-9).

A cell is usually modeled a M/M/N queue: We assume exponential (Markov - M) arrival times for both handover (with arrival rate $\lambda_h$ calls/seconds) and new calls (with arrival rate $\lambda_n$). We also assume an exponential (Markov - M) service time for both handovers to the other cells and call ends in the cell (with service rate $\mu$) [39]. The number of channels in the cell is denoted with $N$. Since call dropping is, in general, experienced as a more severe system failure than call blocking, the users already in the system (arriving from another cells by HO) should be given higher priority than new users. We can assume that *N-K* channels are reserved for HO calls i.e. the new call users are not admitted by AC in the cell if there are already *K* users in the cell (see Figure 2-12):

```
At  each new call or HO request:
if new call user
    if number of users in the cell < K
        allow the access of the user to the cell
    else
        refuse the user
else (HO user)
    if number of users in the cell < N
        allow the access of the user to the cell
    else
        refuse the user
    end
end
```

**Figure 2-12: An example of a load (number of users) based AC algorithm**

The state transition diagram of the queue/cell is represented in Figure 2-13 as a Markov chain [48]. The state of the queue/cell (oval symbols) is the number of active calls (users) or occupied channel. We assume that each user occupies one and only one channel. The transitions between states are triggered by call arrivals (new calls and handovers from other cells) and call departures (handover to other cells and call ends in the cell).

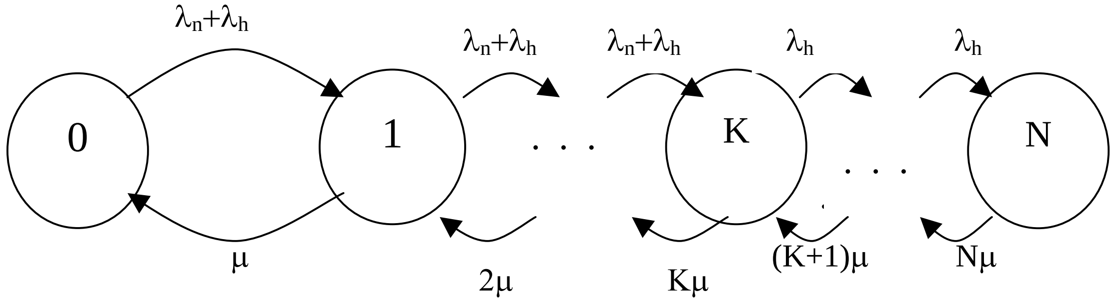



**Figure 2-13: Markov Transition Diagram with N-K reserved channels for HO calls**

New calls and HO calls are admitted in the cell if there are more than *N-K* free channels, otherwise only HO calls are admitted in the cell. By appropriate choice of the number of reserved channels *N-K*, a trade-off can be made between HO failure probability and new call dropping probability. If there are not enough free channels to accommodate new calls we say that **hard blocking** has occurred. If no channels are reserved for HO and setting $\lambda_h + \lambda_n = \lambda$ we obtain the following formula for the blocking probability $P_b$:

$$P_b = \frac{\frac{\rho^N}{N!}}{\sum_{k=0}^{N} \frac{\rho^k}{k!}} \tag{2.23}$$

where $\rho = \lambda/\mu$ is the total offered load (HO and new users) and $N$ is the number of channels in the system (cell). Formula (2.23) is famous **Erlang B formula**, which is widely used for **blocking probability calculation** in telecommunication networks [48]. Networks are usually designed so that some fixed blocking probability (say 2%) is not exceeded. According to (2.23), the higher the number of available channels $N$, the higher the ratio $\rho/N$ i.e. “channel exploitation” can be achieved for the same blocking probability. For example, if 10 channels are available and the blocking probability is 2%, only 30% of channels are used on average; if 30 channels are available about 60% of channels can be used on average for the same blocking probability of 2% [2]. This “large scale effect” is called **“trunking efficiency”** [2] and plays an important role in the design of telecommunication networks.

In the case of **mixed services**, for example, real-time (RT) service like speech and non-real-time (NRT) service like packet data, different services must share the same channels. A similar Markov state transition model, as the one in Figure 2-13, but in two dimensions, one for RT and one for NRT users, can also be used for mixed services (see Figure 2-14). In the case of Figure 2-14 the number of RT users and NRT users characterizes cell states. If there are no free channels available RT services are blocked but NRT services can be stored in the cell queue. Often a very large cell queue is assumed, which can store almost all NRT data. Note in Figure 2-14 that more than N (total number of channels) NRT users can be admitted in the cell (by queuing them in the cell queue). In a RT service model (see Figure 2-13) the probability that all channels are occupied (see (2.23)) is always kept below some predefined limit (hard blocking probability, say 95%) in order to provide the required service quality. However, in the case of relatively high NRT traffic load, all channels could be used all the time and possibly some data stored in the queue. When a RT user arrives, it should then get the next free channel or some channel used by a NRT user is made free for the RT user, since RT users have more stringent delay requirements than NRT users.

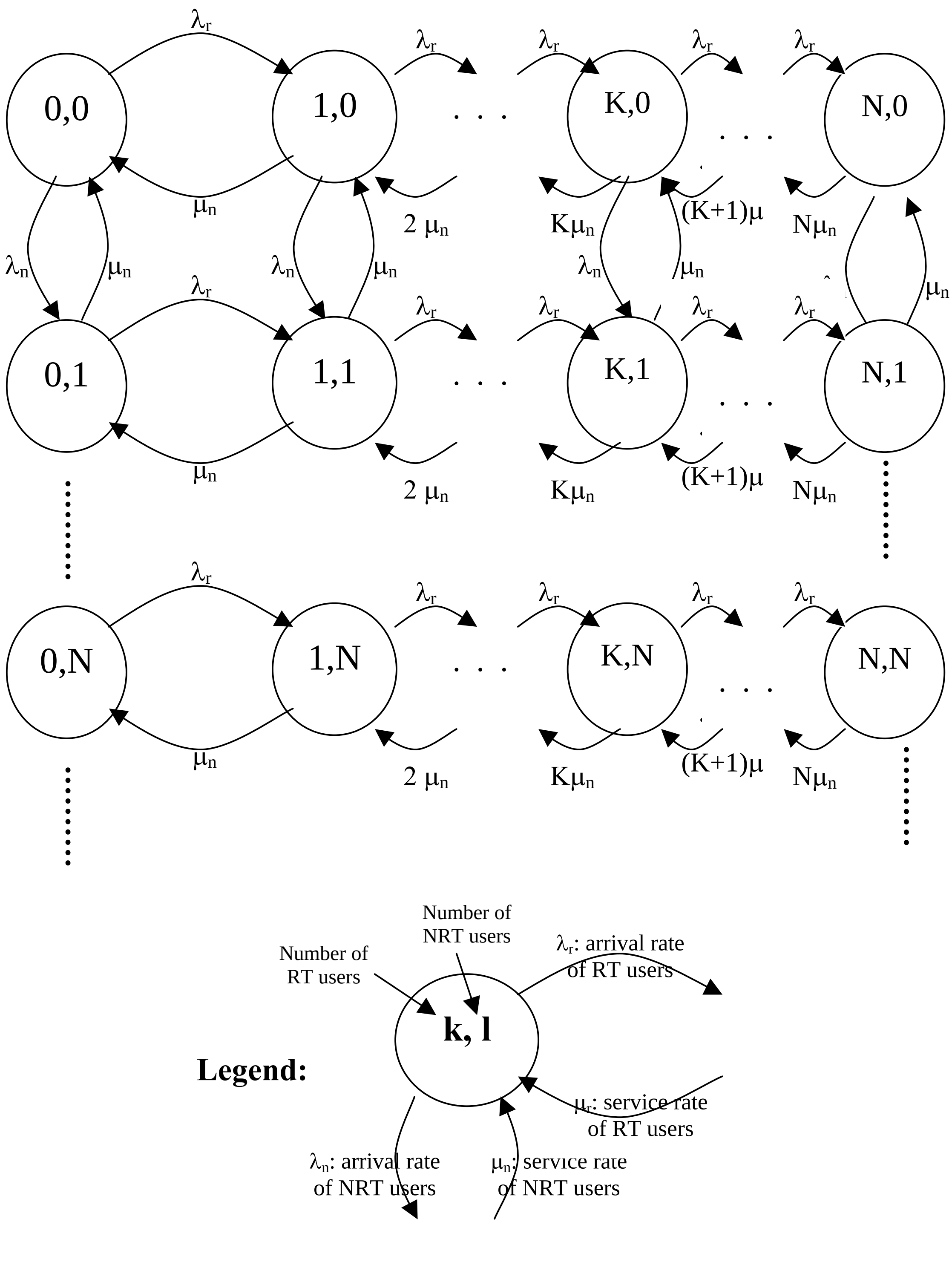



**Figure 2-14: Markov Transition Diagram of a cell with a real-time and non-real-time services**

According to Figure 2-14 different performance characteristics like the blocking probability of RT users, mean wait and service time for NRT users etc. can be derived, in a similar way as is done above in the case of only one service type. For example if the serving rate is the same for RT and NRT users ($\mu_r = \mu_n = \mu$) we can use the following formula for the blocking probability of the RT users [2]:

$$P_{RT}\,(blocking\,) = \frac{NP_b\,(\rho)}{N - \rho_n + \rho_n P_b\,(\rho)} \tag{2.24}$$

where $\rho = (\lambda_r + \lambda_n)/\mu$ is the total offered load (RT and NRT users), $\rho_n = \lambda_n/\mu$ (NRT offered load) and $P_b(\rho)$ is Erlang B formula from (2.23). The waiting probability for NRT users is the same as the blocking probability for RT users. The steady state exists only if $\rho_n$ *(NRT load)* $<$ $N$ (total number of channels).

Some problems which can arise by the **use of teletrafic theory equation in mobile networks** are:

- How to **determine arrival and service times** in the cells: Cells have an **irregular geometry,** defined by propagation conditions, HO parameters and interference from the other users (cells). Furthermore, **traffic behavior** of wireless, especially packet data users, is not well known.
- “Classical” teletraffic formulas, like Erlang B formula, do not say anything about **channel quality**. In fixed network we can assume that the main problem is to obtain a channel and once the channel is obtained the signal quality will be good enough to successfully finish the call. But in mobile networks if the channel gain is too low or the interference on the channel is too high, the users’ data cannot be decoded properly and should be retransmitted (packet data) or users might be dropped (speech). Relatively high interference from other users in the system on some channels could make the channels unusable (**soft blocking**).

In interference limited systems, blocking of users, due to the lack of free channels i.e. hard blocking seldom occurs. If the **interference on the free channels is too high** it has no sense to use the channels (**soft blocking**), since the users cannot achieve the required CIR even if they transmit with the maximal possible power (see (2.3)). Due to the use of modern coding and modulation technologies and RRM algorithms, channels can be re-used more tightly (re-use factor low), which increases interference in a system. Consequently, modern wireless systems are usually interference limited (like in UMTS) i.e. soft blocking occurs more often than hard blocking [38]. That is why an important issue of admission control and other RRM algorithms is to take interference on the channels into account in their decisions.

#### 2.6.2.2 Interference Issues

Since modern wireless systems are often interference limited, **new users should not be admitted into the system (cell) by AC if their admission would raise interference to other users already in the system over a certain limit**. This is very important for CDMA systems like UMTS, where the same band is usually shared by all users and each user experiences interference from virtually all other users in the system [29].

**Interference estimation** is an important issue for AC and other RRM algorithms. The interference information is obtained by measurements of base stations (in uplink) and mobile stations (in downlink). In the case of a high number of relevant interferers (usually more than 10), we can apply the central limit theorem and assume that interference distribution is log-normal [6], [83]. In this case it is enough to estimate only mean and variance in order to determine interference probability distribution.

An interference based AC can **protect the system from too much interference** if it checks, for example, if the interference rise in UL due to a new user $\Delta I$ would cause total interference increases over a certain limit $I_{thr}$ [38] (see Figure 2-15):

***At each new call or HO request:***

*Estimate interference increase due to new user in a cell:* $\Delta I$

*If* $I_{old} + \Delta I \geq I_{thr}$

    *Allow the access of the user to the cell*

*else*

    *Refuse the user*

*end*

**Figure 2-15: An example of an interference based AC algorithms in UL**

$I_{old}$ is the interference (in dB) before admission of a new user, $I_{thr}$ in general, depends on the service type of the new user (speech, data), environment (urban, suburban) or the load of the cell. Estimating correctly the expected increase of the interference $\Delta I$, is not an easy task [38], [87], since $\Delta I$ changes due to users' movement and changes in interference from other users.

The algorithms in Figure 2-15 can be used in UL where all users in the same cell experience the same interference in UL. In DL, different users experience different interference, but total DL power of a BS is usually limited in order to limit potential interference to other cells. AC in DL could then check if the power increase in DL due to a new user $\Delta P$, would cause that total power of the base station increase over a certain limit $P_{thr}$ [38] (see Figure 2-16).

***At each new call or HO request:***

*Estimate power increase due to new user in a cell:* $\Delta P$

*If* $P_{old} + \Delta P \geq P_{thr}$

*Allow the access of the user to the cell*

*else*

*Refuse the user*

*end*

**Figure 2-16: An example of power based AC algorithms in DL**

$P_{old}$ is the power used by the cell in DL before admission of a new user. $P_{thr}$, in general, depends on the service type of the new user, the environment or the load of the cell. $\Delta P$ is also not easy to estimate, since users' power changes with users' movement and interference from other users. Some hints for estimation of $\Delta P$ can be found in [38].

## 2.7 Smart Antennas and RRM algorithms

Although the topic of this work is investigation of RRM algorithms, we describe shortly smart antennas (SA) technology, which is gaining in importance in optimization of modern wireless systems. Furthermore, the results of our simulations and the performance of some RRM algorithms significantly differ with and without the smart antennas and interesting trade-offs can be made between SA deployment and the use of some RRM algorithms.

SA help reduce interference to the other cells in DL and eliminate interference from the users in other cells in UL [69]. The spatial beam-forming scheme aims to **point beams in the direction of paths belonging to the desired user while minimizing the energy transmitted in the other directions**. The "normal" omni-directional antennas transmit power in almost all directions, thus generating much more interference to other users than SA (see Figure 2-17). At the receiver side, SA can suppress interference from other users by **suppressing the signal from directions of interfering users**.

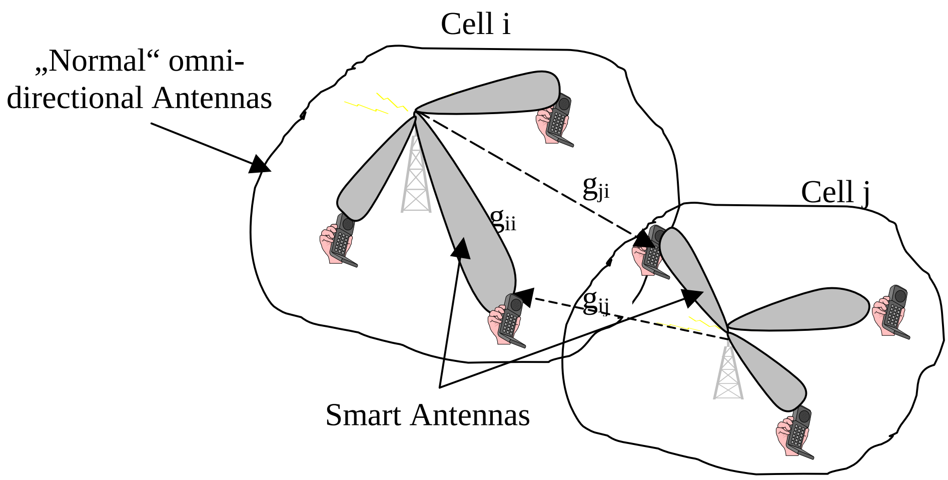



**Figure 2-17: Smart antennas deployment**

SA perform beam-forming by extensive adaptive signal processing algorithms [73]. The signals from each antenna element are multiplied by complex weights, which are selected so that the antenna beams are oriented in the direction of the desired users and, at the same time, the signals from undesired interferers are attenuated. SA are mainly used at BS, due their relatively large size and costs. If the channels in UL and DL are not correlated, channel information should be transmitted from mobile to BS in order to form beams optimally, which required an additional signalization overhead.

From the system-level point of view, the role of SA is to **reduce overall interference in the system and improve system capacity**. This is done at the expense of technology deployment. RRM algorithms also try to improve system capacity by interference reduction or redistribution, but at the cost of complexity, time consumption and signalization overhead. Providers can also make interesting trade-offs between SA deployment and the use of RRM algorithms. As will be shown in section 4.6, the deployment of SA sometimes makes use of some RRM algorithms unnecessary.

## 2.8 Summary

In this chapter we described the tasks of RRM and gave a brief overview of some RRM algorithms. The most practical RRM algorithms have the following features in common:

- **Distributed** i.e. no central controller is needed; BS or MS make decisions on their own about cell, channel and power allocation.
- **Measurement based:** RRM decisions like cell, channel and power choice are taken according to interference (interference based DCA, CIR-based PC) and channel gain (Handover, CIR-based PC) measurements.
- **State and service dependent:** Algorithms parameters and performance depend on service type (RT or NRT), environment (urban, sub-urban or rural) and actual state (traffic, interference, total power available). In order to provide an optimal network performance, algorithms and their parameters should be adjusted according to service type and actual state of the cells and users.
- **Cooperation is required.** For example: Minimum interference or channel segregation DCA could be applied only if other cells apply the same algorithms, CIR-based PC converges for most of the time when each user sets its $CIR_{thr}$ as low as possible etc.

In the next chapter we try to abstract these main features of the RRM algorithms and give a more general framework of RRM algorithms description and design based on “pricing” and game theory concepts. The RRM algorithms described in this chapter should be then obtained as special cases of this more general framework.

# 3 Radio Resource Management by "Pricing" and Game Theory

*"Every individual necessarily labours to render the annual revenue of the society as great as he can. He generally neither intends to promote the public interest, nor knows how much he is promoting it...He intends only his own gain, and he is in this, as in many other cases, led by an **invisible hand** to promote an end which was no part of his intention... By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it"*

***Adam Smith***

## 3.1 Introduction

In this chapter we try to generalize the ideas behind RRM algorithms described in chapter 2 and put them into more general context. Our approach is inspired by a free market economy where resource allocation occurs in a distributed manner. People try to maximize only their own "utilities" (do not have to be equal to money amount) using prices of the resources as signals of their relative "scarceness". In some sense, "optimal" social allocation can be "enforced" by a government and other state institutions using laws, taxes, subsidies and punishments. Similarly, **"optimal" resource allocation in wireless networks can be achieved by allocating the resources to the users for which users' "utility" is maximized, taking resource "prices" set by network into account**. In this way, **signalization overhead is drastically reduced** - each user maximizes only its own utility without taking into account the utility of other users. Furthermore, we can also **use well-established tools in economics like decision and game theory**, which help to analyze theoretically resource allocation among competing agents in a distributed manner. An "economic" approach for resource allocation in fixed networks like the Internet and ATM based on decision and game theory has already been proposed in [63], [65], [17], [95], [101] and for mobile networks in [80], [62], [79], [92], [91], [94], [102], [105], [78]. But these works almost ignore **a wealth of existing "heuristically" algorithms** like those described in chapter 2, which are not based on decision and game theory. Furthermore, the **dependence of the resource prices from network load** is not extensively investigated.

In our work **we try to encompass both "worlds" in our framework, the old one based on rather heuristic algorithms and the new "microeconomic" one, based on game theory**. Furthermore, in our work the "players" are "**bounded rational**" i.e. have a limited information available and do not necessary have to make optimal decisions but **"satisfactory" decisions** ("good enough"). The decisions are made according to the heuristics or "**rules of thumb**": for example "take the channel with lowest interference, when the load is relatively low" or "set CIR-threshold to lowest level for a service, when the load is high" etc. That is why we propose **a special form of the users' utility functions, which enables us to**

**implement such rules and obtain the popular algorithms like those described in chapter 2 as special cases**.

Instead of using complex algorithms, searching for a global optimum by each allocation of a radio resources, **our RRM algorithms just allocate those resources to users which satisfy users' constraints and for which users' utility functions takes a maximum**. Like a government, in order to be effective, should encourage compliance from the majority of the governed, so the users' utility function must be defined so that it reflects the basic wireless users' needs like maximizing signal quality (data rate) and minimizing resource (power, cells, channels) costs. That is why **our utility function increases with channel gain and decreases with interference and resource "prices"**. The **resource "prices" are set by network (or by each cell independently from other cells) according to network's (cell's) state** (load and interference). By setting prices the network (cell) maximize its utility (number of satisfied users and total data rate) for the given state, taking into account that each user gets those resources for which its utility (known to the network/cell), with the "prices" as parameter, is maximized. The base stations can signalize "prices" in broadcast or dedicated channels each time the state of their cells or network (significantly) changes. For example, the "prices" can be changed with a frequency of several hours i.e. increased during "rush hours" and decreased in other times. The "prices" can also change spatially, for example, "hot-spot" cells in the middle of the city would have higher admission "prices" (influencing HO algorithms) than other cells. **The "prices" can, but do not have to be, "real" prices with monetary meaning**. They can be regarded just as **parameters of the users' utility function** without any monetary meaning.

Due to the specific form of our utility function, **we also obtain many "existing RRM algorithms as special cases by appropriate parameter ("price") settings in our model**. The further advantage of our model is that network providers can estimate by simulations or network statistics for each relevant network (cell) state an "optimal" resource allocation and then using the **"backward deduction"** enforce by "prices" the optimal (satisfactory, "good enough") RRM algorithms for the state in a distributed manner with low complexity and signalization overhead. The algorithms we use are relative simple "rules" like "take the channel with minimal interference" or "take the cell with highest channel gain", like those described in chapter 2, but their use is state dependent.

We also investigate **how to enforce cooperation in the networks without "prices"** like in ad-hoc networks or among the networks from different providers. We show, using Game theory results, how **"cooperative" RRM algorithms can be the rational choice** for wireless users or networks even without explicit resource “prices”, provided the gain or cooperation is higher than from defection and “defecting” users (networks) are “punished” by the other users (networks).

In the following, we start in section 3.2 with the definition of a wireless game. Further, we define in section 3.3, users’ utility functions on the basis of users’ needs and constraints. In section 3.4 we define the utility and constraints of a wireless network. Based on users’ and network utilities we describe our “pricing” approach to RRM in section 3.5. In section 3.6 we describe how “prices” could be set using the concept of “backwards deduction” and in section 3.7 what are relevant cells’ or network’s states and state-dependent price settings. In section 3.8 we provide mathematical “guidelines” or heuristics for developing RRM algorithms: In subsections 3.8.1 and 3.8.2 we investigate an impact of the “prices” on the system dynamics and in subsection 3.8.3 we provide a “geometric picture” of tasks and the role of RRM algorithms for resource allocation in an equilibrium state. In section 3.9 we show how RRM algorithms from chapter 2 can be obtained as special cases of our “pricing” framework. Some advantages of our “pricing” framework are listed in section 3.12. We discuss how measurement decisions can also be done on the basis of expected change in utility function in section 3.10.

Further, we investigate users’ cooperation without “prices” which might arise in ad-hoc networks on the basis of game theory in section 3.13. In section 3.11 we discuss complexity issues of our algorithms. Finally, we summarize in section 3.14.

## 3.2 Wireless Game Model

We define Radio resource management as a game in extensive form [26] with the following characteristics:

- **The set of players** - We define two kinds of players: **users** or users’ part of RRM algorithms and **cells** (network) or the network part of RRM algorithms. Cells are responsible for setting the ”prices” for resources and users are responsible for the actual allocation of resources taking into account resource “utility”. This classification is **rather functional and not according to the place of execution**. RRM algorithms can be executed in network (BS, RNC, Node B) or in MS, or one part of the algorithms can be done in MS and one part in network.

- **The set of possible actions or moves of the players** - Users can choose between **cells** and providers (intra- or inter-system handover), **channels** within a cell (channel allocation), **power** to be used (scheduler and power control). Cell choices are the **“prices”** for cell (system) entrance (AC), channel “prices” (DCA) and power “prices” (Scheduler, PC). The decisions about “prices” (by cells) are made so that a simple heuristics (**“rules of thumb”**) by resource selection (by users) is enforced: For example, “if the load is low take the channel with minimum interference”, “if the load is high use “fair-throughput” scheduling” etc.

- **The order of actions –** In order to enable a full decentralized implementation, the actions of the players are executed **asynchronously and sequentially**. All decisions in a network occur in general at different time points and independently from each other.
- **What each player knows when he takes an action –** A wireless user has, in general, information obtained by his own **measurements** (channel gain, interference) and information obtainable from the cell by means of **signaling** (“prices” for cell access, channel and power). In the cells following data is available: interference from other cells, number of users of each service type and their arrival rates, channel gains, delays and data rate of the users in the cell, etc. Further, **the players in our model are “bounded rational”** i.e. have limited computational capacity and make their action based on the limited information available, like local measurements of channel gain and interference. The players also do not try to “outguess” each other i.e. they do not make “circular” assumption of the type “if the other players know that I know that they know that I know …”, as sometimes assumed in “classical” game theory [100].
- **The probability distributions of any exogenous event –** We regard users’ movement as an exogenous event in a wireless game. That is why **arrival and departure of the users in the cells and their channel gains** can be regarded as exogenous events too.
- The users’ and network's decisions are **state dependent (Markov Games** [26]**)-** Users make **resource allocation decisions according to users’ states** i.e. power consumed, data rate achieved so far, channel gain, interference and “prices” for resources. Cells make **“price” decisions according to the cells’ states** i.e. the traffic load and interference in the cells.
- **What the “players” maximize** i.e. **the users’ and network's utility functions.** In our model the players make not absolutely optimal but **“satisfactory” decisions**, which provide **“good enough” performance under given load**. A search for globally optimal solutions would require too much signalization overhead and would be of very limited duration due to fast changes in a fading environment.

In the next sections we define the users' and network's utility functions in such manner that we obtain some of the RRM algorithms described in 2 as special cases of our "wireless game".

## 3.3 Users' Utility and Constraints

The "happiness" of wireless users in general increases with signal quality (data rate) and decreases with resource (cell, channel, power) costs. Resource costs can be summarized in power costs, which measure value of power in a certain cell on a certain channel. Power costs reflect "subjective" costs of the users i.e. how the users value their own battery power, but also, more importantly, "social" costs due to interference to other users. These "social" costs should be taken into account by base stations (RRM algorithms in BS) by determining power costs in the cell.

In Game theory the "happiness" or the "gain" of the players (users) is summarized in users' utility function [100]. We differentiate between the so-called Bernoulli utility $u$ i.e. the utility at the certain moment and Von-Neumann-Morgenstern utility function $U$ i.e. the expectation of the Bernoulli utility function [64] over the time. The form of the utility function depends on service type but, in general, **all services try to maximize data rate *R* and, at the same time, minimize power consumption**. In the following, we assume that utility function is separable and linear in power costs and data rate. Using Shannon's formula for data rate R (per unit bandwidth) for the user sending with power $p$, with channel gain $g$ and interference $I$, and setting power costs proportional to consumed power we propose the following Bernoully utility function $u$:

$$\boxed{u(p) = R - \lambda p = \log\left(1 + \frac{pg}{I^{'}}\right) - \lambda p} \tag{3.1}$$

where $\lambda$ *r*epresents "price" per unit power and is calculated relative to value of the data rate and amount of bandwidth B i.e. if a user weights data rate with factor $w_R$ and power with factor $w_p$, $\lambda$ is set to $w_p/(w_R B)$. Note that if $u$ is a utility function than each linear transformation of $u$ is also a utility function [100]. If the data rate of the service is fixed, like for speech service, or the users have already achieved their maximal data rate, the utility as defined in (3.1) decreases with increase in power consumption i.e. the goal of the users is to minimize power costs maintaining the fixed data rate.

Equation (3.1) represents users' utility according to power $p$ and can be used to evaluate PC decisions. In order to evaluate decisions of the other RRM algorithms like Scheduler, AC, HO and DCA, we need a utility function representation, which is independent from the users' powers, since power changes during the time according to a PC algorithm. In the following, **we try to express users' utility function in terms of channel gain, interference and "prices"**, since these are the usual values that can be obtained by signalization ("prices") or measurements (interference and channel gain). To that aim we investigate three cases of PC (see section 2.2):

**Constant power PC**. $p=P_{max}$. From Shannon's formula for data rate we obtain the following expression:

$$R \sim \log\left(1+\frac{P_{\max} g}{I^{'}}\right) \tag{3.2}$$

In the case of **"Water filling" PC** (see section 2.3), power $p$ of the users is set according to following expression:

$$p = \left(\frac{1}{\lambda} - \frac{I^{'}}{g}\right)^{+} \tag{3.3}$$

Where $\lambda$ is a Lagrange multiplier, which can be interpreted as "cost" of the power. Setting (3.3) into Shannon's formula for data rate we obtain the following expression:

$$R \sim \log\left(\frac{g}{\lambda I^{'}}\right) \tag{3.4}$$

In the case of **CIR-based PC** power of the users is set according to following equation:

$$p = CIR_{thr}\frac{I^{'}}{g} \tag{3.5}$$

Consequently (if PC converges, see subsection 2.2.3):

$$R \sim \log\left(1+CIR_{thr}\right) \tag{3.6}$$

Also, in all three cases of PC: Constant power, water-filling or CIR-based, the **data rate increases with increase in channel gain and decrease in interference**. **Power required to achieve certain data rate or certain signal quality (CIR) also decreases with increase in channel gain and decrease in interference**. The (power-dependent) utility function defined in (3.1) also increases with increase in channel gain, decrease in interference and decrease in power costs $\lambda$. This motivated us to define Bernoulli's utility function as follows:

We define the (power independent) **Bernoulli's utility function *u* (utility at certain moment) of a wireless user as logarithm of the ratio of channel gain *g* and product of interference *I* and "price" per unit power** $\lambda$ i.e.

$$u = \log\left(\frac{g}{\lambda I}\right) \tag{3.7}$$

The utility definition like this in (3.7) is a practical choice because:

- The utility is **directly proportional to maximal achievable data rate**.
- The **larger the utility the less power is needed** to achieve a certain data rate or CIR.
- It is a **decreasing risk averse** [47] function (in g/λI) because log-function is a concave function (a twice differentiable function f(x) is concave if $a*f(x_1)+(1-a)*f(x_2) \leq f[a*x_1+(1-a)*x_2]$ *for all* $0 \leq a \leq 1$) like most of the utility functions used in practice. This feature means that users with high $g/(\lambda I)$ are readier to take a risk than the users with low $g/(\lambda I)$: The higher $g/(\lambda I)$ the lower the difference in the utility of having certainly $\log[g/(\lambda I)]$ and the utility of taking a risk and playing a lottery whose expected gain is $\log[g/(\lambda I)]$.
- The utility function **favor diversity**, since the concave functions like logarithms fulfill Jensen's inequality [15] i.e. the utility of the expected value is always better than expected value of the utility:

$$\log\left[E\left\{\frac{g}{\lambda I}\right\}\right] \geq E\left\{\log\left(\frac{g}{\lambda I}\right)\right\} \tag{3.8}$$

  This is important since diversity is one of the most often used concepts in wireless communications to improve signal quality and increase data rates.
- **Calculation of the expected utility reduces to sum of mean values** in dB(m) for channel gain and interference (plus “price” in dB). Both average channel gain (slow fading) and interference (as a sum of a large number of log normal variables) are independent lognormal variables [6], [83]. Furthermore, the channel gain, interference as well many other quantities in wireless communication are usually represented in logarithmic measure.

Finally, Scheduler, AC, HO and DCA algorithms uses for their decisions not the utility in certain moment $u$ (Bernoulli utility function) but the expected value of the $u$ i.e. Von-Neumann Morgenstern utility function $U$:

$$\boxed{U = E\{u\} = E\left\{\log\left(\frac{g}{\lambda I'}\right)\right\}} \tag{3.9}$$

**Users’ Constraints**

The users’ constraints are set according to service type: For example a Non-Real Time (NRT) service like packet data has to achieve at least 10% of its nominal data rate during the session in order to be satisfied [19]. Also, the minimum data rate $R_{min}$ (say 38.4 kbps for 384 kbps service) should be achieved during a certain time period (session length) i.e.:

$$E\{R\} \geq R_{\min} \ \textit{within session length} \tag{3.10}$$

Real-time RT service like speech requires that the probability that channel quality (CIR) is lower than a certain threshold $CIR_{thr}$ (say 12 dB) lies below certain threshold (outage probability) $P_{out}$ (say 2%) [19]:

$$\Pr\left(CIR < CIR_{thr}\right) < P_{out} \tag{3.11}$$

Since CIR is proportional to data rate ($R \sim \log(1+CIR)$), we can combine constraints for both RT and NRT users as follows: Probability of achieving certain data rate $R_{min}$ must be above certain threshold ($1-P_{out}$) within certain time period $t_d$ in order that the users are satisfied:

$$\Pr\left(R > R_{\min}\right) > 1 - P_{out}, \quad within \;\; t_d \;\; seconds \tag{3.12}$$

Naturally, RT users have more stringent delay constraints $t_d$ (several milliseconds) than NRT users, but NRT users usually need higher minimal data rates $R_{min}$.

## 3.4 Network Utility and Constraints

Two main factors define the capacity of a wireless networks:

- **Number of satisfied users**: The first goal of a network provider is to maximize the number of satisfied users in its network. Users are regarded as satisfied if they achieve their minimum data rate $R_{min}$ (with certain probability) within predefined time $t_d$ (3.12). Constraints $R_{min}$ and $t_d$ depend on service type. For example, a RT services like speech have more stringent delay constraints than a NRT service like packet data. Speech data should be transmitted each 20 ms (interleaving period), but the data rate $R$ for speech is relatively low (about 10 kbps) and depends on required speech quality. For NRT data delays are measured in seconds or even minutes as in the case of web-browsing but on average data rate $R$ should be at least 10% of maximal data rate in order that NRT users are satisfied (38.4 kbps for a 384 kbps service).

- **Total data rate achieved:** If the “prices” depend on data rate consumed by the users, the more data rate users achieve, the higher revenue for the provider. In general, revenue per data rate depends on service priority and the art of the service, some services like real-time videos may be charged more per kbps than others like web-browsing or e-mail.

In the following, we consider as an example the mixed service environment with two service types: Speech users (as a representative of RT services) and packet data users (as a representative of NRT services). We assume an additive utility function of the provider i.e. the utility of the sum of the users is equal to the sum of the utilities obtained from each user alone. Since different optimization goals according to the relative importance (priorities) of different users’ types can be applied, we introduce $w_s$, $w_d$ and $w_i$ as the weight factors defining the weight of satisfied speech users, the weight of satisfied data users, and the weight of achieved data rate by the user i respectively. We define the following optimisation problem for a network provider:

**Objective**: Maximize weighted sum of satisfied speech user $N_s$, satisfied data user $N_d$ and sum of achieved data rates per user $R_i$:

$$U_{network} = w_s N_s + w_d N_d + \sum_{i=1}^{N_d} w_i R_i \tag{3.13}$$

Under the following **constraints**: Total bandwidth used (sum of bandwidth used per user $B_i$) is lower or equal to available bandwidth $B_{total}$ (number of used channel is lower or equal number of available channels). Power of user $p_i$ does not exceed the maximal allowable power for the user $P^i_{max}$ and total sum power does not exceed some limit $P_{total}$ (important in DL). Further, speech and data users are satisfied - achieve minimum data rate $R_{min}$ within predefined time $t_d$ i.e.:

$$\begin{aligned}
&\sum_{i=1}^{N} B_i \le B_{total} \qquad N = N_s + N_d \\
&p_i \le P^i_{\max} \quad i = 1..N \\
&\sum_{i=1}^{N} p_i \le P_{total} \\
&\Pr(R_i \ge R^i_{\min}) > 1 - P_{out} \ \ within\ t^i_d\ \ seconds \ \ i = 1..N
\end{aligned} \tag{3.14}$$

A network provider should make an important trade-off between **efficiency** i.e. maximizing total throughput and **fairness** i.e. maximizing the number of satisfied users. Some services like speech service usually have a fixed data rate and the main issue of the provider is to maximize the number of speech users, who achieve their minimum signal quality (data rate) for most of the time. On the other hand, in the case of other services like NRT data a trade-off between data rate and the number of satisfied users should be made. Whether a provider gives more importance to efficiency or fairness depends on revenues obtained pro satisfied user and pro bits of data rate i.e. mathematically from weights $w_s$, $w_d$ and $w_i$ in (3.13). Usually the data rate issue is subordinate to the issue of users' satisfaction. Users often pay some fixed part of monthly bill independently of data consumed - if they are unsatisfied it is usually worse for providers than if the users' data rate is moderate, as long it stays above some limit. In this work **we usually assume the utility of the network is higher from having one more satisfied user than of having more data rate or better signal quality from already satisfied users** i.e. fairness has more priority than efficiency.

In the following section we describe how network (or each cell) can maximize its utility in a distributed manner by setting the “prices” appropriately and letting each user to maximize its utility with “prices” as parameters.

## 3.5 Resource Allocation by “Pricing”

In this section we describe our “pricing”-based framework for RRM. We first define users’ utility function as a linear function of channel gain and interference with weights (“prices”) as parameters. Further, we describe our RRM concept based on adaptively setting of weights in the users’ utility function (“pricing”) according to network’s (cell’s) state.

### *3.5.1 Users’ Utility Function with Weights*

For the purpose of more flexibility and easier handling of the utility function, we use a slightly changed form of the utility function given in the equation (3.9). Note that **“price” of power per unit bandwidth (λ in** (3.9)) **could be a function of the channel gain g and interference I**. For example, we can set $\lambda = g^{1-W_G} I^{-1+W_I} 10^{C}$ in (3.9) and obtain expected utility as the sum of the mean values of *I* (mean value of interference in dB), *G* (mean value of channel gain in dB) with the cost factor *C* (comprising other factors like load, average delay etc.):

$$\boxed{U = W_G G - W_I I - C} \tag{3.15}$$

where $W_G$ and $W_I$ and are the weights for the interference and channel gain respectively and *C* represents the additional costs. This “linear” form enables easy and flexible use of the utility function and simple parameter handlings. Furthermore, **due to the particular form of the utility function, we can obtain (as will be shown later) some algorithms from chapter 2 as special cases of our “pricing” framework**.

### *3.5.2 A “Pricing” Framework for RRM*

In Figure 3-1 our RRM model based on “pricing” is represented. Cells set “prices” for the radio resources (cells, channels and power) in order to maximize network utility. **The “prices” are set by a network** (Operation and Maintance Centar (OMC)) **or, in order to reduce signalization, in each cell** (BSC, RNC, Node B) independently from other cells, **according to the state of the network (cell)**. The users’ part of RRM then allocates those resources to users which maximizes the users’ utility function (3.15), taking “prices” set by the cells as parameters of the utility function.

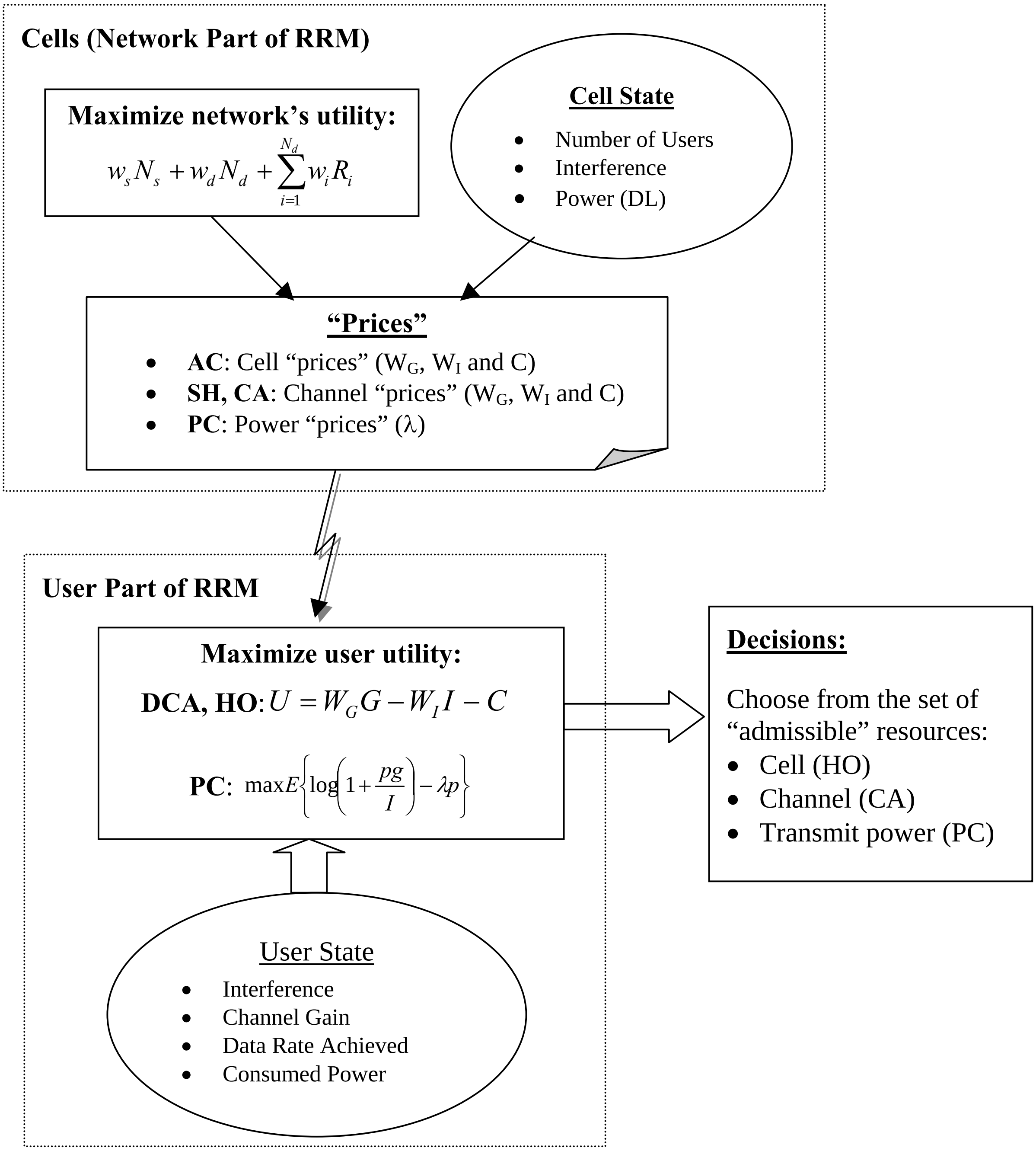



**Figure 3-1: A "pricing"-based model for resource allocation in mobile networks**

In this way, **we can realize distributed RRM algorithms, where cells and users take actions that maximize their utilities, using "prices" as mediators between cells' and users' utilities**.

In Figure 3-2 a schematic RRM algorithm based on pricing is represented. As can be seen from Figure 3-2 the pricing-based RRM consist of cells' (network's) specific part that set "prices" (weights) for the resources according to network state and users' specific part that allocates resource(s) with maximal utility taking into account "prices" set by cells' part.

| *Cells' (network's) specific part of RRM:* | *Users' specific part of RRM:* |
|---|---|
| *DCA, Scheduler and AC:*<br><br>*Set prices (weights) $W_G$, $W_I$ and C for the resources (channel, cell, scheduling) according to the state (load, interference) of the cell.*<br><br>*PC*<br><br>*Set power" prices" $\lambda$ for water-filling PC or $CIR_{thr}$ for CIR-based PC according to cell state (load, interference, available power)* | *DCA, Scheduler and HO:*<br><br>*At each resource request*<br>*For each resources in the cell*<br>*If resource "admissible" i.e.:*<br>*resource free and*<br>*$Pr(CIR<CIR_{thr}) < P_{out}$*<br>*find resource's utility:*<br>*$U(resource) = W_G G - W_I I - C$*<br>*end*<br>*Allocate resource with maximal utility U*<br><br>*PC (set power)*<br>*$P = (1/\lambda - g/I)^+$, water-filling PC or*<br>*$P = CIR_{thr}*g/I$, CIR-based PC* |

**Figure 3-2: A pricing-based RRM algorithms**

Naturally, users get only "admissible resources" i.e. free resources for which minimal users' requirements like outage probability or minimal expected data rate are satisfied (3.12).

We can also use the same "pricing" model if the users' **utility function is not known**. In that case a provider could assume that users choose among the resources which satisfy users' constraints those with **minimal costs**. The cost for a resource ("Cost function") is inverse proportional to resource utility $U$ i.e. ***Cost Function* $= -U = C - W_G G + W_I I$**. As in the case of utility maximization a provider can "enforce" in the distributed maner optimal resource allocation by setting the weights $C$, $W_G$ and $W_I$ and assuming that users **minimize the Cost Function**.

In the language of Game theory we talk about the "Stackelberg game" [26]. This is a "hierarchical" game where one agent, called leader (in our case a cell or BS) moves first and takes his actions in order to maximize his utility (3.13) knowing the utilities (3.15) of the other players (users) called the followers, who move after they have observed the action of the leader. This kind of game usually has **more efficient Nash equilibriums** [26] than in the case when the sequences of players' moves is not predefined. The network can, for example, determine according to simulation or network statistics "optimal" RRM algorithms for each state and enforce it by "prices" as a leader in a Stackelberg game. The users then maximizing their utilities (with network's "prices" as parameters) at the same time maximize the network utility.

It should be stressed that **we do not search for "absolutely optimal" solutions for each network state**, since such search would be **too costly** (in signalization overhead) and of **short duration** due to fast changes in a fading environment. Instead, we search for "good enough" or "satisfactory solution" which can be achieved by use of heuristically "rules of thumb" and with limited information available ("bounded rationality" assumption).

By setting the (sub-) optimal "prices" cells use some kind of **"backwards deduction"** which consist of the following two steps:

1) Determine **"optimal" resource allocation for the state** i.e. the allocation which maximizes networks' or cells' utility.
2) **Choose "prices" so that users maximizing their utilities take the resources according to the optimal resource allocation for the state**.

"Backwards deduction" will be explained on the example of DCA algorithms in the following section.

### 3.6 "Backwards Deduction" and Load Dependent "Prices"

In this section we describe a "backwards deduction" on the example of dynamical channel allocation algorithms (DCA). The idea is to **imagine an "optimal" channel allocation for the given cell's state and than set the "prices" ("weights") in the users' utility function so that user's maximizing their utility function, with weights as parameters, chose the channels so that the desired "optimal" allocation is established**.

The optimal channel allocation is load dependent. For example, for lower loads it would be optimal that neighbor cells uses different channels as oft as possible in order to minimize inter-cell interference. To this purpose, Minimum Interference DCA can be used in order to avoid the use of channels with higher interference, which are probably oft used in neighbor cells. On the other hand, in the case of higher loads, almost all channels must be used almost all time. In this case **ARP DCA**, where each **user gets the channel which is just "good"**

**enough to achieve its minimal CIR, but not better**, in order to left better channels for the users with higher path-loss, would maximize the number of satisfied users in the cell. The example of the use of the "backward deduction" for "high load" is represented in Figure 3-3.

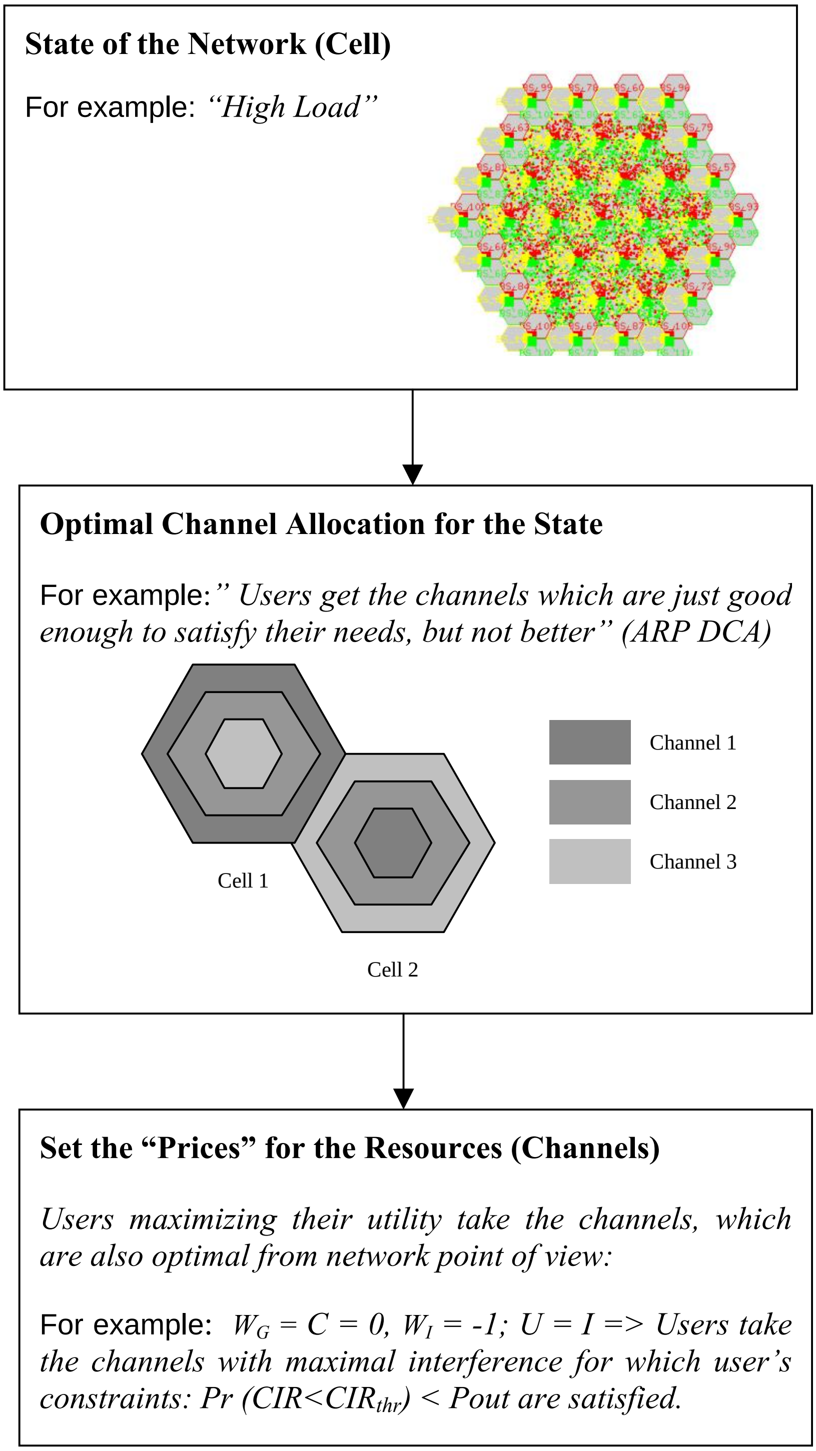



**Figure 3-3: "Backward Deduction" for a "high" load**

Since optimal channel allocation is load dependent, **for different loads different DCA should be used**. The load dependent "price" setting can be realized by **choosing the weights in the utility function according to the channel quality** (weights $W_I$ and $W_G$) **and load** (weight $C$) as represented in Figure 3-4. Both weights $W_I$ and $W_G$ can be set to 1, when channel quality (*G-I*) is below a predefined level (*C*), or to –1, when channel quality is greater than *C*. The weight *C* is set inversely proportional to the load in a cell: For lower loads, *C* is set to a higher value and users often take the "better" channels (channels with lower interference) i.e. use minimum interference DCA, since utility decreases with channel interference for (all) most of the channels. For higher loads, it might be desirable to allocate those channels which just fulfill minimal users' requirements (like minimal CIR) i.e. use ARP DCA. To that purpose, cells can set *C* relatively low in the case of higher load in order to enforce users to "take" worse channels i.e. channels, which just fulfill their minimal requirements. Also, if gain *G* is channel independent (all channels use the same frequency as in the case of TDMA), by load dependent setting of the weight *C* as in Figure 3-4, cells can switch between minimum interference DCA for lower loads and ARP DCA algorithm for higher loads.

*If* $G-I < C$ $(C \sim 1/load)$

$W_G = W_I = 1$,

$U = G - I - C$

*Else*

$W_G = W_I = -1$;

$U = I - G - C$

*End if*

*Allocate channel for which |U| is maximal!*

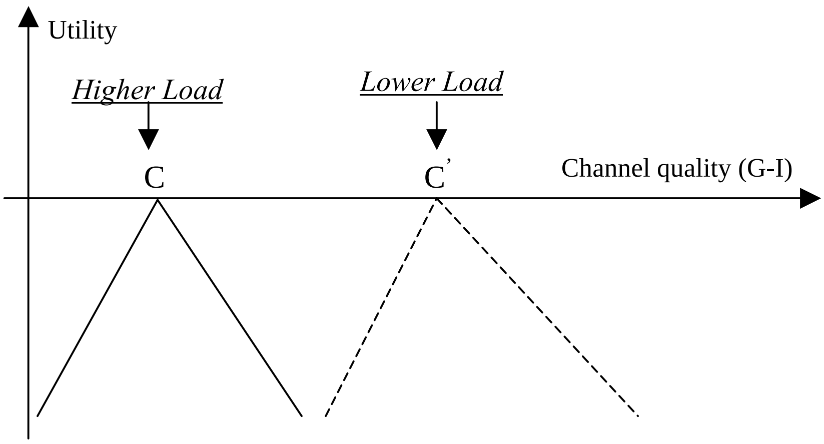



**Figure 3-4: Load dependent cost ("*C*") settings**

If the users minimal quality requirements cannot be satisfied on any free channels, the user should not start the call at all.

Let us take the example of optimal (maximize the number of satisfied RT users) DCA in the case of higher loads. We assume that outage probability is maximal 2% (see constraints of RT users (3.11)). We also assume the Gaussian distribution of interference in the case of higher load with mean value $I = I(channel)$ and standard deviation $\sigma_I = \sigma_I(channel)$. Then DCA set the weight *C* to the value $C = CIR_{thr} + 3\sigma_I(channel) - P_{max}$ and $W_G$ and $W_I$ to 1, as in Figure 3-4 left. When a user needs a channel, the highest utility (0) is obtained for the resource (channel) on which the I distribution is such that the user constraints are just satisfied i.e. $U = P_{max} + G - I(channel) - CIR_{thr} - 3\sigma_I(channel) = 0$ i.e. the probability that CIR ($P_{max} + G - I(channel)$) is lower than $CIR_{thr}$ on the channel is just about 2%. In this way, users get (approximately) those channels, which they need to fulfill their minimum quality requirements but not better, in order to leave “better” channels for the users with worse conditions (lower channel gains). This is also a kind of ARP DCA as described in subsection 2.5.2.4. As discussed above, ARP DCA might be useful in the case of higher loads in order to maximize the number of satisfied users. In the case of lower-medium loads, cells can set $W_G$ and $W_I$ to 1 and $C$ to 0 and users would, by maximizing their own utility ($G$-$I$), always “take” the free channel with lowest interference i.e. apply Minimum interference DCA.

The same reasoning can be extended to costs of the resources other than channels like power, cells, layers in the case of hierarchical cells, systems like GSM and UMTS in the case of mixed systems or even codecs with variable data rate like AMR codecs in the case of link adaptation. “Resource quality” is then defined according to channel gain, interference and load of the resources as well as according to power level (in the case of power control) or codec data rate (in the case of link adaptation). The same idea can then be used as in the case of DCA in Figure 3-4: Set “prices” for the resources so that users maximizing their utility take “good” resources in the case of lower loads and “worse” resources (but where minimum users’ requirements still can be satisfied) in the case of higher loads. In this way “better” resources are left to the users with worse conditions (lower channel gains and data rates) in the case of higher loads in order to maximize the number of satisfied users (“fairness”) and the data throughput is maximized in the case of lower loads (“efficiency”).

If the speed of the users is relatively high, the **“optimal resources” for the users could change relatively frequently due to fast changes in channel gain**. In that case HO (intracell, intercell or intersystem) or codec change (in the case of link adaptation) must be often performed in order to **“reshuffle” the users to the new “optimal” resources**. If the signalization overhead for “reshuffling” i.e. HO or codec change is relatively high in a system, simple random channel allocation or default codec selection might be a better choice for the system than the algorithm described above.

In the following sections we formalize the principle of “backward deduction” and state dependent price decisions to use it also for the other RRM algorithms in a unified manner.

### 3.7 Relevant States and State dependent "Pricing" Decisions

In this section we investigate how "prices" should be set in a cell in order to optimize resource usage in the cell in each (relevant) cell's state using "backward deduction" heuristics described above.

The state of the users or cells conveys all information from users' (cells') history necessary to take actions (Markov property) i.e. make "price" and resource allocation decisions. We define cells' and users' states as follows:

- The **cells' state** is defined by the (statistics of) **the number of users** in the cells and **interference.**
- The **users' state is** defined by the **channel gain and interference** of the users as well as **power consumed** and **data rate** achieved by the users.

For interference characterization, we can use interference distribution which can be obtained by cell statistics, or in the case of Normal interference distribution [83], [6], we only need mean value $m_I$ and standard deviation $\sigma_I$. As a measure of the quality of interference estimation $d_I$ - ratio between interference standard deviation $\sigma_I$ and mean $m_I$ can be used. $d_I$ in general decreases with increase in the number of interfering users. For example, if the interference stems from $n$ users on the same channels with independent and identically distributed interference contributions, than standard deviation of the total interference increases with $\sqrt{n}$ and mean increase with n. Consequently, $d_I(n)$ decreases with $\sqrt{n}$ and interference estimation becomes better with increasing number of interfering users, i.e.:

$$d_I = \frac{\sigma_I}{m_I} \sim \frac{1}{\sqrt{n}} \tag{3.16}$$

This means that interference based RRM algorithms like Min I DCA or CIR-based PC, make less reliable interference estimation in the case of (very) low loads.

As already described, the state of a cell is defined by (a statistic of) traffic and interference in the cell. In order to simplify analysis, we reduce two-dimensional state description (traffic, interference) to one-dimensional **load** defined by the number of "useful" channels - channels free and have relatively low interference (below a certain threshold). We differentiate between three possible cell states: **"low", "medium" and "high"** load. This state classification is defined rather in "**fuzzy**" manner according to the states found to be "relevant" by simulations. A state is "relevant" if a relevant performance improvement (at least 5-10%) can be achieved by using another RRM algorithm in this state in comparison to the other states. It turns out (see simulation results in 4) that **for the most RRM algorithms and situations only 3 states are relevant: low, medium and high**. These states are also typical in the praxis according to users behavior during the day: **low** load (for example late in the night), **high** load “busy hour” (for example at noon and after the work) and normal or **medium** loads (in all other times).

A strategy is a mapping from the state space into action space i.e. **a strategy defines for each state which actions should be taken**. For example, cell strategy defines which “prices” (parameters of users’ utility function) are set for each state of the cell. Since we assume that the cell knows the utility function of the users, setting “prices” means enforcing the users to apply those Scheduler, DCA, PC, HO, AC algorithms or “rules”, which are "optimal" for the given state of the cell. In this work we discuss time invariant or **stationary strategies** i.e. the same actions are taken in the same states independently from the time step.

In our work we do not search for absolutely “optimal” but for “satisfactory” strategies assuming “bounded rational” (with limited information available) users, which make the decisions based on the “rules of the thumb” like: “Take the channel with medium interference if the load is low”. We propose the following basic strategies for each of the relevant states (“low”, “medium” and “high”):

**Low load:** In the case of low loads, the “prices” for resources (cell, channels and power) can be set so that users can apply **“greedy”** algorithms (**maximize users’ CIR or data rate**): “Send with maximal power”, “make HO to best cells (with highest channel gain)” or “take best channels (with minimal interference)”. All the users in the system should be satisfied (enough free channels with low interference) and provider revenue should be maximized by maximizing users’ data rate (see (3.13)). Since interference estimation can be relatively bad (low load – relatively high interference fluctuations, see (3.16)) and almost all channels have low interference, for very low loads the channels could also be allocated randomly.

**Medium load:** In the case of medium load **the “prices” for cells, channels and powers increases with load**. The higher the load, the lower CIR-threshold of CIR-based PC (as long CIR-threshold $\geq CIR_{min}$) and the lower water-filling level in the case of water-filling power control. Also, some over-loaded cells can prevent new users from starting the call in the cells by increasing the cell “prices”. Channel “prices” can be set proportional to interference in order to enforce users to choose channels with lower interference more often. Interference estimation is good i.e. interference fluctuations are relatively low due to moderate load.

**High load**: In the case of high load the “prices” for resources i.e. cells, channels and power are such that **users take the resources which provide them with minimal required signal**

**quality** (data rate) **in order to leave other "better" resources (with lower load and interference) free for users with worse conditions** (channel gain, interference). In this way the number of satisfied users in the system i.e. provider revenue should be maximized (if $w_s >> w_i$ and $w_d >> w_i$ in (3.13)). "Prices" for cell entrance (AC) for the new users are high – almost only users already in the network (HO users) can access the cells. In order to provide each user with a satisfied signal quality (CIR), those users who have better channel gain (G) should take channels with higher interference and vice versa i.e. use ARP DCA (see section 2.4).

In following section we describe heuristics behind "price" decisions i.e. the influence of "prices" on system dynamics and equilibrium states.

## 3.8 Heuristics behind RRM Decisions

In this section we describe **impact of "price" settings on system dynamics** i.e. minimization of the maximal channel eigenvalues (in the case of PC) **and the achievable equilibrium states** i.e. the shapes and choice of points in the "capacity regions" of the network or cells.

### 3.8.1 Impact of RRM decisions on System Dynamics

In subsection 2.2.3 we saw that the largest eigenvalue of the channel matrix has a crucial role in convergence and maximum achievable CIR-values of the iterative CIR-based PC. The lower the maximal eigenvalue of the channel matrix the higher achievable CIR and faster convergence of the iterative PC. In this subsection we investigate the influence of RRM algorithms on eigenvalues on channels i.e. how maximum eigenvalues can be minimized on as many channels in the system as possible by setting weights ("prices") in users' utility functions.

In subsection 2.2.3 we have shown that iterative CIR-based power control is represented by the following matrix form (2.8):

$$\mathbf{p}(\mathrm{n}+1) = \mathbf{H}\mathbf{p}(\mathrm{n}) + \boldsymbol{\eta} \tag{3.17}$$

where $\boldsymbol{\eta}$ is the Nx1 noise vector with the element $\eta_i/g_{ii}$ at the i-th position. **H** is NxN "normalized link gain matrix" or "channel matrix" such that:

$$h_{ij} = \begin{cases} CIR_i \dfrac{g_{ij}}{g_{ii}} & if \;\; i \neq j \\ 0 & if \;\; i = j \end{cases} \tag{3.18}$$

where $CIR_i$ denotes the CIR threshold of the user i. In similar manner we can represent "water-filling" PC (see section 2.3) in matrix form by the following equation:

$$\mathbf{p}(n+1) = \mathbf{Ap}(n) + \mathbf{b} \tag{3.19}$$

where **b** is the Nx1 "control" vector with the element $1/\lambda_i$ - $\eta_i/g_{ii}$ at the i-th position and **A** NxN "normalized link gain matrix" such that:

$$a_{ij} = \begin{cases} -\dfrac{g_{ij}}{g_{ii}} & if \;\; i \neq j \\ 0 & if \;\; i = j \end{cases} \tag{3.20}$$

In the same manner as in subsection 2.2.3 we conclude that the convergence of the "water-filling" PC given by equation (3.19) is determined by the largest eigenvalue (in absolute sense) of the matrix **A** called spectral radius of **A** and denoted ρ(**A**).

We can combine equation (3.17) and (3.19) in the following general form:

$$\mathbf{p}(n+1) = \mathbf{Sp}(n) + \mathbf{c} \tag{3.21}$$

Where **S** is the "channel matrix" and **c** is "control vector". **S = H** and **c** = η in the case of CIR-based PC. In the case of "water-filling" PC: **S = A** and **c** = **b.** From subsection 2.2.3 we know that convergence behavior of PC is determined by the largest eigenvalue of the matrix **S** (spectral radius of **S**) denoted with ρ(**S**). ρ(**S**) must be lower than 1 in order that PC converges and the lower ρ(**S**) the higher speed of convergence of PC.

**If the largest eigenvalue of the matrix S is lower than 1, the PC in (3.21) is a "contraction operator" in the power space** i.e. the distance between the two consecutive power vectors decrease as time increases [31]. PC converges to an equilibrium (if the largest eigenvalue of the matrix **S** is lower than 1) even when the users **asynchronously just increase (decrease) their powers for a step** (say 1 or 2 dB) **if their actual CIR is lower (higher) than required CIR-threshold** (like in the case of CIR-based PC in UMTS, see [21]).

Note that the equation (3.21) describes power changes due to iterative PC of all users on **one** orthogonal channel in wireless systems. We have N such equation in the system, one for each one of N orthogonal channels. Also, in order to provide convergence of iterative PC and increase the convergence speed, **efficient RRM algorithms should minimize maximal eigenvalues of the matrix S in the equation (3.21) on as much channels as possible**. Now we define the influence of RRM algorithm on the matrix **S**:

- **Channel allocation decides to which matrixes** S($k$) (channel matrix of channel $k$, $k = 1 \dots N$) **a row and a column are added due to allocation of new users on the channel $k$**. When users left the cell or ended their calls the row and column of the matrix is deleted.

- **Handover and admission control change also matrixes** S($k$) ($k = 1 \dots N$). For example, **path-loss-based HO** (without HO margin) **ensures that ration $g_{ij}/g_{ii}$** (see (3.18) and (3.20)) **is always less than one** (up to HO-margin). If this ratio gets higher than one (plus HO-margin) for some base station j, path-loss based HO is performed from base station i to the base station j, which causes that some channel matrixes change.

- **Power Control** decides about entries of the control vector **c**(k) ($\lambda$) in the case of water filling PC, and elements of matrix **S**(k) (***$CIR_{thr}$***) in the case of CIR-based PC (see (3.18)).

**The heuristics behind RRM algorithms is to select base station, channel and *$CIR_{thr}$* in a way that maximal eigenvalues of the matrixes S(*k*) are minimized for as many channels *k* in the system as possible**, since the lower the eigenvalue ρ(**S**($k$) the more users can achieve their required signal quality and the faster the convergence of the PC on the channel $k$.

We use the following linear algebra results [98] for the upper bounds on the eigenvalues of the matrix **S**:

$$\rho(\mathbf{S}) \le \max_i \left\{ \frac{(CIR^i_{thr}) \sum\limits_{j,j\neq i}^{N} p_j g_{ij}}{p_i g_{ii}} \right\} = \max_i \left\{ \frac{(CIR^i_{thr})}{CIR_i} \right\} \tag{3.22}$$

And following bound:

$$\rho(\mathbf{S}) \le \max_i \left\{ \frac{(CIR^i_{thr}) \sum\limits_{j,j\neq i}^{N} g_{ij}}{g_{ii}} \right\} \tag{3.23}$$

Note that $CIR^i_{thr}$ is the required CIR-threshold for the user i in the case of CIR-based PC and $CIR^i_{thr} = 1$ if water-filling PC is used. According to (3.23) water-filling PC has an advantage over the CIR-based PC since **its convergence is not influenced by choice of water-filling levels**. **In the case of CIR-based PC choice of *$CIR_{thr}$* has an influence of the maximal eigenvalues** (see (3.23)) **and thus on PC convergence**.

In equation (3.22) users' powers ($p_i$, $i = 1 \dots N$) appear and in equation (3.23) they do not. Powers can assume, naturally, only non-negative values. The advantage of (3.22) in comparison to (3.23) is that relatively low signalization overhead is required to estimate maximal eigenvalue of the channel matrix **S**. For each user possible eigenvalue can be estimated as a ratio between required CIR-threshold and CIR on the channel. The term in nominator of (3.22) represents the interference, which can be measured at the receiver and signalized back to the transmitter. The channel gain $g_{ii}$ of the user i can also be measured at the receiver and signalized back to the transmitter or, in the case of a TDD systems, obtained directly at the transmitter assuming that the time elapsing between UL and DL transmission is smaller than channel coherence time. Own power $p_i$ (and required $CIR_{thr}$ in the case of CIR-based PC) is/are naturally available at the transmitter itself. Furthermore, statistics from the cell and mobile can be used to estimate expected channel gain and interference on the channel. In this way **maximal possible eigenvalues on the channels can be estimated in a distributed manner with relatively low signalization overhead**.

The advantage of (3.23) is that eigenvalue estimation is independent from users' powers. The disadvantage of (3.23) is that channel gains from other users are not known in a distributed RRM implementation. But equation (3.23) can be used for principal investigation of possible eignevalues of channel matrix **S**. For example, if HO does not provide that link to the own base station i is always stronger than to a neighbor base station j i.e. if $g_{ij} \geq g_{ii}$ for some i and j, than $g_{ij}/g_{ii} \geq 1$ and according to (3.23) maximal eigenvalue might become higher than 1 and PC might not converge.

RRM can influence resource allocation and thus maximal eigenvalues in a cell by setting the weights ("prices") in users utility function (3.15). In Figure 3-5, an impact of the components of the utility function on eignevalues of the matrixes ρ(**S**) is represented.

**PC**

$$U = W_G G_{mean} - W_I I_{mean} - C$$

$$\rho(\mathbf{S}) \leq \max_i \left( \frac{(CIR_i) \sum_{j, j \neq i}^{N} (p_j) g_{ij}}{(p_i) g_{ii}} \right)$$

**Figure 3-5: The role of utility function in minimizing maximal eigenvalues**

The roles of the components of the utility functions on maximal eigenvalues are (see Figure 3-5):

- **Channel gain** ($W_G$, $G_{mean}$) Maximizing channel gain (minimizing path-loss) decreases the denominator in (3.22) and thus reduces the maximal eigenvalues on the channels. This in turn increases the convergence speed of PC as well as the highest CIR achievable on the channels. That is why many HO algorithms are path-loss based (minimize path-loss i.e. maximize channel gain). This is especially important **in the case of re-use 1** (all cells can use all channels) and **iterative PC**, when **HO should provide that link to its own base station i is always stronger than to a neighboring base station j**. If $g_{ij} \geq g_{ii}$ for some $i$ and $j$ than $g_{ij}/g_{ii} \geq 1$ and according to (3.23) maximal eigenvalue might (in (3.23) an upper limit is given) be higher than 1. In that case PC would not converge. In order to prevent this, cells can set weights as follows: $W_G=1$, $W_I=0$ *and* $C=0$ (see 3.9.4). In that case the utility function (3.15) reduce to $U = G$ i.e. the users part of HO select always the cell with maximal channel gain i.e. performs path-loss based HO (see 2.6.1).

- **Interference** ($W_I$, $I_{mean}$): The lower the interference, the lower the nominator in (3.22) and the lower the possible eigenvalues on the channel. That is why many DCA algorithms are interference based i.e. take the channel with lowest interference. But, since maximal eigenvalues should be minimized on **as much channels as possible,** for higher loads ARP DCA (2.5.2.4) could help **maximizing the number of channels with maximal eigenvalues lower than 1.** For example, by assuring that each user gets a channel where possible CIR is slightly above its CIR-threshold, maximal eigenvalues can be kept below 1 (and thus insure convergence of PC) on as much channels as possible (see (3.22)).

- **Cost (*C*):** It is not only the goal to minimize eigenvalue on one channel but on as many channels as possible in order to support as many users as possible i.e. **"balance" eigenvalues on different channels**. In the case of higher cell load the users and interference on the channels should be distributed equally to achieve approximately the same eigenvalues (but low enough to enable achieving $CIR_{thr}$ for as many users as possible) on all channels. **Term "*C*" in the utility function enables us to "enforce" load dependent distribution of users among the channels (or cells)**. *C* could be set (by a network part of DCA) according to the load of the channel and $W_G$ and $W_I$ could be set to zero. **AC can also set a higher cost *C* for the new users than for the users already in the network**. In this way some new users would not access the network in order to protect the users already in the network from too high interference. This could be especially useful in the case of **hierarchical networks** or **inter-system HO**, where the **users could be shifted from overloaded layers (systems) to the layers (systems) with lower loads by appropriate choice of the parameter *C***. Since the layers or systems use orthogonal channels, among the layers (systems) we do not have a problem that it should be provided for each $i$ and $j$ that $g_{ij} \geq g_{ii}$ as in the case within one layer or system and re-use 1 (see above).

### 3.8.2 Smart Antennas and Eigenvalues of Channel Matrixes

As described section 2.7 Smart antennas (SA) decrease gain from interfering users ($g_{ij}$ and $g_{ji}$ in (3.18) and (3.20)) and increase gain to the desired user ($g_{ii}$), also the ratio of these two gains ($g_{ij}/g_{ii}$) decreases due to SA deployment. This means that **each non-zero entry in channel matrixes decreases due to SA** (see (3.18) and (3.20)). In this way the row or column sum is reduced and **the largest eigenvalue of the matrix is also reduced** due to SA according to (3.22) or (3.23). Also, SA have similar effects as RRM algorithms – reduce largest eigenvalue of channel matrix for as many channels as possible. The advantage of SA is that the reduction of the largest eigenvalue on one channel is independent from other channels and from users' powers i.e. **decrease of eigenvalue (interference) on one channel is not paid by increase of eigenvalue (interference) on another channel (layer, system)** as in the case of DCA or AC.

Reducing the gain of interfering users ($g_{ij}$) and increasing the gain of desired user ($g_{ii}$ in equation (3.23)) by SA has an effect of **"decoupling" the users**. The relative influence of one user on another is decreased, and **"social" behavior of the users and thus sophisticated RRM algorithms become less important in presence of SA than without SA**.

### 3.8.3 Impact of RRM decisions on Capacity Region

In this subsection we investigate the "equilibrium states" i.e. after the users reach their equilibrium powers due to PC. We provide a "geometrical" insight into the role of different RRM algorithms by showing the **impact of RRM decisions on the "shapes" of the achievable capacity regions and choice of the points in the capacity regions**.

To this aim we obtain from (3.13) as special case the network utility *U(network)* as maximum of the weighted sum of data rates (as long the users are satisfied) i.e.:

$$U(network) = \max\left[\sum_{i=1}^{N} w_i R_i\right] \tag{3.24}$$

Under constraint:

$$\sum_{i=1}^{N} P_i \le P_{\max} \tag{3.25}$$

$$\Pr(R_i \le R_i^{\min}) < outage_i \quad i = 1..N$$

where $w_i$, $R_i$ and $P_i$ is the weight (or priority), data rate and power of user $i$ respectively. $P_{max}$ is the maximal power allowed in the cell, $R_i^{min}$ is the minima data rate required for the user $i$ and $outage_i$ is the maximal outage probability for the service class of the user $i$.

We assume that an optimal resource allocation should at least satisfy the following requirement ("Pareto optimality" condition [64]): **An allocation is Pareto optimal if there is no other allocation in a given set, which can make one user better off without making some other users worse off**. We call the set of possible data rates in a cell the cell capacity set. Consequently, a data rate allocation from the cell capacity set is Pareto optimal, if there is no other data rate allocation in the cell capacity set which makes at least one user better off (with higher data rate) without leaving the other users worse off (with lower data rate). The set of all Pareto optimal allocations (points) in the cell capacity set is called **Pareto boundary**. **The cell capacity cell is a convex set** (set C is a convex set if $\forall x, y \in C,\ \lambda x + (1-\lambda) y \in C \quad \forall \lambda \in [0,1]$) as shown in [58] for ergodic capacity and in [59] for outage capacity. **According to the supporting hyperplane theorem** [64] **the network utility** (equation (3.24)) **achieves its maximum only for the points on the Pareto boundary.** Furthermore, each Pareto optimal point is the result of maximization of (3.24) for some weights (priorities) $w_i$ ($i=1,N$) setting. Weights $w_i$ also define a family of hyperplanes in N dimensional space according to the equation (3.24) i.e. each value of the equation (3.24) for fixed $w_i$ (i=1,N) defines one hyperplane. The hyperplane for which the equation (3.24) has a maximum is called the supporting hyperplane [56] and represents a tangent on network's (cell's) capacity set (see Figure 3-6).

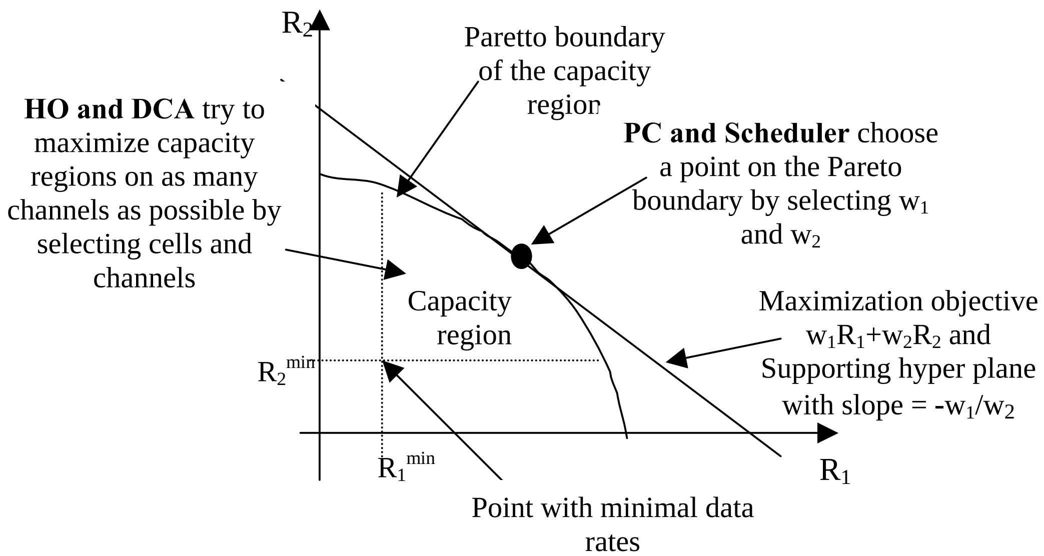



**Figure 3-6: An example of the Capacity region, Pareto boundary, Supporting hyper plane in $R^2$ and the role of RRM algorithms**

Since the setting of weights $w_i$ is the task of Scheduler (see 2.4 and 3.9.2) and data rates $R_i$ depend on CIR-threshold (set by outer loop PC, see 2.2.4), we can say that **"optimal"**

**Scheduler and PC (which maximize cell's or network's utility) should choose a point on the Pareto boundary in the cell's capacity set.**

This is an example of the application of **The Second Welfare Theorem** from the economic theory [64]**,** which states that **each Pareto allocation** can be achieved in a **distributed manner** (in a competitive market) **by appropriate price settings**, if the users utility functions and production set are **convex**.

In our case the conditions of The Second Welfare Theorem are fulfilled: The users' utility function defined in (3.1) is linear (also both convex and concave) in data rate and the "production set" i.e. the set of the possible data rates in a network or the cell is also convex as showed in [58] and [59]. The precondition for the application of The Second Welfare Theorem is that **there are no externalities** i.e. there is a market and a price for each commodity, which might be "good" or "bad" for some users. That is why **it is important that the power in mobile networks has a "price", otherwise the power is externality** (generate interference to other users) and Pareto optimal allocation in a distributed manner is not possible. Note that even if there is the "price" for the power it is oft not possible in praxis to "search" the Pareto optimal allocation in a distributed manner by a long process of "tatonement" among the users, which **would require enormous signalization effort** in a wireless network. As stated in section 3.2, we assume that **the users do not search for absolutely "optimal" but "satisfied" solutions**. The Second Welfare Theorem provides us with the result that in the ideal case the "optimal" resource allocation is possible with the "pricing" approach.

Above discussion is done for a "snapshot" of a system, when the capacity set is fixed (after PC convergence) and users do not move, arrive in or left the system. But the **capacity sets** (available rates, see Figure 3-6) **on the channel changes during the time**. With each allocation of a channel to a user by DCA or change of the cell by HO, the shape of the capacity set changes. **The role of a DCA and HO is to select the "optimal" channels and cells for the users in order to keep the capacity regions as large as possible on as many channels as possible**. The "shape" of the capacity region is also important because the shape defines the minimal data rates achievable for each user. The goal is that in each dimension i (for each user i) the shape of the capacity set is such that $R_i^{max} \geq R_i^{min}$, where $R_i^{max}$ and $R_i^{min}$ are maximal achievable and minimal required (respectively) data rates for the user i.

Note that AC influence HO decisions by setting cell "prices" and so AC has also an impact on the size and shape of the capacity regions.

## 3.9 RRM algorithms based on “Pricing”- Framework

In this section we show some examples of the RRM algorithms based on the “pricing” framework. We show also how RRM algorithms from chapter 2 can be obtained as special cases of the above “pricing” framework, by appropriate weights (price) settings in the utility function.

### 3.9.1 Power Control

As described in section 2.2, PC sets the power according to Water-filling or CIR-based rule i.e.:

$$p = \begin{cases} \left(\dfrac{1}{\lambda} - \dfrac{I^{'}}{g}\right)^{+}, & if\ water - fiiling\ PC\ is\ used \\ CIR_{thr}\dfrac{I^{'}}{g}, & if\ CIR - based\ PC\ is\ used \end{cases} \tag{3.26}$$

The water-filling level ($1/\lambda$) or $CIR_{thr}$ are set for each user according to cell’s and user’s state i.e. according to required data rate (signal quality) for the user’s service, load in the cell or network and available power of the user and the cell (in DL).

As described in section 3.3 and 3.5.2, $\lambda$ can be interpreted as a “price” per unit power and unit bandwidth. A user has an incentive to increase its power as long as the increase in its utility function $U$ (per unit power) are greater or equal to the prices for the power i.e.:

$$\lambda \le \frac{dU(p)}{dp} \qquad \lambda = max\,(cell\ price, own\ price\ for\ battery\ power) \tag{3.27}$$

The price of the power is the maximum between the cell’s power price and user’s own power “price”, which shows how user values its own battery power.

Setting $U(p)$ proportional to data rate (Shannon’s formula) we obtain from (3.27) the following equation for the maximal power price $\lambda_{max}$:

$$\lambda_{\max} = \frac{dU(p)}{dp} = \frac{d\left[\log\left(1+\dfrac{pg}{I^{'}}\right)\right]}{dp} = \frac{1}{\dfrac{I^{'}}{g} + p} \tag{3.28}$$

Also, according to equation (3.28) the maximal price, which is user ready to pay of a unit of power, decreases with the increase in interference and power, and the power price increase with increase in channel gain $g$. This was expected, because the lower interference and the higher the channel gain more “valuable” the power for the user is i.e. the higher data rate (better signal quality) can be achieved. Consequently, **the better user’s channel (high channel gain and low interference), the more is user ready to “pay” for an additional power.** Since users’ utility is concave in power (logarithm is a concave function), the higher the power the lower the relative increase in data rate per unit power.

From the network's point of view, the price per unit power is inverse proportional to the increase in the cell's utility per unit power i.e.

$$\frac{1}{\lambda} = \frac{dU_{network}(p)}{dp} \tag{3.29}$$

Consequently, the higher the network utility of additional power allocated to a user, the lower is the price to the power. Consequently, the water-filling level (inverse proportional to power "price") or CIR-threshold of a user should increase proportional to the increase in the utility of the network by allocating additional power to the user. Since networks utility function $U_{network}$ is defined in (3.13), this means that with additional power allocated to a user, the number of satisfied users $N_s$ or data rates $R_i$ increases in (3.13) (or at least stay on the same level) and constraints (3.14) are satisfied. But, when by allocating additional power for a user the data rate increase for the user can not compensate decrease in data rates of the other users (due to additional interference) or, even worse, it causes decrease in the number of satisfied users, the water-filling level should becomes zero or negative and the user should not get any additional power.

According to microeconomic results [64] the price change by **"tatonement" process** can be roughly described by the following equation:

$$\frac{d\lambda(t)}{dt} \sim excess\ \ demand \sim \max\left[(used\_power - total\_power), (outage - max\_outage)\right] \tag{3.30}$$

Equation (3.30) says that the "price" decreases if the amount of "free" resources increases and vice versa. Unfortunately, economic theory can not help us very much by determining exact "trajectory" of the price change. According to [64] **economist are good in describing stabile (equilibrium) states, but can not say very much about system dynamics**. "There are a lot of ways to be in disequilibrium, but only a few to be in equilibrium" [64].

Summarizing, in the case of lower loads the users' and network's utility increase with the increase of users' power and power price should be low (see Figure 3-7). Consequently, the users' water-filling level or $CIR_{thr}$ increases. The higher the load, the lower relative increase in users' and network's utilities i.e. the lower the water-filling level or $CIR_{thr}$ should be. Finally, at the high load, water-filling level or $CIR_{thr}$, decreases to the minimal level needed for the user to satisfy its minimal data rate (signal quality) requirements, in order to minimize interference in the network and maximize number of satisfied users. The increase/decrease of water-filling level or $CIR_{thr}$ can be performed by outer loop PC (see 2.2.4).

*For each user:*

*Power_price ~max[1/(increase in network's utility), value of the battery power)*

*Case* **Load**:

**Low:** *Increase power (water-filling level or $CIR_{thr}$) as long as increase in user's utility higher than the power price,*

**Medium-High**: *Decrease the water-filling level or $CIR_{thr}$ proportional to load i.e. inverse proportional to power price.*

**High:** *Set the water-filling level or $CIR_{thr}$ to minimal values needed to achieve the required data rates or signal quality.*

*End*

**Figure 3-7: Load dependent PC**

### 3.9.2 Scheduling

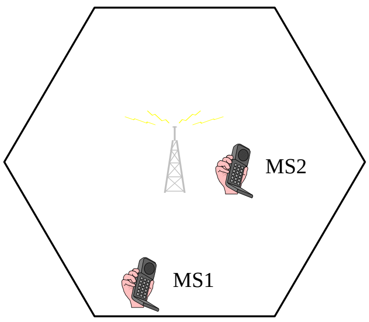
The scheduling algorithms based on "pricing" select users with maximum utility $U_i = W_G G_i - W_I I_i - C_i$ (see (3.15)) to be scheduled for the next transmissions. Note that scheduling is usually done only for the services with not stringent delay constraints i.e. for non-real time (NRT) users. By the appropriate "pricing" i.e. setting of users' utility function parameters $W_G$, $W_I$ and $C$ different scheduling algorithms from section 2.4 can be implemented. For example:

**"Fair" scheduling: $W_G$,= -1, $W_I$ = -1 and $C_i$~ $R_i$ => $U = I – G – k*R_i$.** The users with worse channel conditions i.e. with lower channel gains $G$ and higher interference $I$ (for example users at the cell border) and with lower data rates already achieved ($R_i$), is scheduled more often. In this way the users with worse channels and lower data rates get more transmission time, which enables them to obtain at least their minimal data rate during the session duration. In general users farther from the base stations (have lower channel gain) are scheduled more frequently than the users near the base station (have higher channel gain) (see Figure 3-8 a)). This algorithm also **maximizes the number of satisfied users in the cell (network)** and could be used when the revenue from the satisfied users is much higher than from data rate i.e. $w_d >> w_i$ in (3.13).

**"Greedy" scheduling: $W_G$,= 1, $W_I$ = 1 and $C_i$ = 0 => $U = G - I$.** The users with better channel conditions i.e. with higher channel gain $G$ and lower interference $I$ (for example, users near to the BS) are scheduled more often. In this way the users with better channels get more transmission time, which enables them and their cells (network) to obtain relatively high total data rates. On the other hand, the users with worse channels are doubly “punished”: By bad channels and lower scheduling times, and sometimes they cannot even achieve their minimum data rates. Since the users near the base stations usually have higher channel gains and lower interference than the users near the cell borders, they are scheduled more frequently than the users farther from the base station as represented in Figure 3-8 b). This scheduling policy **maximizes total data rate in the cell (network)** and could be used when the revenue from data rate is much higher then from the satisfied users i.e. when in (3.13) $w_i >> w_d$ (at least for some *i*).

**“FIFO” scheduling: $W_G$,= $W_I$ = *0* and $C_i$ ~-*T* => *U* = *T*.** The users with the longest time spent in the queue (*T*) are scheduled for the transmission. Also, **all users get the same portion of the scheduling time** independently from their channel gain (distance from the base station) (see Figure 3-8 c)). With these scheduling algorithms, the data rate and number of satisfied users are between “Fair” and “Greedy” scheduling and could be used in the case of (uncritical) low load or when the number of satisfied users and total data rate are of equal importance.

**“Priority-based” scheduling: $W_G$,= $W_I$ = *0* and $C_i$ = -*priority* => *U* = *priority*** The users with the highest priority are scheduled first. The priority can be set according to the users’ contract with a provider and/or service type requirements. For example, the users with harder delay constraints should be given higher priority than users with less stringent delay constraints. Users can also have priority according to their data rate: 128 kbps users can get 1/3 of priority of 384 kbps users in order to provide both classes of users with their required data rates.




a) „Fair“ scheduling:

| | | | | | | |
|---|---|---|---|---|---|---|
| MS1 | MS1 | MS2 | MS1 | MS2 | MS1 | MS1 |

a) „Greedy“ scheduling:

| | | | | | | |
|---|---|---|---|---|---|---|
| MS2 | MS2 | MS2 | MS1 | MS2 | MS2 | MS2 |

c) „Round-Robin“ scheduling:

| | | | | | | |
|---|---|---|---|---|---|---|
| MS2 | MS1 | MS2 | MS1 | MS2 | MS1 | MS2 |

**Figure 3-8: Some scheduling algorithms [38]**

Also a mixture of the above scheduling algorithms are possible, for example, setting $w_{G,}=-1$, $w_I=-1$ and $c(i)= 1/T_i$, the user(s) with the worst channel conditions and/or the longest time spent in the queue are scheduled first.

### 3.9.3 Dynamic Channel Allocation

After intracell/intercell HO or at a new call begin of a user, a channel should be allocated to the user. To this purpose DCA evaluates the “utility” of each “admissible” channel for the user in the cell according to the user’s utility function (see equation (3.15)):

$$U(c) = W_G G(c) - W_I I(c) - C(c) \tag{3.31}$$

Where $c$ denotes channel, $I(c)$ is the average interference (in logarithmic measure) on the channel $c$, $G$ is the average channel gain (in logarithmic measure). An admissible channel is a free channel on which user’s constraints (3.12) can be satisfied. The set of admissible channel is called the admissible set and denoted with A.

Not that channel gain $G$ in (3.31) is usually user specific i.e. depends on distance between the mobile user and the base station, but can also be channel specific if different channels have different frequencies. $C$ is the cost of the channel that comprises other factors, which might influence choice of the channels like channel load i.e. the number of users using the channel currently in the cell or the number of codes already used on a channel ((timeslot, frequency)-pair in a TDD system). $W_I$ and $W_G$ are the weights of the interference and channel gain respectively.

**User specific DCA allocates a channel to the user** from the set A of admissible channels **with maximum utility** i.e.:

$$channel = \arg\max_{c \in A} \left[ U(c) \right] = \arg\max_{c \in A} \left[ W_G G(c) - W_I I(c) - C(c) \right] \tag{3.32}$$

**The “cell specific” DCA set “prices” (weights $W_G$, $W_I$ and $C$) according to the state of the cell**.

By appropriate parameter settings in (3.32), we obtain different DCA algorithms from chapter 2 as special cases (see Figure 3-9):

**Random DCA**: *$W_G = W_I = 0$, $C$ = random => $U$ = random (take channel randomly)*

*Example of usage: In the case of very low or very high loads, when all channels are equally "good".*

**Priority-based DCA**: *$W_G = W_I = 0$, $C$ = -priority => $U$ = priority*

*Example of usage: In the case of low-medium loads to establish "channel segregation" i.e. neighbor cells should use different channels.*

**Min I DCA**: *$W_G = C = 0$, $W_I = 1$ => $U = - I$ (take channel with minimum interference I)*

*Example of usage: In the case of low till medium loads, when neighbor cells can use different channels for the most of the time.*

**ARP DCA**: *$W_G = C = 0$, $W_I = -1$ => $U = I$ (take channel with maximum interference I)*

*Example of usage: In the case of high loads, in order to provide as much users as possible, with at least minimum signal quality.*

**Figure 3-9: Different DCA algorithms as special cases of "price" settings**

Note that in Figure 3-9 for ARP DCA, maximal utility is obtained for a (**admissible**) channel with maximum interference on which still. Too "bad" channels (where minimal CIR can not be achieved, see (3.11)) are not taken into account i.e. they do not belong to the set A of "admissible" channels (see (3.32)).

### 3.9.4 Admission Control and Handover

In this subsection we describe how admission control (AC) and handover (HO) can be modeled by the use of our "pricing" framework for RRM.

The role of **AC** is to **admit or reject** new users in the cell and to indicate, by **cell "prices",** how **desirable the admission of the users in the cell is.**

**Handovers** decide about **system (layer), cell (intercell HO) or channel (intracell HO) change** taking into account AC "prices" for the cells (systems, layers, channels).

With each HO, the shape of the capacity sets (available rates, see Figure 3-6) on the channels from and where HO is done, is changed. The role of AC is to influence by "prices" the shape of the capacity sets (one on each orthogonal channel) and consequently the achievable rates in the cells. Furthermore, AC should prevent admitting of new users when the systems is on its limits i.e. if, for example, the maximal eigenvalues on all channels are close to 1 (see subsection 3.8.1). HO selects the "optimal" cell for a user or initiate channel change in order to keep the signal quality (data rate) for the user at least at the minimal required level as long as possible.

According to our "pricing" framework **the task of the AC is to set "prices" weights $W_G$, $W_I$ and $C$ in the users' utility function (3.15) and the role of HO is to select the cell (system, layer), which maximize users' utility function**. Also, AC represents "cells' specific" part of the algorithms and HO the "users' specific" part. In the following, we describe some AC and HO algorithms based on our "pricing" framework.

#### *3.9.4.1 Admission Control*

Admission Control (AC) has the following tasks to fulfill by an access of the user in the cell:

**Check if the service constraints are satisfied** for the user in the cell i.e. check equation (3.12) for each user asking for an admission according to the service type of the user. For example, in the case of a RT service AC checks if probability that CIR fails below certain threshold ($CIR_{thr}$) is lower than a certain quality threshold for the service $P_{out}$ (see also (3.11) and section 3.3):

$$\Pr(CIR < CIR_{thr}) < P_{out} \tag{3.33}$$

Above probability can be calculated knowing the probability density function (statistics) of the CIR in the cell. Equation (3.33) also means that the maximal available transmit power $P_{max}$ is not enough to reach $CIR_{thr}$ . By setting the parameters in users' utility function (3.15) as follows: $W_G = W_I = 1, C = - P_{max},$ we obtain maximal expected users utility as $U_{max} = P_{max} + G - I$. Assuming log-normal CIR-distribution in the cell equation (3.33) reduces to the following condition

$$U_{\max} \geq CIR_{thr} + k(Pout) * \sigma_{CIR} \tag{3.34}$$

where k depends on the required outage probability, for example for $P_{out} = 98\%$, k $\approx$ 3.

For a NRT service usual requirement is that the expected data rate must not be below certain threshold $R_{min}$ - say 10% of a nominal data rate or that the error probability lies below certain threshold $BER_{thr}$, which is service dependent i.e.:

$$\Pr(BER) \geq BER_{thr} \Rightarrow U_{\max} \geq CIR(BER_{thr}) \tag{3.35}$$

where $CIR(BER_{thr})$ is the CIR required to achieve certain $BER_{thr}$ and depends on coding and modulation used. $CIR(BER_{thr})$ can be estimated by simulations or real network statistics, or even calculated in certain cases.

**Also, the necessary conditions for admitting users in the cell reduces to check of user's maximal possible utility $U_{max}$ (utility with maximal power) against certain threshold**, as represented in (3.34) and (3.35). This threshold can be even increased by AC i.e. admission "prices" for the cell can be increased according to, for example, following values:

- **The number of users in the cell**: the higher the number of users in a cell the higher the AC "prices".
- **Interference** in the cell: The higher the interference in a cell the higher the AC "price" for the cell.
- **Service type of the user**: Admission "price" for a service decreases with service priority.
- **HO type: Handover users (users already in the system) should have generally lower AC "prices" than new call users** especially in the case of higher loads, since call dropping is usually regarded as a worse system failure than call blocking.
- **Signalization overhead** needed to perform HO. The higher the overhead, the higher the admission price.

Taking (3.34) and (3.35) into account, the role of AC can be finally described by the following rule: **Admit the user if its utility $U = G - I - C > 0$,** where $C$ is a function of cell load, interference, service type of the user and kind of HO (new call or existing call handover).

With the increase of "prices" for cell by AC, the effective radii of the cells decrease (less users access the cell due to HO) and vice versa. In this way the **"cell breathing"** effect [110] can be achieved, where **the effective cell area changes with the cell "prices"**. Since the "prices" can be set proportional to load, the cells with lower loads would have large effective radii and vice versa. Note that cell "prices" do not have to be signalized directly, cells can just reduce the power of their beacons proportionally to the cells' "prices" (measured in dB). In this way the gain $G$ to the cell as estimated by all users is reduced, which would have the similar effect on utility function evaluation as sending the "prices" directly.

#### *3.9.4.2 Handover*

The role of handover is **to select the cells where to start calls** in the case of new users, and to perform **cell (channel, system, layer) change** from the old to the new (target) cells (channel, system, layer) if the users are already in the system. The selection of the cells at the new calls request can be regarded as special case of intercell HO, where only target cells are selected.

The handover of a user is performed to the cell with the highest expected utility *U(cell)* for the user, from the cells which belong to the user's "admissible set" A (cells whose AC admit the user):

$$cell = \arg\max_{cell \in A_c} [U(cell)] = \arg\max_{cell \in A} [W_G G(cell) - W_I I(cell) + C(cell)] \tag{3.36}$$

By appropriate parameter settings in (3.36), we can obtain different HO algorithms from subsection 2.6.1 as special cases (see Figure 3-10):

**Pathloss-based HO**: $W_G = 1, W_I = C = 0 \Rightarrow U = G$ *(path-loss based HO)*

> *Example of usage: In the case of re-use 1 and PC in order to provide the convergence of PC.*

**CIR (quality-based) HO**: $W_G = W_I = 1, C = 0 \Rightarrow U = G - I$ *(CIR-based HO)*

> *Example of usage: When re-use > 1 and/or no PC is used or inter-system/inter-layer HO in order to maximize signal quality of the users.*

**Load-based HO**: $W_G = W_I = 0, C = load \Rightarrow U = - load$ *(load-based HO)*

> *Example of usage: By inter-system or inter-layer HO in order to balance the load.*

**Figure 3-10: Different HO algorithms as special cases of "price" settings**

Note that "too bad" cells (where minimal CIR can not be achieved) are previously eliminated by AC i.e. they do not belong to the set of "admissible" cells (see subsection 3.9.4.1).

In order to perform HO, **the expected gain of HO should be higher than HO costs** i.e. the difference between user's utility in the old cell and the new cell should be greater than handover costs (HO-margin, see subsection 2.6.1):

$$U(cell\,new) - U(cell\,old) > HO - margin \tag{3.37}$$

The parameter *HO-margin* reflects signalization and overhead costs for performing HO. For example, HO-costs can be contained in the parameter *C* of the users' utility function (3.15) for the candidate cells. For the new incoming users, **costs of the new call** (increase in interference to other users) can also be comprised in the parameter *C* in *U(cell_new)* like in the case *HO-margin*. In general, costs of the new call should be higher than *HO-margin*, since **users already in the system should have higher priority than new incoming users**. This is especially important for higher loads to provide at least users already in the system with sufficient signal quality.

### *3.9.5 Summary of the State-Dependent Algorithms*

The state-dependent algorithms choice, as described above is summarized in Table 3-1**:**

**Table 3-1: Load dependent RRM Algorithm Choice**

| Algorithms/Load | Low | Medium | High |
|---|---|---|---|
| **DCA** | Random | Minimum I | ARP |
| **PC** | Maximum power | Water-filling level or CIR-threshold inverse proportional to load | Use CIR-based PC with minimal required CIR-target |
| **AC and HO** | HO according to signal quality. Admission "prices" zero for new and HO users | Admission "prices" for new users increase with load | No new users are admitted in the system, make HO to layers or systems with lower loads |
| **Scheduler** | Round robin (to minimize scheduling overhead) or Greedy scheduling (to maximize data rate) | Scheduling priority increases with channel quality and time in the queue and decreases with users' data rate | Schedule the users with minimal data rates first ("fair" scheduling). |
| **Link adaptation** | Take codec or modulation schema, which enables highest data rate. | Take the codec (modulation schema) according to channel quality of the user. | Take codec (modulation schema) with minimum data rate for the service |

Table 3-1 can be built or changed according to simulation results (see chapter 4) or network statistics. In Table 3-1 we assumed that maximizing the number of satisfied users is much more important than maximizing total data rate. Users should apply a **"greedy" policy in the case of lower loads** i.e. users get resources according to their maximal "possibilities" (channel gain and interference) in order to maximize total network throughput. **In the case of high loads a "fair" policy should be applied**: Everyone gets resources according to their (minimal) needs in order to maximize the number of satisfied users. Between these to extremes, "prices" of resources increase with load and a trade-off is made between the number of satisfied users and total data rate. If provider preferences and utility function are different than we assumed, the same "pricing" framework can be used but with different "prices".

What are the numerical values for “large”, “medium” and “low” load, depends on the concrete system. Anyway, the load should be defined relative to the average number of unused channels in the systems, which depends on (hard) blocking probability (see Erlang B formula (2.23)). If the system contains packet data NRT users too, then all channels might be occupied at all times (see Figure 2-14 and (2.24)) and load can be defined according to the ratio of arrival rate and service rate or NRT and RT users.

The same reasoning can be in a straightforward manner enlarged to more possible values for load i.e. for more than three states. In the case of a higher number of states for both cells and users so-called Markov-Decision strategy [40] can be used to select “prices” and resources for each state. A Markov-Decision strategy takes into account the gains (utility) of each state and transition probabilities between states, to find ”optimal” decisions for each state. In order to calculate “optimal” strategies we need a representation of state (interference, channel gain, load) changes as a Markov Process. A method for derivation of a Markov chain from process statistics is described in [18].

## 3.10 Value of Measurements

Our “pricing”- and utility - based concept enables us not only to make PC, AC, HO and DCA decisions but also to **decide if measurements should be taken or not**. In general, measurements should be done when the utility of measurements i.e. expected capacity or data rate improvements due to use of measurements is higher than the costs of measurements.

This means, more formally, that if the difference between network utility function (see (3.13)) before and after the measurements is higher than measurements costs, the measurements should be done, otherwise they should not. For example, if the utility of ARP DCA, which needs channel gain and interference measurements, minus the utility of Random DCA is grater than measurement costs for channel gain and interference, the measurements should be done (i.e. ARP DCA used), otherwise not (Random DCA should be used). Utility, for example number of satisfied users or total data rat (see (3.13)) can be, for example, estimated by simulation (see the chapter 4). Measurement costs can be estimated by network (provider) according to: measurements duration, bandwidth used for the measurements, amount of signaling etc. In the same manner each user can decide to perform measurements or not according to the value of their utility function before and after the measurements and costs for measurements.

### 3.11 Complexity Considerations

RRM algorithms based on "pricing" by each allocation of resources only search for the resources (channels, cells), which maximize the user's utility function. For example, our DCA algorithms based on "pricing" would require at most N(N-1) comparisons to allocate N channels to N users i.e. the complexity of the algorithms would be $O(N^2)$. If we "searched" all possible combinations of channels and users in a cell with N users and N channels, we would need N!-1 comparisons in order to find the "best" channel allocation to the users; also the complexity of the algorithms would be O(N!). If we also take other cells into account and try to find an "optimal" channel allocation over K cells, where in each cell there are N possible users the complexity of the algorithms would be $O(N^K)$ for each channel. Since our RRM algorithms based on "pricing" and game theory at each resource allocation request only maximize utility of a single user without taking other users (in own or in other cells) into account (only indirectly over "prices"), they have a much lower, "polynomial" search complexity.

On the contrary, "global" search algorithms have non polynomial (NP) search complexity of order O (N!) or $O(N^K)$, because the search is done not only over the set of possible resources (as in our case) but also over the set of possible users and cells. Furthermore, a global search would require an enormous signaling overhead in order to transport all relevant information (like users' channel measurements) to a centralized controller and resource allocation decisions from a centralized controller to concerned users and cells, whereas "our" RRM algorithms take all decisions locally (in one cell). Finally, it is not clear how to define "an optimization" criteria for a global search. We could try to allocate resources so that each user gets a "sufficient" signal quality. But, the channel changes in a mobile environment can be very large and fast (over 20 dB in a few milliseconds channel [45], [46], [47]), which would require a new channel allocation again with huge signaling and computational requirements.

On the other hand, according to "pricing" and game theory based approach a network provider or a user selects "an optimal" RRM algorithms for each situation (state) in the network (cells), which does not change for a longer time (usually several hours). What an optimal RRM algorithm is and for which state, can be estimated by simulations and/or network statistics, stored in a table and used during the system operations. This approach has a further advantage that the algorithms are not very sensitive to the parameter changes: The parameters' ($W_G$, $W_I$ and $C$) values are usually 1 or 0 and they serve to effectively switch between different algorithms (for example, Random, Min interference, ARP DCA). Of course, we cannot exclude that some global optimization algorithms would bring better performance than our algorithms. But if with our algorithms the system capacity is very close to the system (hard blocking) limit (see chapter 4), the costs (algorithms complexity and signalization overhead) of such global algorithms would then highly overweight their eventual gains.

### 3.12 Advantages of a “Pricing” Framework

The “pricing” framework provides us with all desired characteristics of RRM algorithms as described in section 1.1 and enables some additional advantages:

**Distributed**: Users try to maximize their utility functions under data and delay constraints independently of other users. Only parameters of the utility function (weights, “prices”) should be provided to the users by the network (base stations). The users then take resources (cell, channel, power) for which their utility function is maximized. There is no need for a complex optimization for all users, which reduce the algorithms complexity (see section 3.11). Although no central optimizing instance is needed, an optimal or sub-optimal solutions could be achieved in a distributed manner (see chapter 4).

**Adaptive and state dependent decisions:** The base stations should set weights in the utility function according to the state of the cell. On this way “prices” and thus RRM algorithms can be changed adaptively according to traffic, interference from the other cells or propagation conditions. For different cell states different weights should be used in order to “enforce” an optimal resource allocation in the cell (network) for each state (see section 3.7). Also users decisions about resource allocation are done according to users’ states (data rates achieved and power consumed so far) as well as users’service requirements (minimal data rate (signal quality) and maximum delay).

**Measurement based:** All information users or base stations need for evaluation of the utility function like channel gain G and interference I can be obtained by local measurements.

**Cooperative**: The network (cells) can enforce cooperation by appropriate weight settings in the utility function. In order to maximize the number of satisfied users, a network (cells) can set the weights in such a manner that each user takes resources that are just good enough for him (for example, according to his channel gain) in order to meet his minimal service requirements (for example minimal data rate) (see section 3.6). In this way the better resources are left for the users with worse conditions (lower channel gain), which enables them to achieve their minimal service requirements (minimal data rate) too.

**Existing algorithms can be obtained as special cases**: By appropriate parameter ("price") settings in the users' utility function, many existing RRM algorithms (like Min I and ARP DCA, path-loss based HO etc.) can be obtained as special cases of this framework, without the need of implementing each algorithm separately (see section 3.9).

**A trade-off between existing algorithms and development of new algorithms possible**: Since our utility function can be parameterized (by "prices") to obtain some RRM algorithms as special cases, we can use intermediate values of parameters ("prices") in order to achieve trade-offs between algorithms i.e. finer algorithms tuning is possible. Furthermore, some new algorithms can be applied based on optimal parameter ("price") settings estimated by simulation and/or real statistics.

**Other decisions in a wireless network** can also be modeled within this framework. For example, **measurements** of channel gain or interference should be done if the gain of measurements is higher than measurement costs (see 3.10).

## 3.13 Cooperation as a Rational Choice

In some networks like ad-hoc networks there are no distinguished “users” like base stations, which can set “prices” for the resources. Thus “cooperative” user’s behavior i.e. the behavior that maximizes the number of satisfied users and system capacity cannot always be “enforced” by “prices” as described in the previous sections. For example, in an ad-hoc network (see Figure 3-11) each user selects its power, channel and peer communication partner in order to maximize only its own utility. **There is nobody who cares about overall network utility** like BS or network provider, as in the case of cellular networks.

Nevertheless, the users can, under certain circumstances, find out that cooperation in the long term brings more (in expected utility sense) than “selfish” behavior. **The game theory provides us with the guidelines how cooperation can emerge in a world of self-seeking egoists, when there is no central authority to police their actions.** In this section, we investigate on the basis of game theory results how users can behave cooperatively in the absence of “prices” set by cells. The results could be interesting for RRM strategies in ad-hoc networks or among the networks from different providers or even in cellular networks if “prices” are not desirable or possible.

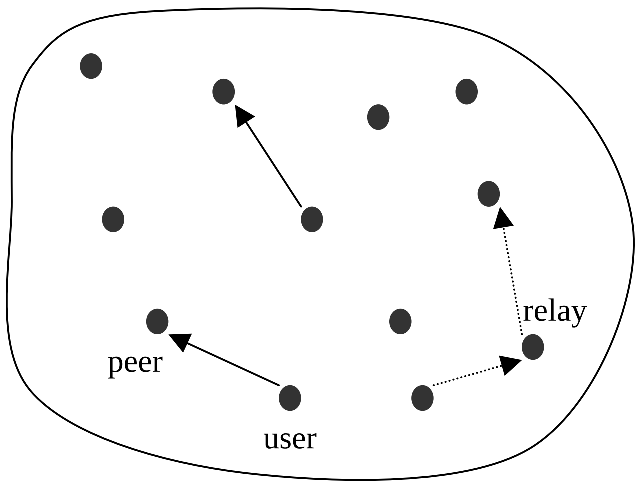



**Figure 3-11: An ad-hoc network**

Note a connection between “pricing” or a Stackelber game described in previous sections and a game without “prices” as discussed in this section. In “pricing” model “rules” like “take the channel with minimum interference when the load is not too high” are “enforced” by the “prices” due to special form of our utility function. In this section we investigate what are “good”, (evolutionary) stable, non-exploitable “rules” in the networks without prices like ad-hoc networks. If data rate or number of satisfied users for certain network state can be increased by “pricing” i.e. by “enforcing” an “optimal” RRM algorithm (“rule”, “strategy”) for the state by the cells, then the probability of user satisfaction or data rate for the same state also increases if the same algorithms are used without "pricing". Then an optimal "**social**" strategy would be to apply the same RRM algorithms for a given state as in the case of "pricing". The advantage of the "pricing" approach is that "optimal" RRM algorithms can be enforced by appropriate "prices" without relying on other users' "rationality" and "cooperation". According to the basic Game theory assumption players "take the actions which maximize own utility, assuming that other players do the same". In the case of ad-hoc networks **users might have initiative to deviate from cooperative behavior ("defect") if it might bring some advantages for them** even **if this behavior leads to "social" undesirable outcomes**. In this section we investigate **how to "design" cooperative RRM strategies in networks without prices**, **which are also "stable" i.e. cannot be "exploited" by "defecting" behavior**.

### *3.13.1 Defecting and Cooperative Behavior*

As in the networks with "pricing" RRM algorithms maximize their expected utility taking into account time (delay) and power constraints i.e.:

$$\boxed{\begin{array}{l} PC: \max E\{u(p)\} = \max E\left\{\log\left(1+\dfrac{pg}{I}\right) - \lambda p\right\} \\ AC, DCA, SH: \max[U(peer, channel)] = \max\limits_{peer,channel} E\left\{\log\left(\dfrac{g}{\lambda I}\right)\right\} \\ under\ constraints: \\ \Pr\left(R > R_{\min}\right) > 1 - P_{out}, \quad within\ \ maximal\ delay\ t_d \\ E\{p\} \le P_{\max} \end{array}} \tag{3.38}$$

In the case of ad-hoc networks, the users or their communication partners define the power "prices" $\lambda$ and not cells (base stations) as in the case of cellular networks. For example, a user might set "price" per unit power only according its own value of the battery power relative to the data rate, without taking into account the "social costs" of generating more interference to other users in the network. Such behavior, which does not take into account well being of the other users, we call "defecting", non-cooperative or "selfish" behavior. Defecting behavior of users in an ad-hoc network would be, for example, that regardless of system load and interference the users always try to maximize its own data rate or CIR for example by:

- Setting their CIR-threshold or water-filling level at the maximum level (outer-loop PC).
- Choosing always channels according minimum interference criteria (DCA).
- Starting always new calls regardless of the network load (HO).

A defecting behavior might bring temporary some advantages (higher data rates or CIR) for a user, when, for example, the user has particularly high channel gain or low interference. But the same user will continue to interact with other users with certain probability, also in the cases when the user has worse channel gain. In that case, **the other users can "punish" the "defecting" users by "defecting" themselves** and preventing the "defecting" users of achieving even its minimal required data rate under bad channel conditions. On the other side if users cooperate and take care of each other well being, the probability of achieving at least minimal data rates or signal quality might be higher and thus user satisfaction. Also, **in the long term it might be better for the user to "cooperate" than to "defect"** in order to achieve its minimal service requirements. Cooperative behavior of the users would be, for example:

- Set CIR-target and water-filling level according to the load in the system: The higher the load the lower CIR-target and water-filling level and thus the lower transmitted power should be used as long as minimum requirements (3.12) on user data rate are fulfilled (see section 3.3).
- Use DCA algorithms according to the load in the system: If load is low-medium use minimum interference DCA in order to reduce interference to other users and enable other users to apply minimum interference DCA too. If load is high, the users can use ARP DCA algorithms in order to enable other users with worse channel gains to obtain at least their minimum signal quality. Note that interference-based DCA algorithms like minimum interference or ARP is only possible if other users (cells, networks) apply some kind of interference based DCA, otherwise the channels cannot be differentiated according to interference (distribution). If other users use channels randomly, all channels would have approximately the same interference and an interference-based DCA could not be applied.
- Do not start new calls when the load in the system is already too high to enable users already in the system to achieve their minimum signal quality requirements. In the case of hierarchical cells make HO to the layers with a lower load although path-loss and interference is lower in the existing layer.

In order to enforce cooperation, we assume that each user values its own satisfaction (achieving its minimal data rate within predefined time, see (3.12)) much more than obtaining possible higher data rate with risk of the increasing probability of being unsatisfied. **It would then be "reasonable" to expect cooperation among the users whenever users with defecting cannot meet their minimum requirements with sufficiently high probability.** In other worlds, if the probability of outage with cooperation is lower than maximal allowable outage $P_{out}^{\max}$ and probability of outage without cooperation ("defecting") is higher than maximal allowable outage cooperation might be a rational choice i.e.:

$$\boxed{\begin{aligned}&P_{out}(cooperation) < P_{out}^{\max} \quad and \quad P_{out}(defecting) > P_{out}^{\max}\\&\Rightarrow\; Utility\,(cooperation) > Utility\,(defecting)\end{aligned}} \tag{3.39}$$

In following we use some game theory results like **Folk theorem** [9] **and Evolutionary stably strategies (ESS)** [93], in order to investigate circumstances under which cooperation arises as a rational choice for mobile users.

### *3.13.2 Folk Theorem and Evolutionary Stable Strategies*

As described in Appendix B it might be "rational" to cooperate in repeated games i.e. both players could obtain higher expected utility from cooperating than from defecting, even if the optimal choice (Nash equilibrium) for both players in the "single-shot" version of the same game is to "defect". This result is summarized in Folk Theorem [51]:

**The Folk Theorem: If there are strategies in the one-shot game that are "better" for all players than the Nash equilibrium strategy**, **and the probability of game continuation is high enough**, **every repeated game has an infinite number of Nash equilibriums**. The expected outcomes are better than the expected outcome of always playing the Nash equilibrium strategy of the one-shot game.

The proof is based on the idea that a player can be "enforced" by the other players to any strategy that brings him more expected gain ("cooperative" strategy) than always playing the Nash equilibrium strategy of the one-shot game ("defecting" strategy). The other players could make a simple threat: "If you defect from the "cooperative" strategy we will return to the Nash equilibrium strategy of the one-shot game and your gain would be lower than when you cooperate" [51]. For a rigor proof of the Folk theorem see [26] or [66].

Also, **whenever cooperation in a network brings more than defecting**, for example (3.39) is satisfied, according to the Folk theorem **users can be enforced by other users in the network to cooperate by "punishment" i.e. using the Nash equilibrium strategy of the one-shot game.**

For example, users in a wireless ad-hoc network might "punish" "defecting" users by generating high interference (sending with maximal power), which might decrease the signal quality of the defecting users to the level where the users cannot fulfill even their minimal data rate and delay constraints. The rational choice of the users in a network with such punishment mechanism would be to prefer cooperation to defecting, since with cooperation they would with high probability be satisfied and with deviation they would with high probability be unsatisfied i.e. the expected gain of cooperation is higher than from defecting (see (3.39)).

As the Folk theorem states, **there are an infinite number of possible Nash equilibriums in repeated games**. Each strategy in a repeated game, which provide a higher expected payoff (played against itself) than the "defect always" (Nash equilibrium strategy of one shot game), is a Nash equilibrium strategy of the repeated game. Which one of these strategies is selected in practice depends on a lot of things some of which have nothing to do with Game theory but are the result of a certain social situation [9]. For example, often used strategies are the **social "focal" points strategies** [51] i.e. the strategies established in a particular society as "good" and desirable, "the custom" strategy i.e. the strategy that was always used in the past, "symmetric" strategies i.e. "do the same as the others do" etc. For example, always driving on the left side or on the right side of the street are both Nash equilibrium strategies with essentially the same outcomes. Which Nash equilibrium strategy is selected depends on **social convention**. The resulting outcomes do not have to be even **Pareto optimal** i.e. provide each player with higher or equal payoffs than any other possible outcome. Players could stick to their strategies anyway because they are too inert or even do not know any better strategy. For example, the strategy to select the side of the street to drive on, by throwing a die, is a Nash equilibrium strategy, given that other players do the same. But, the Nash equilibrium resulting from the "random" selection of the street side is not Pareto optimal, since always driving on the left or always on the right makes all players better off. For example in "wireless games" CIR-based PC or Minimum Interference DCA can be used as Nash equilibrium strategies in repeated games because they are "symmetric" and relative simple ("use as much power as needed to achieve CIR-target", "take the channel with lowest interference") but robust against "exploitation" by other "defecting" strategies (see subsections 3.13.4 and 3.13.5).

In order to prevent "exploitation" of a RRM strategy it would be desirable to make it "evolutionary stable" (see [93] and Appendix B):

The strategy $S$ is called **evolutionary stable** if [93]:

1) The expected utility of the strategy $S$ played with itself $E(S,S)$ is greater than the expected gain $E(S,O)$ from all other strategy $O$ played with the strategy $S$ i.e. $\boldsymbol{E(S,S) > E(S,O)}$ for all other strategies $O$ or
2) If $E(S,S) = E(S,O)$ than $E(S,O) > E(O,O)$ for all other strategies $O$.

The expected utility of the strategy played against the other strategies measures a relative "fitness" of the strategy in a population of different strategies.

In [5] following characteristics of a "good" ESS strategy (like "TIT FOR TAT", see Appendix B) are listed:

- **Be nice at the beginning**
- **Punish those who defect themselves**, even if the "punishment" costs more than no punishment, in order to discourage motivation for "defecting".
- "**Forgive"** i.e. if a "defecting" player returns to the cooperative behavior then cooperate too.

It is important to note that "nice" strategy like "cooperate always" opens the door for "exploitation" by defecting strategies like "defect always". Therefore, a sort of "punishment" mechanisms should be build in each strategy in order to discourage defecting and make the strategy an ESS.

In wireless networks, the "defecting" users could be punished by reducing users' CIR by their peers, if the CIR of "defecting" user exceed certain threshold i.e. applying some threshold based PC (see subsection 3.13.4), or, if it is not possible, to send with maximum power on the channels used by "defecting" users. It is important to "punish" also those users who themselves do not "punish" when required i.e. users should "guard" each other (this tactic is well known in all dictatorial systems).

Taking into account results of the Game theory like "Evolutionary stable strategies" or "Folk theorem", an effective and stable RRM strategy should:

- **Cooperate when other users cooperate** i.e. use so much from resources (power, channels) as needed, but not more. Adapt resource usage to the load in the system.
- **Punish defecting users** in order to encourage cooperation i.e. send with maximum power (generate more interference) or take the "best" channels (with lowest interference) in the presence of "defecting" users.
- **Using some kind of threshold to "detect" defecting users** (as oft done in repeated games, see [51]). For example, if a user's signal quality is "too high" decrease his CIR by "threat" that more interference will be generated to the "defecting" user's signal (see subsection 3.13.4).

### 3.13.3 What the Players should know in the Wireless Game

In order to promote cooperation the players should be able to **“recognize” the defecting users** and to have **enough information about interference statistics** and about **previous behavior of other players** i.e. the players should have a “long enough” memory about previous game outcomes. That is why it is important how we define the players (users) in a wireless game. Since single services (users) can live for a relatively short time but a mobile station (MS) itself lives much longer, it would be desirable to define as a player in a wireless game the **(service type, MS)-pair**. In that case the service on a MS can be awarded or punished according to the behavior of all previous services of the same service type on the same MS. The probability that the same users (MS) “meet each other” (or play the same game again) would be increased, which “enlarges shadow of the future” and increases gain of the cooperation (see Repeated Prisoners’ Dilemma in Appendix B). We can also assume that the users know everything that their MS knows i.e. statistics (channel gain, interference etc.) of all previous usages of the same MS, which helps them to make better decisions and to know whom to “punish”.

It would be advantageous that **users “live” so long in the network that their own CIR or interference statistics can be representative for the network statistics** (“ergodicity”) assumption. In that case the users could judge according its own statistics about network load i.e. estimate the outage probability in the network. This estimation can be in turn used to adjust RRM strategy according to load, for example CIR or water-filling level settings, channel choice etc.

### 3.13.4 Evolutionary Stability of Power Control Algorithms

In this section we investigate some PC algorithms (“rules”, “strategies”) like Water-filling PC, CIR-based PC or Max power PC on their **“evolutionary stability”**, as defined in Appendix B and 3.13.2.

In the following we assume that the ***E*** **expected utility of the strategy is measured in the probability to achieve required CIR** i.e. outage probability (see (3.11)). If outage probability of two strategies is the same, the second criteria is the consumed power with a strategy i.e. the lower the expected power consumption the higher the utility $E$.

For example, we show below that Water-filling PC is not an evolutionary stable strategy (ESS) since its expected utility played against itself is lower than from (“defecting”) Max power PC strategy against the Water-filing strategy:

- A “Max Power” user increasing its power in comparison to a “Water-filling” user, who sends with power lower or equal to maximal power. Consequently, interference to “Water-filling” user is increased,
- On the other hand, “Water-filling” users behaves cooperatively and decrease its powers in presence of higher interference (see (2.16)) caused by “Max Power” users.

Thus, non-cooperative “Max Power” users have a higher probability of obtaining the higher CIR than “water-filling” users since they **use higher power** and **experience lower interference** from water-filling users. Consequently expected utility *E* of the “water-filling” PC played against itself is lower than from (“defecting”) Max power PC strategy against the water-filing strategy i.e.:

*E(“Water-filling”, “Water-filling”) < E(“Max Power”, “Water-filling”)*

Also, according to 1) in subsection 3.13.2 **Water-filling PC is also not an evolutionary stable strategy**.

On the other hand, Max Power PC can not exploit CIR-based PC:

- With use of Max Power instead of CIR-based PC interference to CIR-based PC user is increased.
- “Max Power” user is automatically “punished” by “CIR-based” user, since CIR-based PC increases its power too, because power of CIR-based PC is proportional to interference (see (2.5)). In turn, interference to “Max Power” user increases too.

Since CIR-based PC sends also with maximum power if this is needed to achieve the required CIR, **with CIR-based PC the probability to achieve the required CIR-threshold is the same as with Max Power PC** (when PC converge at all). But **CIR-based PC needs lower power to achieve the CIR-threshold than Max Power PC**, since CIR-based PC sends with so much power needed to achieve the CIR-threshold, but not more. This can be easily seen from (2.3): When all users increases their power for a factor $a$ ($a > 1$), then CIR of the users stays almost the same (since noise can be neglected in interference limited systems), but the total power increases and user utility *E* decreases. This power allocation is not Pareto optimal, since all users can be better of by decreasing the power for the same amount ($a$).

Also taking into account the probability to achieve the required CIR and consumed power, the utility *E* of CIR based PC played against itself is higher than the gain of Max Power PC against CIR-based PC i.e.:

*E(“CIR-based”, “CIR-based”) > E(“Max Power”, “CIR-based”)*

**Thus, a population of CIR-based PC can not be “invaded” by Max Power PC users.**

In the case of “Water-filling” and “CIR-based” PC together, both PC can profit from each other:

- “Water-filling” users can profit from “CIR-based” users, since “CIR-based” users do not send with higher power than needed to achieve the required CIR.
- “CIR-based” users could also profit from “Water-filling” users, since “Water-filling” users do net send at all when the channel is too bad.

This performance of “Water-filling” and “CIR-based” PC together is further investigated by simulations in subsection 4.3.2.

Note that a “defecting” PC could be not only Max power PC but also “water-filling” of CIR-based PC, which sets its CIR-threshold or water-filling level too high. In order to prevent the users to use Max power PC or to set their CIR-thresholds or water-filling levels too high, some kind of punishment is needed, as discussed in subsection 3.13.1. **“Punishment” can be performed by the communication partners of the “defecting” users** (in ad-hoc networks) or cells (in a cellular networks). A communication partner (receiver) of a “defecting” user (transmitter) can make for example a simple “threat”: **“I will decrease your signal quality (CIR)**, **if your CIR is for a certain threshold higher than my average CIR**”. The “punishment” can be performed, for example, **by generating some amount of noise in the receiver to the signal of the “defecting” users** or by simple **not decoding the data of the “defecting” users**. This in turn can be signalized to the “defecting” user (assuming that users have a backwards signalization channel), which encourage him to decrease its CIR-target or water-filling level and consequently its transmit power (see Figure 3-12).

*For each user i and its peer communication partner j:*

*if CIR(i) - CIR(j) > threshold1*

*Decrease power of the user i i.e.*

*Decrease CIR-target$_t$(i) or water-filling level(i);*

*elseif CIR(j) - CIR(i) > threshold2 AND P(i) < $P_{max}$*

*Increase power of the user i i.e.*

*Increase CIR-target$_t$(i) or water-filling level(i);*

*end if*

**Figure 3-12: A threshold-based “punishment”**

The “punishment” strategy is further investigated by simulations in subsection 4.3.3.

#### *3.13.5 Cooperation in a Channel Allocation Game*

We can model the channel allocation as a two-player game: The player itself and the rest of the network modeled as the second player. The **players** can be **mobiles** in ad-hoc networks or **cells** belonging to the same or different providers. **The possible strategies for the players are the choices of different channel allocation algorithms**. The player chooses the channel allocation algorithm with maximum expected gain (see (3.38)), given the algorithms applied by other users. When gain and costs are same on all channels and difference in interference on different channels is significant, maximizing of DCA part of (3.38) means minimizing interference. Also, **the myopic choice of the users would be to use Min I DCA algorithm**. If the users cannot distinguish channels according to interference due to relatively high interference variance (as in the case of lower loads), they could apply random channel allocation (see Figure 3-13). For **higher loads**, users could gain more by cooperation than by defecting ((3.39) is satisfied) i.e. **the users should apply "cooperative" ARP DCA if the other users "cooperate" too and apply ARP themselves**. Defective behavior should be "punished" by defecting too i.e. applying Min I or random DCA in order to "discourage" the other users from defecting. Also, as in the case of DCA based on "pricing", the same "optimal" DCA algorithm for a given load can emerge as a "rational choice" in networks without centralized controllers provided the other users "cooperate" too.

```
If significant difference in interference on channels
      // i.e. other cells apply some interference based DCA algorithms
      if load is "high" AND other users apply ARP DCA
          Apply ARP DCA
      else
          Apply Min I DCA
      end if
 else   // no significant difference in interference on different channels
    Apply random DCA
end
```

**Figure 3-13: A "cooperative" DCA algorithm**

Provided that the expected gain $E$ of a DCA algorithms ("fitness") is measured in the number of satisfied users (as the most important criteria) and power consumed (as a less important criteria) (see above), Min I DCA is an ESS (see Appendix B) in a pool consisting of Random DCA. Min I and ARP DCA.:

- *E(Min I, Min I) > E(Random, Min I),* since Random DCA can not benefit from lower interference on some channels as Min I DCA does, and
- *E(Min I, Min I) > E(ARP, Min I)*, since ARP DCA can not benefit from leaving some “good” channels for users with a higher path-loss (they are used by Min I regardless of the path-loss).

**In order to make ARP an ESS stable strategy some sort of “punishment” should be built in**. Users, who use “better” channels (with lower interference), although their path-loss is relatively low, should be punished, for example, by using Min I or Random DCA against the users.

Note that Min I DCA needs interference measurements and **ARP DCA needs interference and channel gain measurements**. Furthermore, the parameters of ARP must be optimized for each load separately. The parameter optimization and measurements overhead for DCA is often avoided in modern mobile systems, which use CDMA (UMTS) or Frequency hopping techniques (GSM). In CDMA or frequency hopping systems, channels consist of different frequencies or timeslots in order to increase diversity. **Due to diversity and powerful coding and modulation technique in some modern systems (UMTS FDD), all channels have almost the same (“satisfactory”) quality and we can use simple Random DCA in the systems**.

#### *3.13.6 On Inter-Network Cooperation*

The main principles described in the previous sections can be applied almost without change to cooperation among network providers. Each provider can be regarded as a single cell maximizing utility according to (3.13). The users themselves maximize their utility according to (3.15) and choose a provider in order to maximize their utility i.e. maximize signal quality and minimize costs. Providers set “prices” according to load, interference, service types etc. and users make HO to the system, which maximizes their utilities i.e. users maximize (3.36) over networks instead of cells.

Further, **cooperative behavior may also arise as a reasonable choice even if each provider only maximizes its own (long term) gain** (see section 3.13). In order to provide most users with at least minimum signal quality a cooperative PC and cooperative DCA is required. **Cooperation by channel allocation** is of great importance in order to protect a network from the **adjacent channel interference**. By cooperative use of frequencies or timeslots, the overall interference can be reduced and both (all) providers can benefit from cooperative channel usage. As in the case of a cellular or an ad-hoc network, an interference-based DCA algorithms is only possible if other networks apply some kind of interference-based DCA, otherwise the channels cannot be distinguished according to interference (distribution).

## 3.14 Summary

In this chapter we described our “pricing” and game theory framework for modeling RRM algorithms. We defined user and network (cell) utility functions. We showed how a network could maximize its utility by load dependent “price” setting. Further, we described heuristics and “geometry” behind RRM algorithms. Finally, we showed using game theory results how user cooperation can emerge as a rational choice in the networks without distinguished “price”-setting users (base stations, access ports) like in ad-hoc networks or among networks from different providers.

# 4 Algorithms and Simulations

*"Computer simulations let us analyze complicated systems that can't be analyzed mathematically. With an accurate computer model, we can make changes and see how they will affect a system."*

***Sheldon M. Ross***

In this chapter we present some RRM algorithms based on the theoretical concepts described in the previous chapters and provide simulation results for the algorithms.

We first describe the simulation model i.e. the simulators, their traffic, mobility and propagation models. Further, we provide the performance metric used to compare the RRM algorithms. Then we present some RRM algorithms like PC, Scheduler, AC, HO and DCA and simulation results for the algorithms in different loads and environments. Finally, we compare performance gains of different RRM algorithms and technologies like SA in respect to reduction in overall interference and analyze possible trade-offs a provider can make by employing the algorithms or SA technologies.

## 4.1 Simulation Model

We used a MATLAB-based snapshot simulator RUNE (Rudimentary Network Emulator) for scheduler and power control simulations in section 4.3. The RUNE simulator is provided in [110] and we modified it for the purposes of our work (see subsection 4.1.1). For all other simulations we used a dynamical, system-level, C++ based simulator (see subsection 4.1.2).

### *4.1.1 Snapshot "RUNE" Simulator [110]*

Rudimentary Network Emulator (RUNE) is a snapshot, MATLAB-based simulator, provided and described in [110]. The RUNE simulator is **based on discrete time steps (snapshots)** i.e. a system is studied at specific, regularly spaced, time instants. In general, the system changes between each time instant i.e. mobiles may have moved, new calls may have been created and others may have been terminated. The advantage of the discrete time steps model is that the whole system can be handled at the same time. The state of the system can be represented by vectors and matrices and treated efficiently with mathematical software like MATLAB. The implementation of **a snapshot simulator is in general simpler and simulations can, in general, run faster** than in an event-based simulator where each event must be treated separately. The disadvantage is **lower accuracy in comparison with an event based dynamical simulator** i.e. the relative order of the events is not always the same as it would be in real systems.

In the RUNE simulator, a relatively small number of cells were chosen to collect relevant statistics. In order to obtain more realistic statistics, surrounded cells were also modeled by use of a “wraparound” technique: The cells under study were placed in a rhombus. The rhombus was then stitched together so that the top meets the bottom and the left side meets the right side. This creates a torus-shaped surface that the cells were placed on. Since the surface had no border, all cells had neighbors on all sides.

Traffic was generated according to Poisson distribution approximated by a binomial distribution. The number of the users who should have left the system was calculated assuming the exponential service time. The users who should have left the system were then selected randomly. The new users were added according to Poisson distribution to keep the average number of users in the system constant.

In order to model mobility in the system with each user a velocity vector (magnitude and direction) and position ((x,y) coordinates) were associated. In each time step the users were moved according to the current velocity and the size of the time step. The velocities of the mobiles were also changed randomly a little bit so that the mobiles accelerate, slow down and change their directions with certain probability.

The propagation loss was modeled as a sum of the antenna pattern, the distance dependent fading, log-normal shadowing and Rayleigh fading. The distance dependent fading was modeled according to the following formula: $G = \frac{C}{r^{\alpha}}$, where C is path gain at a distance of one meter from the transmitter antenna and $\alpha$ is a parameter which determines how power decays as a function of the distance from the base station. For free space propagation $\alpha$ is 2 and in a typical urban environment $\alpha$ ranges from 3 and 4.

The log-normal shadowing was modeled as $G = 10^{X}$, where X is normal distributed with mean 0 and variance $\sigma$ (typically 8-10 dB). With each geographical point in the system, a specific amount of shadow fading was associated. In this way it was ensured that the shadowing was correlated in space and that the amount of shadow fading will always be the same in the same position. In order to keep implementation expenditure moderate, a smaller shadowing map was repeated many times over the system area.

In a similar manner a map was used to define Rayleigh fading for each geographical point within the simulation area. Because of fast changes of Rayleigh fading, two maps were used and the fading in a point was obtained as the sum of both maps. In this way memory requirement for storing map was reduced.

In the RUNE simulator random channel allocation and path-loss-based HO was used. We also implemented and simulated water-filling PC additionally to constant (maximum) power, C-based and CIR-based PC, which were already available in the simulator.

#### *4.1.2 System Level Dynamical Simulator*

We used a dynamical, C++ based simulator for system-level simulations of the most RRM algorithms. The simulator is an **event-triggered** i.e. everything that happens in the model is triggered by a certain event. For example, users' arrivals and call termination, updating position of mobiles, handover etc., but also lapse of a certain time like radio frame or HO period are events, which are placed in an event queue and performed one event at a time. The events also have priorities, which regulate the order of event execution in the case of simultaneous arrival of events. The **basic simulator time interval is a radio frame** (with 10 ms duration for the TDD system). Within the radio frame, events are performed according to their priorities and position in the event queue. The advantage of an event-triggered simulator is **more accurate modeling of time events and no sampling error** - all events are modeled according to their time arrival as in real systems. The disadvantage is implementation complexity in comparison with a snapshot simulator and, in general, increase in simulation time.

We implemented in the simulator users' movements, propagation models and interference calculation as well as RRM algorithms like handover (HO), power control (PC) and DCA. Channel models, user mobility and traffic behavior was implemented according to **[19]** for both urban (Micro) and sub-urban (Macro) environments. We simulated basically UMTS TDD mode where channels are frequency (f), timeslot (TS), and code (c) triples (f,TS,c)-triples). Since the inter cell interference is the same for all codes of a f-TS-pair we use the term "channel" as f-TS-pair. Frame duration was 10 ms and we used one frequency, 8 timeslots per link (uplink and downlink) and 16 codes per timeslot. The maximum transmit power was 42 dBm and the number of base station was 72 (Micro) and 57 (Macro) (see Table 4-1 for main parameter values).

**Table 4-1: Overview of system parameters**

| | Urban | Rural |
|---|---|---|
| **Number of base station sites** | 72 | 19 |
| **Number of sectors per site** | 1 | 3 |
| **Number of cells** | 72 | 57 |
| **Number of reference cells** | 6 | 3 |
| **Antenna height MS [m]** | 1.5 | 1.5 |
| **Antenna height BS [m]** | 30 | 30 |
| **Frame duration [ms]** | 10 | 10 |
| **Number of slots per frame (DL)** | 8 UL + 8 DL | 8 UL + 8 DL |
| **Number of codes per slot** | 12 | 12 |
| **Carrier frequency [GHz]** | 2 | 2 |
| **Bandwidth [MHz]** | 5 | 5 |
| **Chip rate [Mchip/s]** | 4.096 | 4.096 |
| **Channel reuse** | 1 | 1 |
| **Shadow fading σ [dB]** | 10 | 10 |
| **Decorrelation length [m]** | 5 | 20 |
| **Speed mean [km/h]** | 3 | 120 |
| **HO-margin [dB]** | 3 | 5 |
| **Max MS power level [dBm]** | 36 | 36 |
| **Max BS power level [dBm]** | 42 | 42 |
| **Noise Power (MS) [dBm]** | -99 | -99 |
| **Noise Power (BS) [dBm]** | -103 | -103 |
| **Power Control** | CIR-based (speech)<br>C-based (NRT data) | CIR-based (speech)<br>C-based (NRT data) |
| **HO type** | Path-loss based | Path-loss based |
| **Scheduling** | First-in-first-out | First-in-first-out |

We focused our simulations on DL, since DL is expected to be a capacity constraining link in the future. On the one hand in DL, a higher traffic load (Internet downloads) is expected, on the other hand signal quality in UL is generally higher due to more powerful processing capabilities and more diversity gains at BS, since BS does not have such tight power, place and costs constraints as MS.

Each radio frame (10 ms), RRM algorithms and interference calculation algorithms were called in the following order: HO, AC, DCA, interference calculation and PC. If required event occurs the RRM module performs its task, otherwise the control is returned to the operating system of the simulator. For example, HO is called periodically by an operating system each HO period (simulator parameter). HO checks if path-loss (plus slow fading, but with fast fading averaged out) of serving cell is at least for HO-margin (3 dB Micro and 5 dB Macro) higher than the path-loss of the candidate cell (path-loss based HO). If yes, AC of the candidate cell is called by HO. If AC allows the entrance of the user in the cell, HO is performed to the cell, otherwise the next best candidate cell with path-loss at least for HO margin lower than the active cell is tried and, so on. If HO is performed to the new cell, DCA in the new cell allocates a channel to the user according to Random, Min I, Priority-based or ARP algorithm (defined by a parameter). In the case of NRT packet data, DCA allocates the channel to the users only after the Scheduler according to the FIFO algorithms schedules the data for the transmission. During the call, PC is called in each frame to calculate required power in the frame for each user. We used CIR-based PC for speech and C-based PC algorithm for NRT data.

#### *4.1.2.1 Traffic Model*

We assume that speech users arrive according to Poisson distribution with mean call time 120 seconds and discontinuous transmission (DTX). Packet data sessions also arrive according to Poisson distribution. The number of packet data requests per session, the reading time between two consecutive packet call requests in a session, the number of packets in a packet call, the time interval between two consecutive packets inside a packet call were modeled as geometrically distributed random variables and the packet size was modeled as a Paretto distributed random variable [19].

In following subsections we describe the simulation environments: Urban (Micro) and sub-urban and rural (Macro) as well as mobility, channel and propagation models used in our dynamic simulator. The implementation was done according to ETSI Selection procedures for the choice of radio transmission technologies of the UMTS [19] where further details can be found.

#### *4.1.2.2 Urban (Micro) Environment*

A Manhattan-like structure (see Figure 4-1) was used for the modeling of outdoor-to-indoor and the pedestrian environment (Micro).

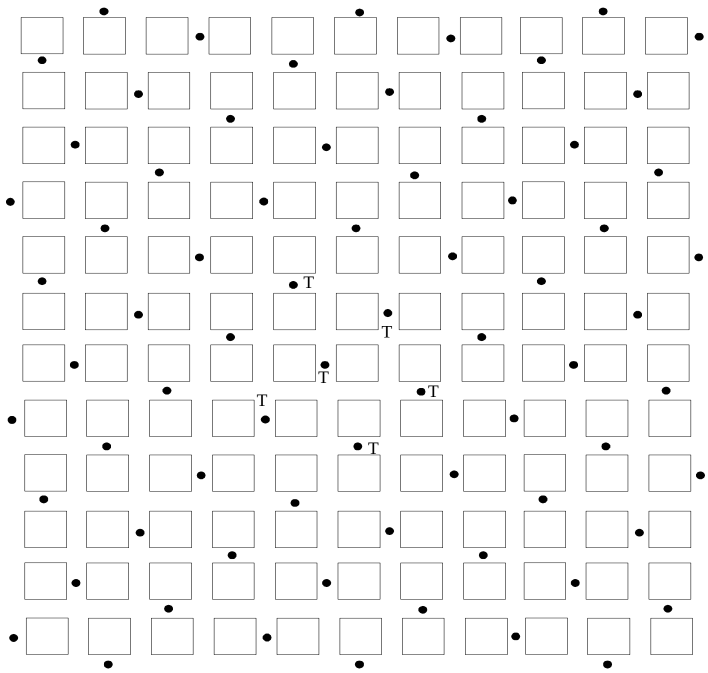



**Figure 4-1: Manhattan-like urban model and deployment scheme**

Parameters for this structure are defined in Table 4-2.

**Table 4-2: Parameter of urban "Manhattan-like" deployment**

| Area | Block size | Street width | Base station - mobile height difference |
|---|---|---|---|
| 6.5 km² | 200 m x 200 m | 30 m | 10 m |

Quality statistics were collected among cells marked with a T on Figure 4-1.

Physical deployment environments are summarized in Table 4-3 below.

**Table 4-3: Deployment model in Micro environment**

| Type | Building Penetration Loss/standard Deviation (dB) | Log-Normal Standard Deviation (dB) | Mobile Velocity (km/h) |
|---|---|---|---|
| **Outdoor** | NA | 10 | 3 |

For the path-loss calculation a three slopes path-loss model was used as described in section 4.1.2.5

Mobility model

Mobiles moved along streets in a Manhattan-like structure defined in Figure 4-2 and might turn at cross streets with a given probability. The positions of the mobiles were updated every 5 meters and speed could be changed at each position update according to a given probability. We used the following parameters in our urban mobility model:

Mean speed: 3 km/h

Minimum speed: 0 km/h

Standard deviation for speed (normal distribution): 0.3 km/h

Probability to change speed at position update: 0.2

Probability to turn at cross street: 0.5

The turning probability is illustrated on the Figure 4-2 below:

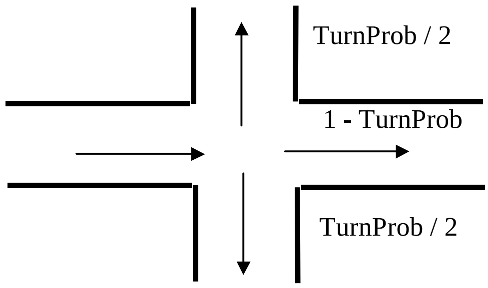



**Figure 4-2:Mobility model in urban (Micro) environment**

Mobiles were uniformly distributed in the street and their direction was randomly chosen at simulation begin.

#### *4.1.2.3 Sub-urban and rural (Macro) Environment*

The cell radius in the Macro environment was 2000 m. The base station antenna height was above the average roof top height of 15 meters. The deployment scheme was a hexagonal cell layout with distances between base stations equal to 6 km (see Figure 4-3 below). Tri-sectored cells were used with the antenna pattern being specified in section 1.5 [19]. The base station may use omni directional and smart antennas in the macro environment.

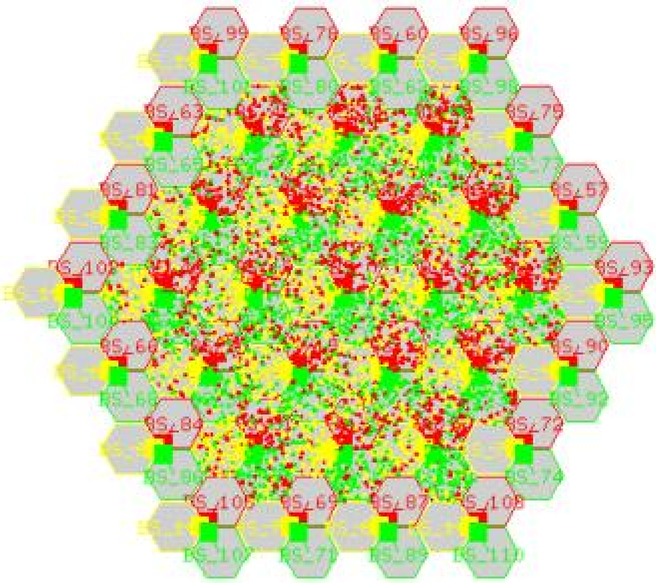


**Figure 4-3: Sub-urban and rural (Macro) simulation environment with 60 cells. The mobile stations are represented by colored dots which indicate their connectivity.**

The log-normal standard deviation for shadowing effects was assumed to be 10 dB.

<u>Mobility model</u>

The mobility model for the Vehicular Test environment was a pseudo random mobility model with semi-directed trajectories. The positions of the mobiles were updated according to the decorrelation length and direction could be changed at each position update according to a given probability.

The speeds of the mobiles were constant and the mobility model was defined by the following parameters:

Speed value: 120 km/h

Probability to change direction at position update: 0.2

Maximal angle for direction update: 45°

Decorrelation length: 20 meters

The mobiles were uniformly distributed on the map and their direction was randomly chosen at the simulation start.

#### *4.1.2.4 Channel Model*

For each terrestrial test environment, a **channel impulse response model was based on a tapped-delay line model** [19]. The number of taps, the time delay relative to the first tap, the average power relative to the strongest tap, and the Doppler spectrum of each tap characterize the model.

Table 4-4 and Table 4-5 describe the tapped-delay-line parameters for urban and sub-urban environments respectively. For each tap of the channels three parameters are given: The time delay relative to the first tap, the average power relative to the strongest tap, and the Doppler spectrum of each tap. A small variation, ±3%, in the relative time delay was allowed so that the channel sampling rate can be made to match some multiple of the link simulation sample rate.

**Table 4-4: Outdoor to Indoor and Pedestrian Test Environment Tapped-Delay-Line Parameters**

| Tap | Channel A | | Channel B | | Doppler Spectrum |
|---|---|---|---|---|---|
| | Rel. Delay (nsec) | Avg. Power (dB) | Rel. Delay (nsec) | Avg. Power (dB) | |
| 1 | 0 | 0 | 0 | 0 | CLASSIC |
| 2 | 110 | -9.7 | 200 | -0.9 | CLASSIC |
| 3 | 190 | -19.2 | 800 | -4.9 | CLASSIC |
| 4 | 410 | -22.8 | 1200 | -8.0 | CLASSIC |
| 5 | - | - | 2300 | -7.8 | CLASSIC |
| 6 | - | - | 3700 | -23.9 | CLASSIC |

**Table 4-5: Vehicular Test Environment, High Antenna, Tapped-Delay-Line Parameters**

| Tap | Channel A | | Channel B | | Doppler Spectrum |
|---|---|---|---|---|---|
| | Rel. Delay (nsec) | Avg. Power (dB) | Rel. Delay (nsec) | Avg. Power (dB) | |
| 1 | 0 | 0.0 | 0 | -2.5 | CLASSIC |
| 2 | 310 | -1.0 | 300 | 0 | CLASSIC |
| 3 | 710 | -9.0 | 8900 | -12.8 | CLASSIC |
| 4 | 1090 | -10.0 | 12900 | -10.0 | CLASSIC |
| 5 | 1730 | -15.0 | 17100 | -25.2 | CLASSIC |
| 6 | 2510 | -20.0 | 20000 | -16.0 | CLASSIC |

In order to accurately model the variability of delay spread also in the "worst case" locations where delay spread is relatively large, we used two channel models (A and B, in Table 4-4 and Table 4-5) as recommended in [19]. Each of the two channel models is expected to be encountered for some percentage of time in a given test environment. Table 4-6 gives the percentage of time the particular channel may be encountered with the associated r.m.s. average delay spread for channel A and channel B for each test environment.

**Table 4-6: Parameters for Channel Impulse Response Model**

| Test Environment | r.m.s. A (ns) | P(A) (%) | r.m.s. B (ns) | P(B) (%) |
|---|---|---|---|---|
| Indoor Office | 35 | 50 | 100 | 45 |
| Outdoor to Indoor and Pedestrian | 45 | 40 | 750 | 55 |
| Vehicular - High Antenna | 370 | 40 | 4000 | 55 |

#### *4.1.2.5 Propagation Model in Urban (Micro) Environment*

The propagation model in an urban (Micro) environment **takes into account the line of sight LOS and non line of sight NLOS** situations. This is a recursive model [7], which calculates the **path-loss as a sum of LOS and NLOS segments**. The shortest path along streets between the BS and the MS has to be found within the Manhattan environment.

The path-loss in dB is given by the formula [7]: $L = 20 \cdot \log_{10} \frac{4\pi d_n}{\lambda}$,

where $d_n$ is the "illusory" distance,

$\lambda$ is the wavelength,

n is the number of straight street segments between BS and MS (along the shortest path).

The illusory distance is the sum of these street segments obtained by recursively using the expressions $k_n = k_{n-1} + d_{n-1} \cdot c$ and $d_n = k_n \cdot s_{n-1} + d_{n-1}$ where c is a function of the angle of the street crossing. For a 90 degree street crossing, the value c was set to 0.5. $s_{n-1}$ is the length in meters of the last segment. A segment is a straight path. Initial value $k_0$ was set to 1 and $d_0$ is set to 0.

The model is extended to cover the micro cell dual slope behavior, by modifying the expression to:

$$L = 20 \cdot \log_{10} \left(\frac{4\pi d_n}{\lambda} \cdot D\left(\sum_{j=1}^{n} s_{j-1}\right)\right) \text{ where } D(x) = \begin{cases} x / x_{br}, x > x_{br} \\ 1, x \le x_{br} \end{cases}.$$

Before the break point $x_{br}$ the slope is 2, after the break point it increases to 4. The break point $x_{br}$ was set to 300 m. x is the distance from the transmitter to the receiver.

We also take into account effects of propagation going above the path-loss according to the shortest geographical distance by using the commonly known COST Walfish-Ikegami Model with antennas below rooftops:

$$L = 24 + 45 \log(d+20)$$

#### 4.1.2.6 Propagation Model in Sub-urban and rural (Macro) Environment

The following model was used for path-loss calculation in the vehicular sub-urban and rural (Macro) test environment [19]:

$$L = 40(1-4\times10^{-3}\Delta h_b) Log_{10}(R) - 18 Log_{10}(\Delta h_b) + 21 Log_{10}(f) + 80 \text{ dB.}$$

Where:

| | |
|---|---|
| R | is the base station - mobile station separation in kilometres; |
| f | is the carrier frequency of 2000 MHz; |
| $\Delta h_b$ | is the base station antenna height, in metres, measured from the average rooftop level. |

The base station antenna height was fixed at 15 meters above the average rooftop ($\Delta h_b$ = 15 m).

Considering a carrier frequency of 2000 Mhz and a base station antenna height of 15 meters, the formula becomes:

$$L = 128.1 + 37.6 \, Log_{10}(R)$$

#### 4.1.2.7 Decorrelation Length of the Long-term Fading

The long-term fading is characterized by a Log-Normal distribution with zero mean (see Appendix B). Since slow fading process does not change rapidly at close geographical distance $\Delta x$, **adjacent fading values are correlated**. Its normalized autocorrelation function $R(\Delta x)$ can be described with sufficient accuracy by an exponential function [32]:

$$R(\Delta x) = e^{-\frac{|\Delta x|}{d_{cor}} \ln 2}$$

Where the decorrelation length $d_{cor}$ depends on the environment [19]. In our vehicular Macro environment, the decorrelation length was 20 meters. In outdoor-to-indoor and the pedestrian environment, the decorrelation length was 5 meters.

#### 4.1.2.8 Smart Antennas Model

The simulations with SA were performed in the Macro environment using 8-element circular arrays with $\lambda/2$ spacing at the base stations [10]. The direction of arrival statistics follows a Laplacian distribution with 15 degrees spreading. An example angular inter-cell interference power pattern for the downlink is shown in Figure 4-4. This angular power profile was calculated for a situation with 8 users on the same timeslot.

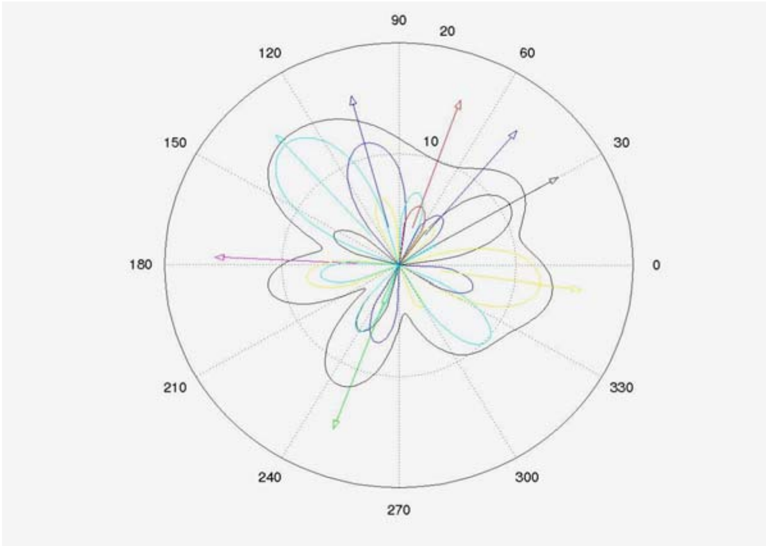



**Figure 4-4**: **A snapshot of the angular distribution of downlink interference power (arrows show the nominal directions of arrival for individual users)**

In TDD systems, the uplink and downlink connections share the same carrier frequency by multiplexing the two directions in time. Therefore, **we could use the spatial covariance matrices estimated on the uplink for the downlink**. These spatial covariance matrices can be estimated by exploiting the knowledge about the spreading codes, properties of the transmitted symbols, the transmitted training sequences, or a combination thereof. Hence, in TDD systems the estimation of the dominant directions of arrival is not required for the downlink beamforming.

#### 4.1.2.9 Interface Link Level / System Level

The receiver-side error statistics of the link-level simulator (which models the behavior of the physical layer) is entered into the system-level simulator via ``actual value interface'' [33], [82]: **The mappings from CIR to SIR, SIR to raw BER and raw BER to user BER are defined by the polynomial regressions of the results from link level simulations**. In the system level simulator, CIR is calculated for each user in each frame. The CIR is then mapped on the SIR after despreading and block-linear zero-forcing algorithm, taking into account the number of users in the timeslot and the number of spreading codes in the timeslot. Further

SIR is mapped on the raw BER possibly taking into account Maximum Ratio Combining (MRC). Whenever MRC is done at the receiver with two diversity branches, the SIR after MRC is modeled as the sum of the input SIRs. Finally, the average raw BER (over interleaving duration) is mapped on the user BER after interleaving and decoding.

Since in TDD the base stations employ joint-detection of users belonging to the same cell, we assumed at the system level that the **intra-cell interference is eliminated by digital signal processing in the receiver**. Link level simulations of physical channel, coder (decoder), modulator and demodulator are done on a sub-chip duration scale whereas the granularity of the system level simulator is the timeslot.

## 4.2 Performance Metric

In this section we define performance metrics used to compare our RRM algorithms. An ultimate goal of all resource management algorithms is to maximize the **number of satisfied users** in the system. In the following, we define when a RT or NRT user is counted in the simulators as the satisfied user. We also use **Grade of Service metric**, which comprises blocking rate and dropping rate in one formula.

### 4.2.1 Satisfied User Definitions

A wireless RT (circuit switched) user was regarded as satisfied if [19]:

1. The user did not get blocked when arriving at the system.
2. The user had sufficiently good quality more than a certain time (fraction) of the session:

$$\textit{Probability } (BER > BER\_Threshold(10^{-3})) < 2\ \%$$

To calculate the above probability, BER was compared with *BER_Threshold* each speech interleaving period (20 ms). If the BER was higher than *BER_Threshold* the interleaving period was marked as “bad”. Finally, the ratio (bad periods)/(total number of interleaving periods in the session) should have been lower than 0.02 in order to count the user as satisfied.

3. The user does not get dropped. A call is dropped if:

$$BER > BER\_Threshold\ (10^{-3})\ \textit{for more than 5 seconds}$$

Again BER was compared with *BER_Threshold* each speech interleaving period (20 ms) and if there were 250 (5 s / 20 ms) consecutive “bad” periods (with *BER* > *BER_Threshold*) the user is dropped.

For NRT (packet data) services, a user was satisfied if the following three constraints were fulfilled [19]:

1. The user was not blocked when arriving at the system.
2. The active session throughput $R$ of the session was equal or greater than some minimal throughput $R_{min}$ (10% of nominal data rate) i.e. if $E(R) \geq R_{min}$ during the session length.
3. The user was not dropped. A packet user was dropped if the number of repeating transmissions of the same packet exceeded predefined threshold (10 re-transmissions in our simulator).

### 4.2.2 Grade of Service (GoS)

In order to be able to compare performance of different RRM algorithms like AC and DCA, which make trade-off between blocked and dropped users, we also use a Grade of Service (GOS) metric, which combine blocking (and unsatisfied user) rate and dropping rate in one formula.

Since call dropping is experienced by a user as a much more severe system failure than call blocking, the relative weight of dropping probability is much greater (usually 10 times) than the weight of blocking probability. GOS is defined in terms of call blocking $P_b$ and call dropping probability $P_d$ for RT users as follows [110]:

$$GOS = P_b + 10P_d \tag{4.1}$$

In the case of NRT packet data users, we define GOS as a combination of unsatisfied users and dropped users criteria:

$$GOS = unsatisfied\_users[\%] + 10 * dropped\_users\,[\%] \tag{4.2}$$

For both RT and NRT users, the higher the GOS, the worse the performance of a system (RRM algorithm) is.

## 4.3 Power Control and Scheduling

In this section we give examples of power control algorithms based on the “pricing” concept. We also investigate performance of different PC algorithms in networks without prices like ad-hoc networks when certain percent of population use one PC algorithms and the other percent use the other PC algorithm(s). We give an example how cooperation might be “enforces” without “prices” by use of a threshold-based “punishment”.

The simulations in this section are done with RUNE simulator from [110] (see also subsection 4.1.1) with parameters as represented in Table 4-7.

**Table 4-7: Overview of system parameters in RUNE simulator [110]**

| Parameter | Value |
|---|---|
| Number of clusters | 36 |
| Cell Radius [m] | 100 |
| Reuse | 3 |
| Channels per Cell | 5 |
| Offered traffic (Erlangs per cell) | 3 |
| Gain at 1 meter [dB] | -31 |
| Noise [dBm] | -118 |
| Distance attenuation coefficient (alpha) | 3.5 |
| Standard deviation for the lognormal fading (Sigma) [dB] | 8 |
| Down link correlation (Raa) | 0.5 |
| Correlation distance (in log-normal map) [m] | 110 |

### 4.3.1 *Pricing-based PC*

In this subsection we investigate how CIR-target of the CIR-based PC (see subsection 2.2.2) should be set in order to maximize (3.24) under constraints (3.25). The algorithms which set CIR-target of the CIR-based PC is called outer-loop PC (see subsection 2.2.4).

If the users have the same service type, the weights $w_i$ in (3.24) are the same for all i and we can assume that the users occupy the same amount of bandwidth (channels) i.e. data rate $R_i$ depends only on $CIR_i$. Consequently, the network (cell) objective (3.24) becomes:

$$U(network) = \max\left[\sum_{i=1}^{N} CIR_i\right] = \max\left[\sum_{i=1}^{N} \frac{p_i g_i}{I_i}\right] \tag{4.3}$$

In order to maximize (4.3) under constraints (3.25), the outer-loop power of each user changes from the user’s CIR-target in following manner:

- If the actual outage (*actual_outage*) in the cell (network) is lower than allowed outage (*max_outage)* the outer loop PC can increase CIR-target of the users for the amount proportional to the difference *max_outage* - *actual_outage*. In order to maximize (4.3) (increase "efficiency") the users with better conditions (higher maximal achivable signal quality $CIR_{max}$) should increase their CIR-targets (and thus power) for a higher amount than users with worse conditions, since the gradient of network utility increases with the increase of users gain and interference i.e:

$$\frac{\partial U(network)}{\partial p_i} = \frac{g_i}{I_i} = G + I \text{ (in log-scale)} \sim CIR_{max}.$$

- If the actual outage (*actual_outage*) is higher than allowed outage (*max_outage)* the outer loop PC must decrease CIR-targets of the user for the amount proportional to the difference *actual_outage* - *max_outage*. In order to decrease outage probability (increase "fairness"), again the users with better conditions, which have higher signal quality, should decrease their CIR-targets (and thus overall interference) proportional to the difference between their $CIR_{max}$ and minimal signal quality ($CIR_{min}$) needed to achieve $R_{min}$. The users with worse signal quality do not have anyway a much room for CIR-target decrease, since they need almost the maximal power to achieve their minimal CIR.

The difference between *actual_outage* - *max_outage* represents "excess demand" (see subsection 3.9.1). Making use of the equation (3.30), we can set power price $\lambda$ proportional to *actual_outage* - *max_outage* i.e.:

$$\lambda = k * (\, actual\_outage - max\_outage) \tag{4.4}$$

To be able to convert the above heuristics in a practical algorithm we need a relatively simple equation for new CIR-target ($CIR_{thr}$) according to old CIR-target, power price $\lambda$ and locally available information like actual power and some constants like maximal power for the user $P_{max}$ and minimal signal quality $CIR_{min}$. We can also increase CIR proportional to power price $\lambda$ as well as the difference between maximal and minimal signal quality of the user ($CIR_{max}$ - $CIR_{min}$). Consequently, we obtain the following equation for $CIR_{thr}$ update:

$$CIRthr = max[CIRmin, CIRthr - \lambda * (CIRmax - CIRmin)] \tag{4.5}$$

Taking into account that maximal achievable CIR for the user with signal gain $G$ and interference $I$ is $CIR_{max} = P_{max} + G – I$ (in log-measure) and that actual power $P$ is set according to $P = min\ (P_{max}, CIRthr - G + I)$, we obtain from (4.5):

$$CIRthr = max[CIRmin, CIRthr - \lambda * (Pmax - P + CIRthr - CIRmin)] \tag{4.6}$$

Note that in above equation we do not need to measure signal gain or interference but just use the actual user power P. When the power "price" $\lambda$ is negative (demand lower than offer), the CIR-target can be increased. In Figure 4-5, an example of CIR-threshold setting according to the "price" for the power is represented.

*Each user:*

*Each $CIR_{thr}$ update cycle do:*

$\lambda = max(cell\_price, own\_price);$

$CIR_{thr} = max[CIR_{min}, CIR_{thr} - \lambda *(P_{max} - P + CIR_{thr} - CIR_{min})]$

*Each cell:*

*cell_price = k*( actual_outage - max_outage)*

**Figure 4-5: A pricing-based PC algorithm**

The user set its CIR-threshold according to equation (4.6). Naturally, if the users value their power more than their cells ("*own_price*"), they don't have to increase their CIR-threshold, even if the outage is below the limit. The algorithms makes use of the equation (3.30), where power price $\lambda$ is inverse proportional to the $CIR_{thr}$. A similar algorithm can also be used for water-filling level settings in the case of water-filling based PC.

We set minimal CIR ($CIR_{min}$) to different values (4, 8 and 15 dB) and compared the algorithm according to Figure 4-5 with **Max Power PC** algorithm, where users always send with their maximal power i.e. 30 dBm (see Figure 4-6). In Figure 4-6 we used for "pricing" PC from Figure 4-5 the following parameter values: *max_outage = 5%, k* = 10 and *own_price* was negligible in comparison to cell price.

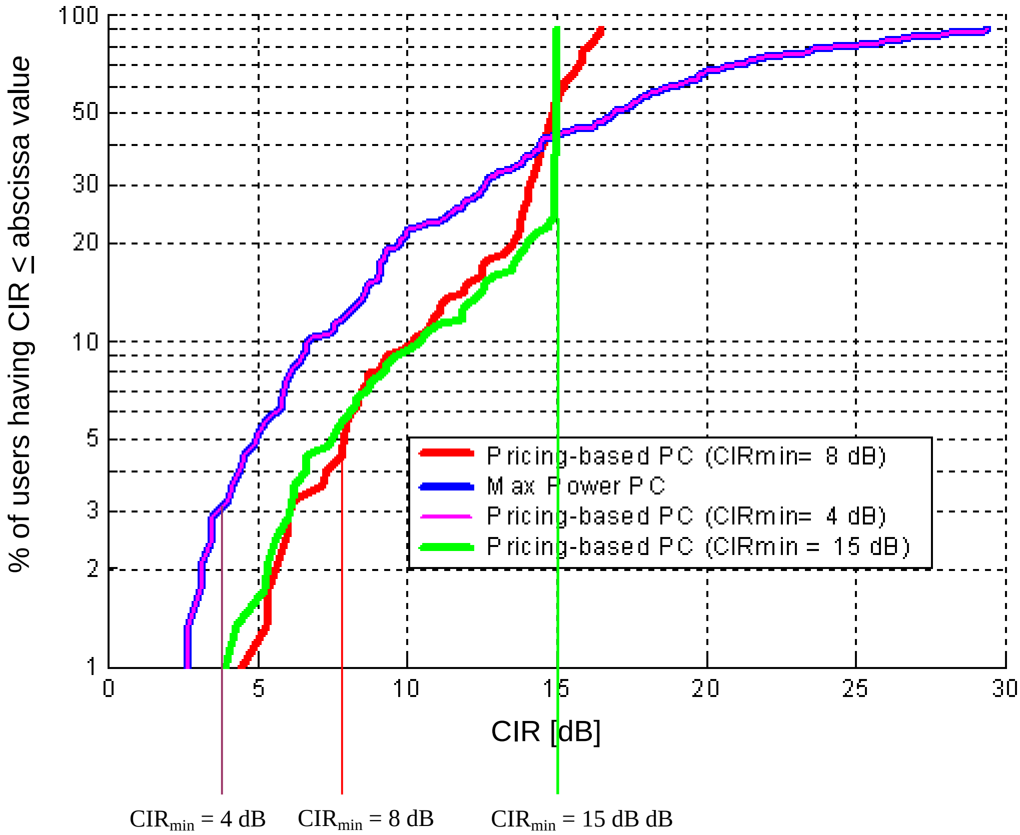



**Figure 4-6: Pricing-based versus Max. Power PC**

As can be seen from Figure 4-6, for higher minimal CIR (8 and 15 dB) Pricing-based PC achieve **much lower percent of unsatisfied users** (5 and 20 % respectively) than Max power PC (10 and 40 % respectively). The users were declared as unsatisfied if their CIR was lower than minimal CIR (see 4.2.1).

On the other side, in the case of Max. Power PC **some users get much higher CIR** and thus can obtain much higher data rates than in the case of Pricing-based PC in the case of higher minimal CIR-targets.

For lower minimal CIR (like 4 dB in Figure 4-6) the Pricing-based PC shows the same CIR-distribution as Max. power PC, since the outage probability for this CIR-targets lies below maximal allowed outage probability (5%) i.e. power "price" is very low.

Also, **if the provider would like to maximize the number of satisfied users it should use some kind of Pricing-based PC in order to keep the outage probability below predefined limit.** Furthermore, Max. Power PC becomes the special case of the Pricing-based PC in the case of relatively low minimal CIR: If the outage probability is relatively low, almost all users can send with maximal power and obtain the higher signal quality.

#### 4.3.2 Power Control Strategies in “Mixed” Populations

In this subsection we investigate different PC strategies like: “Max. Power”, “CIR-based” and “Water-filling” PC in the case when the power “prices” are not predetermined by the network.

We investigate performance of the PC strategies in “mixed” populations i.e. certain percent of users (say 50%) apply one PC strategy (say “Max Power”) and the other users apply the other PC strategy (say “CIR-based”). This situation might arise **in ad-hoc networks**, when there are no distinguished users like base stations which might determine the power “prices” and thus, indirectly the power strategy.

As the performance merits we use the following criteria:

- The **percent of satisfied users** i.e. the percent of users having the CIR equal or above certain threshold (in our case 8 dB).
- **Power consumption**: The less power is transmitted to obtain the same CIR (distribution), the better the performance.

We assume that from the above two criteria **the percent of satisfied users is more important**. In the figures below we provide CIR-distribution and power distribution for the users with different PC algorithms. Note that 100% on the y-axis is never reached because some users are blocked due to lack of the channels (hard blocking) and do not have any CIR. Also in this subsection all “Water-filling” users have the same water-filling level (2W in Figure 4-12 and 20 mW in all other figures with Water-filing PC).

In following we investigate the performance of different PC strategies in mixed populations.

##### 4.3.2.1 CIR-based PC versus Max Power PC

In Figure 4-7 CIR-distribution of CIR-based and Max Power PC alone (“pure” population) and together (“mixed” population: 50% users apply CIR-based PC and 50% apply Max Power PC) is represented.

As can be seen from Figure 4-7, as expected, in the “pure” populations (all users apply the same PC strategy) the number of satisfied users (users having CIR $\geq$ 8 dB) is higher when all users apply CIR-based PC than when all users apply Max Power PC. This is because the lower interference is generated with CIR-based PC: Users send with so much power as they need to achieve the desired CIR-threshold (8 dB) but not more (see subsection 2.2.2).

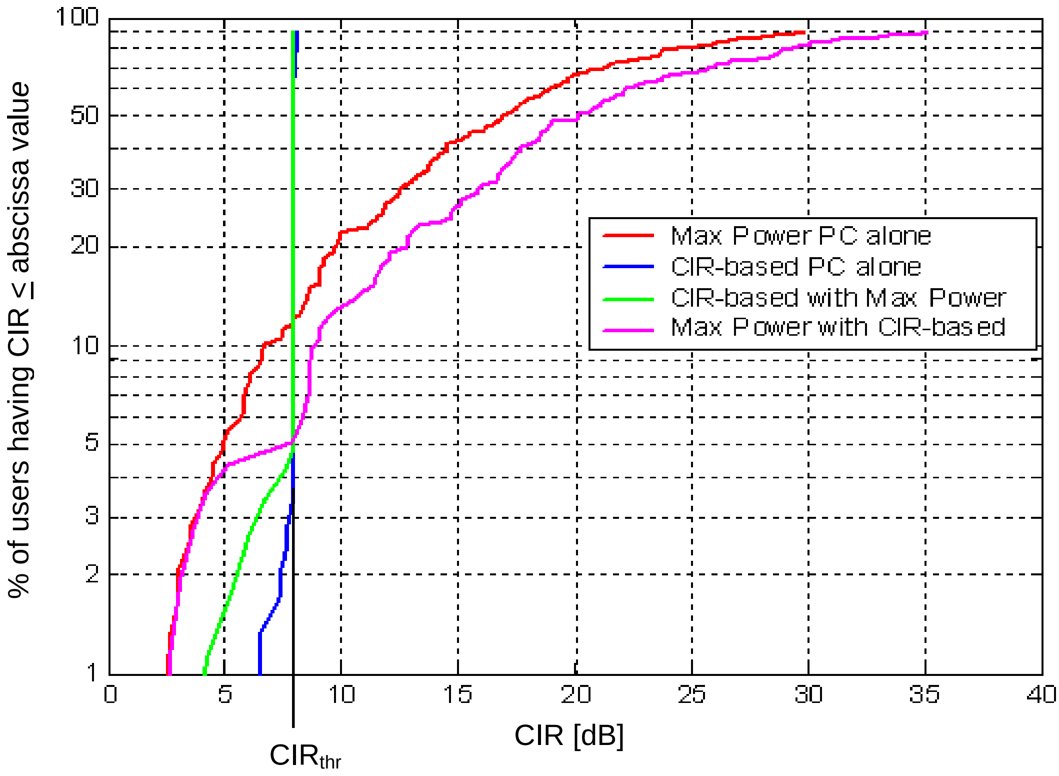



**Figure 4-7: CIR-based and Max Power PC alone ("pure" population) and in "mixed" population (50% CIR-based and 50% Max Power PC users)**

Even in the **"mixed" population** (50% of the users apply CIR-based PC and 50% of the users apply Max Power PC) **users applying CIR-based PC achieve at least the same or better performance (in number of satisfied users) than Max Power PC**. But performance of CIR-based PC in the "pure" population i.e. when all users apply CIR-based PC is better (the number of users having CIR above 8 dB is higher).

This indicates that (as discussed in subsection 3.13.4) **CIR-based PC is an Evolutionary stable strategy** in comparison to Max Power PC i.e. the expected gain (the number of satisfied users) of the CIR-based PC played against itself is higher than the gain of the MaxPower PC played against CIR-based PC. Furthermore, the power consumption of CIR-based PC is much lower than with Max Power PC (see Figure 4-11), since the users send with the power needed to achieve the desired CIR-threshold, but not more.

In the terms of evolutionary biology **a population of "CIR-based" users in an ad-hoc network can not be invaded by "Max Power" users** i.e. the users in such mixed population have the incentive to apply rather "CIR-based" strategy than "Max Power" strategy, if the performance merit is the probability to be satisfied (to have the CIR equal or above certain threshold) and total power consumed.

*4.3.2.2 Water-filling PC versus Max Power PC*

In Figure 4-8 CIR-distribution of Water-filling and Max Power PC alone and together (50% users apply Water-filling PC and 50% apply Max Power PC) is represented.

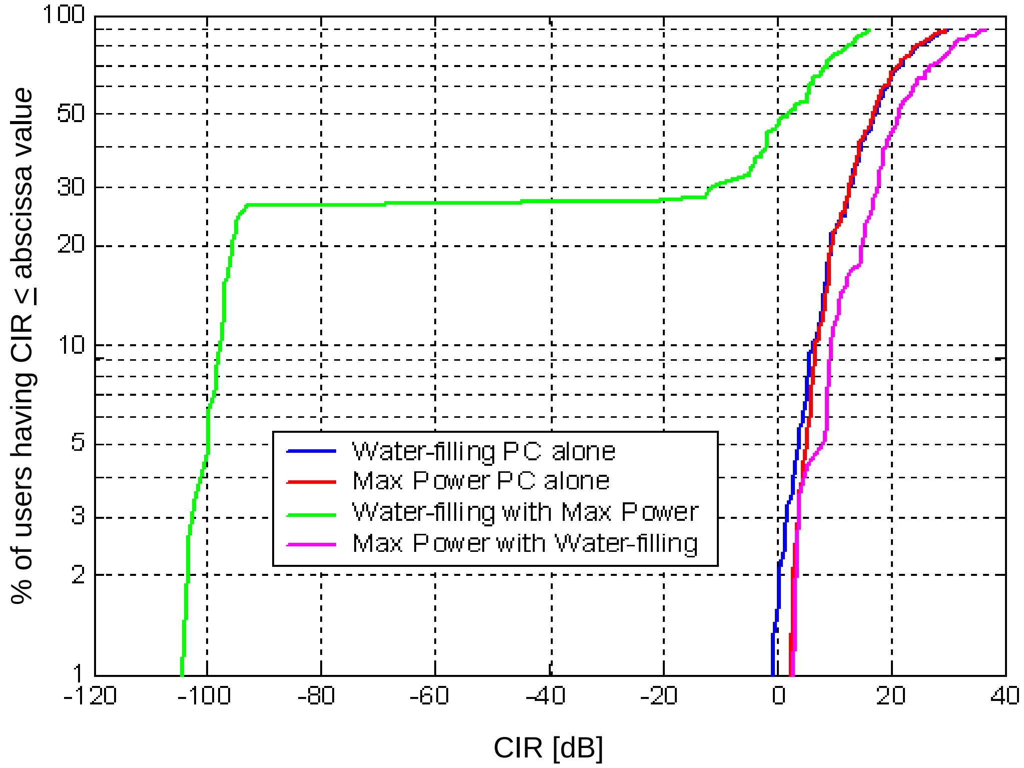



**Figure 4-8: Water-filling and Max Power PC alone ("pure" population) and in "mixed" population (50% Water-filling and 50% Max Power PC users)**

As can be seen from Figure 4-8, as expected, in the "mixed" populations (50% users apply Water-filling PC and 50% apply Max Power PC) the number of satisfied users (users having CIR $\geq$ 8 dB) and probability of achieving better CIR values is much higher when Max Power PC is used than with Water-filling PC. This is because the **Water-filling PC is very cooperative** – it decreases power when the interference (the power of the other users) increases.

This indicates that (as discussed in subsection 3.13.4) **Water-filling PC is not an Evolutionary stable strategy** in comparison to Max Power PC i.e. the expected gain (the number of satisfied users and data rates) of the Max Power PC played against Water-filling PC is higher than the gain of the Water-filling PC played against itself. This is because **Max Power PC exploits cooperativeness of Water-filling PC**, it sends with the highest possible power and in turn get reduced interference due to decreased power by Water-filling PC users. On this way CIR of the Max Power PC users (experience lower interference and use higher powers) is much higher than CIR of water-filling users (experience higher interference and use lower powers).

In the terms of evolutionary biology **a population of "Water-filling" users in an ad-hoc network can be invaded by "Max Power" users** i.e. the users in such mixed population have the incentive to apply rather "Max Power" strategy than "Water-filling" strategy if the performance merit is the probability to be satisfied or to achieve higher CIR.

#### *4.3.2.3 CIR-based PC versus Water-filling PC*

As can be seen Figure 4-9 from **both CIR-based PC and Water-filling PC are better of in the mixed population** (50% CIR-based PC users and 50% Water-filling PC users) than in "pure" population (only one PC strategy is used).

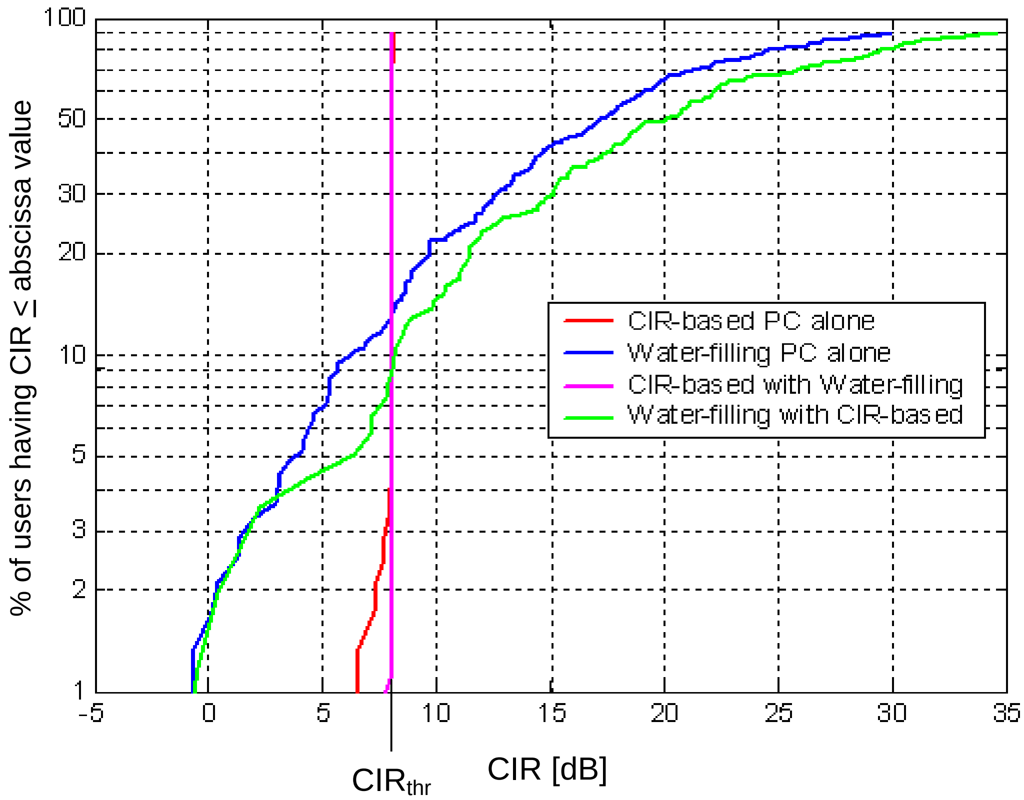



**Figure 4-9: CIR-based and Water-filling PC alone ("pure" population) and in "mixed" population (50% CIR-based and 50% Water-filling PC users)**

This mutual beneficial behaviour of CIR-based and Water-filling PC is caused by following reasons:

- "CIR-based" users profit from "Water-filling" users since **"Water-filling" users do not send any power at all when the channel is too bad** (interference too high and/or channel gain to low). On this way interference to "CIR-based" users is reduced and they can obtain required CIR (8 dB) with higher probability than in the population consisting of “CIR-based” users alone.

- “Water-filling” users profit from “CIR-based” users since **“CIR-based” users send only so much power needed to achieve desired CIR-threshold** (8 dB), but not more. On this way interference to “Water-filling” users is reduced and they can obtain higher CIR than in the population consisting of “Water-filling” users alone.

Also, **a mix of “CIR-based” and “Water-filling” users is desirable**, which is of special importance in the case of **mixed services** (like speech and packet data users). In the case of mixed services **speech users might apply CIR-based PC and packet data users might apply “water-filling” PC and both might profit from each other.**

#### *4.3.2.4 CIR-based, Water-filling and Max Power PC*

According to Figure 4-10 **CIR-based PC strategy is the most “robust” strategy in the population mix consisting of “CIR-based”, “Max Power” and “Water-filling” users**.

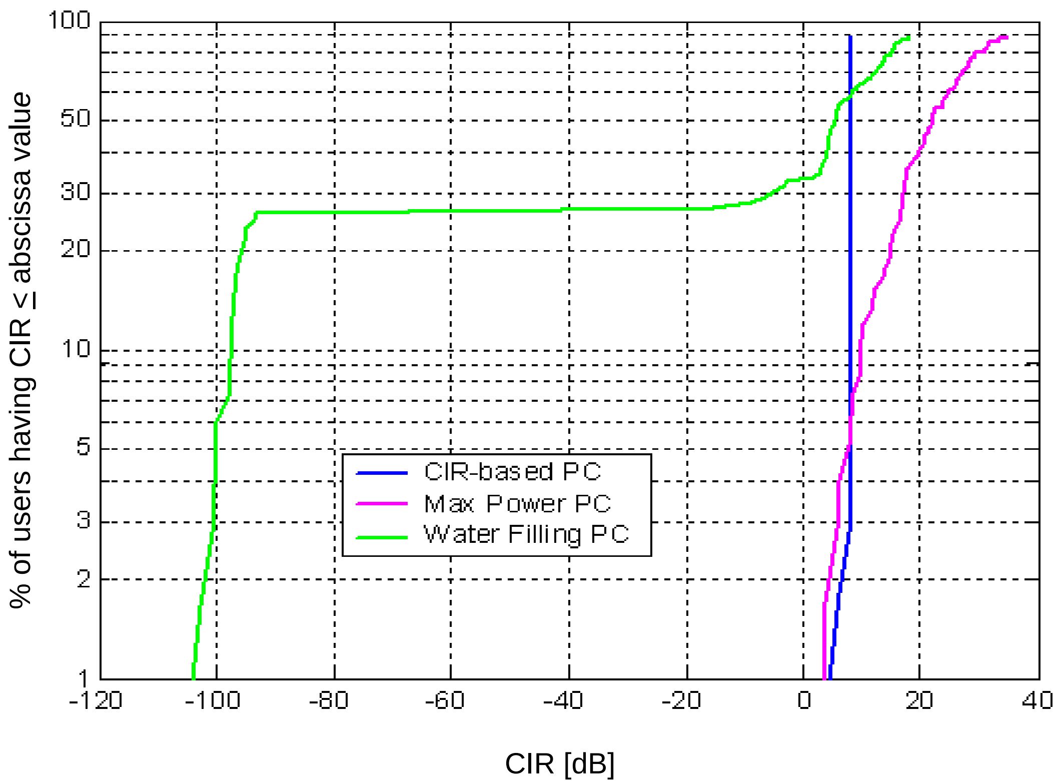



**Figure 4-10: CIR distribution for different PC strategies (each PC strategy is used by 1/3 of population)**

CIR-based PC users achieve their required CIR-targets (are satisfied) if it is possible by using the power lower or equal the maximum available power. In general “CIR-based” PC users need much lower power than Max-Power users as can be seen from Figure 4-11.

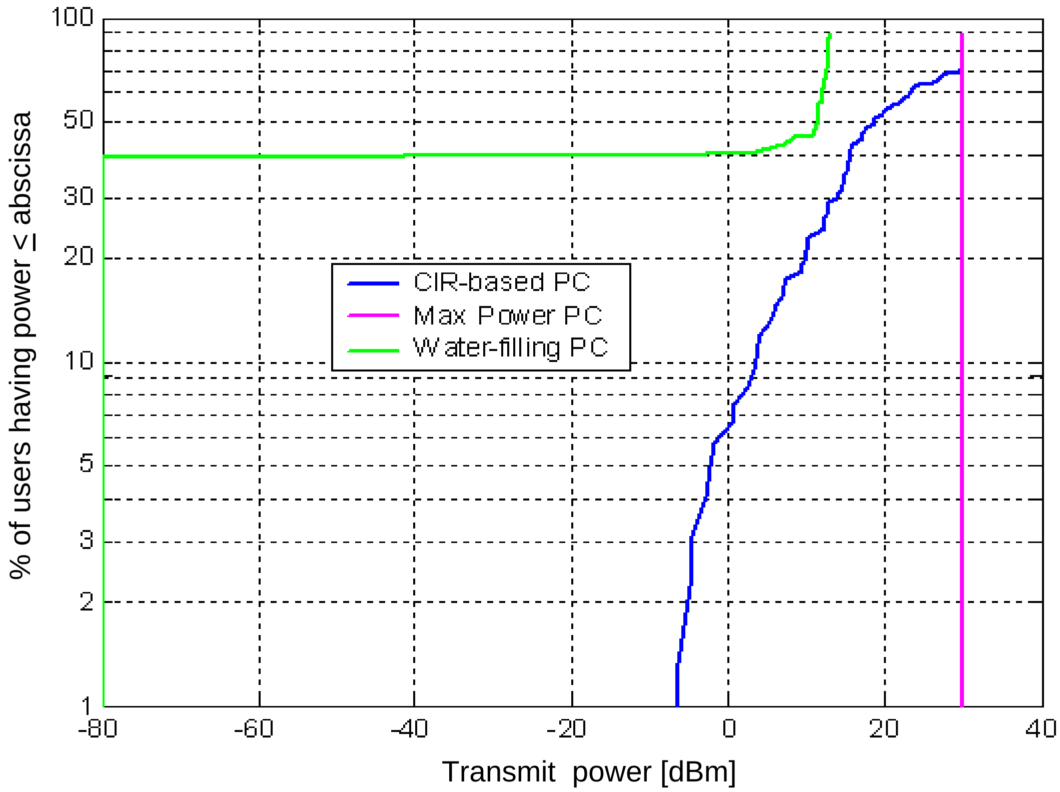



**Figure 4-11: Transmit power distribution for different PC strategies**

Water-filling users consume at least power from all three PC strategies but on the costs of signal quality reduction. This is because **“defensive” character of “water-filling” PC to decrease power if interference increases**. This make Water-filling PC strategy **subject to exploitation by Max Power PC** strategy as described in subsections 3.13.4 and 4.3.2.2 (see also Figure 4-8 and Figure 4-10.

In order to avoid the exploitation by non-cooperative “Max Power” users, “Water-filling” users could **set their water-filling level very high** and obtain the same performance as Max Power PC (see Figure 4-12). In Figure 4-12 water-filling level is set to 2 W instead to 2 mW as in the case of Figure 4-8 and Figure 4-10. When the “water-filling level is very high as in the case of Figure 4-12 all “water-filling” users send with maximal power i.e. Max Power PC becomes a special case of Water-filling PC. On this way “water-filling” users can protect themselves against “Max Power” users in the case of lower loads. In the case of higher loads CIR-based PC can be used as a “security” strategy.

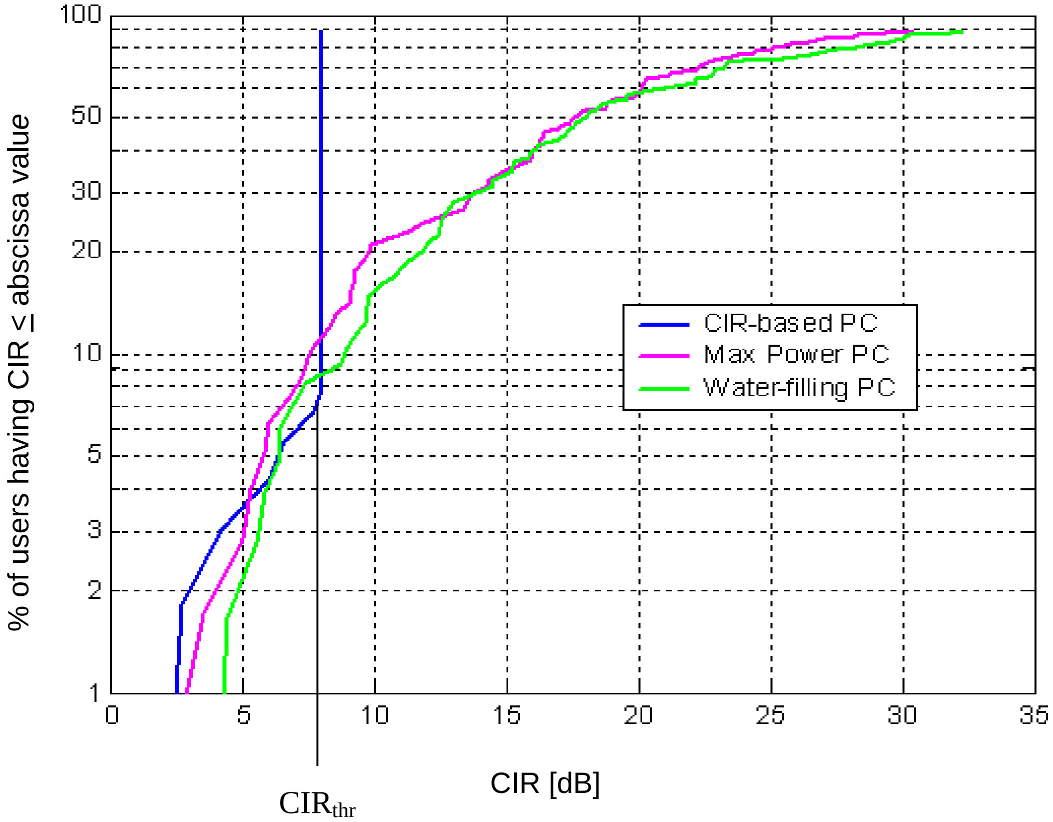



**Figure 4-12: CIR distribution for different PC strategies (water-filling level 2 W)**

Summarizing, in order to **protect itself in a non-cooperative environment** users should apply **CIR-based PC** or **set "water-filling" level to very high values**. CIR-target or water-filling levels should be set **according to load**:

- For **lower loads CIR-target or "water-filling" level could be very hig**h i.e. users should be able to achieve the **maximal data rates** required by the service type.
- For **higher loads** CIR-target or "water-filling" level should be set so that the users could be able to achieve the **minimal required data rates**.
- For the intermediate loads CIR-targets and "water-filling" levels should be set **inverse proportional to load**.

We can also use some sort of "punishment" to enforce cooperation in networks without "prices". This issue is discussed in the next subsection.

#### 4.3.3 Threshold-based PC

The problem with the Pricing-based PC is that in the networks without price-setting controllers (base station) like in ad-hoc networks, it is difficult to “enforce” optimal user behaviour for each network state without “prices”. Furthermore, some users might “defect” and send with maximal power also in situation when a lot of users are unsatisfied (see previous subsection). In order to “enforce” cooperation we can use a threshold-based “punishment” as discussed in subsection 3.13.4. “Defecting” users can be “punished” by their communication partners by decreasing their signal quality (CIR), if the CIR of the “defecting” user is (much) higher than CIR of its partner (“threshold”).

For example, a “Threshold-based PC” algorithm is a simple but practical application of threshold-based “punishment” (see Figure 4-13). For each user i we select the user j for comparison. The user j can be for example a randomly chosen user (by the BS) in a cellular network or the peer communication partner in ad-hoc networks. If the CIR of the user j ($CIR_j$) lies below minimal CIR ($CIR_{min}$) or $CIR_i$ is higher than $CIR_j$ for some parameter (“*unfair*”), the $CIR_{thro}$ of the user i is decreased for the amount defined by parameters “*punish1*” and “*punish2*” respectively. Usually, *punish1*> *punish2* (see Figure 4-13).

```
For each user i
  take randomly user j;
   if CIR(j) < CIR_min
         CIR_target(i) = max(CIR_min, CIR_target(i) – punish1);
   elseif CIR(i) - CIR(j) > unfair
         CIR_target(i) = max(CIR_min, CIR_target(i) – punish2);
    end if
end for
```

**Figure 4-13: A threshold-based PC**

In this way **the defecting users i.e. the users who obtain “unfair” high signal quality (CIR), can be punished by their communication partners or cells** (in a cellular networks). This can be achieved by decreasing of the CIR of the defecting users, for example, **by generating by its communication partner more noise to the signal of “defecting” user** (for amount of *punish1* or *punish2* in the case of Figure 4-13) or just **not decoding the data from “defecting” user**. This in turn can be signalized to the “defecting” user (assuming that users have a backwards signalization channel), which encourage him to decrease its CIR-target and consequently its transmit power. On this way the percentage of satisfied users is increased due to overall interference reduction but on the cost of the reducing the highest achievable CIR, which is approximately equal $CIR_{mean}$ + *unfair* (see Figure 4-13).

The CIR distribution for Threshold-based PC(according to Figure 4-13 (*CIRmin* = 8 dB, *punish1* = 10 dB, *unfair* = 5 dB and *punish2* = 3 dB) in comparison with Pricing-based and Max. power PC is represented in Figure 4-14.

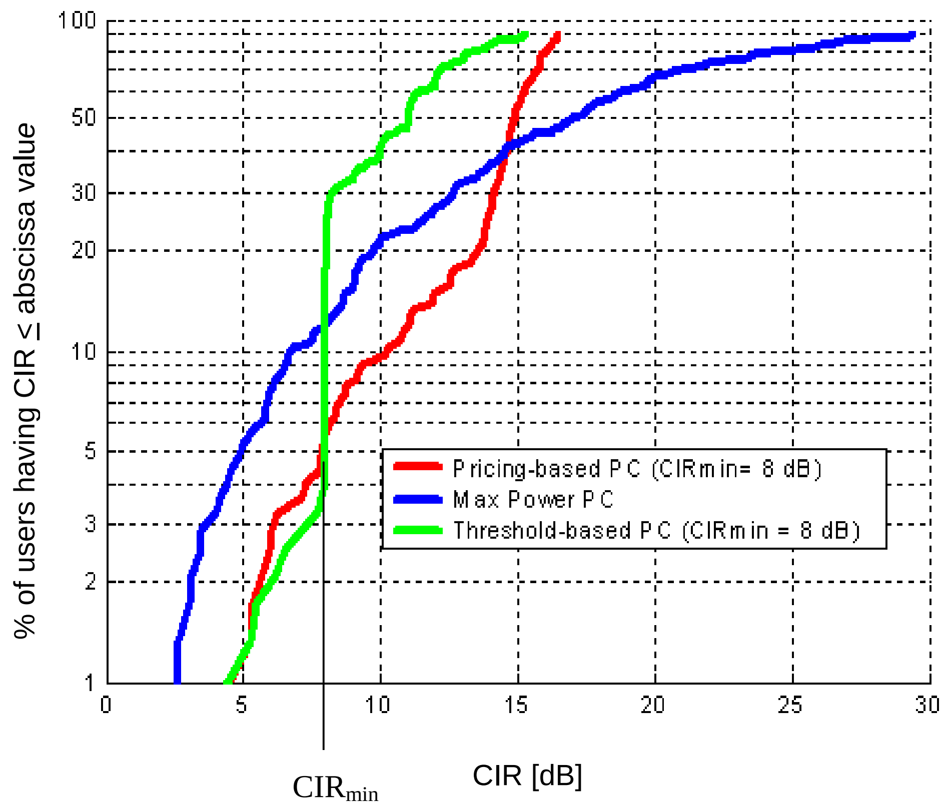



**Figure 4-14: Threshold-based PC versus Max. Power and Pricing-based PC**

As can be seen in Figure 4-14 threshold-based PC has less "smooth" CIR-distribution curve than Pricing-based PC i.e. some users have worse CIR than with "pricing". This is due to "punishment" of defecting users – their CIR is drastically reduced (down to $CIR_{min}$) when there are many unsatisfied users (having CIR lower than $CIR_{min}$ = 8 dB). However, **with "threshold-based PC", the percent of unsatisfied users is highly decreased** (from 10% to 5%) **in comparison with "defecting" Max Power PC.** That is why threshold-based PC can be used in networks without "prices" like **in ad-hoc networks** in order to **increase the number of satisfied users**.

## 4.4 Admission Control and Handover

### *4.4.1 Admission control*

We investigated **interference based AC** (AC "price" increase with the interference) for both service types: RT speech and NRT data. Interference based AC for RT speech means that the users are not admitted to a cell if interference in the cell is above a certain (interference) threshold (see subsection 2.6.2.2).

In Figure 4-15, dropping ratio as a function of AC interference threshold for new users and HO users in the case of RT speech service is represented. In our "pricing" terminology, the AC interference threshold is inversely proportional to the "price" of cell access. As can be seen from Figure 4-15, dropping ratio increases with increase of AC interference threshold for the new call users, because with the increase of the interference threshold AC becomes less restrictive i.e. more new call users are admitted in the cell (system). Consequently, the number of users in the system is increased and overall interference is increased which in turn implies that more users are dropped due to bad signal quality.

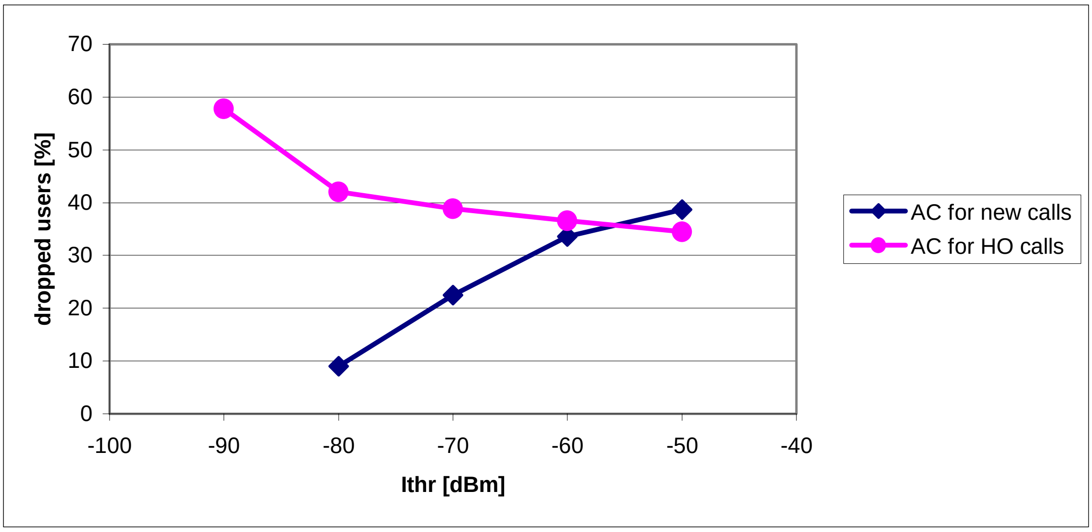



**Figure 4-15: Dropping ratio as a function of AC interference threshold (RT speech)**

On the other hand, in the case of users already in the system (HO users) too restrictive AC (low interference threshold) causes many users to stay connected with their "old" cells instead of making HO to the new cell with higher channel gain. These users then have worse signal quality (CIR) due to lower channel gain and, if PC is used, generate more interference in the system by sending with more power than needed if the user was connected to the BS with lowest path-loss. Consequently, the number of dropped users increases with decrease of AC interference threshold.

The decrease in the dropping rate by more restrictive AC (lower interference threshold) in the case of new users is at the expense of increase in the blocking rate of new users i.e. the number of users who were denied access to the system due to interference higher than a threshold is increased (see Figure 4-16).

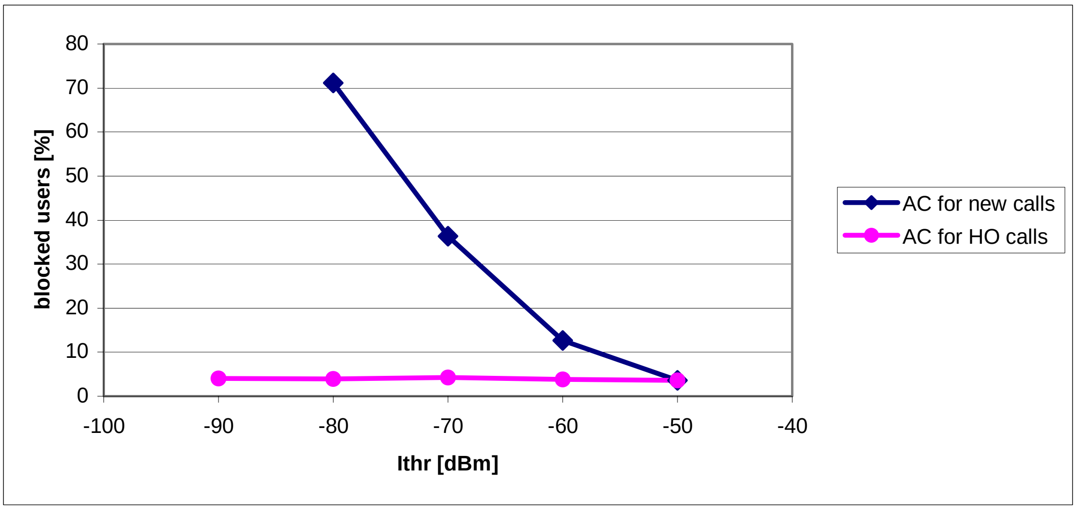



**Figure 4-16: Blocking ratio as a function of AC interference threshold (RT speech)**

Naturally, AC influences the blocking rate only in the case of new users, since HO users are already in the system and cannot be blocked. That is why the blocking rate for HO users (see Figure 4-16) stays constant (Ithr for new users was set to –50 dBm in the case of AC simulation for HO users).

Consequently, in the case of new users **a trade-off should be made between the dropping and blocking rate.** As a measure of AC performance we use a “grade of service (GOS) i.e. a weighted sum of dropping and blocking probability (see (4.1)), the higher the GOS, the worse the performance. Since call dropping is experienced as a much more severe network failure than call blocking, we weight dropping probability 10 times higher than the blocking probability.

In Figure 4-17, GOS as a function of AC interference threshold for new call and HO users are depicted. As can be seen from Figure 4-15, Figure 4-16 and Figure 4-17 the dropping probability is the determining factor for GOS due to the much higher weight of call dropping in comparison with call blocking.

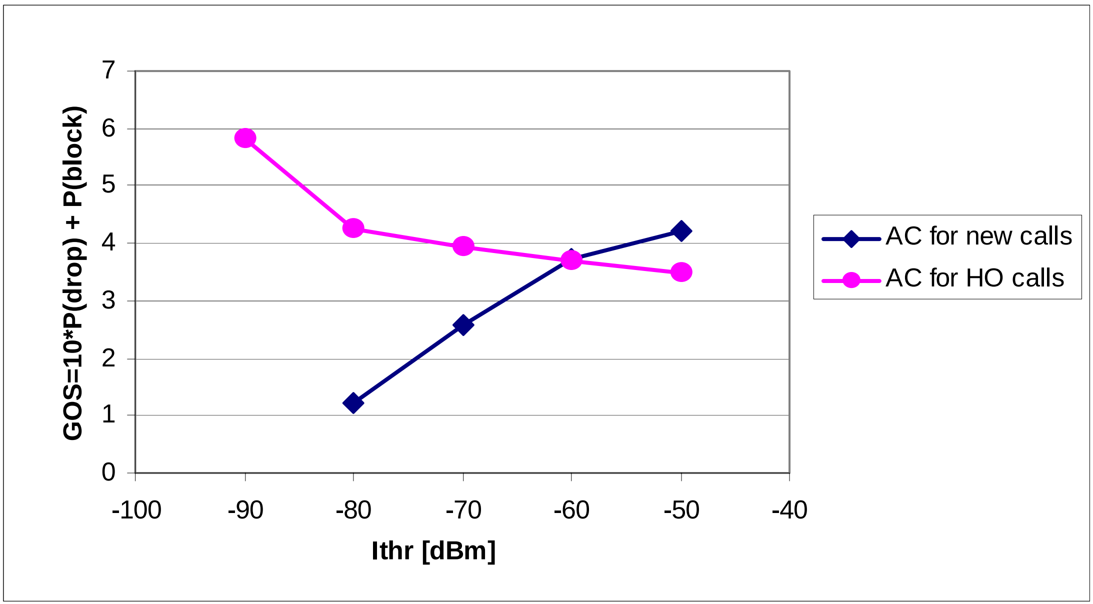



**Figure 4-17: GOS as a function of AC interference threshold (RT speech)**

According to Figure 4-17, t**he AC interference threshold should be relatively low (restrictive AC) for the new calls in order not to disturb the users already in the system**. **For HO calls**, **the AC threshold should be relatively high** (less restrictive AC) i.e. AC should not prevent users already in the system from making HO **in order to reduce overall interference in the system and thus maximize the number of satisfied users**. Figure 4-15, Figure 4-16 and Figure 4-17 are made for the case of relatively high load, when some gain can be achieved with AC at the system limit by protecting the active users at the cost of new users. For lower loads, AC should naturally also be less restrictive for new users.

The results of interference based AC in the case of NRT data are represented in Figure 4-18. As defined in subsection 4.2.1 an NRT user is satisfied if it achieves at least 10% of its nominal data rate (in this case 38.4 kbps for a 384 kbps service). An NRT user is dropped if it needs more than a certain number (in this case 10) retransmissions. The role of AC in the case of NRT users is to prevent users from sending the data from the queue if interference on all free channels lies above a certain threshold. If the interference threshold is set too restrictively, the users would be held in the queues too long. In this way, the users cannot achieve their minimal data rates and the number of unsatisfied users can drastically increase if the interference threshold is higher than a certain limit (-35 dBm in the case of Figure 4-18).

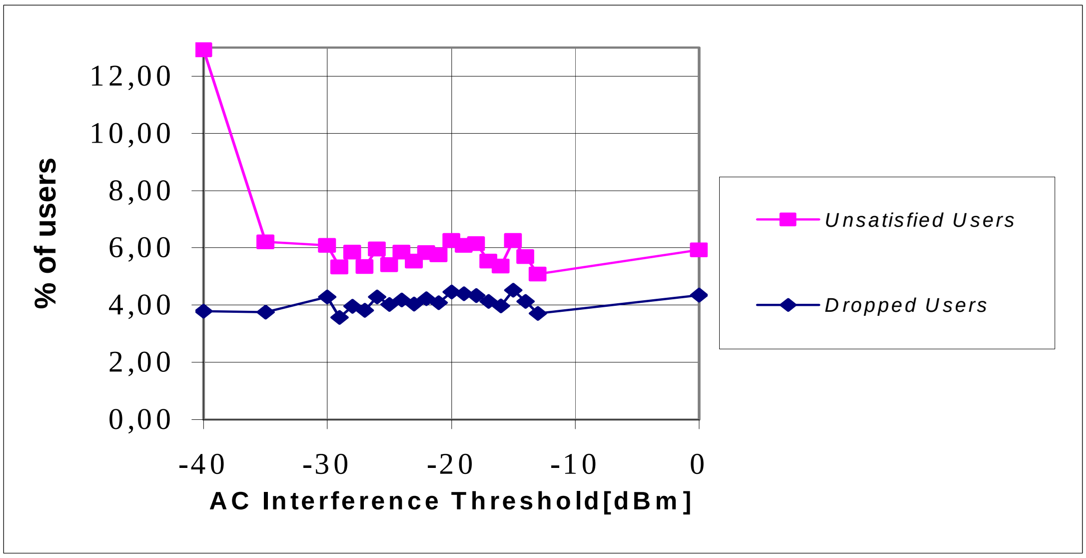



**Figure 4-18: Interference-based AC for NRT data**

On the other hand, the decrease of interference threshold has almost no impact on the ratio of dropped users (see Figure 4-18), since keeping the data in the queue, if interference is over a certain threshold, does not increase the number of retransmission, at least for this range of interference thresholds.

### *4.4.2 Handover*

In this subsection we investigate by simulations some basic trade-offs that must met by Handover (HO) algorithms. In our dynamic simulator we adopt re-use factor 1 (all cells use all channels) and CIR-based PC. According to subsection 3.8.1 we should use path-loss-based HO, since path-loss-based HO helps keeping maximal eigenvalues on the channels lower than 1 in order to ensure the convergence of the PC. With pathlos-based HO, the change of the cell (HO) is performed **when the channel gain of a candidate cell is for HO-margin (parameter) or more better than the gain of the active cell** (see subsection 2.6.1). The parameter **HO-margin can be interpreted as the "price" for HO** (in dB): HO is performed when the utility of HO is equal or above the "price" for HO (see (3.37)). The "price" for HO depends on the provider's relative utility from HO i.e. increase in system capacity (the percent of satisfied users) and HO overhead i.e. increase in the number of HO as a function of the parameter HO-margin.

As can be seen from Figure 4-19, the percent of satisfied users, as expected, decreases with the increase of the parameter HO-margin. With higher HO-margin, users stay connected longer with a "worse" cell (the cell with lower channel gain) and consequently need more power to achieve their required signal quality (CIR), which in turn increases overall interference in the system.

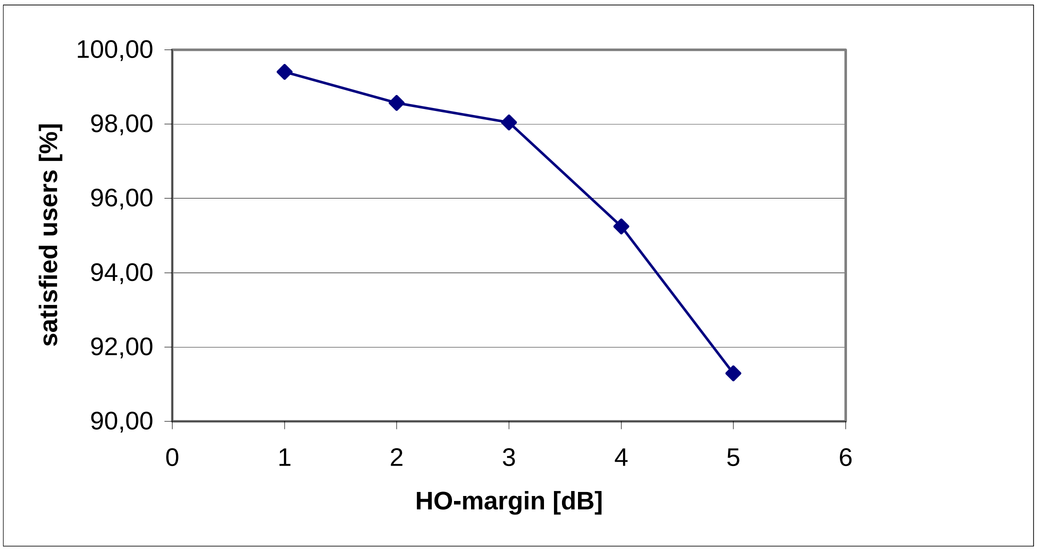



**Figure 4-19: Ratio of satisfied users as a function of HO-margin**

The "price" for increasing the percent of satisfied users by decreasing the HO-margin is the increase in the number of HOs in the system. As can be seen from Figure 4-20, the mean number of HOs increases for about 5 additional HOs when HO-margin decreases for 1 dB.

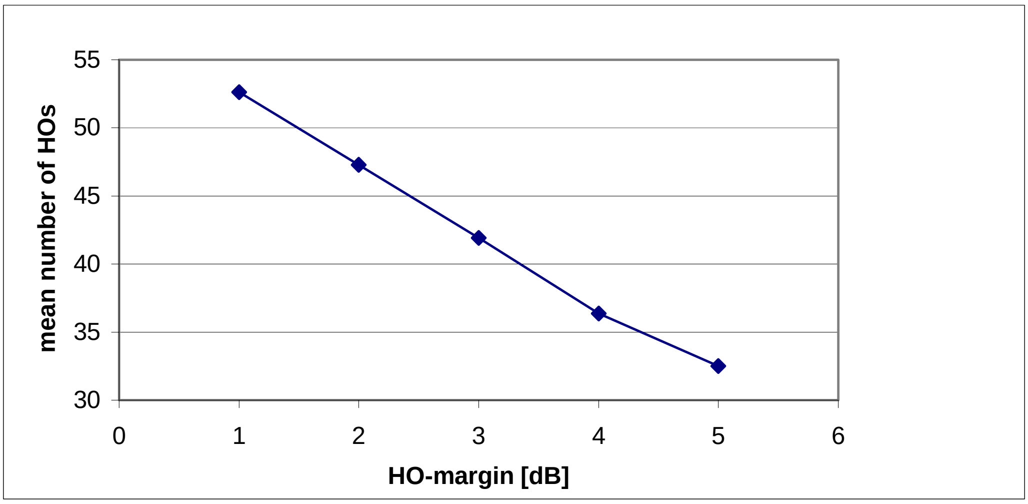



**Figure 4-20: Mean number of HOs according to HO-margin**

A **provider should make a trade-off between the percent of satisfied users and the signalization overhead due to the larger number of HOs.** In order to solve these trade-offs, a provider might need a **unified scale** for measuring **signalization overhead for HO** and gain of HO. For example, the costs for a HO can be measured by an **additional interference** due to measurement and signalization overhead needed for HO execution, and gain of HO can be measured by **decrease in overall interference due to HO**. Finally, HO-margin with best gain (capacity) and cost (overhead) ratio should be chosen.

## 4.5 Dynamic Channel Allocation (DCA)

The role of DCA is to influence by channel allocation the shape of the capacity sets (available rates, see Figure 3-6) and minimize maximal eigenvalues (in the case of PC) on as many channels as possible (see Figure 3-6). This could be accomplished by flexibly combining of optimal DCA algorithms for each network state. This combining is performed by the use of "pricing" concept as described in section 3.5.

We show in this section how, by using appropriate "price" politics i.e. weight settings in the users' utility function, "optimal" channel allocation algorithms can be "enforced" for each load in the cells. We obtain as special cases three DCA algorithms: Random, Minimum Interference (Min I) and Autonomous Reuse Partitioning (ARP) DCA (see section 2.4) as "optimal" DCA for certain loads.

In this section we analyze also simulation results of some distributed, measurement based dynamic channel allocation algorithms described in 2.4 and 3.13.5. Distributed DCA means that cells (or users) make channel allocation decisions independently from other cells (or users). For channel allocation decisions, cells (or users) utilize statistics (like channel gains and interference) obtained by measurements of mobiles and BS in the cell. We also assume re-use factor 1 i.e. each channel can be used in each cell. A channel selected by DCA is a (frequency, timeslot) – pair since we simulated a UMTS TDD system (see section 2.1) and within the channel up to 16 codes can be allocated to different users. Each user needs at least one code (depends on service) in order to establish communication with a BS.

We have analyzed different DCA algorithms under different environments (urban (Micro) and sub-urban/rural (Macro), with different service types (speech and NRT 384 kbps packet data), with and without PC and SA.

### *4.5.1 Dynamic Channel Allocation*

The following Dynamic Channel Allocation (DCA) algorithms are obtained by appropriate "price" settings in the user's utility function (see subsection 3.9.3) and investigated by simulations:

***Random Channel Allocation (RCA)***: Chooses channel **randomly** from the set of free channels.

***Minimum Interference (Min I):*** Chooses the channel with the **lowest average interference** from the set of free channels.

***Priority-based (Channel Segregation):*** Chooses the channel with the **highest priority** from the set of free channels. Priorities are established according to previous experience on the channels in a cell: At allocation request, the cell tests the interference on a free channel with the highest priority against an interference threshold. If interference on the channel is lower than the threshold a cell uses the channel and increases the priority of the channel in the cell, otherwise the cell decreases the priority of the channel (see subsection 2.5.2.2). In this way "channel segregation" should be established i.e. neighbor cells should learn to use different channels with different frequencies [28], [1]. If one cell often uses some channels its neighbors should seldom use the channels and vice versa. In Figure 4-21, it is represented how channel segregation emerges during simulation: At the beginning of the simulation (Figure 4-21 left) almost all channels are used equally like in neighbor cells (BS_1-3 are neighbor base stations). After a certain simulation time (10 second or 10 000 radio frames), due to application of priority-based DCA a channel segregation is established i.e. the channels ((frequency timeslot) – pairs) often used by a cell are used relatively seldom by its neighbors and vice versa (Figure 4-21 right).

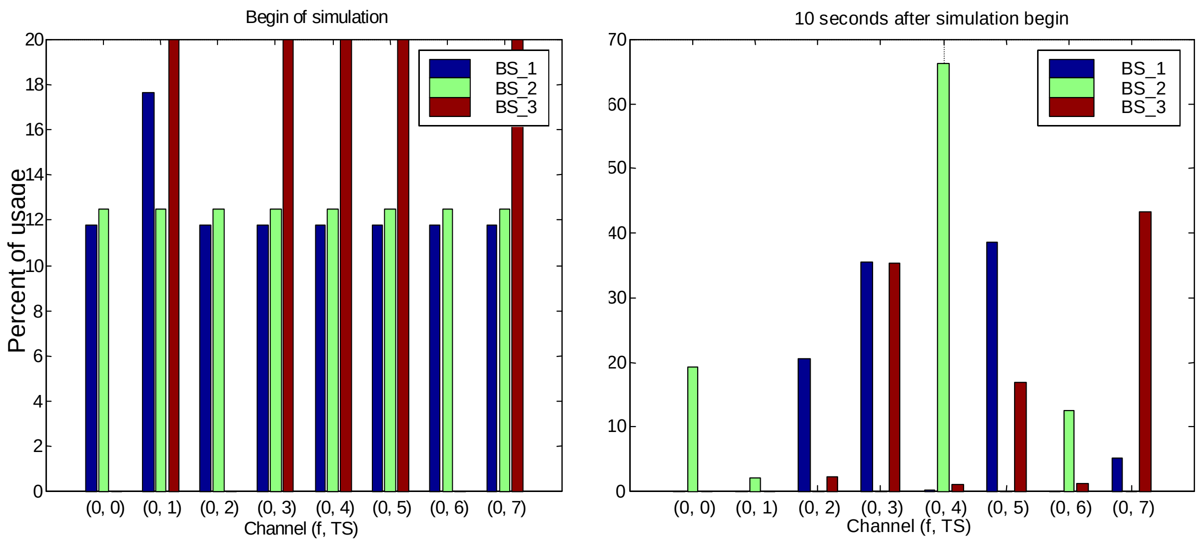



**Figure 4-21: Establishing of channel segregation in the case of lower loads: Left channel usage at the simulation begin, right channel usage after 10000 frames (10s)**

In the case of higher loads, almost all channels must be used in all cells, so that channel segregation is worse than for lower loads (compare Figure 4-21 rights and Figure 4-22).

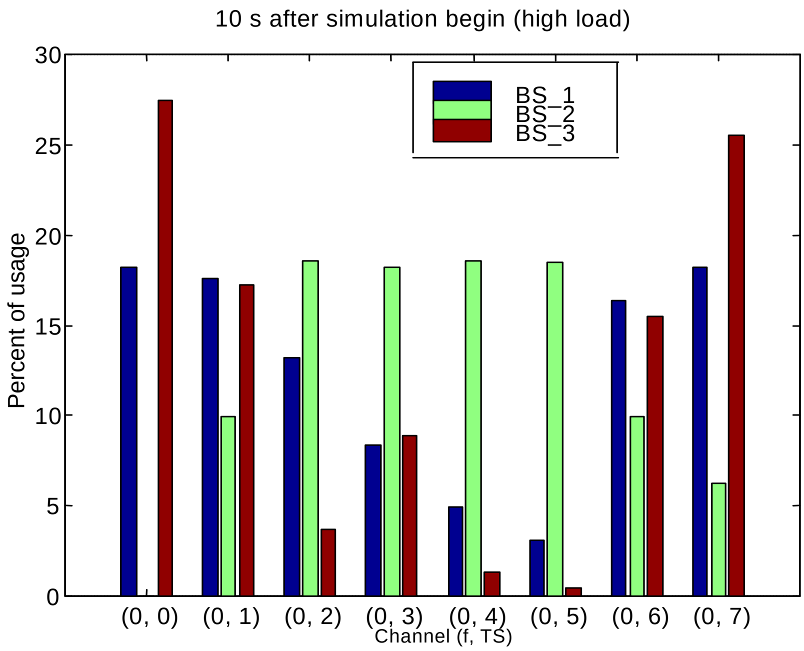



**Figure 4-22: Channel segregation in the case of higher loads**

***Autonomous Reuse Partitioning (ARP):*** The idea of ARP is that the users far away from the base station (high pathlos) should use channels with lower interference and users near the base station (low pathlos) should use the channels with higher interference (see 2.5.2.4). In this way the number of users in the system with sufficient signal quality should be maximized. In order to simplify implementation, we used the following ARP algorithms: Each time a user needs a channel DCA allocates to the user the free channel k for which the expression $|G_i - I_{ik} - C|$ takes the maximal value. $G_i$ is channel gain between the user i and the base station (in DB) and $I_{ik}$ is the (average) interference experienced by the user i on the channel k (in dBm and always negative) and $C$ is a simulation parameter. An “optimal” value of the parameter $C$ changes with load and each time the load is changed, the optimal value of the parameter $C$ has to be found for the load. Note that if we take the parameter $C$ high enough the channel with minimum interference would always be selected since G is channel independent in a TDD system and I is always lower than 0. This means that minimum interference DCA is a special case of the reuse partitioning as defined above, and reuse partitioning DCA can achieve at least the same performance as minimum interference DCA.

In Figure 4-23 the results of the comparison of the above algorithms are depicted [11].

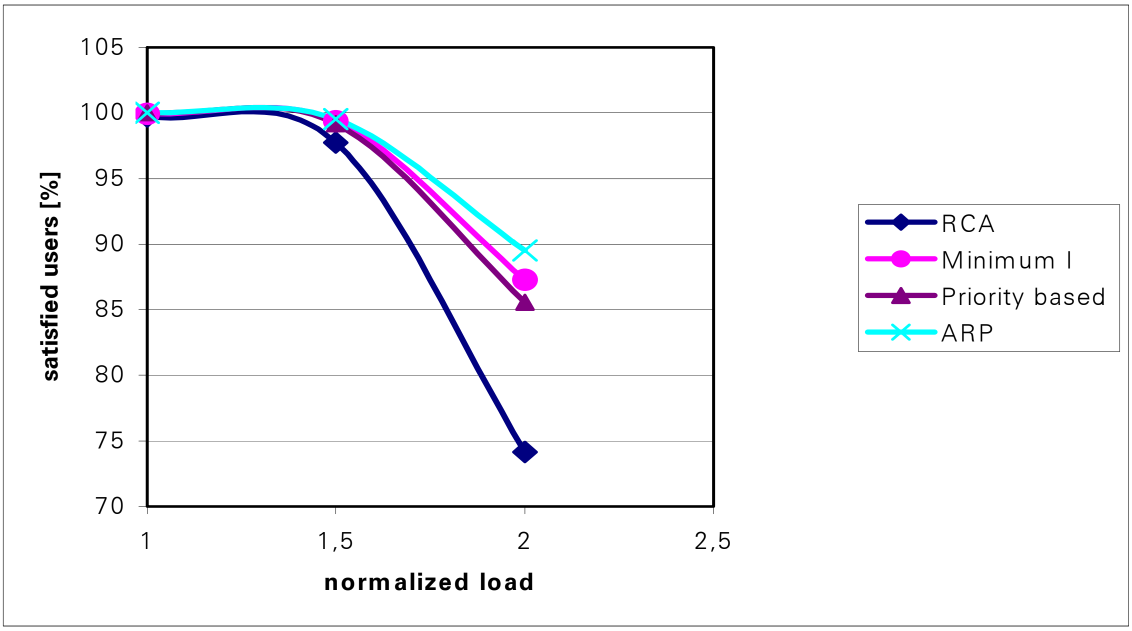



**Figure 4-23: Comparison of DCA algorithms**

As can be seen from Figure 4-23, the best performance e.g. the maximum number of satisfied users is achieved with ARP algorithm and the worst performance with random allocation (RCA). The disadvantage of the ARP algorithm is that the optimal value of the parameter $C$ should be found for each load separately, whereas minimum interference and random allocation DCA do not need any parameters at all. Furthermore, ARP needs measurements of interference and channel gain for its decisions, Minimum interference or Priority-based DCA need interference measurements only and random DCA needs no measurements at all.

Taking into account not only performance (in number of satisfied users) of the algorithms, but also, costs of measurements and parameter optimizations needed for algorithms, we can say that an "optimal" DCA algorithm is load (state) dependent.

Cells can, by appropriate "pricing" (setting the parameters $W_G$, $W_I$ and $C$ of the users' utility function (3.31)), enforce the "optimal" DCA algorithm according to the state of the cells, for example (see Figure 4-24)**:**

| *Cell's part of DCA* | *User's part of DCA* |
|---|---|
| | *At each allocation request find the channel with maximum utility:* |
| *// Set weights ("price"s) for the channel c according to the load of the cell:* | $Max_c(U(c)) = max_c(W_G G(c) - W_I I(c) - C(c))$ |
| | *Among the channels satisfying constraint:* |
| | $Pr(CIR \leq CIRthr) < P_{out}$ *or* $E(r) \geq R_{min}$ |
| | *If "prices" are set as depicted on the left side the following algorithms are chosen:* |
| *case* ***Load*** | *case* ***Load*** |
| ***Low:*** | ***Low:*** |
| $W_I = W_G = 0, C(c) = random$ | $Max_c(random(c)) \Rightarrow$ *Random DCA* |
| ***Medium:*** | ***Medium:*** |
| $W_G = C(c) = 0, W_I = 1$ | $Max_c(-I(c)) = Min_c(I(c)) \Rightarrow$ *Min I DCA* |
| ***High:*** | ***High:*** |
| $W_G = C(c) = 0, W_I = -1$ | $Max_c(I(c)) \Rightarrow ARP$ |
| *end case* | *end case* |

**Figure 4-24: "Pricing" based DCA**

**Low load:** In the case of a low load, interference estimation is relatively bad i.e. interference variance is relatively high in comparison to interference mean ($d_I$, is relatively high, see (3.16) and Table 4-10). Interference based DCA like Min I or ARP might not be optimal for very low loads, since it relies on (for low loads unreliable) interference estimation. But, due to low load, almost all channels would have relatively low interference and we can use simple Random Channel Allocation. In order to "activate" Random DCA cells can set "prices" (weights) to following values (see Figure 4-24 lefts): $W_I = W_G = 0, C(c) = random$. Consequently, users trying to maximize their own utilities choose Random DCA (see Figure 4-24 rights).

**Medium load:** In the case of medium load, with Random DCA the channels with lower and higher interference would be equally probably used (which increase overall interference) and Channel Segregation among neighbor cells could not be established due to increased load (almost all channels must be used in all cells). But since interference estimation is relatively good ($d_I$ from (3.16) is relatively low), we can use Min I DCA in order to allocate channels with lower interference more often, which saves power and reduces overall interference. To "activate" Min I DCA, cells can set "prices" to the following values $W_G = C(c) = 0, W_I = 1$. Consequently, users trying to maximize their utility choose Minimum Interference DCA.

**High load**: In the case of a high load, we can use ARP DCA (see subsection 2.5.2.4), in order to provide as many users as possible with at least minimum signal quality (or data rate). This makes sense, since if revenue of satisfied users is much higher than from data rate ($w_s >> w_i$ *or* $w_d >> w_i$, see (3.13)), maximizing the number of satisfied users brings the most gain to providers. To "activate" ARP DCA, cells can set "prices" to the following values $W_G = C(c) = 0$, $W_I = -1$. Consequently, users trying to maximize their utility choose the free channel with maximum interference, which satisfy users' constraints (see 3.3).

We investigated our pricing based DCA from Figure 4-24 on the example of NRT 384 kbps packet data service and compare it with Minimum Interference (DCA min I), Random Channel Allocation (RCA), Autonomous Reuse partitioning (ARP) DCA. We used **Grade of Service (GOS)**: GOS = unsatisfied users + 10*dropped users, as an evaluation criteria (see section 4.2). As can be seen from Figure 4-25, the best (lowest) GoS for all loads is achieved with the "pricing" based DCA.

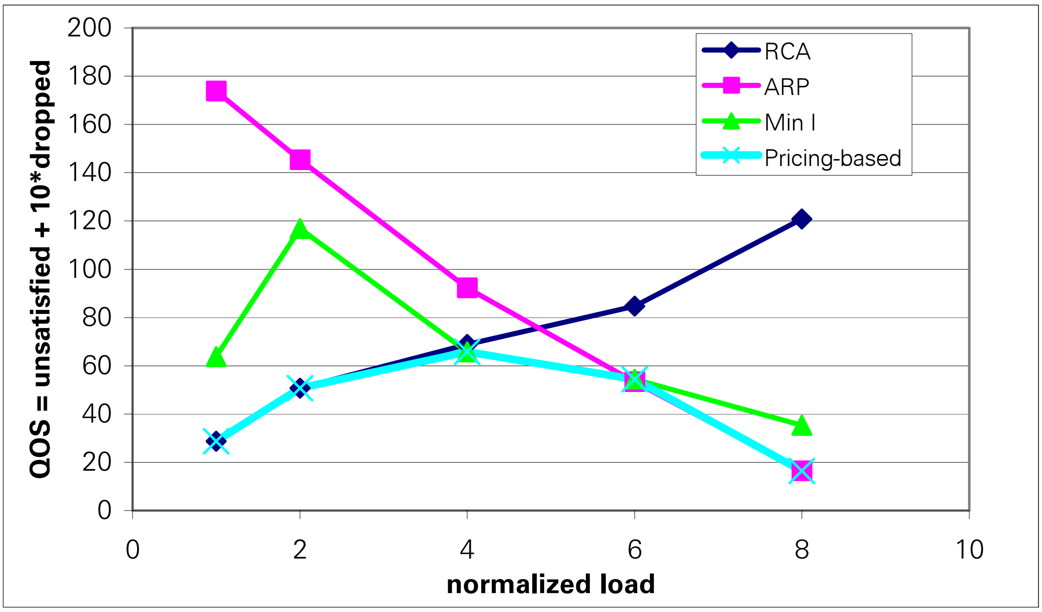



**Figure 4-25: GoS statistics for packet data with different DCA algorithms**

As special case of the "pricing" based DCA we obtain for lower loads RCA, because interference based DCA algorithms have relatively unreliable interference information for lower loads (relatively high variance/mean ratio). For medium loads, the best results are achieved with minimum interference DCA by using channels with lower interference more often. For higher loads, the better performance can be achieved with ARP which try to provide each user with "satisfactory" but not the best channels in order to maximize the number of satisfied users in the system. ARP parameters are load dependent and should be optimized for each load separately.

By using **"Pricing"–based** a network provider can **flexible combine different DCA algorithms** and **activate "optimal" DCA algorithm for each load** simply **by appropriate parameter ("price") settings** in the users' utility functions (see Figure 4-24), **without need for software changes**.

### *4.5.2 PC and DCA*

It is also interesting to see dependence of DCA and other RRM algorithms like (CIR-based) PC (see Figure 4-26).

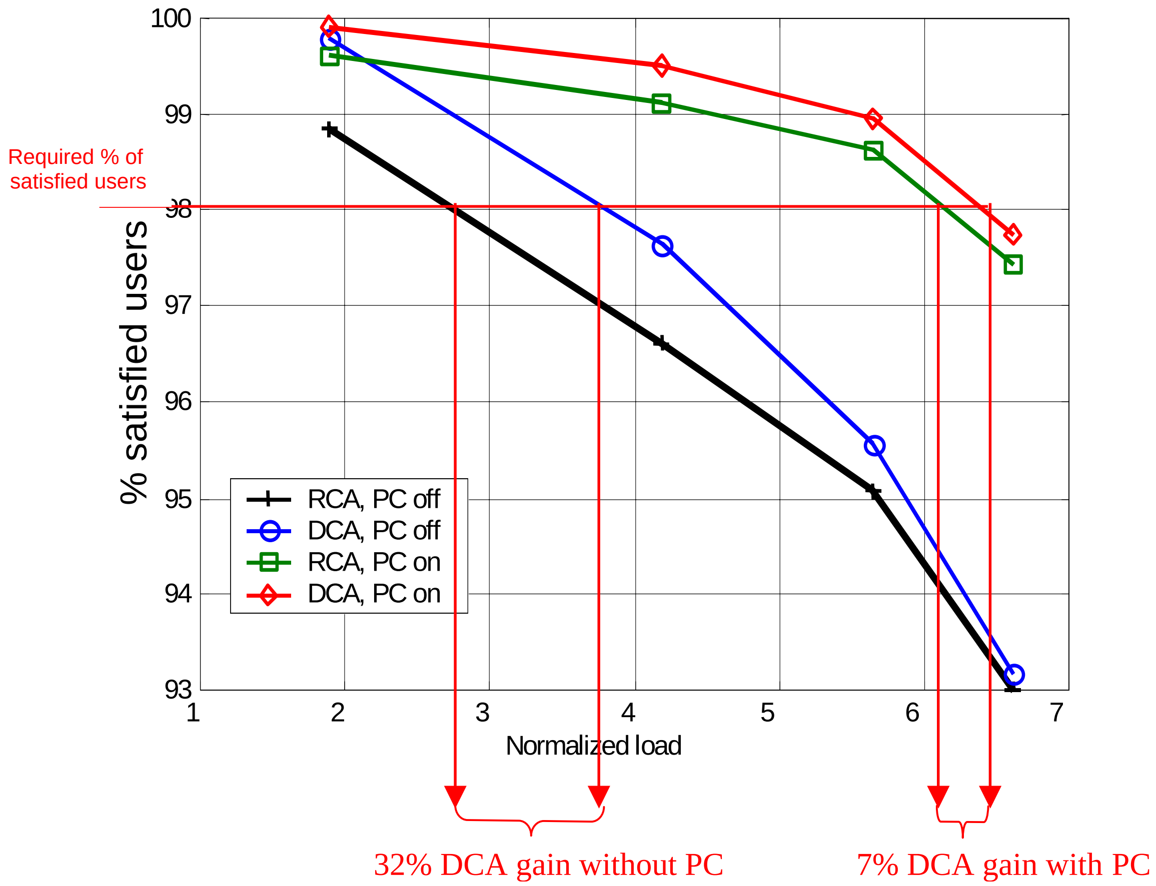



**Figure 4-26: Gain of PC and DCA in urban (micro) environment**

From Figure 4-26 can be seen that gain from ARP DCA is about 36% without PC and only about 5% in the case with PC. The **gains from PC and DCA are also not additive**. PC enables in the case of RCA use of more than 80% of timeslots, thus leaving little room for further improvements by DCA [11]. This is because DCA is based basically on the principle of "channel segregation among the cells": Neighbor BSs should use different channels as often as possible. Since more then 80% of channels can be used as a merit of PC, there are hardly channels left "to be segregated" between neighbor BSs.

But besides capacity gain, **DCA also saves the transmit power** (Txpwr) of the BS and mobiles as can be seen from Figure 4-27. This could be of especial importance for mobile stations because battery life can be increased [11].

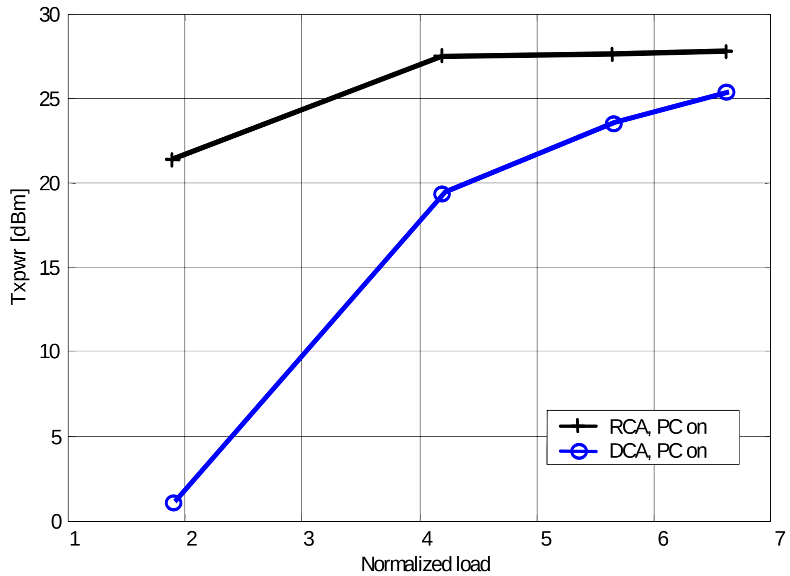



**Figure 4-27: Reduction of transmit power due to DCA in urban (micro) environment**

With DCA, channels with lower interference are used more frequently in the case of low-medium load. Consequently, on average less power is needed to achieve required signal quality (CIR) using CIR-based PC (see (2.5)).

Consequently, **if PC is used, the relative gain due to DCA is reduced** and a provider can use only minimum interference DCA to save power or apply random allocation to reduce algorithm complexity. In the case of **higher loads, ARP channel allocation can be used in order to maximize the number of satisfied users**. A provider does not need to implement each of the algorithms separately but can activate "optimal" DCA algorithm for each load simply by appropriate parameter ("price") settings in the users' utility functions.

### 4.6 Smart Antennas and RRM Algorithms

In this section we analyze capacity gain of RRM algorithms with and without Smart antennas (SA) (see 2.7). As described in subsection 4.1.2.8, SA were simulated only in the Macro environment. As can be seen in Figure 4-28 and Table 4-8**, the capacity gain due to smart antennas can be up to 180% in comparison with the same system without SA**, whereas the capacity gain due to PC is about 70%, and the additional capacity gain due to DCA is about 2%. These results also confirm the results provided in [27] and [52], where the gain of SA is estimated to be 40-200%. In fact, the capacity achieved with smart antennas and PC is very close to the hard blocking limit i.e. the number of users in the system is constrained by the number of channels and not interference. Consequently, almost no significant improvements by, for example, by sophisticated RRM algorithms are possible.

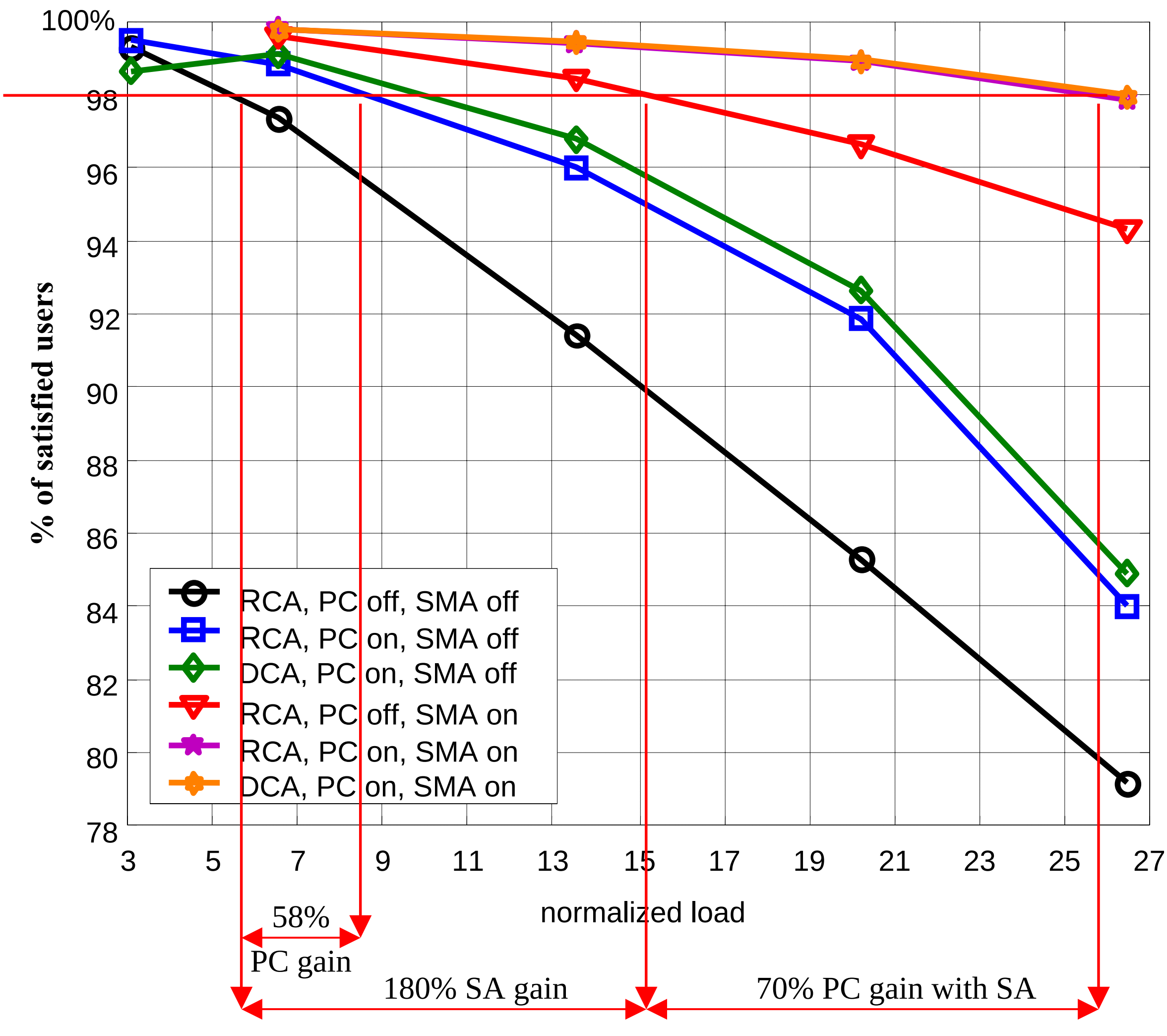



**Figure 4-28: Gains of RRM algorithms with and without smart antennas**

**Table 4-8: Relative (in comparison to the case when the algorithm or technology is switched off) gains of RRM algorithms and SA**

| Scenario | Gain in capacity [%] |
|---|---|
| RCA, PC off, SMA off | Reference gain |
| RCA, **PC on**, SMA off | 58 |
| **DCA**, PC on, SMA off | 14 |
| RCA, PC off, **SMA on** | 180 |
| RCA, **PC on**, SMA on | 70 |
| **DCA**, PC on, SMA on | 2 |

**The high capacity gain achieved with SA is due to highly reduced overall interference by use of SA in the system**. Instead of radiating power in all directions (like omni-directional antennas), **SA radiate power (almost) only in the direction of desired users** (see section 2.7). In this way, interference to other users in systems is highly reduced. Furthermore, **interference reduction is almost independent of the number of users in the system** and (much) higher than interference reduction of any RRM algorithm (see also Figure 4-29). Also, SA effectively **"decouple" the users** from each other and **interference reduction to some users is not paid by interference increase to other users** (like in the case of DCA, see section 4.5). Consequently, signal quality CIR (see (2.3)) and probability to achieve required CIR (to be satisfied) increases with SA.

In order to make a decision whether to deploy SA or not, a provider should compare relatively high capacity gains achieved by SA with the costs of implementing an SA technology. In the early phase of network (system) operation, PC and possibly DCA would be enough to provide the desired capacity. Note that relative gain of DCA is much higher without SA (14%) than with SA (2%). If capacity becomes a problem in a later phase of the system operation, the provider can then employ SA. SA could bring about 50% capacity gain with PC and up to 180% capacity gain without PC. Furthermore, **the employment of SA reduce overall interference so much that parameter optimization of other algorithms (like PC) become less important and some RRM algorithms like a sophisticated DCA might become unnecessary**.

## 4.7 Performance Gain of RRM Algorithms and SA due to Interference Reduction

Since modern wireless systems are interference limited, we expect that the gain of each RRM algorithm or SA can be measured in average interference reduction due to deployment of the algorithm or SA. In Figure 4-29 and Table 4-9 mean interference in DL for speech service as a function of relative load, RRM algorithms or SA is represented.

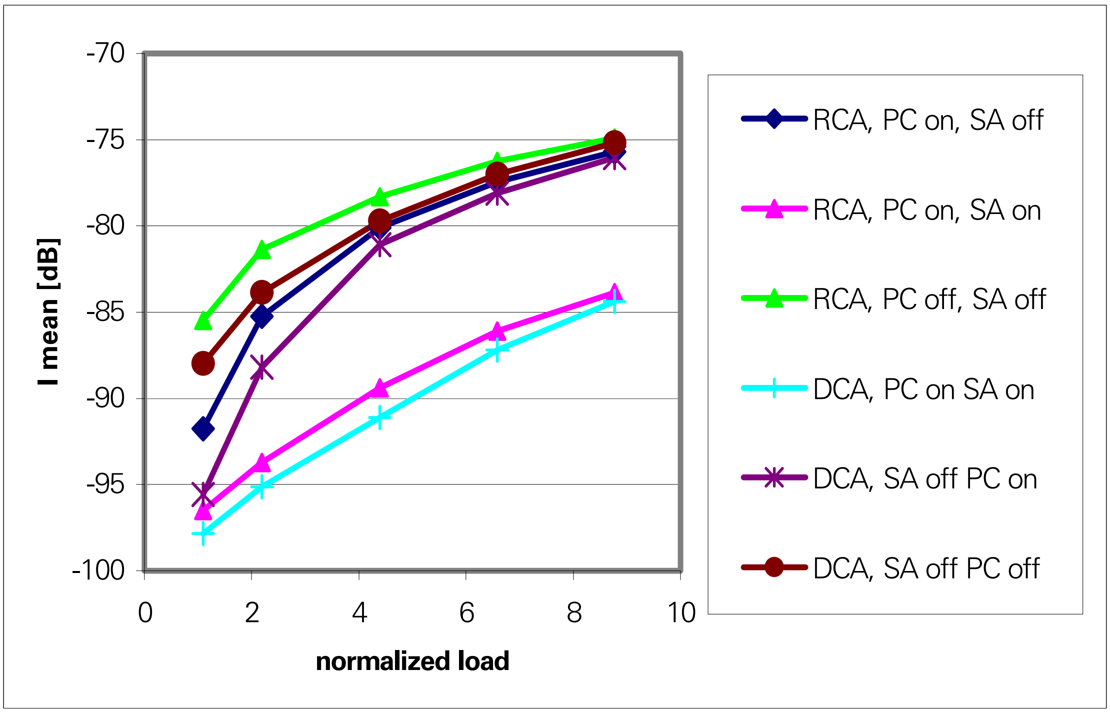



**Figure 4-29: Mean Interference as a function of RRM algorithms and SA**

**Table 4-9: Mean interference values for different RRM algorithms and SA**

| Load / Algorithms | 1 | 2 | 4 | 6 | 8 |
|---|---|---|---|---|---|
| RCA, PC off, SA off | -85,47 | -81,36 | -78,32 | -76,24 | -74,92 |
| **DCA**, PC off, SA off | -87,97 | -83,88 | -79,70 | -77,02 | -75,18 |
| RCA, **PC on**, SA off | -91,77 | -85,25 | -80,08 | -77,44 | -75,70 |
| **DCA**, PC on, SA off | -95,57 | -88,19 | -81,10 | -78,12 | -76,03 |
| RCA, PC on, **SA on** | -96,50 | -93,73 | -89,39 | -86,09 | -83,86 |
| **DCA**, PC on, SA on | -97,82 | -95,12 | -91,12 | -87,20 | -84,37 |

According to Figure 4-29 and Table 4-9, reduction of mean interference due to different RRM algorithms or SA is load dependent i.e. for lower loads relative interference reduction is higher and vice versa. The contribution of different RRM algorithms and SA to interference reduction could be explained as follows:

- **PC reduces overall interference because the users do not send with maximal power but with a power needed to achieve their required CIR**. Reduction of interference due to PC is 1-7 dB (see Table 4-9). The lower the number of users the greater the gain (interference reduction) of PC because at lower loads the interference is also low and the power needed to achieve required CIR is much lower then maximal power. When the number of users is high, the interference is also high and almost all users must send with maximal power.

- DCA reduce interference in the system **by more frequent use of the channels with lower interference** (minimum interference DCA). The gain of DCA in interference reduction is up to 4 dB and depends on interference i.e. load in the system. For lower loads, a good “segregation” can be achieved i.e. neighbor cells use different channels for most of the time (see Figure 4-21 right). The higher the number of users, the lower DCA gain in reduction of interference, because almost all channels are occupied almost all the time and there are low possibility for channel “segregation” between neighbor cells.

- SA reduce overall interference in the system b**y transmitting the power only in the direction of the desired user (several degrees in space) and not all directions** (360 degrees) as in the case of omni-directional antennas. The interference reduction gain of SA is about 7-10 dB and is **almost independent of the system load**. This is because the capability of SA to reduce interference by transmitting the power always (approximately) in the direction of desired users is almost load independent.

As can be seen from Figure 4-30, **the reduction of overall interference has a direct impact on the number of satisfied users** (for RT speech service), because signal quality CIR is inverse proportional to interference (see (2.3)). For example, reduction of interference of 5 dB (from –80 dBm to –85 dBm) brings about 4 percent more satisfied users (from 94% to 98%). This means that in the case of medium load (about 3 users per channel) application of DCA and PC which bring about 5 dB gain according to Figure 4-29 in comparison with FCA and no PC, would bring the system to the required level [19] of satisfied users (98%). Without these RRM algorithms, the network does not fulfill requirements (only 94% satisfied users). It is interesting to note that the curve of satisfied users as a function of mean interference is pretty good approximated with log-normal distribution for all RRM algorithms.

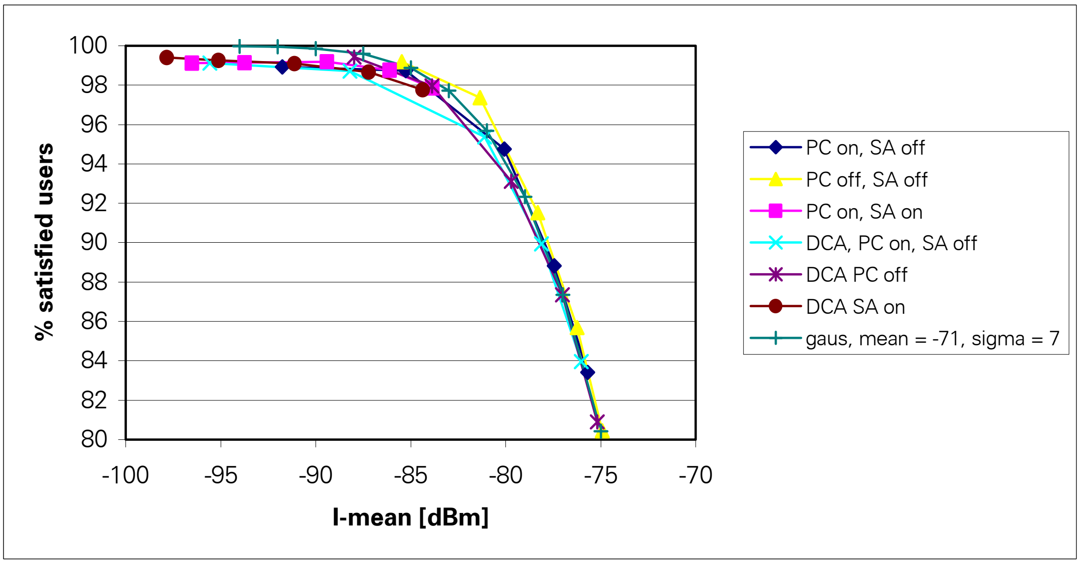



**Figure 4-30: Log –normal Approximation of Interference Distribution**

But, as can be seen from Table 4-10, the effect of RRM algorithms and SA is not only reduction in the interference mean value but also an increase in standard deviation i.e. increase in their ratio ($\sigma/\mu$). This is because **RRM algorithms and SA introduce a "structure" in interference distribution**: Some channels are more used than others due to DCA, some directions (users' locations) are preferred to others due to SA, and users transmit with different powers due to CIR-based PC. The ratio $\sigma/\mu$ is especially high for lower loads, because with arrival and departure of each new user interferences changes in percents are relative high. That is why interference estimation is relatively bad at very low loads and the algorithms, which rely on interference estimation, like PC and DCA cannot achieve such high gains as at higher loads (see Figure 4-28).

**Table 4-10: σ/μ ratios for different loads and RRM algorithms with/without SA**

| Load / Algorithms | 1 | 2 | 4 | 6 | 8 |
|---|---|---|---|---|---|
| RCA, PC off, SA off | 0,49 | 0,40 | 0,32 | 0,31 | 0,30 |
| **DCA**, PC off, SA off | 0,57 | 0,44 | 0,33 | 0,31 | 0,30 |
| RCA, **PC on**, SA off | 0,84 | 0,51 | 0,36 | 0,33 | 0,31 |
| **DCA**, PC on, SA off | 1,12 | 0,63 | 0,38 | 0,33 | 0,32 |
| RCA, PC on, **SA on** | 1,53 | 1,01 | 0,63 | 0,51 | 0,46 |
| **DCA**, PC on, SA on | 1,79 | 1,15 | 0,72 | 0,55 | 0,48 |

Combining Figure 4-29 and Figure 4-30, we obtain a **unified scale** for measuring performance of the RRM algorithms or mobile technologies like SA: **The gain of an algorithm or technology can be measured in reduction of overall interference due to implementation of the algorithm or technology** (see Figure 4-29). The interference reduction in turn results in an increase in the number of satisfied users in the system (see Figure 4-30). The capacity gain of an algorithm (technology) is also directly proportional to the reduction of interference due to implementation of the algorithm (technology). The provider can then compare gains of different algorithms or technologies (like SA) according to their reduction of overall interference.

Since the provider utility $U$ is proportional to percent of satisfied users (see 3.4) and the percent of satisfied users decreases as interference increase i.e. $dU(I)/dI \leq 0$ (see Figure 4-30), interference is "bad" commodity according to microeconomic terminology [64]. This means that the provider would pay for decrease of interference and the price for interference reduction per dB would be proportional to $|dU(I)/dI|$. The slope of the curve $U(I)$ ("price for interference reduction" ~ $|dU(I)/dI|$) depends on the interference in the network. If the interference is near 98 percentile of satisfied users or higher (see Figure 4-30), the slope of the $U(I)$ curve (~ % of satisfied users) is steeper (approximately 2.5% per dB) than for the lower loads (approximately 0% per dB). This means that **the provider has an incentive to pay higher prices for interference reduction techniques in the case of higher (average) loads than in the case of lower loads**. This also means that the **"price" for the power also increases with load** (see subsection 3.9.1).

Also, **investment in SA or RRM algorithms are more likely in the case when the network comes to its capacity limit**. Which algorithm(s) (or technology/technologies) will be finally chosen, depends on the benefit/costs ratio of the algorithms (technology). For example, investments in RRM algorithms are in general cheaper than in SA, since these changes concern only software in the network, but capacity gains from the RRM algorithms are lower than from SA (see Figure 4-28). Like for SA and RRM algorithms the same trade-off could also be implied to other technologies like coding or modulation. A better coding or modulation technique would require lower signal quality (or CIR-target) in order to achieve the same BER. Consequently, less power is needed if CIR-based PC is used with the coding or modulation technique and overall interference in the system could be reduced.

## 4.8 Summary

In this chapter we gave performance metrics and some numerical results for RRM algorithms. Below are the main conclusions of our simulations:

- PC should make a **trade-off between maximizing data rate (efficiency) and maximizing the number of satisfied users (fairness)**. These trade-offs could be achieved by setting the power (CIR-threshold or water-filling level) for the users according to **power “prices”**. Power “prices” can be set according to the network load and interference or, more directly, **according to outage probability in the network**. We showed that **CIR-based PC, in contrast to Water-filling PC, is an “evolutionary” stable, non-exploitable (by Max Power PC), strategy.** On the other hand, **CIR-based PC and Water-filling PC strategies may profit from each other**. To encourage cooperation a **“punishment” mechanism** like a “threshold-based PC” might be used. With a “threshold-based PC” receivers in an ad-hoc network prevent senders from obtaining much higher data rates or signal qualities than defined by a certain threshold, which in turn increases the number of satisfied users by forcing the transmitters to reduce their powers and reduce overall interference.

- **HO makes a trade-off between the number of handovers and system capacity.** With the increase in **HO-margin,** the number of handovers is reduced but the system capacity is reduced too. **AC should make a trade-off between the number of blocked and dropped users.** This trade-off is also **load dependent**: In the case of the high load, AC should give priority to the “HO users” i.e. the users already in the system over the new users; for low-medium loads AC should be less restrictive i.e. admit almost all users in the system (cell).

- An optimal DCA algorithm depends on the state (load) in network. **For lower loads, Random, for medium loads Min I and for higher loads ARP are “optimal” DCA algorithms.** For each cell state, an **“optimal” DCA algorithm for the state can be “enforced” by appropriate (state dependent) “pricing”**. With PC and (even more) with SA, an overall interference in the system is reduced so much that almost all channels have relatively low interference in all cells. Consequently, **almost no capacity gain can be achieved by “smart” channel allocation in present of PC and SA for speech services**. For NRT data services there is still capacity gain obtainable with a “good” DCA, even in the case of PC and SA.

- The **most gain in interference reduction and consequently in number of satisfied users can be obtained with SA and than with PC and then with DCA**. The SA gain is almost **load independent** in contrast to the gains by the RRM algorithms like PC and DCA.

# 5 Conclusions

*"Thus our happiness will never consist, and must never consist, in complete joy, in which nothing is left to desire, and which would dull our mind, but must consist in a perpetual progress to new pleasures and new perfections."*

***Gottfried Wilhelm Leibnitz***

In our opinion two main contributions of this work are:

- **Methodological contribution**: We showed how RRM algorithms can be developed and evaluated by using "pricing" and game theory methodology.
- **Simulation results**: We provided simulation results for capacity gains and analyzed some trade-offs of different RRM algorithms.

In the following, we describe these two kinds of contributions and give some suggestions and hints for further works.

## 5.1 Methodological Contributions

We showed that application of "pricing" and Game Theory on RRM algorithms brings following advantages:

- It enables a **unified framework** for the design, analyses and comparison of RRM algorithms. Many of the **existing, rather heuristic, algorithms can be regarded as a special case of a "pricing" and game theory framework** described in this work.
- The algorithms are **distributed and self-adaptive**. Each unit (BS or MS) optimizes only its own gain (expected utility), given locally available information like channel gain, interference and "prices" (parameters of the utility function). Thus, **signaling is significantly reduced** due to decentralized optimization. Furthermore, the RRM is also adaptive: We **do not need to implement different algorithms for each state**. Instead, we can use the same algorithm with different parameters ("prices") for different states.
- Since we assume **"bounded rationality"** (limited information available), **"satisfactory" solutions** and use of **heuristic rules**, our RRM algorithms are **suitable for practical applications with low computation and signalization overhead.**
- The proposed approach enables an e**conomic evaluation of benefits and costs of different RRM algorithms and technologies like SA**.

## 5.2 Simulation Results

### 5.2.1 *Power Control*

A network has to find a **trade-off** between two basic issues:

- **Total data rate maximization**
- **Maximizing of the number of satisfied users**

A possibility to make the trade-offs between total data rate and the number of satisfied users is a **"pricing-based" PC**, where the CIR-threshold or water-filling level is set according to the prices for the power in network. The **prices are set according to the network's load and interference or according to outage probability in the network**.

We showed in the networks without centralized controllers, like in ad-hoc networks, that **CIR-based PC, in contrast to Water-filling PC, is an "evolutionary" stable i.e. non-exploitable (by Max Power PC), strategy.** Further, **CIR-based PC and Water-filling PC may profit from each other**, which is especially useful **for mixed services** where one service type (RT) can apply CIR-based PC and the other service type (NRT) could apply Water-filling PC. **To encourage cooperation in the networks without prices a "punishment" mechanism like a "threshold-based PC" should be used**. With a "threshold-based PC" receivers **prevent senders from obtaining much higher data rates or signal qualities than defined by a certain threshold**, which in turn forces the transmitters to reduce their powers and reduce overall interference.

### 5.2.2 *Channel Allocation*

Our results show that optimal channel allocation is state (load, interference) dependent:

- **Random DCA should be used in the case of very low loads**, because **all channels are "good" enough** and no signalization of measurement overhead is needed.
- **Minimum interference-based or Channel Segregation DCA should be used in the case of low-medium loads** in order to minimize interference and power consumption in the system.
- **Autonomous Reuse Partitioning (ARP) DCA should be used in the case of higher loads** (user with high path-loss gets channel with low interference and vice versa) in order to provide almost all users with appropriate signal quality (data rate). **If PC and SA are used, the Random DCA could be also used in the case of higher load and RT service like speech,** because overall interference due to PC and SA is highly reduced and almost no further improvements could be achieved with a more sophisticated DCA like ARP.

Nevertheless, **even with PC and SA capacity can be improved for NRT users with a “fair” DCA strategy like ARP**. For all services **significant reduction in transmit power can be achieved in the case of low-medium loads with minimum interference DCA** .

#### *5.2.3 Admission Control and Handover*

Handover (HO) algorithm must make **a trade-off between the number of handovers (signalization overhead) and system capacity** (reduction of interference due to fast HO). This trade-off is summarized in the parameter **HO-margin** in the case of path-loss based HO. In one layer system with CIR-based PC, **path-loss-based HO minimize maximal eigenvalues of the channel matrixes i.e. improves speed of the PC iteration cycles and maximize achievable CIR values on the channels.** In the case of hierarchical cells or inter-system HO (for example GSM to UMTS or vice-versa), HO algorithms can **make (by “prices”) trade-offs between channel gain, interference and system load.**

Admission control (AC) should find a **trade-off between percent of blocked and dropped users**. AC should differentiate between new users and the users already in the systems and **increase the “prices” for cell (system) access for new users in the case of high loads**, since call dropping (of existing users) is generally regarded as a more severe system failure than call blocking (of new users).

#### *5.2.4 Smart Antennas*

Our results show that most capacity gain comes from the deployment of the SA. The capacity gain due to smart antennas can be up to 180% in comparison with the same system without SA (see section 4.6). The high capacity gain due to SA is due to relatively **high reduction of generating interference by use of SA** in comparison to RRM algorithms (see section 4.7) and this interference reduction is almost **independent of the number of users in the system**. SA effectively **“decouple” the users** from each other and **interference reduction to some users is not paid by interference increase to other users** as in the case of some RRM algorithms.

We also showed that **gains of SA and RRM algorithms can be measured on unified scale as an amount of total interference reduction in the system**. If overall interference is reduced, the number of users who can be served with sufficient signal quality (system capacity) is also increased.

Relatively low interference level due to SA has as a consequence that **some RRM algorithm like a sophisticated DCA might be superfluous in presence of SA**

## 5.3 Further Work

We tried with this work to make a step further (however small) on the way to establishing the role of Game theory in RRM research, like, for example, linear algebra plays in signal processing research.

In our opinion, we have only scratched the surface of the iceberg. Some future investigations that could be built upon this work are for example use of **real prices** instead of “fictive” prices to “enforce” optimal resource allocation in real networks **when user utility function is not known, for example by auctions** or investigation of **“evolutionary” stable RRM algorithms** in networks without prices like ad-hoc networks.

We hope that this work will serve as help and motivation for further research in exciting field of radio resources management, or even in game theory, since mathematics can also profit from real world problems.

## *Appendix A: Basics of Mobile Radio Propagation*

In this chapter we describe some basic features of mobile radio propagation and channel modeling according to [41], [81] and [72]. An understanding of the mobile radio channel is an essential part of the understanding of the operation, design, and analysis of any mobile radio system. An illustration of how the signal strength at the mobile station antenna may look like according to distance from the base station is shown in Figure A1. The signal strength as a global mean value decreases with the distance (path-loss). Superimposed on this global mean, slow variations are present due to shadowing effects and fast variations due to multipath fading.

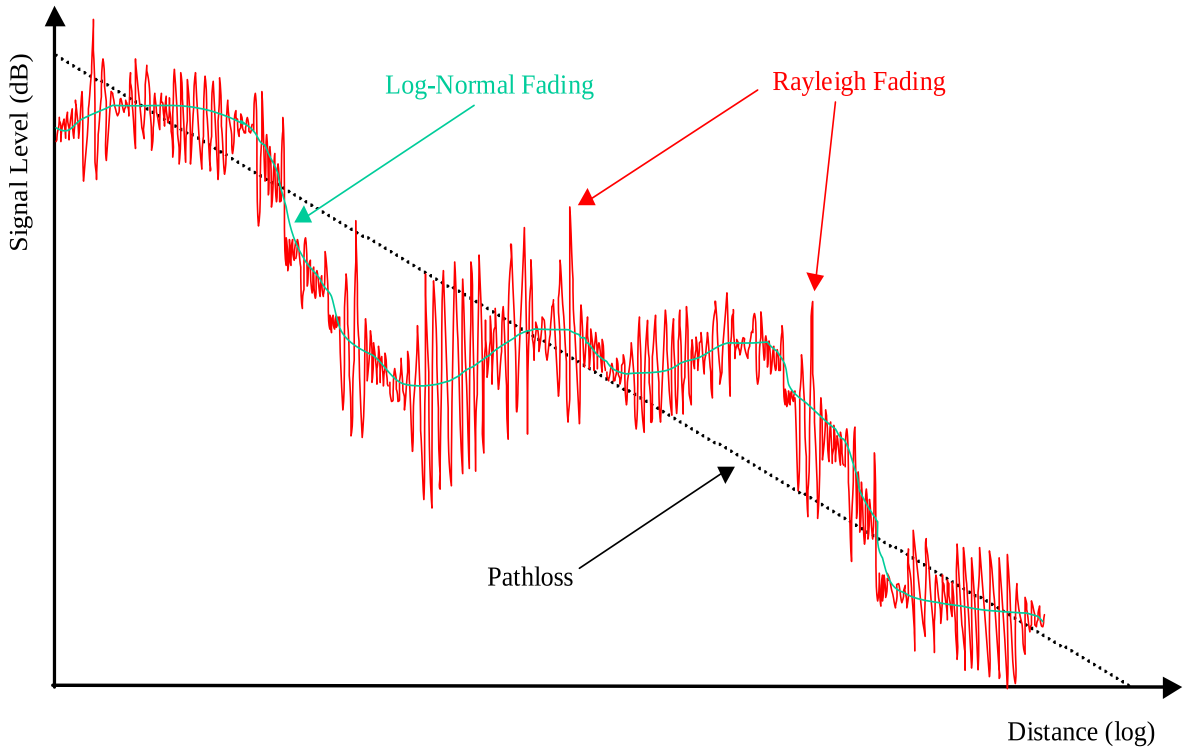



**Figure A1: An example of received signal strength variations as a function of distance between transmitter and receiver in a mobile radio environment**

The noise in a radio channel can be classified into two classes:

- Additive generated within the receiver, such as thermal and shot noise in passive and active devices; and from external sources such as atmospheric effects, and interference from other transmitters and
- **Multiplicative** caused by effects such as reflection, absorption, scattering, diffraction, and refraction.

For complex input signal **x**, complex output signal samples **y** is obtained by the following equation:

$$\mathbf{y} = \mathbf{ax} + \eta \tag{A.1}$$

Where **a** is multiplicative interference represented by samples of complex circularly symmetric process with the single-dimensional distribution of the power $g = |\boldsymbol{a}|^2$ and uniformly in [-π, π] and independently from g distributed phase. $\eta$ stands for additive interference i.e. sum of thermal noise and interference from other users.

This section describes the multiplicative interference i..e. variation of signal strength due to path-loss, slow fading (shadowing) and fast fading. The propagation model is depicted in Figure A2. Interference from other users is described in previous chapters, since it has great impact on design and performance of RRM algorithms. Thermal noise is usually negligible in comparison with interference from other users (interference limited systems).

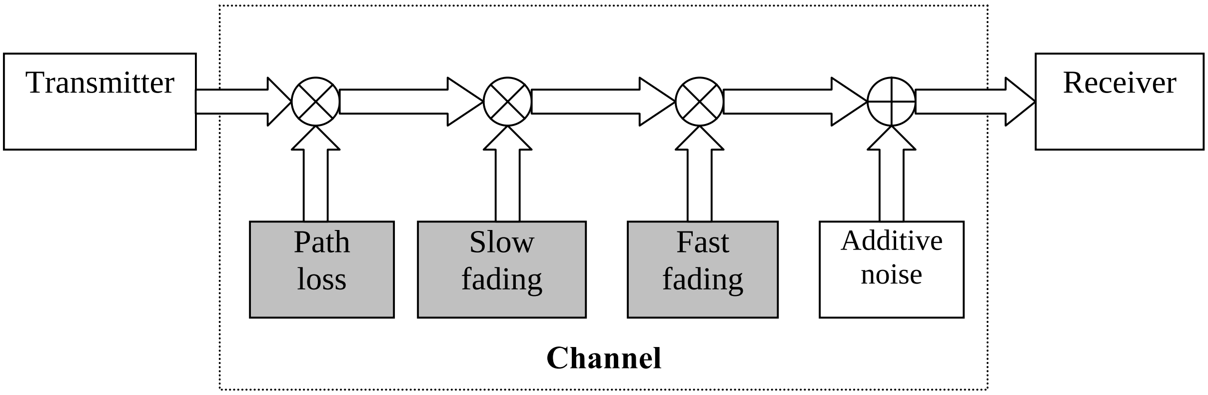



**Figure A2: Propagation model**

Usually multiplicative noise (attenuation) is expressed in dB. In that case the total channel attenuation $L$ (in dB) is the sum of the path-loss attenuation $L_{pl}$, slow fading attenuation $L_{sf}$ gain and fast fading attenuation $L_{ff}$:

$$L(dB) = L_{pl}(dB) + L_{sf}(dB) + L_{ff}(dB) \tag{A.2}$$

$L$ is usually between (50 and 80 dB). Channel gain g is naturally inverse to channel attenuation and is usually measured in linear scale i.e.:

$$g = 10^{-\frac{L(dB)}{10}} \tag{A.3}$$

In the following, we describe separately each of the components of the propagation model from Figure 2: Path-loss, slow fading (shadowing) and fast fading.

### *A.1 Path-loss*

Path-loss is an overall variation in signal strength mostly due to the distance between the transmitter and the receiver. This is caused by the outward spreading of the electromagnetic waves from the transmitter antenna. According to Hata model path-loss can be expressed as the sum of channel attenuation at the distance of 1 km $L_0$ (which depends on antenna heights and gains and frequency) and distance dependent factor, which increase with the distance R between transmitter and receiver [81]:

$$L_{pl}(dB) = L_0 + B\log R \tag{A.4}$$

Constant B depends on the environment and usually lies between 30 and 50 (large in urban than in suburban or open area).

### *A.2 Slow Fading*

Superimposed on the path-loss is a more rapidly changing effect called slow fading or (log-normal) shadowing. It is caused by the changing nature of the obstructions within the path between the transmitter and receiver, such as tall buildings or dense forest. It has significant variations over distances of hundreds of meters, with a typical variation of over 20 dB.

Assuming independent contribution to the signal attenuation along the propagation path from different obstructions like buildings, hills, forest etc., the resulting total signal attenuation is the product of single attenuations.  We can take logarithm of the attenuations product to get the sum of attenuations (logarithms) and apply central limit theorem (assuming a relative large number of obstructions between transmitter and receiver) [81]. Finally, we obtain the lognormal distribution of the underlying signal power. This means that the signal measured in decibels has a normal distribution i.e. probability density function (pdf) of slow fading random variable $L_{sf}$ with standard variation $\sigma_{sf}$ is given by the following equation.

$$pdf(L_{sf}) = \frac{1}{\sqrt{2\pi}\sigma_{sf}} \exp\left[-\frac{L_{sf}}{2\sigma_{sf}^{\ 2}}\right] \tag{A.5}$$

Slow fading occurs over distances comparable to the size of the buildings and hills in the region of the mobile, and is usually in the range of hundreds of meters. The standard deviation of the shadowing distribution $\sigma_{ls}$ is a location variability parameter that varies with frequency, antenna heights, and the environment. It is smallest in open areas and greatest in (sub) urban areas. This parameter is usually in the range 6—10 dB.

### *A.3 Fast fading*

Fast fading or fast signal strength variation at the mobile receiver is caused by constructive and destructive interference between multiple electromagnetic waves arriving at the receiver. This can introduce variations as large as 35 to 40 dB.

When objects, which reflect and scatter the transmitted signal, surround a transmitter and receiver, multipath propagation arises. Consequently, several waves arrive at the receiver via different routes. Each of the waves has a different phase, which could be considered to have an independent uniform distribution. This means that the phase associated with each wave is equally likely to take on any value.

According to the central limit theorem, the sum of enough independent random variables very closely approaches a normal distribution. If the real and imaginary parts of the multipath are such random variables, then the magnitude of a complex gaussian random variable is a Rayleigh-distributed random variable [41] i.e. probability density function (pdf) of fast fading random variable $r$ ($L_{ff}(dB) = 10log(r)$) with standard variation $\sigma_{ff}$ is given by the following equation.

$$pdf(r) = \frac{r}{\sigma_{ff}^{2}} \exp\left[-\frac{r^{2}}{2\sigma_{ff}^{2}}\right] \tag{A.6}$$

Rayleigh distribution fits well with fast fading measurements in the case of a non line of sight (NLOS) situation i.e. when no direct link (sight) over air between mobile and base exists. In a line of sight (LOS) scenario, the received signal is made of both the multipath component, plus a coherent LOS component. This leads to the other distribution called “Rician Distribution” [72].

# Appendix B: Basics of Game Theory

### *B1 Prisoners' Dilemma Game*

In this chapter we describe some basic Game Theory results in the example of Prisoners' Dilemma game[5]. For this game, the story goes as follows: Two prisoners accused of the same crime are kept in jail separated from each other in order to prevent mutual arrangements. Each prisoner has two possible options (decisions, actions): Commit or confess. A prisoner does not know the decision of his accomplice before taking his decision. The outcome of the game depends on what the both prisoners do

- If both confess they both go to the jail for 5 years.
- If one prisoner confesses and the other does not, the guy who confessed gets no punishment at all and the other guy gets 6 years.
- If both do not confess they both go to jail for 3 years.

This game can be represented in matrix form as represented in Table B1. The rows are possible actions of prisoner 1 and columns are possible actions of prisoner 2. The game outcome depends on actions of both players. In Table B1 gains for each possible action pair are given– right the gains of prisoner 1 and left the gains of prisoner 2. In order to obtain positive number as the gains we define the gain as: maximal number of jail years (6) - actual number of jail years, since the game is not changed if the same constant (in this case 6) is added to all possible outcomes of the game. The game as represented in Table B1 is the game in **normal or matrix form**.

**Table B1: The Prisoners' Dilemma Game**

| Player2 / Player1 | Cooperate | Defect |
|---|---|---|
| Cooperate | **3, 3** | **0, 6** |
| Defect | **6, 0** | **1, 1** |

As can be seen from Table B1, the optimal strategy (with highest gain) of player 1 is to defect whatever player 2 might do, since the values in the row "defect" are always greater than corresponding values of the row "cooperate" (6>3 and 1>0). In the same manner, the optimal strategy of player 2 is to "defect" too, since the values in the column "defect" are always greater than corresponding values of the column "cooperate". In this case, we say that the

**dominated strategy** (defect) exists for both players. Some games have a dominated strategy for only one player and most have no dominated strategy at all.

But if the optimal strategy for both players is to defect, no matter what the other player does, the pair of the strategy (defect, defect) is a stable outcome of the game – equilibrium i.e. no player can do better given the action(s) of other player(s). In game theory this equilibrium is called **Nash equilibrium**. More formally, Nash equilibrium is defined as follows:

**Nash equilibrium Definition** [70], [26]: The Nash equilibrium of the (two players) game is a pair (σ, τ) of strategies (σ denotes the strategy of the player 1 and τ the strategy of player 2) such that the payoff of the player 1 with the strategy σ against the strategy τ of player 2 $\pi_1(\sigma, \tau)$ is greater than the payoff of any other possible strategy s for player 1 given the strategy τ of player 2. Also payoff of player 2 with the strategy τ against the strategy σ of player 1 $\pi_2(\sigma, \tau)$ should also be greater than the payoff of any other possible strategy t for player 2 given the strategy σ of player 1 i.e.:

$$\begin{aligned} &\pi_1(\sigma,\tau) > \pi_1(s,\tau) \quad \forall s \neq \sigma \quad and \\ &\pi_2(\sigma,\tau) > \pi_2(\sigma,t) \quad \forall t \neq \tau \end{aligned} \tag{B.1}$$

The notion of Nash equilibrium can be defined in a similar way for N players (N>2) where the payoff of each player with his equilibrium strategy is greater than the payoff with any other strategy given the equilibrium strategy of all other players. If the strong inequality sign in B1 is replaced with the weak inequality sign ( $\geq$ ), we talk of "weak" Nash equilibrium.

For the Prisoners' Dilemma game, the strategy pair (defect, defect) satisfies condition (B.1) of the Nash equilibrium.

The strategies "defect" or "cooperate" are called pure strategies. But also **mixed strategies** are possible i.e. strategies where each pure strategy is played with some probability (possible 0). Most of the games have no equilibrium in pure strategies. But Nash [70] has proved an important theorem in Game Theory that all matrix games with a finite number of actions have at least one equilibrium, if mixed strategies are allowed.

The bad news from the Prisoners' Dilemma game is that the "rational" choice for both players is not to cooperate although both players would be better off if they both chose to cooperate: The payoff from (defect, defect) is (1,1) which is worse for both players than the pay off from (cooperate, cooperate) = (3,3) (see Table B.1 above).

As shown above, this is true for one-shot games i.e. if the game is played only once, since (defect, defect) is the Nash equilibrium of the one-shot Prisoners' Dilemma game. By using the recursive argument, it can also be shown that the same is true i.e. (defect, defect) is a "rational" outcome in each stage of the game, if the game is played finitely number of times. At the last stage of a finitely repeated Prisoners' Dilemma game, only one game is left to be played. Each player reasons before this last stage as in the one–shot Prisoners' Dilemma game and the equilibrium outcome is (defect, defect) as shown above. Further, in the last but one game each player knows what the outcome of the last game will be (defect, defect) and decisions should be made only for this one (last but one) game. Again we have the one–shot Prisoners' Dilemma game, and according to the same reasoning as above the equilibrium outcome is (defect, defect) for this last but one game and so on … Also in the finitely repeated Prisoners' Dilemma game each stage has the outcome (defect, defect). This is a disappointing conclusion since the strategy (cooperate, cooperate) is "better" for all players than the strategy (defect, defect)!

### *B2 Repeated Games*

The good news from the Prisoners' Dilemma game is that the rational choice for both players is to cooperate if the game is played **infinite number of times** or if the **probability *p*** ($0 \leq p \leq 1$) **of game continuation is high enough**. To show this, consider the "**TIT FOR TAT**" strategy [5] in which a player cooperates the first time and each time after that plays what the other player has played in the previous game stage. This means that the outcome of the "TIT FOR TAT" strategy played against the "TIT FOR TAT" strategy or against the "cooperate always" strategy would be (cooperate, cooperate) in each stage of the game. The expected "cooperative" payoff "C" of the "TIT FOR TAT" strategy against the "TIT FOR TAT" or the "cooperate always" strategy in the repeated Prisoners' Dilemma game with continuation probability p and payoffs according to Table B.1 would be:

$$C = 3 + 3p + 3p^2 + \ldots\ 3p^N + \ldots = 3/(1-p)$$

The expected payoff D of the "defect always" strategy against the "TIT FOR TAT" strategy would be:

$$D = 6 + 1p + 1p^2 + \ldots\ 1p^N + \ldots = 6+p/(1-p).$$

Also the expected gain of "cooperate always" (or play "TIT FOR TAT") against "TIT FOR TAT" would be larger than of "defect always" against "TIT FOR TAT" whenever:

$$3/(1-p) > 6+p/(1-p) \text{ i.e. } p>3/5$$

Also it would be better to cooperate in the repeated Prisoners' Dilemma against the player who plays "TIT FOR TAT" whenever $p > 3/5$. It can also be shown that the strategy pair ("TIT FOR TAT", "TIT FOR TAT") represents a Nash equilibrium for the repeated

Prisoners' Dilemma game with payoffs according to Table B.1, if the probability of game continuation p is larger than 3/5.

Similar results, i.e. that long-term cooperation brings more than defecting can also be obtained for some other repeated games and is formalized in the Folk Theorem [26], [66]:

**The Folk Theorem:** If there are strategies in the one-shot game that are "better"for all players (like "cooperate" in the Prisoners' Dilemma game) than the Nash equilibrium strategy (like "defect" in the Prisoners' Dilemma game), and the probability of game continuation is high enough, every repeated game has an infinite number of Nash equilibriums (like ("TIT FOR TAT", "TIT FOR TAT") in the the repeated Prisoners' Dilemma game). The expected outcomes are better than the expected outcome of always playing the Nash equilibrium strategy of the one-shot game ($3/(1-p) > 6+p/(1-p)$ *if* $p>3/5$).

The proof is based on the idea that a player can be "enforced" by the other players to any strategy that brings him more expected gain ("cooperative" strategy) than always playing the Nash equilibrium strategy of the one-shot game. The other players could make a simple threat: "If you defect from the "cooperative" strategy we will return to the Nash equilibrium strategy of the one-shot game" and your gain would be lower than when you "cooperate". Put in other worlds [51]: "Each player is told by the others to stick to the agreement or everyone will gang up on her. Then no single player, acting alone, has any incentive to deviate; the condition necessary for a Nash equilibrium". For a rigor proof of the Folk theorem see [26] or [66].

As the Folk theorem states, there are an infinite number of possible Nash equilibriums. Each strategy in a repeated game, which provide a higher expected payoff than the "defect always" strategy against the "cooperate" strategy of other users and "punishes" the users who defect by defecting itself, is a Nash equilibrium strategy of the repeated game. Which one of these strategies is selected in practice depends on a lot of things some of which have nothing to do with Game theory but are the result of a certain social situation. For example, often used strategies are the social "focal" points strategies i.e. the strategies established in a particular society as "good" and desirable, "the custom" strategy i.e. the strategy that was always used in the past, "symmetric" strategies i.e. do the same that others do etc. For example, always driving on the left side or on the right side of the street are both Nash equilibrium strategies with essentially the same outcomes. Which Nash equilibrium strategy is selected depends on social convention. The resulting outcomes do not have to be even **Pareto optimal** i.e. provide each player with higher or equal payoffs than any other possible outcome. Players could stick to their strategies anyway because they are too inert or even do not know any better strategy.

For example, the strategy to select the side of the street to drive on, by throwing a die, is a Nash equilibrium strategy, given that other players do the same. But the Nash equilibrium resulting from the "random" selection of the street side is not Pareto optimal, since always driving on the left or always on the right makes all players better off.

Note that "TIT FOR TAT" is a relatively simple strategy: It requires only the knowledge of the action of the other player in the previous stage of the game in order to make decisions for the next stage. The state of the player is "punish" if the opponent has chosen "defect" in the previous stage; otherwise the state is "do not punish". Now the strategy "TIT FOR TAT"is state dependent: If the state is "punish" play "defect", if the state is "not punish" play "cooperate". This is a simple example of state dependent strategies. In general, the state of the player can depend on the complete previous history of the game. Naturally, state definition is context dependent and it is a matter of modeling art to make a "good" definition of the state for a particular problem. If the action depends only on the last state, we talk of a **Markov game**. "TIT FOR TAT" is also a Markov game. By the appropriate choice of states, many games can be converted into Markov games. The strategy of the user is then defined by the actions the user should take for each possible state of the game.

It is interesting to note that in spite of its simplicity the "TIT FOR TAT" strategy achieved the best average score in the computer "Olympiad" organized by Axelrod [5]: Computer programs applying different strategies were submitted from many scientists from allover the world and were led by Axelrod to play the infinitely repeated Prisoners' dilemma game against each other. Although some very complicated strategies were submitted, the relatively simple "TIT FOR TAT" strategy, showed on average, the best performance.

***B3 Evolutionary Stable Strategies***

An interesting question is: which features should a strategy possess in order to "survive" in environments with a mixture of different other strategies like in Axelrod's "Olympiad". Biologist like Maynard Smith [93] investigated this question in order to describe the "survival of fittest" within the framework of evolutionary theory. They found out that a "successfully" strategy should be an Evolutionary Stable Strategy (ESS), which is defined as follows:

The strategy S is called evolutionary stable if [93]:

1) The expected gain of the strategy S played with itself $E(S,S)$ is greater than the expected gain $E(S,O)$ from all other strategy $O$ played with the strategy $S$ i.e.: $\boldsymbol{E(S,S) > E(S,O)}$ for all other strategies $O$ or

2) If $\boldsymbol{E(S,S) = E(S,O)}$ **than** $\boldsymbol{E(S,O) > E(O,O)}$ for all other strategies $O$.

The expected gain of the strategy played against the other strategies measures a relative "fitness" of the strategy in a population of different strategies.

For example in a population consisting of "TIT FOR TAT", "cooperate always" and "defect always", "TIT FOR TAT" is an evolutionary stable strategy since *E("TIT FOR TAT", "TIT FOR TAT") > E("defect always", "TIT FOR TAT")* (see above). Although *E("cooperate always", "TIT FOR TAT") = E("TIT FOR TAT", "TIT FOR TAT")*, "cooperate always" is not resistant against the invasion of "defect always" mutants, since *E("cooperate always ", "cooperate always") < E("defect always ", "cooperate always")* (see above).

We could say that "good" strategies (like "TIT FOR TAT") should [5]:

- Be nice at the beginning
- Punish those who defect themselves, even if the "punishment" costs more than no punishment, in order to discourage motivation for "defecting".
- "Forgive" i.e. if a "defecting" player returns to the cooperative behavior then cooperate too.

It is important to note that only "nice" strategy like "cooperate always" opens the door for "exploitation" by defecting strategies like "defect always". Therefore, a sort of "punishment" mechanisms should be build in each strategy in order to discourage defecting and make the strategy an ESS.

# Appendix C: List of Acronyms

| | |
|---|---|
| AC | Admission Control |
| ATM | Asynchronous Transfer Modus |
| BER | Bit Error Rate |
| BS | Base Station |
| CIR | Carrier to Interference Ratio |
| CSI | Channel State Information |
| DCA | Dynamic Channel Allocation |
| GSM | General System for Mobile Communications |
| HO | Handover of Handoff |
| MS | Mobile Station |
| NRT | Non Real Time (Service) |
| OMC | Operation and Maintance Centar |
| PC | Power Control |
| QoS | Quality of Service |
| RNC | Radio Network Controler |
| RRM | Radio Resource Management |
| RT | Real Time (Service) |
| SA | Smart Antennas |
| SH | Scheduler |
| SIR | Signal-to-Interference Ratio |
| UMTS | Universal Mobile Telecommunication System |
| WLAN | Wireless Local Area Network |
| 4G | 4. Generation of Wireless Networks |

## References

*"Knowledge is of two kinds.*
*We know a subject ourselves, or we know where we can find information on it."*
***Samuel Johnson***


[1] Y. Akaiwa, H. Andoh: "Channel Segregation-A Self Organized Dynamic Channel Allocation Method: Application to TDMA/FDMA Microcellular System", IEEE Journal on Selected Areas in Communications, Vol. 11, No. 6, pp. 949-954, Aug. 1993.

[2] H. Akimaru and K. Kawashima: *"Teletraffic, Theory and Applications"*, Springer 1993.

[3] M. E. Anagnostou, J. A. Sanchez-P., I. S. Veniers: "A Multiservice User Descriptive Traffic Source Model", IEEE Transaction on Communications, Vol. 44, No. 10, pp. 1243-1246, Oct. 1996.

[4] J. Anderson: "A simulation study of some dynamic channel assignment algorithms in high capacity mobile telecommunications system", IEEE Trans. Veh. Tech. vol. 22, No. 4, pp.1210-1217, Nov. 1973.

[5] R. Axelrod and W. D. Hamilton: "The evolution of cooperation", Science, No. 211, pp. 1390-1396, Mar. 1981.

[6] A. A.Beaulieu, A. A. Abu-Daya and P. J. MacLane: "Estimating the Distribution of a Sum of Independent Log-Normal Random Variables," IEEE Trans. Communication Vol. 43, No. 12, Dec. 1995.

[7] J.E. Berg: "A recursive Method for Street Microcell Path-loss Calculations", PIMRC Vol 1, pp 140-143, 1995.

[8] E. Biglieri, J. Proakis, S. Shamai: "Fading Channels: Information-Theoretic and Communications Aspects", IEEE Trans. Inform. Theory, Vol. 44, pp. 2619-2690, Nov. 1998.

[9] K. Binmore: *"Fun and Games"*, D. C. Heathand Co., Toronto, 1992.

[10] M. Bublin, G. Diernhofer, C. F. Mecklenbräuker, T. Pacic, J. Plogsties, P. Slanina: "Simulation of smart antennas in 3G mobile systems", Proceedings 4th European Personal Mobile Communications Conference EPMCC 2001,Vienna, Feb. 2001.

[11] M. Bublin, G. Ostermayer: “Comparison of Interference Based Dynamic Channel Allocation Algorithms in mobile Networks”, in Proc. of the IEEE SoftCOM'03, Split-Rijeka-Trieste-Venice, Croatia-Croatia-Italy-Italy, Oct. 2003.

[12] M. M.-L. Chang, J. C.-I. Chuang: “Performance Evaluation of Distributed Measurement-Based Dynamic Channel Assignment in Local Wireless Communications, “ IEEE Journal on Selected Areas in Communications, Vol. 14, No. 4, pp. 698-710, May 1996.

[13] H. Chen and D. D. Yao: *"Fundamentals of Queuing Networks"*, New York, Springer-Verlag, 2001.

[14] M.Chu and W.Stark: “Effect of Mobile Velocity on Communications in Fading Channels”, IEEE Trans. On Vehicular Techn, Vol.49, No.1, Jan. 2000.

[15] T. M. Cover and J. A. Thomas: *“Elements of Information Theory”*, Wiley, New York, 1991.

[16] D.C.Cox and D.O.Rudnik: ”Increasing channel occupancy in large-scale mobile radio systems: dynamic channel reassignment”, IEEE Trans. Veh. Tech. 22(4), pp. 218-222, Nov. 1973.

[17] Z. Dziong: *"ATM Network Resource Management"*, McGraw-Hill, New York, 1997.

[18] R. Eier: *“Analyse und Synthese von diskreten Zufallsprozessen mit Hilfe von Markoffschen Ketten“*, Habilitationsschrift der TU Wien, Wien, 1972.

[19] ETSI: “Selection Procedures for the choice of radio transmission technologies of the UMTS”, UMTS 30.03 version 3.2.0, ETSI, April 1998.

[20] ETSI GSM Technical Specifications 05.08: “Radio Subsystem Link Control”, Mar. 2000.

[21] ETSI: “Physical Layer Procedures (FDD)”, 3GPP Technical Specification 25.214, Mar. 2000.

[22] ETSI: "Radio Resource Management Strategies", 3GPP Technical Specification 25.922 version 3.1.0, Mar. 2000.

[23] J. S. Evans and D. Everitt: “Effective Bandwidth-Based Admission Control for Multiservice CDMA Cellular Networks”, IEEE Trans. On. Veh. Tech., Vol. 48, No. 1, Jan. 1999.

[24] N. Feng, S. C. Mau and N. B. Mandayam: “Pricing and Power Control for Joint network-Centric and User-Centric Radio Resource Management”, 39th Annual Allerton Conference on Communication Control and computing, Alerton, IL, Oct. 2001.

[25] G. J. Foschini and Z. Miljanic: “A Simple Distributed Autonomous Power Control Algorithm and Its Convergence”, IEEE Trans. Veh. Tech., Vol. 42, No. 4, pp. 57-62, Nov. 1993.

[26] D. Fudenberg and J. Tirole: *“Game Theory”*, The MIT Press, Cambridge 1991.

[27] J. Fuhl, A. Kuchar, E. Bonek: “Capacity increase in cellular PCS by smart antennas”, IEEE 47th Vehicular Technology Conference (VTC´97), vol.3, pp. 1962-1966, May 1997.

[28] Y. Furuya, Y. Akaiwa: "Channel Segregation, a Distributed Adaptive Channel Allocation Scheme for Mobile Communication Systems", IEICE Transactions, Vol. E 74, No. 6, June 1991.

[29] K.S. Gilhausen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, C. E. Wheatley: "On the Capacity of a Cellular CDMA System", IEEE Trans. Veh. Tech. Vol. 40, No. 2, pp. 303-312, May 1991.

[30] A. Goldsmith and P. Varaiya: “Capacity of Fading Channels with Channel Side Information”, IEEE Trans. Infrom. Theory, Vol. 43, pp. 1986-1992, Nov. 1997.

[31] S. Grandhi, J. Zander, and R. Yates, "*Constrained power control*," International Journal of Wireless Personal Communications, vol. 1, no. 4, 1995.

[32] M. Gudmundson: “Correlation Model for Shadow Fading in Mobile Radio Systems,” Electronics Letters, vol. 27, No 23, pp. 2145-2146, Nov. 1991.

[33] S. Haemaelaeinen, P. Slanina, M. Hartman, A. Lappetelaeinen, H. Holma, O. Salonaho: “A Novel Interface Between Link and System Level Simulations”, Acts Mobile Communications Summit '97, Aalborg/Denmark, pp. 599-604, Oct. 1997.

[34] S. W. Halpern: “Reuse Partitioning in Cellular Systems”, Proc. of Vehicular Technology Conference VTC '83, pp. 322-327, 1983.

[35] S. V. Hanly and D. N. Tse: ”The multi-access fading channel: Shannon and delay limited capacity”, in 33th Annu. Allerton Conf. Communications Control and Computing, Monticello, Il, Alerton House, Oct. 1995.

[36] S. V. Hanly and D. N. Tse: "Multi access fading channels - Part II: Delay-limited capacities", IEEE Trans. Inform. Theory, vol. 44, pp. 2816-2831, Nov. 1998.

[37] J. D. Herdtner, E. K. P. Chong: "Analysis of a Class of Distributed Asynchronous Power control Algorithms for Cellular Wireless Systems", IEEE Journal on Selected Areas in Communications, Vol. 18, No.3, March 2000.

[38] H. Holma and A. Toskala (editors): " *WCDMA for UMTS*", revised edition, John Wiley & Sons, New York, 2001.

[39] D. Hong, S. S. Rappaport: "Traffic Model and Performance Analysis for Cellular Mobile Radio Telephone Systems with Prioritized and Nonprioritized Handoff Procedures", IEEE Trans. Veh. Tech., Vol. 35, No. 3, August 1986.

[40] R. A. Howard: *"Dynamic programming and Markov processes"*, 2. print., Cambridge, Mass., M.I.T. Press, 1962.

[41] W.C. Jakes et. al.: "*Microwave Mobile Communications"*, New York, Wiley & Sons, 1974.

[42] D. G. Jeong, W. S. Jeon: "CDMA/TDD System for Wireless Multimedia Services with Traffic Unbalance Between Uplink and Downlink", IEEE Journal on Selected Areas in Communications, Vol. 17, No. 5, pp. 939-946, May 1999.

[43] S. Jordan A. Khan: "A performance bound on dynamic channel allocation in cellular systems, equal load", IEEE Trans. Veh. Tech., vol 43, No. 2, pp. 333-344, April 1994.

[44] S. Jordan: "Resource allocation in wireless networks", Journal of High Speed Networks, vol. 5, pp. 23-34, 1996.

[45] T. J. Kahwa, and N. D. Georganas: "A hybrid channel assignment scheme in large-scale, cellular structured mobile communication systems", IEEE Trans. Comm. 7(8), Oct. 1989, pp. 1172-1179.

[46] T. Kanai: "Autonomous Reuse Partitioning in Cellular Systems," IEEE VTC, 1992 pp. 782-785.

[47] R. L. Keeney and H. Raiffa: "*Decisions with Multiple Objectives, Preferences and Value Trade-offs,",* Cambridge University Press, New York, 1993.

[48] L. Kleinrock: *"Queuing systems - Theory"*, Vol I, New York, John Wiley and Sons, 1975.

[49] L. Kleinrock: *"Queuing Systems - Applications"*, Vol II, New York, John Wiley and Sons, 1975.

[50] R. Knopp and P. A. Humblet: “Information Capacity and Power Control in Single- Cell Mulituser Communications”, in Proc. Int. Conf. Communications (ICC’97), Montreal, Que., Canada, pp. 331-335, June 8-12, 1997.

[51] D. Kreps: “*A Course in Microeconomic Theory*”, Prentice Hall, March 1991.

[52] A. Kuchar, J. Fuhl, E. Bonek: “Spectral efficiency enhancement and power control of smart antenna system”, European Personal Mobile Communications Conference (EPMCC´97), Bonn, Deutschland, pp. 475-481, Sept. – Oct. 1997.

[53] X. Lagrange and P. Godlewski: “Teletraffic Analysis of Hierarchical Cellular Networks”, Proc. Veh. Tech. Conf., pp. 882-886, July 1995.

[54] A. Lapidoth and S. Shamai: “Fading channels: How perfect need “perfect side information” be”, Proc. IEEE Inform. Theory Workshop, pp. 38, South Africa, June 1999.

[55] W. C. Y. Lee: "Overview of Cellular CDMA", IEEE Trans. Veh. Tech. Vol. 40, No. 2, pp. 291-302, May 1991.

[56] D. Luenberger: *"Optimization by Vector Space Methods"*, John Wiley, New York, 1969.

[57] D. Luenberger: *“Introduction to Dynamic Systems”*, John Wiley, New York, 1979.

[58] L. Li and A. Goldsmith: "Capacity and optimal resource allocation for fading broadcast channels: Part I: Ergodic capacity", IEEE Trans. Inform. Theory. Vol. 47, pp 1083-1102, March 2001.

[59] L. Li and A. Goldsmith: "Capacity and optimal resource allocation for fading broadcast channels: Part II: Outage capacity", IEEE Trans. Inform. Theory. Vol. 47, pp 1103-1127, March 2001.

[60] X. Liu and A. Goldsmith: “Optimal Power Allocation over Fading Channels with Stringent Delay Constraints”, IEEE ICC 2002.

[61] V. H. Mac Donald: "Advanced Mobile Phone Service: The Cellular Concept”, The Bell System Technical Journal, vol. 58, No. 1, pp. 15-41, Jan. 1979.

[62] A. B. MacKenzie and S. B. Wicker: “Game Theory and the Design of Self-Configuring, Adaptive Wireless Networks”, IEEE Comm. Magazine, pp. 126-132, Nov. 2001.

[63] J. MacKie-Mason, H. R. Varian: “Pricing the Internet”, In *Public access to the Internet*, B. Kahn, J. Keller (eds.), Prentice Hall, Englewood Cliffs, New Jersey, 1995.

[64] A. Mas-Colell, M. D. Whinston and J. R. Green: *"Microeconomic Theory"*, Oxford University Press, New York 1995.

[65] L. McKnight, J. Bailey (eds.): *“Internet Economics”*, MIT Press, Cambridge, 1997.

[66] R. B. Myerson: *“Game Theory, Analysis of Conflict”*, Harvard University Press, Cambridge, 1991.

[67] P.E. Mogensen, P. Zetterberg, H. Dam, P. Leth-Espensen, F. Frederiksen: “Algorithms and Antenna Array Recommendations” (Part I), ACTS 020: Technology in Smart Antennas for Universal Advanced Mobile Infrastructure (Part 2), TSUNAMI(II), CEC Deliverable AC020/AUC/A1.2/DR/P/005/b1, 27 Oct. 1996.

[68] M. Nagshineh and I. Katzela: ”Channel assignment schemes for cellular mobile telecommunication systems”: A comprehensive survey, IEEE Personal Commun., vol. 3, pp. 10-31, June 1996.

[69] A.F. Naguib, A.J. Paulraj, and T. Kailath: “Capacity Improvement with Base-Station Antenna Arrays in Cellular CDMA”, IEEE Trans. Vehicular Technology, vol. 43, no. 3, pp. 691-698, Aug. 1994.

[70] J. F. Nash. “Noncoperative Games”, Ann. Math., Vol. 54, pp. 289-295, 1951.

[71] R. Negi: *“Power Adaptation Strategies for Delay Constrained Channels”*, Dissertation, Stanford Univ., Septembar 2000.

[72] J. D. Parsons: *“The Mobile Radio Propagation Channel”*, Second edition, J. Wiley & Sons Ltd, Baddins Lane, 2000.

[73] A.J. Paulraj and C.B. Papadias: “Space-time processing for wireless communications”, IEEE Signal Processing Magazine, pp. 49--83, Nov. 1997.

[74] G. P. Pollini: “Trends in Handover Design”, IEEE Comm. Mag., Vol. 34, No. 3, March 1996.

[75] R. Prasad, W. Mohr, W. Konhäuser (Editors): *"Third Generation Mobile Communication Systems"*, Artech House 2000.

[76] J. Punt, D. Sparreboom: “Decentralized DCA algorithms”, FRAMES internal document, Apr. 1998.

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

[78] M. Radimirsch, K. Jobmann: "Energy efficiency versus throughput in wireless data networks – a game theoretical model", COST273, 2nd workshop, Paris, France, May 2003.

[79] C. U. Saraydar, N. B. Mandayam and D. J. Goodman: "Efficient power control via pricing in wireless data networks", IEEE Trans. on Communications, February 2002.

[80] C. U. Saraydar, N. B. Mandayam and D. J. Goodman: "Pricing and power control in a multicell wireless network", IEEE JSAC, Vol. 19, No. 10, pp. 1883-1892, Oct. 2001.

[81] S. R. Saunders: *"Antennas and Propagation for Wireless Communication Systems",* J. Wiley & Sons Ltd, Baddins Lane, 1999.

[82] R. Schönherr: *"Schnittstelle zwischen System- und Linklevelsimulation in Mobilfunk"*, Diplomarbeit, Institut of Computer Technology, Vienna University of Technology 1998.

[83] S. C. Schwartz and Y. S. Yeh: "On the Distribution Functions and Moments of Power Sums with Log-Normal Components," Bell Syst. Tech. J., Vol 61, No. 7, 1982.

[84] C. E. Shannon. "A Mathematical Theory of Communication", Bell Syst. Tech. J., Vol 27, pp. 379-423, July 1948.

[85] C. E. Shannon: "Communication in Presence of Noise", Proc. IRE, vol. 37, pp. 10-21, Jan. 1949.

[86] S. J. Shenker: "Fundamental Design Issues for the Future Internet", IEEE JSAC, vol. 13, no. 7, pp. 1176-1188, Sept. 1995.

[87] S. M. Shin, C. H. Chao and D. K. Sung: "Interference-Based Channel Assignment for DS-CDMA Cellular Systems", IEEE Trans. On Veh. Tech., Vol. 45, No. 1, pp. 233-239, Jan. 1999.

[88] H. Simon: *"Administrative Behavior"*, 3. ed., The Free Press, New York,1976.

[89] H. Simon: *"The Sciences of the Artificial"*, 2. ed., The MIT Press, Cambridge, MA, 1981.

[90] H. Simon: *"Models of Bounded Rationality"*, The MIT Press, Cambridge, MA, 1997.

[91] V. A. Siris, B. Briscoe and D. Songhurst: "Economic Models for Resource Control in Wireless Networks", In Proc. of IEEE PIMRC, Lisbon, Portugal, Sept. 2002.

[92] V. A. Siris: “Cell Coverage based on Social Welfare Maximization”, In Proc. of IST Mobile and Wireless Telecommunications Summit, Thessaloniki, Greece, June 16-19, 2002.

[93] J. M. Smith: *“Evolution and the Theory of Games”*, Cambridge University Press, 1982.

[94] L. Song and N. B. Mandayam: “Hierarchical SIR and Rate Control on the Forward Link for CDMA Data Users under Delay and Error Constraints”, IEEE JSAC, vol. 19, No. 10, pp. 1871-1882, Oct. 2001.

[95] B. Stiller, P. Reichl and S. Leinen: „ Pricing and Cost Recovery for Internet Services: Practical Review, Classification, and Application of Relevant Models“, Netnomic – Economic Research and Electronic Networking, Vol. 2, No. 1, Jan. 2000.

[96] S. Tekinah and B. Jabbari: “Handover Policies and Channel Assignment Strategies in Mobile Cellular Networks”, IEEE Comm. Mag., Vol. 29, No. 11, 1991.

[97] D. N. Tse and S. V. Hanly: "Multi access fading channels - Part I: Polymatroid structure, optimal resource allocation and throughput capacities", IEEE Trans. Inform. Theory, vol. 44, pp. 2796-2815, Nov. 1998.

[98] R.S. Varga: *“Matrix Iterative Analysis”*, Englewood Cliffs, Prentice-Hall, New York, 1962.

[99] A. J. Viterbi, A. M. Viterbi, E. Zehavi: "Other-Cell Interference in Cellular Power-Controlled CDMA", IEEE Trans. Comm. Vol. 42, N0. 2/3/4, February/March/April 1994, pp. 1501-1504.

[100] J. Von Neumann and O. Morgenstern: *“Theory of Games and Economic Behavior”*, Princeton University Press, Princeton 1945.

[101] X. Wang and H. Schulzrinne: “An Integrated Resource Negotiation, Pricing and QoS Adaptation Framework for Multimedia Applications”, IEEE JSAC, Vol. 18, No. 12, , pp. 2514-2529, December 2000.

[102] H. Wang , N.B. Mandayam: "Energy and Delay Constrained Power Control for File Transfers over Fading Channels" in Proceedings of Asilomar, CA, Nov. 2001.

[103] W. B. Yang, E. Geraniotis: “Admission Policies for Integrated Voice and Data Traffic in CDMA Packet Radio Networks”, IEEE JSAC, Vol. 12, No. 4, pp. 654-664, May 1994.

[104] R. D. Yates: "A Framework for Uplink Power Control in Cellular Radio Systems", IEEE Journal on Selected Areas in Communications, Vol. 13, No. 7, pp. 1341-1347, Sept. 1995.

[105] W. Yu: *“Competition and Cooperation in Multi-User Communication Environments”*, PhD thesis, Stanford. Univ., June 2002.

[106] J. Zander: "Radio Resource Management in Future Wireless Networks: Requirements and Limitations", IEEE Communications Magazine, Vol. 35, No. 8, Aug.1997

[107] J. Zander, H. Eriksson: "Asymptotic Bounds on the Performance of a Class of Dynamic Channel Assignment Algorithms", IEEE Trans. Veh. Tech. Vol. 11, No. 6, pp. 926-933, Aug. 1993

[108] J. Zander: "Performance of Optimum Transmitter Power Control in Cellular Radio Systems", IEEE Trans. Veh. Tech. Vol. 41, No. 1, pp. 57-62, Feb. 1992.

[109] J. Zander: "Distributed Cochannel Interference Control in Cellular Radio Systems", IEEE Trans. Veh. Tech. Vol. 41, No. 3, pp. 305-311, Aug. 1992.

[110] J. Zander, S. L. Kim: *“Radio Resource Management for Wireless Networks“*, Artech House, Norwood MA, 2001.