Politecnico di Milano
V Facoltà di Ingegneria
Corso di laurea in Ingegneria delle Telecomunicazioni
Dipartimento di Elettronica e Informazione

Radio Planning of Energy-Aware Cellular Networks

Advisors: Prof. Antonio CAPONE
Prof. Brunilde SANSO'

Thesis by:
Silvia BOIARDI Student ID 736038

Session: 2009-2010

Abstract

The constant development and the increasing importance on everyday life of the Information and Communication Technology industry fostered the sensitiveness toward ICT energy consumption problems. In an attempt to reduce the environmental impact of the communication sector, wireless access networks have recently received great attention and energy-aware models have been proposed for both cellular and WLAN networks. At our knowledge, all the suggested approaches focus on the network management phase, aiming at powering on and off network devices depending on traffic levels; however, the strong impact of the design stages on an effective energy-efficient operation has not been considered yet. In order to delve into this issue, here a joint design and management optimization approach is proposed. The model tries to reduce energy consumption while guaranteeing users Quality of Service constraints and minimizes installation (Capex) and operational (Opex) expenses in charge of network providers. It is shown that, when energy costs are included in Opex and energy management strategies at the design stages are taken into account, more energy efficient and versatile topologies are obtained than when Capex only are considered.

Sommario

L'influenza del settore Information and Communication Technology (ICT) sulla produttività e la crescita economica è in continuo aumento, e proprio a causa della sua sempre maggiore diffusione non è più possibile trascurarne l'impatto energetico. Per quanto concerne le reti cellulari, ad esempio, sono stati proposti alcuni approcci e modelli attenti al fattore energetico che mirano a ridurre il consumo di potenza abbassando nel contempo i costi sostenuti degli operatori di rete. Finora, tuttavia, l'attenzione è stata posta esclusivamente su questioni di carattere gestionale quali l'accensione e lo spegnimento di elementi di rete in base ai livelli di traffico, non considerando che un comportamento realmente efficiente dal lato energetico dipende in gran parte dalle decisioni prese in fase di design. Per colmare questa lacuna proponiamo qui un approccio congiunto di ottimizzazione di design e gestione, il cui scopo è quello di limitare lo spreco di potenza, garantendo allo stesso tempo i vincoli di qualità della connessione per gli utenti, e di minimizzare i costi di installazione (Capex) e quelli operativi (Opex) a carico dei gestori di rete. Si mostra quindi che, includendo i costi energetici in Opex e adottando strategie di gestione energetica in fase di design, è possibile raggiungere topologie di rete più efficienti e versatili di quelle ottenibili considerando solamente i Capex.

Acknowledgements

After these nine intense months, it is a pleasure to thank those who contributed in some way to this thesis. First and foremost I would like to express my gratitude to my supervisors, Prof. Capone and Prof. Sansò, who patiently answered all my questions and assisted me with their knowledge and advices. They even nursed me when I was ill: one simply could not wish anything better.

I cannot forget the great help I received for the programming aspects of the project. A big thank to Massimo and Ilario, researchers at DEI who introduced me to the AMPL world, and to Pierre and Luc, fine technicians at GERAD who always tried their best to solve my computer problems. I take the opportunity to thank my flatmates and fellow students Daniele and Filippo, who spent time to give me hand when it was needed.

A special mention should be made of Ginette, secretary at École Polytechnique, Valérie, Marilyne, Marie, Carole and all the staff of GERAD. With their kindness and helpfulness, they gave me professional support and assistance for all the bureaucratic matters relating to my stay in Canada.

Lastly and most importantly, I wish to thank the people who encouraged me every time I needed moral support: my parents, who were always nearby despite the geographical distance, and Fabio, for his endless patience and his ability to get a smile out of me in bad times. I am also tempted to individually name all my friends that trusted me, but by fear of leaving someone out I will just say *thank you very much to you all.*

Contents

List of Figures

List of Tables

Acronyms

AMPL	A Modeling Language for Mathematical Programming
AP	Access Point
BS	Base Station
BSS	Basic Service Set
Capex	Capital Expenditures
CDMA	Code Division Multiple Access
CS	Candidate Site
CSMA	Carrier Sense Multiple Access
CSMA/CA	Carrier Sense Multiple Access - Collision Avodance
CTMC	Continuous-Time Markov Chain
CTS	Clear To Send
DCF	Distributed Coordination Function
ESS	Extended Service Set
ETSI	European Telecommunications Standards Institute
FAP	Frequency Assignment Problem
FDD	Frequency Division Duplexing
FDMA	Frequency Division Multiple Access

FM	Frequency Modulation
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
HSPA	High Speed Packet Access
IA	Infrastructure Activation
ICT	Information and Communication Technology
IEEE	Institute of Electrical and Electronics Engineers
IG	Instance Generator
IP	Internet Protocol
LAN	Local Area Network
LTE	Long Term Evolution
MEP	Maximum Efficiency Problem
MF-MEP	Multiple Frequency Maximum Efficiency Problem
MI-FAP	Minimum Interference Frequency Assignment Problem
MILP	Mixed Integer Linear Programming
MIP	Mixed Integer Programming
MOP	Minimum Overlap Problem
MS	Mobile Station
MS-FAP	Minimum Span Frequency Assignment Problem
NAM	Network Allocation Map
NAV	Network Allocation Vector
NIC	Network Interface Connector
NP	Non-Deterministic Polynomial-Time

OF	Objective Function
Opex	Operational and Management Expenditures
PC	Power Control
PoE	Power over Ethernet
QoS	Quality of Service
RoD	Resource on-Demand
RTS	Request To Send
SCP	Set Covering Problem
SF	Spreading Factor
SINR	Signal to Interference and Noise Ratio
SIR	Signal to Interference Ratio
S-MF-MEP	Simplified Multiple Frequencies Maximum Efficiency Problem
TDMA	Time Division Multiple Access
TP	Test Point
UMTS	Universal Mobile Telecommunications System
UTP	Unshielded Twisted Pair
W-CDMA	Wideband Code Division Multiple Access
WLAN	Wireless Local Area Network
WMN	Wireless Mesh Network

Chapter 1

Introduction

The interest in the area of green networking is growing more and more, given the importance of environmental and energy related issues for the Internet and Communication Technologies sector. In fact, it has been reported [12] that the ICT percentage with respect to the world energy expenditures range from 2% to 10%. Of particular concern is the consumption of the cellular wireless system, both for its increasing pervasiveness that pushes for more wireless infrastructure and for the well known fact that Base Stations are particularly energy-hungry, representing over 80% of the power used in the radio segment [26].

Despite the novelty of the problem, the research community has already produced interesting models and approaches to deal with the issues of energy-savings in cellular networks. Most studies have focused on the *network operation* aspects, in particular on *management* issues mainly aiming at switching off parts of the network when the traffic load decreases [7, 21, 11]. However, an important issue that has not been explored in the literature is that an effective energy-efficient operation depends on the type of Base Stations deployed and the coverage structure of the network and hence on radio planning decisions taken during the design phase. Since the coverage of the service area must be ensured at all times, the level of flexibility offered by the network topology is essential in order to be able to switch some Base Stations on and off to dynamically adapt the network capacity to the traffic load without violating coverage constraints.

While concerns for climate change is the main push for research on efficient technology developments, network operators are as well interested in reducing the energy consumption of their networks for economic reasons. Indeed, as reported in [1], in a mobile network about 80-90% of the overall energy consumption is in charge of the operator, while in wired networks the percentage decreases at about 30% (the other 70% is distributed among end users). Two are the cost categories that mobile operators have to meet: capital investments related to radio equipment, license fees, site buildouts and installation of equipment, commonly identified as Capital Expenditures (*Capex*) and running costs, such as transmission, site rentals, marketing terminal subsidies and operation and maintenance (Operational and Management Expenditures, *Opex*) [18]. The challenge in terms of energy-aware modeling is to be able to convey both type of costs and energy issues in a single modeling framework.

The objective of this thesis is therefore to fill that gap and present, for the first time, an energy-aware joint design and management problem aiming at maximally limiting energy consumption while reducing both Capex and Opex categories. Among other things, the impact of having such a modeling approach when compared with more traditional Capex optimization at the planning stages is evaluated.

This work is organized as follows. First of all, we want to give the reader an overall view of wireless networking; so, in Chapter 2 we present some general concepts. After a brief introduction on general mathematical modelization, we broadly describe cellular and WLAN technologies. With regard to cellular networks, we discuss two different planning approaches, used depending on the system technology. When talking about second generation systems (GSM) a two-step planning is adopted, which consists first in selecting location and configuration for the deployed BS (*coverage planning*) and then in assigning available frequencies to each station in order to satisfy all traffic demands (*capacity planning*). If third generation systems (UMTS) are considered, a joint coverage and capacity planning is preferable since CDMA is used and all transmissions share the whole bandwidth. For both cases, we provide exhaustive explanations and problems formulations. Then, concerning WLANs, we describe single and multiple channel planning, giving several examples of currently used models.

In Chapter 3 we introduce the energy consumption problem which, together with costs reduction, motivates the efforts in the energy-aware network management. Again, we review the most popular and innovative power-saving techniques for cellular and WLAN networks. In the first case we start explaining single network management expedients, such as powering off some Base Stations when traffic requirements are low; this can be done statically, using a predefined BSs sleep scheme, or dynamically, minimizing the number of active stations according to real traffic variations during the day. We also report a new idea which allows the network sharing between two or more competitor operators in order to reduce the energy waste. Moving on with high-density WLANs, we cover some basic concepts regarding Resource on-Demand and Mesh WLAN networks, geared to complete the brief look on energy saving networking.

Chapter 4 is devoted to the introduction and presentation of our model. First, we considered it necessary to make some preliminary remarks. We expose and discuss the common radio planning and coverage formulation that is at the base of our approach; then, we dedicate some paragraphs to explain the mathematical and physical assumptions we employed. We define a daily traffic pattern based on real measurings, which reflects mobile users'habits and states the active users percentages in every time period, and describe four Base Stations categories. In order to enable the energy management mechanism, we also allowed the Base Stations to adjust their emission power by introducing five power levels; this way, each BS can work at its maximum power, use just a fraction of it or even enter the stand-by mode. After having shown the radio propagation model and the path loss values we adopted for calculating the power received by Test Points, we discuss our optimization model, which aims at minimizing at the same time the equipment deployment costs (Capex) and the price of operational and management expenditures (Opex).

The resolution approach and numerical results are presented in Chapter 5. Given that an important part of the presentation is to be able to test the model with appropriate instances, which is not a straightforward matter, we explain here the Instance Generator and the three test scenarios it has produced. Also, we give a brief description of the tools used in order to obtain mathematical optimization and graphical representations. Finally, we show our results and discuss how they support our approach.

Chapter 6 concludes the thesis and presents some ideas for further work.

Chapter 2

Radio Planning

2.1 Solution-Oriented Modeling and Mathematical Optimization

Because of the growth of the telecommunication industry and the uninterrupted development of new technologies, wireless network optimization is becoming more and more a critical issue. Although planning and optimization campaigns are not frequent, they could enable big medium- and long-term benefits in terms of quality and costs. The tools supporting this kind of actions are limited, specially when network sizes become very large.

In this contest, solution-oriented modeling approaches turn out to be the best to use. Unlike engineering-driven automatic optimization and metaheuristics, which put most effort in system description and cannot assert the quality of the achieved result, solution-oriented modeling attacks the basic structure of the problem reaching an appropriate model through an iterative process.

To better highlight this difference, it is useful to make a distinction between a *system model* and an *optimization model*. A *system model* is the outcome of an exhaustive system analysis using all available engineering expertise. The analysis clarifies what are goals to be reached, important parameters and important quality indicators of a solution. In contrast, an *optimization model* is simpler and more suitable for optimization. It states the optimization goals as well as the relevant constraints in mathematical terms, but it focuses on the decisive features that a solution has to have, temporarily ignoring the other ones (technical features could

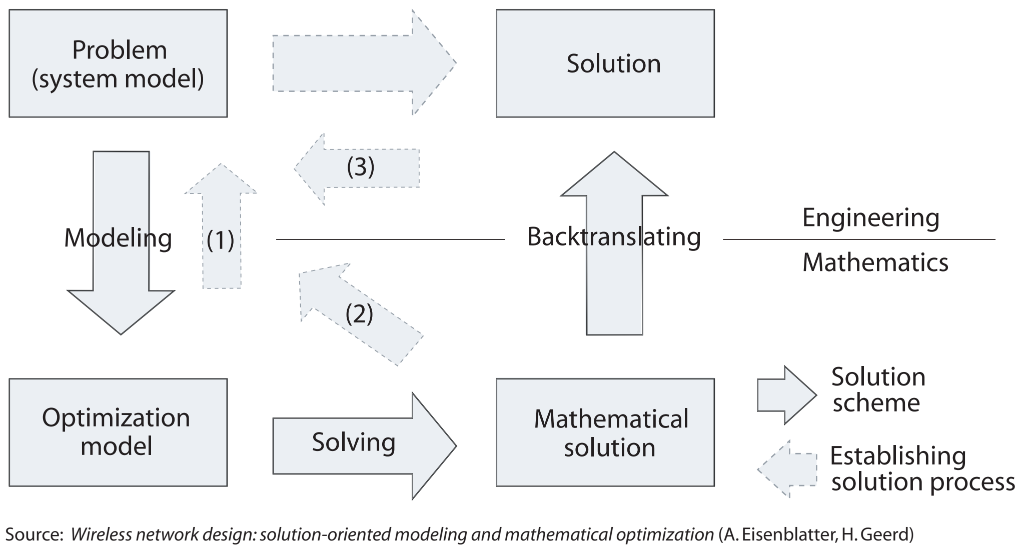


Figure 2.1: Solving a real-world problem with mathematical methods.

be achieved by minor changes to a solution obtained without considering these constraints in the optimization model).

In Fig. 2.1, the problem solving cycle of modern applied mathematics is clearly depicted. A solution-oriented modeling approach translates the planning problem (system model, corresponding to the upper level of the picture) into a formal optimization model, which is solved with the available mathematical solution methods (as shown in the lower level). Then, the mathematical solution is interpreted and transformed into a good solution of the system model, that is, in terms of the real world problem.

Some difficulties can occur. The system model often needs to be refined in order to find an appropriate optimization model; in turn, optimization models have to be simplified if they are too complex to be solved for big size instances (back arrow (1) in Fig. 2.1) . Moreover, experiments are sometimes required to find the aspects that have to be retained in the model and to discard those that may be left aside (back arrow (2) in Fig. 2.1). Early solutions in the process help determine which part of the system or optimization model has to be corrected (back arrow (3) in Fig. 2.1). Finally, a computational test on realistic data should be conducted in order to test the fitness of the optimization model for practical use: if the solutions are not satisfactory, the cycle has to be repeated.

A key point in the solution-oriented modeling approaches is the way to solve the optimization model. The more a model becomes complex, the harder it approaches to good solutions: for this reason, the tendency is to keep the model as simple as possible. If a direct optimal solution algorithm to the problem is not within reach, mathematical optimization tends to operate on linear objective and linear (in)equality constraints using models called *Mixed Integer Linear Programming* (MILP). The general form of a rational MILP is:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \geq b \\ & x \in \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}, \end{aligned}$$

where vectors b and c are in $\mathbb{Q}^m$ and $\mathbb{Q}^n$, respectively, A is a rectangular matrix in $\mathbb{Q}^{m,n}$, and $n = n_1 + n_2$ with $m, n_1, n_2 \in \mathbb{N}_0$. The first n_1 variables can be continuous, while for the last n_2 variables only integral values are feasible.

As solving an MILP is NP-hard, no algorithm is known to solve MILP for which the running time is bounded by some polynomial in the size of the input data. Nevertheless, very powerful MILP solvers are now available. They get optimal or probably good solutions with reasonable computational effort, using the only instances of the NP-hard problem that are "easy" to solve.

The methods underlying such solvers can obtain great results thanks to an interesting feature. Not only the best known solution improves over time (as it occurs in search methods), but also an increasing lower bound on the objective function value exists. No feasible solution can have an objective function value below the lower bound, that is to say that once the upper and lower values meet, a provable optimal solution is found. Even if the computation is terminated prior to this point, the gap between the best solution and the best lower bound allows the quality of the solution to be certified. This cannot be done by stochastic search methods.

As shown in [8], there are four basic models that can show state-of-the-art MILP solvers behavior. For these models, even instances with a large number of variables and constraints can be solved to optimality or close to optimality with reasonable computational effort. When it is possible to choose an optimization model “close” to one of these model, the good solution properties are preserved in the network design problem.

Set covering. Given a family of sets over a common ground set, the task is to choose some of these sets such that each element of the ground set is contained in at least one selected set. Usually, the number of selected sets is to be minimized. This model can well describe coverage problems.

Facility location. Given a set of facility locations and a set of customers who are to be served from the facilities, the problem is to distribute facilities in the region where there are customers. The average proximity of the facilities to customers is to be minimized. A typical example of facility location problem is cellular network design.

Assignment. In case of bipartite assignment, elements in one set (customers) are to be assigned to elements in another set (facilities) subject to a linear utility function.

Knapsack. Given items of different values and volumes, the task is to find the most valuable set of items that fit in a knapsack of fixed volume. The knapsack problem is the simplest form of an integer linear program since it has one constraint and a linear objective function, all with positive coefficient. If the problem has more than one constraint, it is called *multiple knapsack*. The capacity of a cell in wireless network design can be described by a knapsack model.

It should be clear that MILP solution-oriented modeling may be the first approach to network planning problems. Even if simple “standard” techniques fail for complex problems, their application could be a starting point for customized formulations. Helping in focusing on the important aspect of the problem, these methods will reduce investments, save operational costs, reduce pollution in the electromagnetic spectrum and improve the use of scarce radio transmission bandwidth.

The following sections report examples on cellular networks [4] and Wireless Local Area Networks [5]. After a brief introduction on the main working principles of the in point network, in both cases the planning problem is faced starting from the basic approaches cited above. Then, the models are refined in order to achieve more and more accurate solutions.

2.2 Cellular Network Planning

The rapid spread of mobile technology has pushed not only for the development of new advanced systems, but also for the investigation of mathematical models and optimization algorithms to support planning and management decisions. The main contribution of optimization in this field is to improve the way the limited resources (e.g., transmission band, antennas) are used, and to enhance the service quality (e.g., bandwidth, transmission delay).

Some distinctions have to be done when talking about optimization issues in second (e.g., GSM) or third (e.g., UMTS) generation cellular systems. With regard to the first one, due to the computational complexity of the planning problem, it is usually split in two different phases: coverage planning, in which antennas are placed so as to maximize service coverage and transmission powers are selected, and capacity planning, in which frequencies are assigned to the transmitters so as to maximize the average quality of the received signals. This approach is no longer appropriate for third generation systems, which require coverage and capacity planning to be addressed simultaneously.

2.2.1 Cellular Technologies

In a cellular radio system, a land area to be supplied with telecommunication services is divided into regularly shaped smaller areas, each covered by a *Base Station* (BS). Every BS can handle radio connections with *Mobile Stations* (MSs) within its service area, called *cell* and defined as the set of points in which the intensity of the signal received from the BS under consideration is higher than that received from the other BSs.

As users move from cell to cell, service continuity is guaranteed by handover procedures. Usually, during handovers a connection is switched from a BS to a new one (hard-handover), but in some cases simultaneous connections with multiple BSs can be used to improve efficiency.

In order to allow many contemporaneous connections between BSs and MSs, in most of second generation systems the radio band is first divided into carriers at different frequencies using *Frequency Division Multiple Access* (FDMA) and then on each carrier a few radio channels are created using *Time Division Multiple Access* (TDMA). According to the *Frequency Division Duplexing* (FDD) scheme, the bidirectional connection is provided by a pair of channels on different carriers used for transmissions from the BS to the MS (downlink) and from the MS to the BS (uplink).

The radio channels obtained in this way are not enough to serve all mobile service users: in order to increase both the coverage and the capacity of the system, radio channels must be reused in different cells. The denser is the channel reuse, the higher is the number of channels available per cell, but on the other hand this generates interference that can affect the quality of the received signals. In fact, due to the capture effect (i.e., the phenomenon associated with FM reception in which only the stronger of two signals at, or near, the same frequency will be demodulated), if the *Signal to Interference Ratio* (SIR) is greater than a capture threshold SIR_{min}, the signal can be correctly decoded.

Although transmissions on the same frequency are the main source of interference, transmissions on adjacent frequencies may also cause interference due to spectrum overlap and should be considered in network planning.

Unlike second generation cellular systems, which were conceived mainly for the phone and low rate data services, third generation systems are able to also support new multimedia and data services. These systems are based on *Wideband Code Division Multiple Access* (W-CDMA) and before each transmission, signals are spread over a wide band by using special codes. Signals transmitted by the same station are encoded by mutually orthogonal codes, while codes used for signals emitted by different stations can be considered as pseudo-random. Thus, in an ideal environment, the de-spreading process at the receiver can avoid the interference of orthogonal signals and reduce that of the others by the *Spreading Factor* (SF), defined as the ratio between the spread signal rate and the user rate. In a real wireless environment, due to multipath propagation, a little interference also persists among orthogonal signals and the SIR is given by:

$$SIR = SF \frac{P_{received}}{\alpha I_{in} + I_{out} + \eta}, \tag{2.1}$$

where $P_{received}$ is the received power of the signal, I_{in} is the intra-cell interference, I_{out} is the inter-cell interference, α is the orthogonality loss factor ($\alpha = 1$ in uplink, usually $\alpha \ll 1$ in downlink) and η is the thermal noise power. A *Power Control* (PC) mechanism (SIR- or power-based) has the task of dynamically adjusting the emitted power according to the propagation conditions in a way that reduces interference and guarantees quality.

For second generation cellular systems a two-phase approach is adopted. First, coverage is planned so as to assure that in the whole service area a sufficient signal level is received from at least one BS. Then, taking into account SIR constraints and capacity requirements, available frequencies are assigned to BSs. Concerning third generation systems, since in CDMA the bandwidth is shared by all transmissions and no frequency assignment is required, a two-step planning is not appropriate. The network capacity depends on actual interference levels, which determine the achievable SIR values. As these values depend on traffic distribution and on BSs location and configuration, coverage and capacity planning must be jointly planned.

2.2.2 Coverage Planning

Given an area to be served, the general *Coverage Problem* consists in determining where to locate the BSs and in selecting their configuration so that each user in the service area receives a sufficiently high signal. The typical goal is that of minimizing the total antenna installation cost while guaranteeing service coverage.

A possible approach considers the problem from a continuous optimization point of view. In this case, a specified number of BS can be installed in any location of the space to be covered, and antenna coordinates are the continuous variables of the problem. However, due to difficulties in the definition of some key parameters such as the signal path loss, these formulations are beyond the reach of classical location theory methods.

The alternative approach to the coverage problem employs discrete mathematical programming models. In the service area, a set of *Test Points* (TPs) representing the users is identified, each one considered as a traffic centroid where a given amount of traffic is requested. The location of BSs is allowed only in a set of *Candidate Sites* (CSs). Although parameters such as maximum emission power or antenna tilt are inherently continuous, the BS configurations can be discretized by only considering a subset of possible values. As described, the coverage problem amounts to an extension of the classical minimum cost set covering problem.

Let $J = \{1, \ldots, m\}$ denote the set of CSs. For each $j \in J$, let K_j represent all the possible configurations of the BS that can be installed in CS j. An installation cost c_{jk} is associated with each pair of CS $j \in J$ and BS configuration $k \in K_j$. Let $I = \{1, \ldots, n\}$ denote the set of TPs.

The propagation information is summarized by the attenuation matrix G, which coefficients $g_{i\,jk}$, $0 < g_{i\,jk} \leq 1$, represent the attenuation factor of the radio link between TP $i \in I$ and a BS installed in $j \in J$ with configuration $k \in K_j$. From the attenuation matrix, it is possible to derive a $0-1$ incidence matrix containing the coverage information and described by the coefficients:

$$a_{i\,jk} = \begin{cases} 1 & \text{if a BS intalled in CS } j \text{ with configuration } k \\ & \text{can cover TP } i, \\ & \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

Once the following binary variables are introduced:

$$y_{jk} = \begin{cases} 1 & \text{if a BS with configuration } k \text{ is installed in CS } j, \\ & \\ 0 & \text{otherwise,} \end{cases} \tag{2.3}$$

the problem of covering all TPs at minimum cost can be formulated as:

$$\min \quad \sum_{j \in J} \sum_{k \in K_j} c_{jk} y_{jk} \tag{2.4}$$

$$\text{s.t.} \quad \sum_{j \in J} \sum_{k \in K_j} a_{i\,jk} y_{jk} \geq 1 \qquad \forall i \in I \tag{2.5}$$

$$\sum_{k \in K_j} y_{jk} \leq 1 \qquad \forall j \in J \tag{2.6}$$

$$y_{jk} \in \{0, 1\} \qquad \forall j \in J, k \in K_j. \tag{2.7}$$

Constraints (2.5) ensure that all TPs are covered at least by one BS, and constraints (2.6) state that in each CS at most one configuration is selected for the installed base station.

Actually, since the problem implies a trade-off between coverage and installation costs, constraints (2.5) are modified by introducing the following variables:

$$z_i = \begin{cases} 1 & \text{if TP } i \text{ is covered,} \\ \\ 0 & \text{otherwise.} \end{cases} \tag{2.8}$$

The model belongs now to the class of maximum coverage problems:

$$\begin{aligned}
\max \quad & \lambda \sum_{i \in I} z_i - \sum_{j \in J} \sum_{k \in K_j} c_{jk} y_{jk} && && (2.9)\\
\text{s.t.} \quad & \sum_{j \in J} \sum_{k \in K_j} a_{i\,jk} y_{jk} \geq z_i && \forall i \in I && (2.10)\\
& \sum_{k \in K_j} y_{jk} \leq 1 && \forall j \in J && (2.11)\\
& y_{jk} \in \{0,1\} && \forall j \in J, k \in K_j && (2.12)\\
& z_i \in \{0,1\} && \forall i \in I, && (2.13)
\end{aligned}$$

where $\lambda > 0$ is a suitable trade-off parameter. Unfortunately, neither of these two discrete models consider the interference or the overlap between cells, which are very important factors in handover procedures.

The full coverage problem can be refined to account for the "shape" of the cells, influenced by the location of the installed BSs. This is done by introducing another set of variables:

$$x_{ij} = \begin{cases} 1 & \text{if TP } i \text{ is assigned to BS } j, \\ \\ 0 & \text{otherwise,} \end{cases} \tag{2.14}$$

it is ensured that there exists at least one configuration of the BS in CS j that allows the communication with TP i. If $K(i,j)$ denotes the set of available configurations for the BS in CS j that permit the connection with TP i, the new formulation of the problem becomes:

$$\min \quad \sum_{j\in J}\sum_{k\in K_j} c_{jk}y_{jk} \tag{2.15}$$

$$\text{s.t.} \quad \sum_{j\in J} x_{ij} = 1 \qquad \forall i \in I \tag{2.16}$$

$$\sum_{k\in K_j} y_{jk} \le 1 \qquad \forall j \in J \tag{2.17}$$

$$x_{ij} \le \sum_{k\in K(i,j)} y_{jk} \qquad \forall i \in I, j \in J : K(i,j) \neq \emptyset \tag{2.18}$$

$$y_{jk} \in \{0,1\} \qquad \forall j \in J, k \in K_j \tag{2.19}$$

$$x_{ij} \in \{0,1\} \qquad \forall i \in I, j \in J : K(i,j) \neq \emptyset. \tag{2.20}$$

The key constraints (2.18) state that a TP can be assigned to a BS only if the configuration of this BS allows that connection. In order to transform the above problem in the maximum coverage variant, the equality constraints (2.16) expressing full coverage have to be transformed into inequalities and a suitable term proportional to the number of connected TPs has to be added to the objective function.

This basic model can be modified so as to assign each TP only to the "closest" (in terms of signal strength) activated BS. A possible way to express this restriction is to consider for a given TP i all the pairs of BSs and configurations that allow connections with i and sort them in decreasing order of signal strength. If $\{(j_1,k_1),(j_2,k_2),\ldots,(j_L,k_L)\}$ is the ordered set of BS-configuration pairs, the necessary constraints are:

$$y_{j_lk_l} + \sum_{h=l+1}^{L} x_{ij_h} \le 1 \quad l = 1,\ldots,L-1. \tag{2.21}$$

According to these constraints, if a BS is activated in a convenient configuration, then TP i cannot be connected to a less convenient BS.

All these models can be solved with known exact heuristic methods.

2.2.3 Capacity Planning

In second generation systems, the coverage planning phase has to be followed by the assignment of an available frequency to each transmitter so that all traffic demands can be served and the quality of the received signals is maximized. The corresponding problem is called *Frequency Assignment Problem* (FAP), and it may assume very different forms depending on spectrum size, objectives and technological constraints.

In the 1970s, when operators had to pay for each single frequency, the purpose was to minimize the total number of frequencies required by non-interfering configurations, which corresponds to solving an appropriate version of the graph coloring problem. In this context, a network R is associated to a graph $G = (V, E)$, where V is defined as the set of antennas (TRXs) of R and $\{i, j\} \in E$ if and only if TRX i and TRX j interfere. Any coloring of the vertices of G such that adjacent vertices have different colors is then an assignment of frequencies to R such that no interfering TRXs receive the same frequency. Solutions to this approach are provided by simple greedy heuristics.

The above model assumes that distinct frequencies do not interfere, but it may not be true: generally, a frequency h interferes with all frequencies $g \in [h-\delta, h+\delta]$, where δ depends upon channel bandwidth, type of transmission and power of signals. So, a modified version of the graph coloring problem was proposed in the 80's. An instance of FAP is now represented by a complete, undirected, weighted graph $G = (V, E, \delta)$, where δ is the *distance vector* and δ_{uv} is the minimum admissible distance between a frequency f_u assigned to TRX u and a frequency f_v assigned to TRX v. The problem becomes that of finding an assignment f such that $|f_v - f_u| \geq \delta_{uv}$ for all $\{u, v\} \in E$ and the difference between the largest and the smallest frequency, called *Span* (f), is minimized. This version of FAP, based on minimum Span assignment and called MS-FAP, has been mainly tackled with heuristics methods, ranging from the simple generalization of the graph coloring heuristics to specific implementations of local search.

The fast-growing traffic demand led up to increase the number of TRXs installed on the same BS (i.e., the number of frequencies assigned to a BS). In the graph model presented so far, each vertex v of G stands for a TRX. However, as for their interferential behavior, TRXs belonging to the same BS are indistinguishable. In order to fit the model to the new problem, it turns out to be necessary to introduce the representation $G = (V, E, \delta, m)$. Each vertex v corresponds now to a BS, while $m \in R^{|V|}$ is a *multiplicity* vector with m_v denoting, for each $v \in V$, the number of frequencies to be assigned to v. The FAP is then the problem of assigning m_v frequencies to every vertex of G so as *every* frequency f_v assigned to v and *every* frequency f_u assigned to u satisfy $|f_v - f_u| \geq \delta_{uv}$ and the difference between the largest and the smallest frequencies assigned (Span) is minimized.

Around the 1990s, due to the ever-increasing number of users, the available band became inadequate to allow for interference-free frequency plans. In addition to this, frequencies were now sold to operators in blocks rather than in single units. For these reasons, the objective of planning was no longer that of minimizing the number of frequencies, but became that of maximizing the quality of service, which corresponds to minimizing the overall interference of the network. This problem so formulated is called *Minimum Interference Frequency Assignment Problem* (MI-FAP).

The basic MI-FAP considers only the interference occurring between couples of interfering TRXs, measured as the number of unsatisfied requests of connection. If v and w are potentially interfering TRXs and f, g two available frequencies (not necessarily distinct), a penalty p_{vwfg} is introduced to represent the interference (cost) generated when TRX v is assigned to frequency f and TRX w is assigned to frequency g. The problem is then that of finding a frequency assignment which minimizes the sum of penalty costs.

In order to describe the 0-1 linear program for MI-FAP, a binary variable for every vertex v and available frequency f has to be introduced:

$$x_{vf} = \begin{cases} 1 & \text{if frequency } f \text{ is assigned to vertex } v, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{2.22}$$

The contribution to the objective value of the interference between v and w can be written as $\sum_{f,g \in F} p_{vwfg} x_{vf} x_{wg}$. The variables $z_{vwfg} = x_{vf} x_{wg}$ have to be introduced for all TRXs $v, w \in V$ and all frequencies $f, g \in F$ in order to linearize the quadratic terms $x_{vf}x_{wg}$:

$$z_{vwfg} = \begin{cases} 1 & \text{if } x_{vf} \cdot x_{wg} = 1, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{2.23}$$

The MI-FAP problem can be formulated as follows:

$$\min \quad \sum_{\{v,w\}\in E} \sum_{f,g\in F} p_{vwfg} z_{vwfg} \tag{2.24}$$

$$\text{s.t.} \quad x_{vf} + x_{wg} \leq 1 + z_{vwfg} \qquad \forall\{v,w\} \in E, f,g \in F \tag{2.25}$$

$$\sum_{f\in F} x_{vf} = m_v \qquad \forall v \in V \tag{2.26}$$

$$x_{vf} \in \{0,1\} \qquad \forall v \in V, f \in F \tag{2.27}$$

$$z_{vwfg} \in \{0,1\} \qquad \forall v,w \in V, f,g \in F. \tag{2.28}$$

Constraints (2.25) enforce $z_{vw\,fg}$ to be one when $x_{vf} = x_{wg} = 1$, while the *multiplicity constraints* (2.26) state that m_v frequencies have to be assigned to each vertex v.

If only co-channel interference is involved (i.e., $p_{vwfg} = 0$ holds wherever $f \neq g$), then the MI-FAP reduces to the *max k-cut* problem: given an edge-weighted graph $G = (V, E, \delta, m)$, find a partition of V into k classes so that the sum of the weights of crossing edges is maximized.

A lot of solution approaches, both exact and approximate, have been proposed for the MI-FAP. Among those heuristic approaches, the most successful seem to be variants of Simulated Annealing, while examples of exact methods are implicit enumeration as well as polyhedral approaches.

2.2.4 Joint Coverage and Capacity Planning

Since third generation systems are based on a CDMA radio access scheme and the signal quality depends on all the communications in the system, the two-phase approach is no longer appropriate.

The general *UMTS network planning problem* can be developed as follows. Given the set $J = \{1, \ldots, m\}$ of candidate sites and the set $I = \{1, \ldots, n\}$ of test points, with the corresponding number u_i of active connections for each TP i, and given the propagation matrix G (providing information about channel attenuation between CSs and TPs), the purpose is that of selecting a subset of CSs where to install BSs together with their configurations so as to optimize an objective function which considers both traffic coverage and installation costs. SIR constraints and power limits are also to be taken into account.

A possible approach to the UMTS network planning problem is based on a *Mixed Integer Programming* (MIP) model, which can apply to power-based as well as to SIR-based power control. In both cases, the MIP formulation involves the location variables y_{jk} and the assignment variables:

$$x_{ijk} = \begin{cases} 1 & \text{if TP } i \text{ is assigned to a BS in CS } j \\ & \text{with configuration } k, \\ & \\ 0 & \text{otherwise,} \end{cases} \tag{2.29}$$

with $i \in I$, $j \in J$ and $k \in K_j$. Since the aim is that of reaching a trade-off between maximizing the total traffic covered and minimizing the total installation costs, the objective function is:

$$\max \quad \lambda \sum_{i\in I} \sum_{j\in J} \sum_{k\in K_j} x_{ijk} - \sum_{j\in J} \sum_{k\in K_j} c_{jk} y_{jk}, \tag{2.30}$$

where $\lambda > 0$ is the trade-off parameter between the two contrasting objectives.

The first three groups of constraints of the formulation are common to the two power control strategies:

$$\sum_{j\in J} \sum_{k\in K_j} x_{ijk} \leq 1 \quad \forall i \in I \tag{2.31}$$

$$\sum_{k\in K_j} y_{jk} \leq 1 \quad \forall j \in J \tag{2.32}$$

$$\sum_{k\in K_j} x_{ijk} \leq \sum_{k\in K_j} y_{jk} \quad \forall i \in I, j \in J. \tag{2.33}$$

Constraints (2.31) and (2.32) state respectively that every TP i can be assigned to at most one BS and at most one BS configuration can be selected for CS j, while constraints (2.33) ensure that a TP i can be assigned to a CS j only if a BS with some configuration k has been installed in j.

Now, the SIR constraints (which express the signal quality requirements) are to be investigated. Considering a SIR-based PC mechanism, the resulting model turns out to be a mixed integer program with nonlinear SIR constraints, since they contain products of assignment variables (x_{ijk} and y_{jk}) and power variables (p_i^{up} and p_i^{dw}, defined respectively as the power emitted by any mobile terminal at TP i and by any BS j).

At this point, in order to simplify the model, a power-based PC mechanism can be assumed, which adjusts all emitted powers so as to guarantee a received power of P_{tar}. Due to this choice, the powers p_i^{up} emitted from any TP i in uplink and the powers p_i^{dw} received at any TP i from the BS they are assigned depend on the value of P_{tar}, as well as on the propagation factor of the corresponding radio links. Taking into account that:

$$p_i^{up} = \sum_{j \in J} \sum_{k \in K_j} \frac{P_{tar}}{g_{ijk}} x_{ijk} \qquad p_i^{dw} = P_{tar}, \tag{2.34}$$

the SIR constraints, once linearized, can be written as follows:

$$(\alpha I_{in} + I_{out} + \eta) \leq \frac{1}{SIR_{min}} + M_{ijk}(1 - x_{ijk}), \tag{2.35}$$

where P_{tar} and M_{ijk} are constants, I_{in} and I_{out} are linear functions in the x and y variables.

Even the linearization of this simplified model yields integer linear programs that are computationally very challenging and cannot be tackled with exact methods. A particular Tabu Search procedure, based on a two-stage approach, provides good quality solutions of relevant-size instances in reasonable computing time. It adopts a two-stage approach: solutions of a simpler model, which considers a power-based PC mechanism and only the uplink direction, are exploited as good initial solutions for the overall uplink and downlink model with a SIR-based PC mechanism. Since this model is a quite accurate approximation of the model with SIR-based power control, the insight gained from solving the former model helps reducing the computing times for handling the overall model without significantly affecting the solution quality.

At present, the goal is to find a reasonable trade-off between an accurate description of the UMTS network planning problem and a computationally tractable model. This is a challenging problem if a representative set of traffic scenarios is considered.

2.3 Wireless Local Area Network Planning

The recent and impressive spread of wireless technologies has allowed the development of *Wireless Local Area Networks* (WLANs), calling for quantitative approaches in the network planning procedure.

The most popular standard for WLANs is the IEEE 802.11. Basically, the IEEE 802.11 based wireless network planning problem consists in selecting the positions of *Access Points* (APs) and in assigning to each of them a channel. A common approach considers feasible positions of traffic concentration points in the service area (Test Points) and feasible positions where APs can be installed (Candidate Sites). The placement of TPs and CSs depends on the traffic distribution and on the characteristics of the area to be covered.

So far, most of radio planning schemes proposed are extensions of those adopted for cellular networks and aim at minimizing installation costs while providing a good signal level for the served area. A model frequently employed in wireless networks design is the NP-hard *Set Covering Problem* (SCP). As mentioned above, for this problem, consisting in selecting a subset of CSs positions able to cover all the TPs with the minimum total installation costs, fast exact algorithms and effective heuristics are known. Sometimes, in order to take into account the fact that in WLANs the data rate varies with the strength of the received signal, facility location models are used instead of set covering ones.

However, wireless network planning is quite different from cellular networks planning and it may be considered separately: first, the installation costs of WLAN APs is much lower than those of cellular networks base stations; second, WLANs weren't devised to provide cellular coverage.

Up to now, the development of coverage planning tools especially conceived for WLANs has been regarded much too expensive with respect to the price of access points. Only in the last period, the growing ability of WLAN systems to provide services is drawing the attention on finding effective methods to determine high capacity and cost-effective solutions to the WLAN coverage planning problem.

Not many works have appeared on this specific issue. All of them focus on the problem of reaching a high coverage level in terms of received signal quality or balanced traffic load among installed APs. Very few of them look at network efficiency as an optimization parameter and no one considers during the planning phase the peculiar channel access mechanism, which can dramatically affect WLAN efficiency.

2.3.1 WLANs Technologies

The IEEE 802.11 standard defines a *Basic Service Set* (BSS) as the main network element. The BSS consists of a single AP connected to a wired backbone network, which provides wireless connectivity to a group of mobile users. Like in a segment of an Ethernet LAN, collisions are limited by the *Carrier Sense Multiple Access* (CSMA) protocol.

This network configuration fits well for small areas to be covered, but one single BSS can hardly provide the required wireless coverage for large facilities like buildings or university campuses. In such cases an *Extended Service Set* (ESS) has to be used, which is composed of multiple APs connected through a wired distribution system. Since some APs may share the same radio resources, inter-domain interference can lead at ESS efficiency degradation. The more different APs' coverage ranges overlap, the more inter-domain interference increases; on the other hand, a certain degree of overlap should be maintained in order to ensure the continuity of service to nomadic users.

Always considering the IEEE 802.11, the basic access method is the *Distributed Coordination Function* (DCF) which uses CSMA with *Collision Avoidance* (CSMA/CA). This model requires each station to listen for other users' transmissions. If the channel is idle the station may transmit, whereas, if it is busy, the station waits until the end of the present transmission and then enters a random back off procedure.

Another mechanism, called *Virtual Carrier Sensing*, can avoid collisions even though some stations are hidden from each other due to the limited transmission ranges. The *Virtual Carrier Sense* enables a station to reserve the medium for

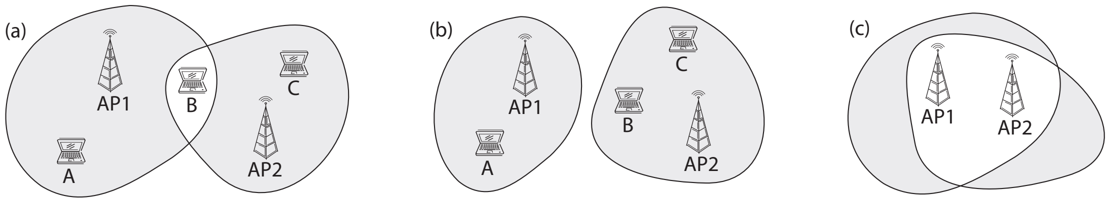


Figure 2.2: Different overlap degrees between two APs' sensing regions.

a period of time using RTS/CTS control frames. The *Request To Send* (RTS) frame containing the length of the reserved period is sent from the sender to the receiver; upon the reception of the RTS, the receiving station responds with a *Clear To Send* (CTS) frame, which also specifies the duration of the period of time for which the medium is reserved. This reciprocal exchange ensures that all the stations overhearing the RTS and/or the CTS packets refrain from accessing the medium for the whole period specified in them.

The carrier sense mechanisms described has been devised to prevent any interference during ongoing transmissions. On the other hand, this limits parallel transmissions in the network. This drawback is displayed in Fig. 2.2a, where the node B is covered by both AP1 and AP2 using the same radio channel: if B is engaged in a communication with AP1, every other user in the intersection of the two coverage regions is forbidden to start any new communication. As a matter of fact, if a WLAN is composed of two APs with non-overlapping coverage areas, the network capacity is the sum capacity of the two APs (Fig. 2.2b), but if the two APs are within the sensing range one another (Fig. 2.2c), then the capacity of the WLAN is the capacity of a single AP. Thus, even if a minimum overlap is required to guarantee service continuity, the degree of overlap among coverage areas highly affects network efficiency. The problem of inter-domain interference can be attenuated by using multiple radio channels within a WLAN.

2.3.2 Single Channel WLAN Planning

In spite of its advantages, using multiple channel can create problems in the network management: handover between different APs can affects the service offered to the user and, for many *Network Interface Connectors* (NICs) available on the market, the connection must be torn down and then re-established, with consequent high delays. For these reasons, it is worth considering the WLAN planning problem also when only one radio channel is available.

As already done for cellular networks planning, let $I = \{1, \ldots, n\}$ denote the set of TPs and let $J = \{1, \ldots, m\}$ denote the set of CSs. The subset $I_j \subseteq I$ represents the set of TPs covered by an AP installed in a given CS $j \in J$, while the subset $J_i \subseteq J$ represents the set of CSs that (once an AP has been installed in) can cover a given TP $i \in I$.

A generic solution of the planning problem is a subset $S \subseteq J$ of CSs where APs have to be installed. The subset $I(S) \subseteq I$ of TPs that can be covered by installing such APs is defined as:

$$I(S) = \bigcup_{j \in S} I_j. \tag{2.36}$$

The decision variables x_j state which subsets are part of the solution:

$$x_j = \begin{cases} 1 & \text{if an AP is installed in CS } j, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{2.37}$$

Moreover, the coefficients a_{ij} define a test points - candidate sites incidence matrix that can resume the information contained in I_j or J_i:

$$a_{ij} = \begin{cases} 1 & \text{if TP } i \text{ belongs to subset } I_j, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{2.38}$$

Set Covering and Minimum Overlap

As mentioned above, the most common approach to the planning problem consists in minimizing the installation costs in terms of weighted number of installed APs, while assuring full coverage for all the TPs. Defining c_j as the cost associated to the installation of an AP in site j, the formulation of this set covering problem can be written as:

$$\min \quad \sum_{j \in J} c_j x_j \tag{2.39}$$

$$\text{s.t.} \quad \sum_{j \in J} a_{ij} x_j \geq 1 \qquad \forall i \in I \tag{2.40}$$

$$x_j \in \{0, 1\} \qquad \forall j \in J. \tag{2.41}$$

Constraints (2.40) force each TP to be covered by at least one installed AP, while constraints (2.41) are the integrality constraints for the binary decision variables.

Although the SCP considers the installation cost as the central parameter to be optimized, it may not be the main issue in WLAN planning. Furthermore, the SCP approach completely neglects the optimization of the network capacity, providing often enough low efficiency solutions.

In order to derive a planning model that provides WLANs with higher efficiency, it is possible to modify the objective function of the SCP maintaining its basic structure. Since the more the coverage regions of different APs overlap the higher is the overall interference in the WLAN, one should try to minimize such overlap degree. A simplified way to minimize the overlap is to minimize the average number of installed APs which cover a given TP. Thus, the first model proposed to get higher efficiency WLANs is the *Minimum Overlap Problem* (MOP), that can be formulated as follows:

$$\min \quad \sum_{i \in I} \sum_{j \in J} a_{ij} x_j \tag{2.42}$$

$$\text{s.t.} \quad \sum_{j \in J} a_{ij} x_j \geq 1 \qquad \forall i \in I \tag{2.43}$$

$$x_j \in \{0, 1\} \qquad \forall j \in J. \tag{2.44}$$

By rewriting the objective function, the MOP formulation can be easily reduced to an equivalent SCP formulation with appropriate costs c_j:

$$\sum_{i \in I} \sum_{j \in J} a_{ij} x_j = \sum_{i \in I} \left(\sum_{j \in J} a_{ij} \right) x_j = \sum_{j \in J} c_j x_j. \tag{2.45}$$

MOP is therefore simply a SCP with the installation costs c_j of each CS equal to the number of TPs covered by the CS itself (i.e., with $c_j = |I_j|$). The MOP approach can provide better solutions that SCP, but it does not really address the problem of maximizing the network efficiency because it does not consider the peculiar access mechanism of WLANs.

Network Efficiency Estimation

Since the capacity of a WLAN depends on many dynamic parameters, it can be hardly defined a priori during the planning phase. At this point, the concept of *balanced share* introduced in [5] turns out to be very useful: it can be used as a simplified estimation of the network saturation throughput to be adopted in the optimization models as a network quality indicator.

The *balanced share* for a given TP i in a given solution S can be defined by the equation:

$$BS(S,i) = \begin{cases} \frac{1}{Int(S,i)} & \text{if } i \in I(S), \\ \\ 0 & \text{otherwise,} \end{cases} \tag{2.46}$$

where $Int(S,i)$ is the number of users interfering with user i in the solution S (i.e., user i's competitors to gain access to the channel). Assuming that:

- each user is connected to a single AP whose capacity is shared by all the users within its coverage range,
- the overall capacity of an AP is equal to 1,
- the fraction of the AP capacity available to a given user is equal to the reciprocal of the number of users in the interference range of the set of APs covering that user,

the number of user i's competitors is also given by:

$$Int(S,i) = \left|I\big(S \cup J_i\big)\right|,$$

where $\left|I\big(S \cup J_i\big)\right|$ is the total number of TPs covered by the same APs covering i. Then, the *balanced share* can be interpreted as the probability that each user has to access the shared resource, under the assumption of uniform and maximal traffic.

Maximum Efficiency Planning

The *balanced share* can be used to plan networks with higher saturation throughput with respect to other planning methods (SCP and MOP). For this purpose, besides the decision variables x a new set of variables has to be introduced to measure $Int\big(S, i\big)$:

$$y_{ih} = \begin{cases} 1 & \text{if TPs } i \text{ and } h \text{ appear together} \\ & \text{in some selected subset,} \\ & \\ 0 & \text{otherwise.} \end{cases} \tag{2.47}$$

The *Maximum Efficiency Problem* (MEP) can be formulated as follows:

$$\max \quad \sum_{i \in I} \frac{1}{\sum_{h \in I} y_{ih}} \tag{2.48}$$

$$\text{s.t.} \quad \sum_{j \in J} a_{ij} x_j \geq 1 \qquad \forall i \in I \tag{2.49}$$

$$y_{ih} \geq a_{ij} a_{hj} x_j \qquad \forall j \in J, i, h \in I \tag{2.50}$$

$$x_j \in \{0, 1\}, y_{ih} \in \{0, 1\} \qquad \forall j \in J, i, h \in I. \tag{2.51}$$

From the definition of variables y it's easy to note that $\big|I(S \cup J_i)\big| = \sum_{h \in I} y_{ih}$. Constraints (2.49) impose the complete coverage, as in the SCP problem, while constraints (2.50) express the interference relation between TP i and TP h (if TP i and h are covered by a common installed AP, then $y_{ih} = 1$).

A network planner may be required to respect a certain cost budget in the APs installation. To this end, both the MOP and the MEP formulations can be modified to account for cost limitations simply by adding the constraint:

$$\sum_{j \in J} c_j x_j \leq B, \tag{2.52}$$

with c_j the cost, as in the SCP problem, and B the budget.

2.3.3 Maximum Efficiency Multiple Channel WLANs

The efficiency of the planned network can be enhanced using a high number of frequency channels: in this way, the distance of interfering stations is augmented and therefore the interference is reduced. The MEP formulation analyzed above for a single channel case can be extended to the case where multiple channels are available. A first, simplified formulation will consider a multi frequency *balanced share*, defined as the average of many single channel *partial balanced shares*, but it does not take into account the assignment of TPs to CSs. To this end, a more accurate but complicated model is derivable, which considers that every user selects as his working frequency the frequency of the AP received with the strongest signal.

Simplified Multiple Frequencies WLAN Planning

In order to express the new optimization problem, the MEP formulation needs some modifications.

Let F denote the set of available frequency channels. Let S be a solution (subset of selected CSs, each one with an assigned frequency) and let $S_f \subseteq S$ be the set of CSs with assigned frequency f. For any frequency f a *partial balanced share* is evaluated, that is equal to the *balanced share* $BS(S_f, i)$. If k_i is the number of frequency covering the user i, the *mean balanced share* could be defined as the mean of the *partial balanced shares* over all the covering frequencies:

$$MBS(S,i) = \begin{cases} \frac{\sum_{f \in F} BS(S_f,i)}{k_i} & \text{if } k_i > 0, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{2.53}$$

The sum over all users of their *mean balanced share* can represent a first approximation of the multiple frequency network efficiency, as in the case of a single channel network.

In order to formulate the problem, new sets of binary decision variables have to be introduced:

$$\bar{x}_{jf} = \begin{cases} 1 & \text{if an AP is installed in CS } j \text{ with frequency } f, \\ \\ 0 & \text{otherwise,} \end{cases} \tag{2.54}$$

$$z_{if} = \begin{cases} 1 & \text{if TP } i \text{ is covered at frequency } f, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{2.55}$$

To calculate the *partial balanced share* another set of binary decision variables is required:

$$y_{ihf} = \begin{cases} 1 & \text{if TP } i \text{ and TP } h \text{ appear together} \\ & \quad \text{in some } S_f, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{2.56}$$

The *Simplified Multiple Frequencies Maximum Efficiency Problem* (S-MF-MEP) can now be formulated as follows:

$$
\begin{aligned}
\max \quad & \sum_{i\in I} \frac{1}{k_i} \sum_{f\in F} \frac{z_{if}}{\sum_{h\in I} y_{ihf}} & & \qquad (2.57)\\
\text{s.t.} \quad & \sum_{f\in F} \bar{x}_{jf} \le 1 & \forall j \in J & \qquad (2.58)\\
& \sum_{jinJ} a_{ij}\bar{x}_{jf} \ge z_{if} & \forall i \in I, f \in F & \qquad (2.59)\\
& z_{if} \ge a_{ij}\bar{x}_{jf} & \forall i \in I, j \in J, f \in F & \qquad (2.60)\\
& k_i = \sum_{f\in F} z_{if} & \forall i \in I & \qquad (2.61)\\
& k_i \ge 1 & \forall i \in I & \qquad (2.62)\\
& y_{ihf} \ge a_{ij}a_{hj}\bar{x}_{jf} & \forall i, h \in I, j \in J, f \in F & \qquad (2.63)\\
& y_{iif} = 1 & \forall i \in I, f \in F & \qquad (2.64)\\
& \bar{x}_{jf} \in \{0,1\} & \forall j \in J, f \in F & \qquad (2.65)\\
& z_{if} \in \{0,1\} & \forall i \in I, f \in F & \qquad (2.66)\\
& y_{ihf} \in \{0,1\} & \forall i, h \in I, f \in F. & \qquad (2.67)
\end{aligned}
$$

Constraints (2.58) state that every CS can be either excluded from te solution or included with a unique assigned frequency. Constraints (2.59) set the variable z to 0 if there is no covering AP at a given frequency for a given TP, while constraints (2.60) state that the variable z must have value 1 if there is a covering AP for the corresponding TP. Constraints (2.61) define for every user i the number of covering frequencies k_i, while constraints (2.62) impose the complete coverage. Constraints (2.63) define the variables y_{ihf}. At the end, constraints (2.64) are needed to avoid that the denominator of the objective function evaluates to 0 in case that a given user is uncovered at a particular frequency.

It is possible to extend the formulation to the uncovered case only relaxing constraints (2.61) to:

$$k_i \ge \sum_{f\in F} z_{if}. \qquad (2.68)$$

If a given TP i is not covered, it will result in $\sum_{f\in F} z_{if} = 0$, and because of constraint (2.62) it will be $k_i = 1$, giving in the objective function a *mean balanced* *share* of 0. Otherwise, if the TP i is covered, then because of the objective function constraint, (2.68) will hold to equality.

Multiple Frequencies WLAN Planning with Assignment

In order to better describe the problem, a more accurate model considers the assignment of TPs to CSs, assuming that:

- a TP works at the same frequency of the AP that hears with the strongest signal,
- a TP h interferes with a TP i if and only if both of them work at the same frequency f and there is an installed AP covering both stations and working at the same frequency f.

Assigning a frequency to the stations changes radically the way the interferers have to be considered. Let $f(i)$ be the frequency assigned to a TP i in a given solution S, and let $S_f \subseteq S$ denote the partitioning of the covered TPs in frequency classes by $T_f = \{i \in I(S) : f(i) = f\}$. The *balanced share* of a TP i is now given by the reciprocal of the number of TPs working at frequency $f(i)$ that share with i a selected covering AP working at frequency $f(i)$:

$$BS(S,i) = \begin{cases} \frac{1}{|T_{f(i)} \cap I(S_{f(i)} \cap J_i)|} & \text{if } i \in I(S), \\ \\ 0 & \text{otherwise.} \end{cases} \tag{2.69}$$

Other parameters and variables have to be added. The previous variables x_j, now linked to $\bar{x}_{jf}$, and y_{ih}, defining the interference between users i and h, are reintroduced to define the way a station selects its own working frequency. The new set of variables l_{ij}, necessary to represent the association of a TP to an installed CS, is defined as follow:

$$l_{ij} = \begin{cases} 1 & \text{if TP } i \text{ is associated to an AP installed in CS } j, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{2.70}$$

Last, the parameter p_{ij} is introduced in order to specify the ordering in which the APs have to be considered (i.e., from strongest to weakest signal). It is defined as the power received in a given TP i of a signal emitted by an AP installed in CS j. The model is based on the assumption that $p_{ij} > 0 \ \forall j \in J_i$ and that $p_{ij} = 0 \ \forall j \notin J_i$.

The *Multiple Frequency Maximum Efficiency Problem* (MF-MEP) is formalized as follows:

$$
\begin{aligned}
\max \quad & \sum_{i \in I} \frac{1}{\sum_{h \in I} y_{ih}} && && (2.71)\\
\text{s.t.} \quad & \sum_{f \in F} \bar{x}_{jf} = x_j && \forall j \in J && (2.72)\\
& \sum_{j \in J} l_{ij} = 1 && \forall i \in I && (2.73)\\
& l_{ij} \le a_{ij} x_j && \forall i \in I, j \in J && (2.74)\\
& x_j + \sum_{l : p_{ij} > p_{il}} l_{il} \le 1 && \forall i \in I, j \in J_i && (2.75)\\
& z_{if} = \sum_{j \in J} l_{ij} \bar{x}_{jf} && \forall i \in I, f \in F && (2.76)\\
& y_{ih} \ge a_{ij} a_{hj} z_{if} z_{hf} \bar{x}_{jf} && \forall i, h \in I, j \in J, f \in F && (2.77)\\
& x_j \in \{0, 1\} && \forall j \in J && (2.78)\\
& \bar{x}_{jf} \in \{0, 1\} && \forall j \in J, f \in F && (2.79)\\
& l_{ij} \in \{0, 1\} && \forall j \in J, i \in I && (2.80)\\
& z_{if} \in \{0, 1\} && \forall i \in I, f \in F && (2.81)\\
& y_{ih} \in \{0, 1\} && \forall i, h \in I. && (2.82)
\end{aligned}
$$

Constraints (2.72) link $\bar{x}_{jf}$ variables to x_j ones. Constraints (2.73) state that each TP must be assigned to a single AP. Constraints (2.74) state that a TP can be assigned to a CS only if the CS is activated and covers the TP, while constraints (2.75) state that if a CS in position k in the preference list of a given TP is selected, then this TP cannot be assigned to any CS having a worse preference order. Thanks to constraints (2.73), (2.74) and (2.75), if a station i is covered by many APs, than it will be $l_{ij} = 1$ for the nearest and $l_{ij} = 0$ for all the others. The quadratic constraints (2.76) define the frequency of work for the TPs, while the cubic constraints (2.77) define the interference among two TPs according to the previous assumption.

2.3.4 Enhanced WLAN Efficiency Estimation

In the definition of *balanced share* considered until now, given a TP, its interferers are those TPs which fall in the transmission range of all the APs covering the TP itself. However, due to the CSMA/CA mechanism, also the TPs that are in the transmission range of the considered TP are interferers.

In order to correct the previous models, for any TPs it is necessary to know the subsets of users that interfere with it and this can be done with an incident matrix described by the coefficients:

$$b_{ih} = \begin{cases} 1 & \text{if TP } h \text{ is within the hearing range of TP } i, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{2.83}$$

In the MEP formulation, a new set of constraints has to be added:

$$y_{ih} \geq b_{ih} \quad \forall i, h \in I. \tag{2.84}$$

In S-MF-MEP, these constraints have to be repeated for any frequency:

$$y_{ihf} \geq b_{ih} z_{hf} \quad \forall i, h \in I, f \in F. \tag{2.85}$$

Finally, in the MF-MEP formulation quadratic constraints are needed to verify that both users are working at the same frequency:

$$y_{ih} \geq b_{ih} z_{if} z_{hf} \quad \forall i, h \in I, f \in F. \tag{2.86}$$

Another problem may appear with hierarchical instances. Since different transmission powers are allowed, it can happen that a user produces interference over an AP without being covered by it. It is the case of a user i covered by a distant large range AP j and that has nearby a low range coverage AP l that do not cover it: user i has to use a high power in order to communicate with AP j, therefore producing interference over AP l.

In order to refine the formulations, for any CS j it is now necessary to know not only the subset of users I_j that are covered, but also the subset of users $I'_j \supseteq I_j$

that produces interference. A new incident matrix is defined by the coefficients:

$$a'_{ij} = \begin{cases} 1 & \text{if TP } i \text{ belongs to subset } I'_j, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{2.87}$$

Then it is only needed to redefine all the constraints that in the previous models define y substituting the matrix a with the matrix a'. In this way, the number of the constraints gets bigger: this means that the optimal solution given by the original formulations always represents an upper bound for the optimal solution of the formulations considering enhanced capacity measure.

Chapter 3

Green Telecommunication Networks

3.1 The Telecommunication Sector's Footprint

The importance of energy related topics is ever increasing. If the global climate change due to increased greenhouse gases concentration levels in atmosphere should represent the main push to investigate on efficient technology developments, network operators are as well interested in reducing the energy consumption of their networks for economical reasons. As a matter of fact, in addition to the minimization of the environmental impact of the industry, savings in both capital and operating expenditures can be realized by the reduction of energy needs.

A recent study reported in [12] shows that in 2007 the *Information and Communication Technology* (ICT) sector was responsible for a fraction of the world energy consumption ranging between 2% and 10%. In particular, as displayed in Fig. 3.1, the carbon generated from materials and manufactures is about one quarter of the overall ICT footprint, while the rest comes from its use. The predominant energy consumers in the ICT field are large data centers and server farms, and telecommunication networks, including wired and wireless telephony networks, as well as the Internet.

Although a growth in developed country markets is also expected in the coming years, the main upturn will involve the rising demand for ICT in developing countries: according to [12], by 2020 almost a third of global population will own a personal computer (currently only 2%), 50% will own a mobile phone and 5% households will have a broadband connection. When a large fraction of developing

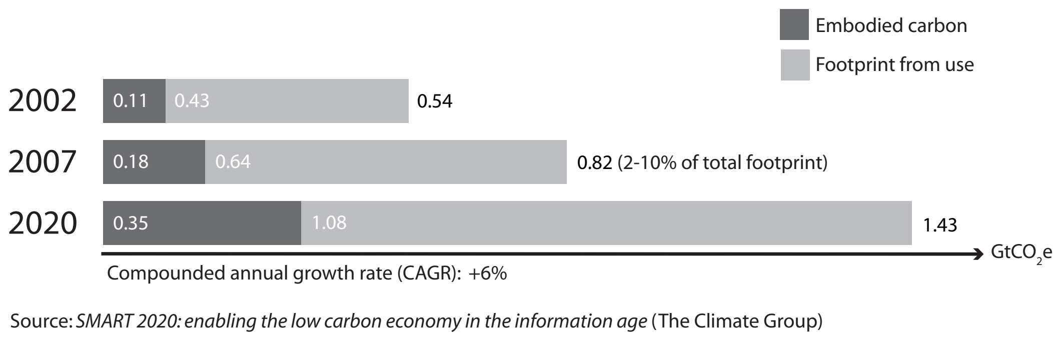


Figure 3.1: The global ICT footprint (ICT includes PCs, telecommunication networks and devices, printers and data centers).

countries' population will be able to afford ICT devices, they will account for more than 60% of ICT's carbon emissions, compared to less than half today (Fig. 3.2).

As demand for telecommunications devices grows, so will the need for the infrastructure that supports it; indeed, the telecoms infrastructure footprint, which was 133 $MtCO_2e$ in 2002, is expected to more then double to 299 $MtCO_2e$ by 2020, a grow rate of 5% pa. Despite that, a decrease in power consumption of telecommunication networks per user is expected, owing to the adoption of efficiency measures. For example, mobile infrastructure technologies now available include network optimization packages which can reduce energy utilization up to 44% and solar-power Base Stations able to reduce carbon emissions by 80%. Natural ventilation is being used by some operators and would diminish the need to cool the network equipment. In addition, competing companies are experimenting with "network sharing", which could reduce the expansion of the infrastructure required by the increasing demand for telecommunication devices.

Concerning this, the literature on general green networking is quickly expanding since the seminal work by Gupta and Singh [13]. Some authors have been involved in virtualization, others in the development of energy-aware Ethernet, others in evaluating the Internet consumption and in proposals to reduce it (for a review see [25]). Most of the work has focused on wireline networks despite the fact that the wireless system is highly responsible for the increase in energy consumption. However, wireless systems engineers have always been concerned with energy issues, since the portability nature of the network, being cellular, ad-hoc or sensor oriented, made it a real challenge in terms of coverage and battery life. Therefore, there is a very large body of literature focused on energy-efficient *de-*

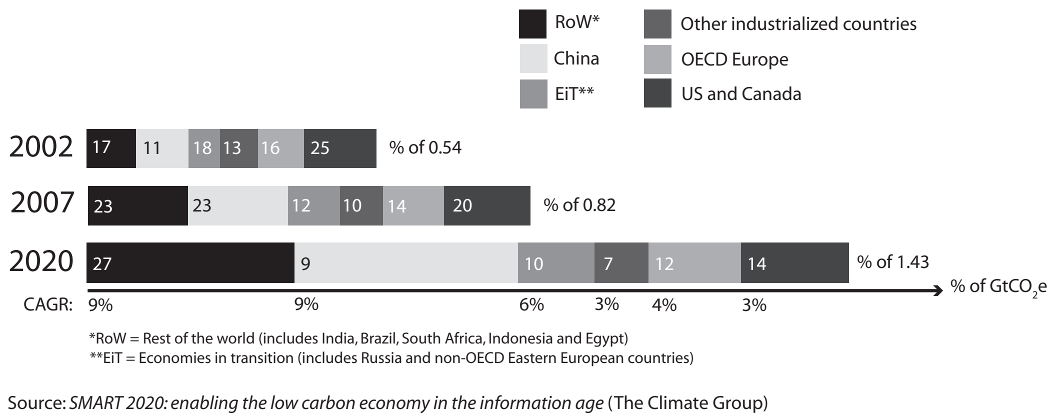


Figure 3.2: The global ICT footprint by geography.

vices or energy-aware *protocols* (an excellent report on the issues affecting wireless energy consumption can be found in [15]), while the literature on green networking planning and operation is recent and scant and mainly deals with management rather than design issues.

3.2 Energy Saving in Cellular Networks

Over 80% of the power in mobile telecommunications is utilized in the radio access network, more specifically by the base stations. In [20] a complete scenario of energy saving opportunities in cellular networks is depicted. There are basically three ways to decrease energy consumption of cellular networks:

Minimizing BS energy consumption. Trying to minimize the energy consumption of the entire cellular network, the first factor to consider is the energy consumption of single BSs. The ways to minimize base stations energy consumption could be divided into three categories:

Improving BS energy efficiency. The main responsible for BS power consumption is the power amplifier in transmission chain, which efficiency depends on the required frequency band, used modulation and operating environment. The ways to improve power amplifier energy efficiency are to use different kind of linearization methods (like digital pre-distortion) or different kind of digital signal processing methods so that the required linear area of the power amplifier is decreased.

Use of system level features. In order to get the right balance between power consumption and performance, it is possible to shut down complete or parts of the BSs during low traffic (for example, during night time).

Base stations site solutions. Another opportunity to save energy consists in taking appropriate precautions in the site choice. For example, by choosing outdoor sites, BSs can be used over wider range of temperature and less cooling or heating is required.

Studying BS deployment strategies. As well as the single BS energy consumption, another factor that has to be taken into account is the number of BSs sites, since the energy expending of the whole network is the multiply of these two. Several features could be used in order to balance between BSs cell size and BSs capacity.

Using renewable energy sources. The most feasible renewable energy sources for base station sites are solar and wind (or hybrid solutions combining of solar and wind). These can be used for several reasons: long distance to electricity grid, unreliable grid and above all in order to reduce the amount of greenhouse gases emissions.

3.2.1 Energy-Aware Management of Individual Cellular Access Networks

Since telecommunication network operators have become interested in energy saving approaches, most attention has been focused on the access segment. This is mainly because there can be found the highest number of elements, so that the energy saved in one access equipment is multiplied by a large factor, with an important contribution to the overall network energy consumption.

From this point of view, in [7], given the network topology and a fixed traffic demand, the authors evaluate the possibility of switching off some nodes in order to minimize the total power consumption, always complying with connectivity and *Quality of Service* (QoS) constraints. However, in [7] no traffic variations in space or time were taken into account. A better approach to tackle the energy consumption problem consists in a *traffic intensity based planning* that reduces the number of active access devices when they are underutilized [21]: in this case the authors show that energy savings of the order of $25-30\%$ are possible for several regular cell topologies.

Since cellular systems are often dimensioned so as to satisfy the quality of service constraints under peak traffic conditions, during low traffic periods they probably result overprovisioned and may waste a significant amount of power. The decrease of the traffic in some portion of a cellular network is due to the combination of two effects:

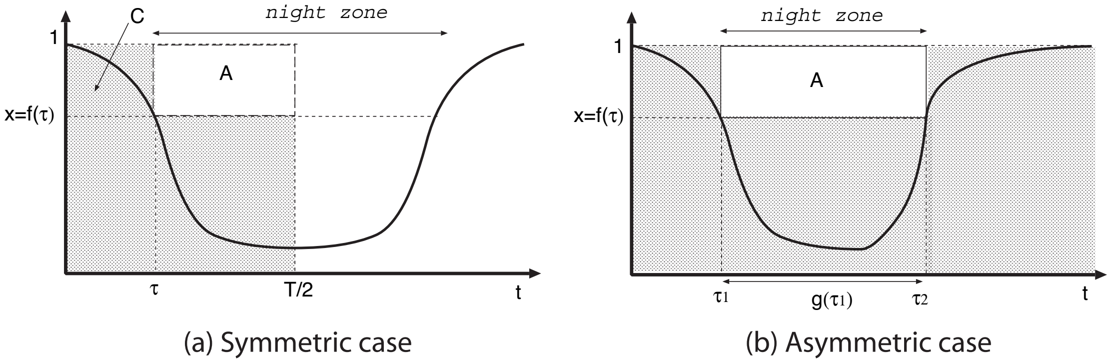
- the typical day-night behavior of users,
- the daily migration of users who move their mobile terminals from residential areas to office areas and back, resulting in the need for large capacity in both areas at peak usage times, but in reduced requirements during the period in which the area is lightly populated.

When a BS is switched on, the energy consumption has a large “floor” level: for this reason, by merely controlling the wireless resources (for example, transmit power), the energy saving is limited, because the energy consumption of processing circuits and the air conditioning largely depends on the on-off states of the BS. Thus, the objective is to *switch off some cells when the load is low.*

Two aspects are to take into account when some BSs are turned off. First of all, the cells that remain active must provide radio coverage over the whole area (including the portions that were taken care of by off BSs) and, in order to increase the radius, cells could require some additional power. Second, the larger cell radius results in an increase of the traffic load, under which quality of service constraints must be guaranteed.

In order to describe this problem with analytical models, assume that during peak traffic periods an area is served by K cells, each one providing QoS for an amount f of traffic (uniform across the cell). The traffic decline could be expressed as xKf (with $x<1$) in the whole area. In this case, only xK cells are necessary to keep the same QoS, as long as electromagnetic coverage is preserved. Hence, $(1-x)K$ cells can be switched off, saving a fraction of energy equal to $(1-x)$ (as to say, the power consumption reduces to a fraction x of the original). Some more mathematical considerations are required so as to choose the fraction x properly.

Let $f(t)$ represent the daily traffic pattern in a cell, with $t \in [0,T], T=$24 h and $t=0$ the peak hour; $f(t)$ is normalized to the peak hour traffic so that $f(0)=1$. In Fig. 3.3 two typical traffic patterns are displayed.



Source: *Optimal energy savings in cellular access networks* (A. Ajmone Marsan, L. Chiaraviglio, D. Ciullo, M. Meo)

Figure 3.3: Possible traffic intensity patterns during a day.

Defined a power-off scheme S such that during the low traffic period (called "night zone") a fraction $x < 1$ of the cells is active, while the remaining $1 - x$ is off, the traffic that the x on cells have to sustain in the night zone is:

$$f^{(S)}(t) = f(t) + \frac{1-x}{x} f(t) = \frac{1}{x} f(t). \tag{3.1}$$

In order to always satisfy the QoS constraint, scheme S can be applied only when the traffic is so low that $f^{(S)}(t) < 1$. Starting from the peak hour, with decreasing $f(t)$, the earliest time instant τ that meets this condition is defined by:

$$f^{(S)}(\tau) = \frac{1}{x} f(\tau) = 1 \tag{3.2}$$

so that:

$$\tau = f^{-1}(x). \tag{3.3}$$

The night zone starts in τ and lasts as long as the traffic intensity is below $f(\tau) = x$.

Symmetric traffic pattern (Fig. 3.3a). Consider a traffic pattern symmetric around $T/2$, such that $f(\tau) = f(t - \tau)$ with $\tau \in [0, T/2]$. The duration of the night zone is $T - 2\tau$. Let W denote the power consumption of a cell. Since for a period 2τ the consumption is W and for a period $2(T/2 - \tau)$ it is a fraction $x = f(\tau)$ of the previous one, that is $Wf(\tau)$, the average energy consumed per cell in a day under scheme S can be written as:

$$C(\tau) = 2W\left[\tau + f(\tau)\left(\frac{T}{2} - \tau\right)\right]. \tag{3.4}$$

At this point, calculating the derivative of the power consumption $C(\tau)$ and letting it equal to zero, it is possible to compute the value $\tau_m \in [0, T/2]$ so that $C(\tau_m)$ is minimized:

$$\frac{dC(\tau)}{d\tau} = 2W\left[1 + f'(\tau)\left(\frac{T}{2} - \tau\right) - f(\tau)\right] \tag{3.5}$$

$$f(\tau_m) - f'(\tau_m)\left(\frac{T}{2} - \tau_m\right) - 1 = 0. \tag{3.6}$$

Looking at the figure, the same result can be achieved by maximizing the rectangular white area A, since the power consumption is proportional to the shaded area:

$$A = \left(\frac{T}{2} - \tau\right)\bigl(1 - f(\tau)\bigr) \tag{3.7}$$

$$\frac{dA}{d\tau} = f(\tau) - f'(\tau)\left(\frac{T}{2} - \tau\right) - 1 = 0. \tag{3.8}$$

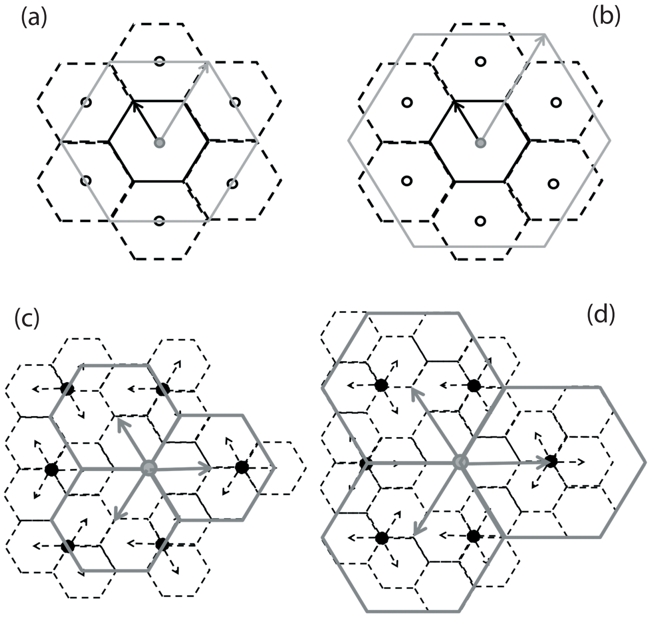
Asymmetric traffic pattern (Fig. 3.3b). Consider now a non-symmetric situation, with τ_1 and τ_2 indicating the two extremes of the night zone and $f(\tau_1) = f(\tau_2)$. In this case, it should be easier to calculate graphically the optimal energy consumption scheme. If $g(\tau_1)$ represents the difference $\tau_2 - \tau_1$, the average energy consumed per cell in a day under scheme S is equal to:

$$C(\tau_1) = W\bigl[T + g(\tau_1) + f(\tau_1)g(\tau_1)\bigr]. \tag{3.9}$$

Again, the derivative of $C(\tau_1)$ put equal to zero gives the value of τ_1 that yields the maximum white area A in the picture:

$$\frac{dC(\tau_1)}{d\tau_1} = W\bigl[-g'(\tau_1) + f'(\tau_1)g(\tau_1) + f(\tau_1)g'(\tau_1)\bigr] = 0. \tag{3.10}$$

It is important to notice that there may be different points that are minimum of the function $C(\tau)$, as to say that there may exist different switch-off schemes corresponding to the same energy saving.



Source: *Optimal energy savings in cellular access networks* (A. Ajmone Marsan, L. Chiaraviglio, D. Ciullo, M. Meo)

Figure 3.4: Hexagonal cells configurations: (a) omnidirectional antennas, 3 cells switched off out of 4; (b) omnidirectional antennas, 6 cells switched off out of 7; (c) tri-sectorial antennas, 3 cells switched off out of 4; (d) tri-sectorial antennas, 8 cells switched off out of 9.

When applying this general rules to real cases, it must be remembered that it is not possible to switch off any fraction of cells, since the access network geometry allow only a few values of x. Consider for example hexagonal cells with the base station located at the center and equipped with omnidirectional antennas. During the night zone, the cells around a working cell are switched off. Two solutions are possible: in the first case, the on cell covers all the six neighboring off cells, so that 6 cells are switched off out of 7 (Fig. 3.4b); in the second case, the on cell covers only half of each neighboring cell, while the rest is covered by another on cell, so that 3 out of 4 cells are switched off (Fig. 3.4a).

If tri-sectorial antennas are used instead of omnidirectional ones, the BS is located at a vertex of a cell. During night, the working cell expands its radius so as to cover the equivalent of 4 or 9 cells. Then, two schemes are now allowed: 3 cells being switched off out of 4 (Fig. 3.4c) or 8 out of 9 (Fig. 3.4d). By comparing these two configurations without refer to a specific traffic profile, it is possible to find regions in which the first scheme is more convenient than the second, and vice-versa. This result indicates that, among the considered options, the best solution is not to switch off the largest possible number of cells; on the contrary, it is important to trade off between the duration of the night zone and the number of the off cells.

3.2.2 Dynamic Base Station Energy Saving

The approach described above gives a predefined BS sleep scheme according to a deterministic traffic variation pattern over time. However, neither the randomness nor the spatial distribution of the traffic is considered. In addition, in [11] it is shown that merely controlling the transmitted power does not allow big energy savings since the energy consumption mainly depends on the on-off states of the BS. Thus, in order to *dynamically minimize the number of active BSs* to meet the traffic variations in the network, the authors take into account the traffic distribution information of the whole network, both in the space and time dimension.

Consider an infrastructured mobile network with a dense deployment of B BSs. The problem is that of minimizing the average energy consumption of the BSs while satisfying the traffic requirement for the users. Let $m(t)$ represent the number of users at time t, and define the binary variable $x_{bi}(t)$ as:

$$x_{bi}(t) = \begin{cases} 1 & \text{if user } i \text{ is associated to BS } b, \\ 0 & \text{otherwise.} \end{cases} \tag{3.11}$$

If the energy consumption of a BS is modeled with two values, $\{0, P_b\}$ respectively when it is in active mode or in standby mode, the optimization problem can be formalized as follows:

$$\min \quad \mathbb{E}\left\{ \sum_{b=1}^{B} P_b \operatorname{sgn}\left(\sum_{i}^{m(t)} x_{bi}(t) \right) \right\} \tag{3.12}$$

$$\text{s.t.} \quad \sum_{b=1}^{B} x_{bi}(t) = 1 \qquad \forall i = 1, \ldots, m(t) \tag{3.13}$$

$$\sum_{i}^{m(t)} \frac{\rho_i(t)}{r_{bi}} \leq W_b^{max} \qquad \forall b = 1, \ldots, B \tag{3.14}$$

$$x_{bi}(t) \in \{0, 1\} \qquad \forall b = 1, \ldots, B, \forall i = 1, \ldots, m(t), \tag{3.15}$$

where $\operatorname{sgn}(x)$ is the sign function, which equals 1 when $x > 0$, otherwise equals 0, and the expectation is taken over time. Constraints (3.13) limit each user to select only one serving BS. Constraints (3.14) correspond to the fact that the bandwidth occupancy of users associated to the BS cannot exceed some limit W_b^{max}, where $\rho_i(t)$ is the rate requirement of user i and r_{bi} is the spectrum efficiency (i.e., rate per unit bandwidth) if user i is served by BS b.

Due to the randomness of $m(t)$ and $\rho_i(t)$, the above model cannot be solved directly. Therefore, by focusing on certain time spots separated by a fixed length period, it is possible to take away the index t and the expectation operator to get a one-shot optimization problem. Still, the problem is very hard to solve because of the binary variable x_{ib} and the sgn function; besides, one must take the traffic variation between consecutive decision times based on the one-shot optimization problem.

Assumption of having only local load information and user specific associate control, a decentralize solution is proposed to tackle the problem. The basic idea of this algorithm is to associate each user to a specific BS by using a BS selection preference function. Specifically, user i selects BS b^* so that to maximize an utility function defined as:

$$b^* = \arg\max_b \quad r_{bi} U(L_b, \alpha_b)/P_b, \tag{3.16}$$

where $L_b = (\sum_i x_{bi}\rho_i/r_{bi})/W_b^{max}$ is the normalized load of BS b, and $\alpha_b \in [0, 1\}$ is a protection margin against outage. The utility function, which can be calculated by each BS, give higher weight to the BSs with relatively high load in order to concentrate traffic to these BSs and let more BSs to standby.

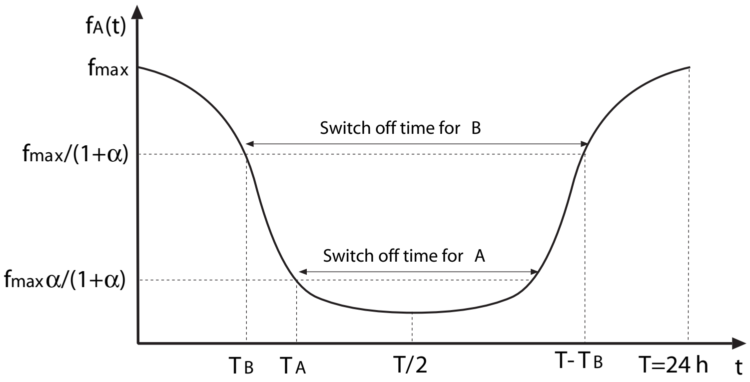
The decentralized algorithm can start with any initial user-BS association state. Then, since the BS selection set of each user is finite, if no two users take action simultaneously, the distributed BS selection will converge to an equilibrium. After the algorithm converges, the BSs with no associated user will enter the standby mode.

3.2.3 The "Network Sharing" Opportunity

Metropolitan areas are normally served by several competing operators, which all provide full coverage of the whole area and dimension networks according to their number of subscribers. In high traffic conditions, the resources of each operator are exploited at capacity and the quality of service is used as a target design objective. However, when traffic is low, the resources of each operator become redundant, and at some point just one of the existing networks would be able to carry all the traffic in the area.

In such a contest, a new and more viable approach is suggested in [23]: in fact, if the operators cooperate, they can switch off their network in turn, at the cost of accepting the competitors' subscribers as roaming customers while their home network is switched off. This implies some increased technical complexity and additional subtle costs, deriving for example by the information gathered by one operator about the behavior of the competitors' customers, but the saved energy can be huge.

In order to analyze in a simple way this scenario, assume an area served by only two operators, A and B, whose access networks are dimensioned according to the traffic demand of their customers but fully cover the area. The daily traffic profiles of the two networks, denoted as $f_A(t)$ and $f_B(t)$, repeats periodically (i.e., the weekend effect on traffic is neglected, which is a conservative assumption as regards energy savings) and the average per-user traffic is the same in both



Source: *Energy efficient management of two cellular access networks* (M. Marsan, M. Meo)

Figure 3.5: Daily traffic profile for network A and possible switch off periods for networks A and B.

networks. If N_A and N_B represent the number of customers in network A and B, as the overall traffic in the networks is proportional to the respective number of users, it holds that:

$$\frac{f_A(t)}{N_A} = \frac{f_B(t)}{N_B} \tag{3.17}$$

and letting $\alpha = N_B/N_A$, it becomes:

$$f_B(t) = \alpha f_A(t). \tag{3.18}$$

Once again, let the traffic profiles be symmetrical with respect to $T/2$, but now let f_{max} identify the maximum traffic level reached in either of the two access networks. Moreover, it is possible to assume without loss of generality that $N_A > N_B$, so that $f_A(t) > f_B(t)$.

As represented in Fig. 3.5, network B can be switched off from time T_B until time $T - T_B$ when:

$$f_A(T_B) + f_B(T_B) = f_{max}. \tag{3.19}$$

Since $f_B(t) = \alpha f_A(t)$:

$$f_A(T_B) = \frac{f_{max}}{1+\alpha} \tag{3.20}$$

so that can be obtained the value of T_B as:

$$T_B = f_A^{-1}\left(\frac{f_{max}}{1+\alpha}\right). \tag{3.21}$$

Similarly, network A can be switched off from time T_A until time $T - T_A$ when:

$$f_A(T_A) + f_B(T_A) = \alpha f_{max} \tag{3.22}$$

as αf_{max} is the maximum traffic managed by network B. Time T_A can be obtained as:

$$T_A = f_A^{-1}\left(\frac{\alpha f_{max}}{1+\alpha}\right). \tag{3.23}$$

It comes from the assumption $N_A > N_B$ that $T_B < T_A$, so the network B, which carries less traffic and requires less capacity, can be turned off for longer periods of time.

At this point, several different switch-off patterns can be taken into account.

Balanced switch-off frequencies

In this case, the two operators alternate in switching off their networks during each 24-hour period. Defining the switch-off frequencies P_A and P_B as the fraction of days in which networks A and B are switched off, if every day one of the two network is turned off:

$$P_A + P_B = 1. \tag{3.24}$$

Balancing the switch-off frequencies implies that $P_A = P_B = 0.5$.

Balanced roaming costs

When one of the two access networks is switched off, its subscribers are accepted as roaming customers by the operator whose access network remains on. Obviously, the traffic generated by roaming customers implies some costs to be charged to the host network. Assume that the amount of roaming traffic carried by network A when network B is turned off is:

$$\int_{T_B}^{T-T_B} f_B(t)dt = \int_{T_B}^{T-T_B} \alpha f_A(t)dt \tag{3.25}$$

and, similarly, the amount of roaming traffic carried by network B when network A is turned off is:

$$\int_{T_A}^{T-T_A} f_A(t)dt. \tag{3.26}$$

If the unit roaming costs are the same for the two operators, in order to balance costs (i.e., it is not necessary that operators pay for the traffic carried by the other operator when the home network is turned off) it has to be:

$$P_A \int_{T_A}^{T-T_A} f_A(t)dt = P_B \int_{T_B}^{T-T_B} f_B(t)dt, \tag{3.27}$$

which leads to:

$$\frac{P_A}{P_B} = \alpha \cdot \frac{\int_{T_B}^{T-T_B} f_A(t)dt}{\int_{T_A}^{T-T_A} f_A(t)dt}. \tag{3.28}$$

Balanced energy savings

The total cost of operating each of the two networks can be estimated as a cost per user C_U (assumed to be the same for the two operators, and computed as the total network capital and operational expenditures divided by the number of subscribers) multiplied by the number of customers, minus the energy saved by switching off the network, represented by the energy cost per unit time C_e multiplied by the switch-off period duration and by the switch-off frequency. These costs are:

$$C_A = C_U N_A - C_e(A)(T - 2T_A)P_A \tag{3.29}$$

$$C_B = C_U N_B - C_e(B)(T - 2T_B)P_B. \tag{3.30}$$

In order to balance energy savings, it is necessary to set:

$$C_e(A)(T - 2T_A)P_A = C_e(B)(T - 2T_B)P_B. \tag{3.31}$$

Under the assumption that every day one of the two network is turned off, it follows that:

$$P_A = \begin{cases} \frac{T-2T_B}{2(T-T_B-2T_A)} & \text{if } C_e(B) = C_e(A) \\ \\ \frac{\alpha(T-2T_B)}{T-2T_A+\alpha(T-2T_B)} & \text{if } C_e(B) = \alpha C_e(A) \end{cases} \tag{3.32}$$

and P_B is obtained as $P_B = 1 - P_A$.

Maximum energy saved

The total energy saved can be expressed as:

$$C_e(A)(T - 2T_A)P_A + C_e(B)(T - 2T_B)P_B. \tag{3.33}$$

Considering $C_e(B) = C_e(A)$, since $T_B < T_A$ the maximum energy savings can be achieved by setting $P_B = 1$, which implies $P_A = 0$, that is to say by turning off the network B every day. In fact, being the unit energy costs equal for the two networks, the maximization of the saved energy requires a longer switch-off duration for the smaller network. On the contrary, if $C_e(B) = \alpha C_e(A)$, the amount that has to be maximized is:

$$(T - 2T_A)P_A + \alpha(T - 2T_B)(1 - P_A), \tag{3.34}$$

that leads to $P_A = 1$ if $(T - 2T_A) > \alpha(T - 2T_B)$, and $P_A = 0$ otherwise.

3.3 Energy Saving in High-Density WLANs

Traditional Wireless Local Area Networks consist of APs that provide simple network connectivity for wireless devices in an area. Each AP is attached to a wired switch on the network and independently executes association, authentication, IP address acquisition and data exchange operations with the WLAN users.

In the last years, the objective of enterprise WLAN deployments was no longer that of just ensuring basic coverage to all areas of the enterprise. Now, large-scale WLANs provide several additional layers of non-interfering APs with overlapping coverage areas, which are grouped into *clusters*: such redundant layers guarantee sufficient capacity for high bandwidth demands and also protect the network against failures and faults.

In order to simplify configuration and organization of this kind of WLANs, vendors have adopted a centralized approach to WLAN management. In centralized WLANs the APs, called *thin APs*, act as simple points of attachment: on the one side they offer wireless connectivity to mobile terminals, while on the other side connect to a wired corporate LAN (normally Ethernet) made of fibers, UTP cables, switches and routers. The thin APs execute time-critical but non-intelligent functions of responding to client probe requests and transmitting periodic beacon frames that provide clients with the AP's capability information. Other more complex functions such as association, authentication, data processing and data acknowledgement are performed at the centralized controller. In this way, APs in centrally controlled WLANs are much simpler and cheaper compared to distributed WLAN APs. Moreover, since their controllers maintain global knowledge about the network state, managing and securing WLANs is a relatively easier task.

The obvious advantage of such redundant WLANs is that the peak capacity they are designed for is sufficient to guarantee bandwidth for most throughput-intensive applications (such as multimedia voice, video streaming, online gaming). On the other hand, peak usage times are rare and sometimes isolated to a portion of the WLANs: as a result, the energy cost of maintaining hundreds of always-on idle APs and wired switches is relevant.

3.3.1 Resource on-Demand WLANs

Since the bandwidth and reliability of all APs in a cluster are required only at peak usage times, powering off some APs in period of low traffic seems to be the best solution. This approach corresponds to what in [16] and [17] is called *Resource on-Demand* (RoD) policy. Thanks to the presence of a single controller that is aware of the global state of the network, RoD WLANs can decide which APs are actually needed to achieve the desired QoS and switch off the other APs, ordering the reassociation of those users that were connected to the APs which are going off.

RoD strategies should follow a set of design requirements:

- *Ensure coverage.* Powering off APs must not create "holes" where users cannot receive services. For this reason, energy-efficient WLANs should ensure that there is at least one AP remains active in all the areas served by the WLAN, to provide service to possible new incoming users.

- *Preserve client performance.* The same QoS should be offered even when part of the infrastructure is off. An expedient consists in avoiding WLAN topologies in which clients are far away from their closest AP or in which an AP is required to supports so many clients that congestion occurs.

- *Maintain redundancy factor.* WLAN redundancy is defined as the excess of APs deployed by WLAN administrators to sustain intense traffic or compensate for equipment failure. In energy-efficient WLANs, it is necessary and sufficient to maintain the same redundancy factor instead of the total redundancy. If an always-on WLAN uses N APs for complete coverage and for sustaining peak bandwidth demand, and $\beta \geq 1$ is the chosen redundancy factor, the total number of APs deployed will be $\beta \times N$; during low traffic periods, if the WLAN controller determines that only $M \leq N$ APs are needed to maintain coverage and support demand, $\beta \times M$ APs can be powered on, such that the same redundancy factor is maintained.

- *Avoid frequent client re-associations.* The turning on and off of APs can cause frequent clients re-associations between APs. These continual disconnections are undesirable because re-association delays can break clients' traffic flows and thereby impact their performance.

- *Guarantee high responsiveness.* Immediate response of the WLAN to client requests is an important feature of both energy-efficient and traditional deployment. An energy-efficient design of a WLAN could lead to delayed responses to users requests for associations; however, already on APs should be able to temporarily accommodate increased demand until new APs are powered on.

Two different classes of operating strategies are described in [17]. In *demand-driven* strategies, the WLAN's central controller periodically collect information from the APs, estimates user demand by the computation of one or more appropriate parameters (such as the number of active users in the network and the volume of offered traffic load) and then determines the best set of APs, switches and routers that will satisfy the estimated user demand. Although this kind of approach allows WLANs to ensure high energy savings and to satisfy end-users performance, the trade-off is in the overhead of assessing user demands and continuously reconfiguring the APs. However, demand-driven strategies are suitable in scenarios where the user demand may vary significantly over time, like in university campuses.

Differently, *schedule-driven* strategies power on and off specific WLAN resources based on pre-determined schedules, which can be derived from WLAN historical usage patterns or can be founded on the administrators' experience. The advantage of using schedules stems from their minimal processing overhead; nevertheless, unlike demand-driven ones, schedule-driven strategies fail in turn on and off the necessary and sufficient set of resources in case of unexpected change in user demand. As a result, this approach is suitable for scenarios where user demand is closely predictable.

Analytical Models for RoD WLANs Assessment

In order to assess the energy efficiency of large RoD WLANs, two simple demand-driven policies are considered in [22]: the *association-based policy*, based on the number of users associated with APs in the cluster, and the *traffic-based policy*, which considers only the users that generate traffic among those which are associated with APs in the cluster.

The evaluation of the system behavior under these two RoD strategies needs a cluster of APs to be modeled as a *Continuous-Time Markov Chain* (CTMC). Let $\{X(t), t \geq 0\}$ denote the CTMC of the associated-based policy, and let $\{Y(t), t \geq 0\}$ denote the CTMC of the traffic-based one. In both cases, the state of the Markov chain is given by vector $\bar{s} = (p, u, c)$, where p is the number of on APs, u is the number of users that are associated with any AP of the cluster and c is the number of associated active users. Also, some assumptions have to be done:

- Users associate with an AP of the cluster according to a Poisson process with rate λ_s and they leave the cluster after an exponentially distributed time with mean $1/\mu_c$.
- Associated users can be *idle*, when they do not generate traffic, or *active*, when they are generating traffic. An idle user becomes active after a time whose probability density function (pdf) is exponential with mean $1/\lambda_c$. The amount of traffic generated to an active user is distributed according to an exponential pdf with mean $1/\mu_c$.
- The AP bandwidth B is fairly shared among all users associated with the AP. Considering long time periods, it is possible to neglect the fact that users at higher distance from the AP might perceive a larger packet loss probability and the fact that the total AP throughput decreases with the number of active users.

- There are A APs in the cluster. No more than M users can be associated with each AP, so that no more than AM users are admitted in the cluster.

Association-based policy

Let $T_h \leq M$ be a threshold. As already said, association-based policy tends to minimize the number of on APs constrained on all active APs having less than T_h user associations. Thanks to the central management of the WLAN, users are evenly distributed among the APs in the cluster: thus, when p APs are on and the number of users is pT_h, T_h users are associated with each AP. As soon as a new user associates with an AP in the cluster, a new AP is turned on so that the number of active APs becomes $p+1$. In order to avoid frequent AP switch-on and switch-off in dynamic scenarios, a hysteresis of amplitude T_l is introduced in the switch-off procedure: in this way, the $(p+1)$-st AP is turned on when the number of users in the cluster exceeds pT_h, but the number of active APs decreases to p again only when the number of users in the cluster becomes $\leq pT_h - T_l$.

The reachable states of the CTMC satisfy the following relations:

$$\begin{cases} 1 \leq p \leq A \\ 0 \leq u \leq MA \\ 0 \leq c \leq u \\ (p-1)T_h - T_l < u \leq pT_h. \end{cases} \tag{3.35}$$

Recalling that the bandwidth is equally divided among the active users associated with an AP, it is possible to say that in state (p, u, c) the average number of connections per AP is c/p and the bandwidth per connection is $B(\bar{s}) = \min(B, Bp/c)$. Under these conditions, the possible CTMC transitions from the generic state (p, u, c) are summarized in Table 3.1, which indicates, for each one, the destination state, the transition rate and the condition over (p, u, c) for the transition to be possible.

Traffic-based policy

Unlike association-based one, the traffic-based policy considers only the users that generate traffic among those which are associated with APs in the cluster. Now, two thresholds C_h and C_l are used. When the number of traffic-generating

Table 3.1: Association-based policy: CTMC transition from state (p, u, c).

Action	Destination	Rate	Condition
		User associates	
AP on	$(p+1, u+1, c)$	λ_s	$u = T_h p \wedge p < A$
AP -	$(p, u+1, c)$	λ_s	$u \neq T_h p \wedge u < AM$
		Connection starts	
AP -	$(p, u, c+1)$	$(u-c)\lambda_c$	$c < u$
		Connection ends	
AP -	$(p, u, c-1)$	$cB(\bar{s})\mu_c$	$c > 0$
		Active user leaves	
AP -	$(p, u-1, c-1)$	$c\mu_s$	$c > 0 \wedge u \neq T_h(p-1) - T_l + 1$
AP off	$(p-1, u-1, c-1)$	$c\mu_s$	$c > 0 \wedge u = T_h(p-1) - T_l + 1 \wedge p > 1$
		Non-active user leaves	
AP -	$(p, u-1, c)$	$(u-c)\mu_s$	$u > 0 \wedge u \neq T_h(p-1) - T_l + 1$
AP off	$(p-1, u-1, c)$	$(u-c)\mu_s$	$u > 0 \wedge u = T_h(p-1) - T_l + 1 \wedge p > 1$

Table 3.2: Traffic-based policy: CTMC transition from state (p, u, c).

Action	Destination	Rate	Condition
		User associates	
AP -	$(p, u+1, c)$	λ_s	$u < AM$
		Connection starts	
AP -	$(p, u, c+1)$	$(u-c)\lambda_c$	$c < u \wedge c \neq pC_h$
AP on	$(p+1, u, c+1)$	$(u-c)\lambda_c$	$c < u \wedge c = pC_h \wedge p < A$
		Connection ends	
AP -	$(p, u, c-1)$	$cB(\bar{s})\mu_c$	$c > 0 \wedge c \neq C_h(p-1) - C_l + 1$
AP off	$(p-1, u, c-1)$	$cB(\bar{s})\mu_c$	$c > 0 \wedge c = C_h(p-1) - C_l + 1 \wedge p > 1$
		Active user leaves	
AP -	$(p, u-1, c-1)$	$u\mu_s$	$c > 0$
		Non-active user leaves	
AP off	$(p-1, u-1, c)$	$(u-c)\mu_s$	$u > 0 \wedge p > 1$

users associated with APs in the cluster is above pC_h, the number of active APs must be $p + 1$; on the contrary, when the number of traffic-generating users in the cluster decreases below $pC_h - C_l$, only p active APs are sufficient to provide service.

The model presented above can be modified in order to fit to the case in which the AP switching is decided by the variable c, that represents the active connections. The reachable states of the CTMC are different with respect to the previous example, and they satisfy the relations:

$$\begin{cases} 1 \leq p \leq A \\ 0 \leq u \leq MA \\ 0 \leq c \leq u \\ (p-1)C_h - C_l < c \leq pC_h. \end{cases} \tag{3.36}$$

Considering this policy, the possible CTMC transitions from the generic state (p, u, c) are reported in Table 3.2.

The two described policies are compared in [22] using a set of parameters based on the steady-state probability of the CTMC models:

- switching rate, defined as the average number of times an AP is switched on (or off) in the time unit;
- average bandwidth per connection, directly related to the QoS perceived by the end users;
- percentage power saving, calculated as the difference between the power consumption of a cluster that always keeps on all its APs (always-on policy) and the power consumption of one of the RoD policies, divided by the always-on policy power consumption and multiplied by 100.

3.3.2 Management Strategies for Energy Saving in WLAN Mesh Networks

As various wireless networks evolve into the next generation to provide better services, *Wireless Mesh Networks* (WMNs) have emerged recently. WMNs are dynamically self-organized and self-configured communication infrastructures, with a high amount of cooperation between many individual wireless stations. Each node in the network (*mesh router* or *mesh client*) operates not only as a host but also as a router, forwarding packets on behalf of other nodes that may not be within direct wireless transmission range of their destinations. Thus, the nodes automatically establish and maintain mesh connectivity among themselves, creating in effect an ad hoc network [2].

This feature brings many advantages to WMNs with respect to other wireless communication technologies. Cellular networks, for example, provide wide area coverage but the service is quite expensive and offers low data rates; on the other hand, traditional Wireless Local Area Networks have limited coverage, reduced mobility and require a wired backbone connecting multiple access points [28]. WMNs have the potential to eliminate many of these drawbacks by offering low-cost, wireless Internet access both for fixed and mobile users, and providing easy network maintenance, robustness and reliable service coverage. Due to their versatility, WMNs can satisfy the needs of different applications: broadband Internet access for those people who do not have the necessary infrastructure (either TV or good quality phone cable), indoor networking, mobile user access outside the coverage of third generation cellular system or WLAN hot spots, connectivity when normal network connectivity is expensive, cumbersome or time consuming [28].

Another recent WMNs implementation consists in providing Wi-fi outdoor coverage: the purpose is to guarantee WLAN coverage to MSs and to move their traffic to and from the wired infrastructure connection at mesh Access Points. In these *WLAN Mesh Networks*, conventional WLAN mesh nodes must be operated using continuous electrical power connections. This requirement may be very expensive and, although power can be supplied through *Power over Ethernet* (PoE), such a solution needs a wired network connection which may not exist.

An alternative is to operate some of the mesh nodes using a sustainable energy source such as solar or wind power: this eliminates the necessity for a fixed power connection, making the nodes truly tetherless and allowing more flexibility in node positioning. In [29] it is shown that WLAN mesh APs which operate using a sustainable energy source could benefit greatly from protocol-based power saving features; however, currently IEEE 802.11 does not provide a mechanism for placing APs in power saving mode and every AP must be available at all time for servicing MSs activity, even if that activity is very low. The authors also discuss some technical challenges associated with achieving AP power saving.

Station QoS

If power saving mode is introduced, it is necessary to use a mechanism which prevents stations from transmitting packets during the periods when the AP is unavailable. Also, the power saving behavior must be consistent with the QoS requirements of the associated stations. From this point of view, two different approaches have been proposed recently. The first one assumes that the AP uses *Network Allocation Vector* (NAV) *blocking* to prohibit channel access to the AP while it is in the sleep mode; unfortunately the NAV blocking method cannot permit reasonable QoS for many types of real-time traffic. In the second proposal, a power saving AP includes a *Network Allocation Map* (NAM) in its beacon broadcasts, which specifies periods of time when the AP is unavailable and assumed to be conserving power. While the allocation of the NAMs and sleep periods is very easy when the traffic supported is connection oriented, more dynamic mechanisms are required for updating NAM durations when supporting non synchronous QoS and for best effort traffic.

Station Power Saving

Obviously, a station in power saving mode that generates a packet for transmission on the uplink should delay its transition to awake mode until some future AP availability period, otherwise the station will consume power unnecessarily. So, when using the NAM scheme previously described, the start and the end of each NAM period should be defined with a *fixed* or *movable boundary*. This technique is used so that end stations that have not seen the most recent AP beacon can know the AP availability period; moreover, the movable boundary allows the AP to dynamically move the limit in order to accommodate traffic conditions.

AP Discovery

Standard scanning procedures are used for the AP discovery. They can be either passive or active. While in *passive scanning* the stations listen on a given channel for a time period sufficient to hear the AP beacon transmissions with a reasonable degree of certainty, in *active scanning* a MS transmits a Probe Request packet on the channel that is being scanned. When this packet is received by one or more APs, they respond with a Probe Response packet. Under ideal conditions an active scan requires a very short time, but clearly the action of a power saving mesh AP can have negative influence on scan latency.

Station Mobility

Mobile Stations move through the WLAN mesh network using standard IEEE 802.11 handoff procedures. Each handoff imposes a temporary disconnection from the network, called *handoff latency*, which should obviously be minimized especially for real time applications. Power saving APs may spend a considerable percentage of time in a sleep mode; for this reason, the AP scanning latency for a MS may be increased beyond levels suitable for real time applications. Various mechanisms have been proposed for reducing latency. These are based on *Infrastructure Activation* (IA): in IA the mesh APs use the quality of their download links to determine which of several power saving modes they should assume. This way, the nodes increase their level of availability when handoff is imminent so that scanning and handoff latency are acceptable.

Chapter 4

The Mathematical Model

4.1 Preliminaries

In order to properly introduce the model in Section 4.2 and 4.3, in this section we explain some preliminary modelling considerations. First, the reasoning behind the joint design and management of energy-aware cellular networks is presented. Then, we show the models and assumptions we used for traffic variations, different types of Base Stations, and channel propagation. However, since traffic values, BSs characteristics and propagation coefficients are input parameters, we can say that our optimization approach is general and not limited to these models. The models are used only for giving concrete examples through numerical results.

4.1.1 Joint Design and Management of Energy-aware Cellular Networks

From an energy saving point of view, a radio coverage obtained using small cells served by low power Base Stations is considered more efficient than adopting macro cells with large radii created by high power Base Stations. This is because, if the cell radius diminishes, the energy consumption usually decreases faster than the growth of the number of BSs required to cover the area, even if differences may be observed in the devices depending on the wireless technology and the components adopted. On the other hand, from a deployment cost standpoint, fixed costs per installation site tend to prevail when powerful Base Stations are chosen.

We claim that, when energy management is considered, the level of flexibility offered by the network topology is also important for adapting the capacity of the cellular system to the varying traffic load, switching on and off some BSs or adapting their emission power. In fact, since not only enough network capacity must be ensured, but also the service area must be fully covered at all time, the availability of a potentially large number of network configurations - consisting in a set of active BSs providing full coverage with different capacity and energy consumption levels - is the key issue that enables energy management strategies in cellular networks. It is just for this reason that a cellular coverage based on small cells only may not be the best option also from the energetic side, since all cells are necessary for guaranteeing full coverage and they cannot be switched off when traffic is low. Differently, a combination of cells with different sizes can potentially offer a flexible coverage topology able to adapt to many traffic scenarios. This is also in line with the need of most operators to mix 2G (like GSM/GPRS) and 3G (like UMTS/HSPA) technologies, and soon also 4G (like LTE) with varying frequencies and coverage ranges.

In order to take into account the energy management when planning the radio coverage of a cellular network, we believe that an optimization approach that jointly considers both the network design based on Capex and Opex costs as well as the power control according to different traffic distributions is absolutely necessary. These two issues are interdependent due to the fundamental role that the management mechanism has in defining the energy cost and hence the Opex.

The problem of planning a wireless access network is approximately that of finding positions and configuration settings for network devices, always matching service requirements stated as given constraints (including budget ones). For different wireless systems, ranging from 2G [24] and 3G [3] cellular networks to Wireless LANs hot spots [5], consolidated models and solution methodologies have already been proposed.

A common approach to the coverage optimization problem resorts to discrete mathematical programming models [8]. In the service area a set of Test Points, representing end users, are identified. TPs can be considered as traffic centroids, where a given amount of traffic (usually expressed in Erlang or bit per second) is requested [30]. Instead of allowing the positioning of BSs anywhere in the area, a set of Candidate Sites where the BSs can be installed is identified. The subset of TPs covered by a sufficiently strong signal is assumed to be known for a BS installed in any CS, since we can evaluate (or even measure in the field) the signal propagation between any pair of TP and CS. The coverage problem results in the classical minimum cost set covering problem [24].

Let S denote the set of CSs where a BS can be installed. For each CS j, $j \in S$, let the set K_j index all the possible configurations of the BS that can be installed in j. Since they may vary with the BS configuration (e.g., its maximum emission power and device type), an installation cost γ_{jk} and a traffic capacity c_{jk} are associated with each pair of CS j and BS configuration k. Let I denote the set of TPs generating a traffic demand d_i. The propagation information can be summarized in the coverage coefficient a_{ijk}, which is equal to 1 if a BS installed in CS j with configuration k can cover TP i, and is equal to 0 otherwise. Such parameters can be deduced from a survey of the site to be planned, or using automatic tools for the prediction of the actual propagation conditions.

Binary decision variables y_{jk} are adopted to define if a BS with configuration k is actually installed in CS j ($y_{jk} = 1$) or not ($y_{jk} = 0$), while assignment variables x_{ijk} define if a TP i is associated to a BS located in CS j ($x_{ij} = 1$) or not ($x_{ij} = 0$).

A frequently adopted formulation of the radio planning problem with coverage constraints is:

$$
\begin{aligned}
\min \quad & \sum_{j\in S}\sum_{k\in K_j} \gamma_{jk} y_{jk} && && (4.1)\\
\text{s.t.} \quad & \sum_{j\in S} x_{ij} = 1 && \forall i \in I && (4.2)\\
& \sum_{k\in K_j} y_{jk} \le 1 && \forall j \in S && (4.3)\\
& x_{ij} \le \sum_{k\in K(j)} a_{ijk} y_{jk} && \forall i \in I, j \in S && (4.4)\\
& \sum_{i\in I} x_{ij} d_i \le \sum_{k\in K_j} y_{jk} c_{jk} && \forall j \in S && (4.5)\\
& y_{jk} \in \{0,1\} && \forall j \in S, k \in K_j && (4.6)\\
& x_{ij} \in \{0,1\} && \forall i \in I, j \in S. && (4.7)
\end{aligned}
$$

Objective function (4.1) aims at minimizing total installation costs. Constraints (4.2) ensure that all TPs are assigned to one BS, while constraints (4.3) that one configuration at most is selected in CS j. Constraints (4.4), stating that a TP i can be assigned to a CS j with configuration k only if it is covered by an actually installed BS, and capacity constraints (4.5), which limit the traffic assigned to installed BSs, are obviously the crucial ones.

Starting from this radio planning formulation, we introduce the following innovative features that lead to an energy-aware design and management model:

- We include in the objective function not only Base Stations installation costs (Capex) but also operational costs (Opex), assuming that their variable part is mainly due to the energy costs (which is largely confirmed by data available from mobile operators worldwide);
- We redefine variables and constraints to include in the model the energy management mechanism that, based on a set of traffic values related to different time periods of the day, can turn on and off Base Stations and modify their emission power;
- In order to define the relative importance of Capex and Opex in the optimization process, we use a tradeoff parameter which multiplies the objective function component of the operational cost;
- We jointly consider radio planning and energy management in order to obtain the network design that minimize deployment and operational costs.

We show that, for realistic size networks, the proposed model can be solved to the optimum and that it allows to get an interesting insight on the coverage and topology characteristics of energy-efficient wireless networks. The mathematical model is presented in Section 4.3.

4.1.2 Traffic Variation Behavior

Intuitively it can be said that traffic intensity varies as a natural effect of users'daily habits. For example, it has been measured that mobile traffic presents its peak between noon and 4 pm and that there is a significant decrease in the late evenings. Moreover, in a typical business area, the traffic pattern is almost the same from Monday to Friday but it decreases during the weekend [14].

In order to account for the main fluctuations, but neglecting the differences that occur between working and weekend days, we consider an approximated daily

Table 4.1: Time periods during one day.

Index	Starting time $o(t)$	Ending time $e(t)$	Duration $h(t)$	Normalized traffic p_t
1	09:00	11:00	2 h	0.9
2	11:00	13:00	2 h	1
3	13:00	15:00	2 h	0.7
4	15:00	18:00	3 h	0.8
5	18:00	23:00	5 h	0.55
6	23:00	09:00	10 h	0.2

traffic pattern based on the measurements presented in [14] and [19]. According to this profile, the whole day is split in time periods, each one gathering smaller intervals (hours) in which the users behavior can be assumed unchanged.

Let T be the ordered set of the different time periods. Let $h(t)$ represent the length of period $t \in T$. Each period is expressed in hours and it lasts

$$h(t) = e(t) - o(t), \tag{4.8}$$

where $e(t)$ and $o(t)$ are, respectively, the end and the beginning times of period $t \in T$. Note that the end of the previous time period $e(t-1)$ is equal to the beginning of the new one $o(t)$, so that there is no time gap between adjacent periods and the summed duration of all periods is equal to the number of hours in a day. Here we have assumed a total of six time periods, which features are displayed in Table 4.1 .

Observing Figure 4.1, the progress of the approximated traffic profile defines active users percentages in every time instance. So, in order to comply with the outlined pattern, our traffic distribution is modelled as follows.

Let us recall the previously defined set of traffic aggregation points, called Test Points (TPs). We assumed that only a group of TPs can provide real traffic, while the others are only used in the network dimensioning phase. For each traffic centroid, we calculate a random traffic value ranging from 0 to 10 Erlangs. This value represents the maximum traffic that the TP can generate. Then, depending

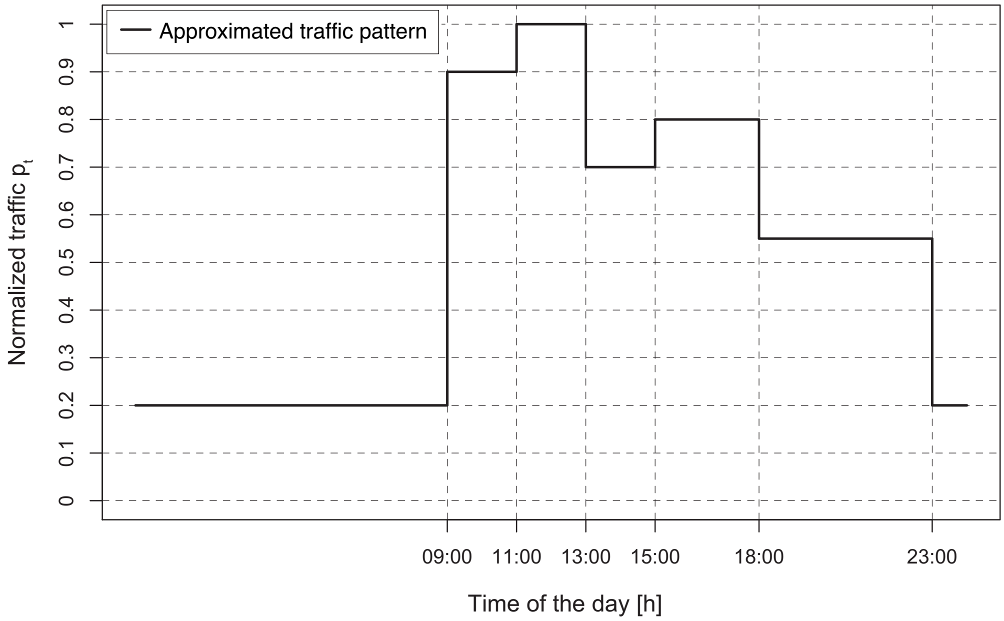


Figure 4.1: Approximated traffic profile of analyzed mobile network.

on the time period, the maximum traffic value is multiplied by the normalized traffic p_t (that is to say, the active users percentage) typical of every instance.

4.1.3 Base Station Categories

The discussion about which Base Station size deployment can offer much economic and energy saving is a common topic in literature [10, 18]. Each BS category has different performances in terms of cell range and reliability and consequently different impact on Capital and Operational and Management Expenditures. In light of that, the total cost structure of a radio access network is closely related to the quantity and types of different access points employed to obtain the required network capacity and coverage.

With respect to energy expenditures, it is often believed that efficiency improvements can be achieved through high density deployments of small and low power BSs, compared to network topologies featuring low density deployment of high power BSs. In fact, although micro sites cover smaller areas, these generally provide much higher *Signal to Interference and Noise Ratios* (SINR) due to shorter propagation distances; besides, accordingly to the coverage area sizes, micro BSs have lower energy consumption [10].

From an economic standpoint, small BSs imply a low cost for equipment, site lease and installation, while macro BSs are clearly more expensive because of the higher output power, capacity and required reliability. On the other hand, the cost per user can be lower in case of bigger BSs due to fixed costs not directly related to the capacity of the BS, which can be divided between many users.

In this work we allowed the joint low and high power Base Stations deployment, which can change in time in order to comply to traffic variations: thus, we tried to achieve the complete coverage of the interested area and limit the energy (and consequently, economic) waste due to low traffic periods.

Four Base Stations types are employed, each one transmitting at a different maximum power. Considering the values provided by GSM standard [9] as reference, we selected the following configurations:

- *Macro BS, type 1.* Emitted power of 55 dBm, corresponding to a maximum coverage radius of about 26200 m;
- *Macro BS, type 2.* Emitted power of 43 dBm, corresponding to a maximum coverage radius of about 9500 m;
- *Micro BS.* Emitted power of 20 dBm, corresponding to a maximum coverage radius of about 1300 m;
- *Pico BS.* Emitted power of 15 dBm, corresponding to a maximum coverage radius of about 825 m.

In order to limit unnecessary power expenditures, we introduced the possibility of using only a percentage (70%, 60% or 50%) of the maximum transmitted power allowed for the specific category. In our model we call *configurations* the different Base Stations types and *power levels* the percentages described above. Moreover, every BS can be switched off and enter in the stand-by mode in case of very low traffic profile.

As displayed in Table 4.2, realistic values of consumed power and capacity are derived for every couple of BS configuration-power level. The former, measured in decibels, takes into account the mean power consumption of the equipment that occurs per sector, such as power amplifier, transceiver, signal generator and

Table 4.2: Transmission and consumption features of every couple of BS configuration - power level.

Config.	Config. cost (Euro)	Power level	Transmit power (dBm)	Consumed power (dB)	Capacity (Erl)	Max coverage distance (m)
C1	58000	P1 - 100%	55	35	45	26146.13
		P2 - 70%	53.5	34.2	45	22910.59
		P3 - 60%	52.8	33.6	45	21639.21
		P4 - 50%	52	32.8	45	20226.22
		P5 - 0%	0	3	0	0
C2	40000	P1- 100%	43	31.2	26	9396.43
		P2 - 70%	41.5	30.7	26	8233.64
		P3 - 60%	40.8	30.1	26	7776.73
		P4 - 50%	40.0	28.8	26	7268.93
		P5 - 0%	0	3	0	0
C3	22000	P1- 100%	20	27.0	14	1321.63
		P2 - 70%	18.5	26.9	14	1158.08
		P3 - 60%	17.8	26.2	14	1093.82
		P4 - 50%	17.0	25.7	14	1022.39
		P5 - 0%	0	3	0	0
C4	15000	P1- 100%	15	22.5	9	862.83
		P2 - 70%	13.5	22.0	9	756.06
		P3 - 60%	12.8	21.3	9	714.10
		P4 - 50%	12.0	20.5	9	667.47
		P5 - 0%	0	3	0	0

AC-DC converter, and the equipment that occurs only once, such as the air conditioning and the microwave link (responsible for communications with the backhaul network). The latter, expressed in Erlangs and computed by setting the desired blocking probability to 0.02, quantifies the maximum traffic volume which can be addressed by a single BS.

It is worth pointing out that the design approach we proposed is general and can be used with any mix of BS types and technologies. The example values above have been used only to obtain numerical results in realistic scenarios.

4.1.4 The Propagation Model

In real scenarios, deterioration of transmitted signal quality is commonly assumed to be due to three different causes: path loss, slow fading (also named shadowing) and fast fading. However, for the sake of simplicity, we concentrate on the effect of path loss and shadowing.

A commonly adopted model indicates that the mean path loss increases exponentially with distance [6, 10, 27]:

$$\overline{PL}(d) \propto \left(\frac{d}{d_0}\right)^n, \tag{4.9}$$

where $\overline{PL}(d)$ is the mean path loss, n is the path loss exponent which indicates how fast path loss increases with distance, d_0 is a close-in reference distance and d is the transmitter-receiver separation distance. Average power path loss is defined as the power loss from the transmitter to the reference distance d_0

$$\overline{PL}(d_0)[dB] = 10\log_{10} P_t - 10log_{10}P(d_0) \tag{4.10}$$

plus the additional path loss described by (4.9) in decibels:

$$\overline{PL}(d)[dB] = \overline{PL}(d_0)[dB] + 10n\log_{10}\left(\frac{d}{d_0}\right), \tag{4.11}$$

where P_t is the power of the transmitted signal. For our model a 1 m reference distance was chosen, and we assume $\overline{PL}(d_0)$ is due to free space propagation from the transmitter to that distance. According to [27], considering antenna gains

Table 4.3: Path loss model values.

Path loss model values	
λ = 0.33 m	f = 900 MHz
d_0 = 1 m	n = 2.7
σ = 13 dB	χ_σ = 6.23 dB
P_{rth} = -102 dBm	$\overline{PL}(d_0)$ = 31.5 dB

equal to system cable losses, leads to 31.5 dB path loss at 900 MHz over a 1 m free space path.

Experimental measurements indicate that path loss is log-normally distributed about (4.11). Hence, path loss at a separation of d meters is better described by

$$PL(d)[dB] = \overline{PL}(d)[dB] + \chi_\sigma[dB], \tag{4.12}$$

where χ_σ is a zero mean log-normally distributed random variable with standard deviation σ, which can be interpreted as representing the variability of power loss of transmitter-receiver configurations located at different places. Assuming that the distribution of large-scale path loss is log-normal for our data, we need to get the path loss exponent n and standard deviation σ as a function of the general surroundings. In [10] a linear regression was used to compute these values in a minimum mean square error sense for some measured data: from those results, we found that for our problem $n = 2.7$ and $\sigma = 13$ dB are the correct values. Once defined the path loss model, the received power at the distance d is calculated as

$$P_r(d)[dB] = P_t[dB] - PL(d)[dB]. \tag{4.13}$$

In order to verify wether or not a receiver is covered by a BS installed in a Candidate Site (CS), we need to know the power level required at the receiver that allows the connection between the two terminals. Chosen the most common GSM 900, constant receiver sensitivity level (P_{rth}) of -102 dBm is assumed for all mobile stations, according to ETSI GSM Technical Specification [9].

In Table 4.3, all the values used in the path loss model are gathered.

4.2 Notational Description

Having described the modeling philosophy as well as the physical details of the system we want to optimize, we need an additional notation to be able to set the mathematical model. For the sake of completeness, some of the notation that was first presented in Section 4.1 is also included here.

The notational presentation is divided in model parameters and decision variables:

Model Parameters

I_c : Set of TPs that need to be covered. We assume that they do not generate any traffic. This first subset of Test Points (*Coverage* Test Points) helps us provide a basic, fixed coverage of the network even in case of very low traffic profile.

I_t : Set of TPs that need to be covered and that generate variable traffic. This second subset of Test Points (*Traffic* Test Points) allows the network to "follow" the traffic variations in different time periods.

S : Set of the available Candidate Sites for the Base Stations.

K_j : Set of possible configurations for a BS located in site j.

L : Set of available power levels for a BS installed with configuration k.

d_{it} : Traffic provided by the Traffic TP i in period t.

c_{jkl} : Capacity of the BS located in site j with configuration k and power level l.

γ_{jk} : Installation cost for a BS located in site j with configuration k. This is composed of two parts: the cost ξ_j due to the characteristics of the chosen site (for example, open spaces or buildings) and the cost τ_k specific for the selected configuration.

ϵ_{jkl} : Power consumption for a BS located in site j with configuration k and power level l.

r_{ij} : Distance between the TP i and the BS located in site j.

β, ϑ : Weight parameters that will be used for trade-off in the objective function.

Finally, to conclude the model parameters, we need to introduce a binary one that summarizes the coverage information for each combination of Coverage or Traffic TP i, CS j, configuration k and power level l:

$$a_{ijkl} = \begin{cases} 1 & \text{if TP } i \in I_c \cup I_t \text{ is covered by a BS installed in } j \\ & \text{with configuration } k \text{ and power level } l, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{4.14}$$

Decision Variables

The problem of finding the subset of powered Base Stations with minimal cost that ensures coverage and capacity demand of all Test Points can be formulated using three decision variables. The first are those selecting which subset of CSs is chosen in any time interval:

$$z_{jk} = \begin{cases} 1 & \text{if a BS is installed in site } j \text{ with configuration } k, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{4.15}$$

Location and transmission characteristics of any installed BS are defined by the second binary variables:

$$y_{jklt} = \begin{cases} 1 & \text{if a BS installed in site } j \text{ with configuration } k \\ & \text{has power level } l \text{ in period } t, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{4.16}$$

Binary decision variables are also used to explicitly indicate the assignment of Traffic TPs to the active BSs:

$$x_{ijt} = \begin{cases} 1 & \text{if TP } i \in I_t \text{ is assigned to a BS} \\ & \text{installed in site } j \text{ in period } t, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{4.17}$$

4.3 Joint Design and Management Energy Aware Model

Having defined all the parameters and variables, we can now describe the modelling objective and constraints.

The Objective Function

$$\begin{aligned} \min \quad & \sum_{j \in S} \sum_{k \in K_j} z_{jk}\gamma_{jk} + \beta \sum_{j \in S} \sum_{k \in K_j} \sum_{l \in L} \sum_{t \in T} \epsilon_{jkl} \cdot h(t) \cdot y_{jklt} + \\ & + \vartheta \sum_{i \in I_t} \sum_{j \in S} \sum_{t \in T} x_{ijt} \cdot h(t) \cdot r_{ij} \end{aligned} \tag{4.18}$$

The objective (4.18) is composed of three major terms: the installation costs (Capex), the price of the operational and maintenance expenditures (Opex) of the Base Stations installed in the area and a final term, introduced to guarantee a better connection quality between users and antennas, that induces the model to try to assign every Test Point to the nearest available Base Station. Note that the model considers that there is a trade off between the three terms that is adjusted by playing with the values of parameters β and ϑ.

Coverage Constraints

$$\sum_{j \in S} \sum_{k \in K_j} \sum_{l \in L} a_{ijkl} y_{jklt} \geq 1 \qquad \forall i \in I_c \cup I_t, t \in T \tag{4.19}$$

$$x_{ijt} \leq \sum_{k \in K_j} \sum_{l \in L} a_{ijkl} y_{jklt} \qquad \forall i \in I_t, j \in S, t \in T \tag{4.20}$$

The two constraints above represent two different types of coverage constraints. Constraints (4.19) provide a minimal and constant (over all scenarios) coverage by ensuring that *all* the TPs are within the service area of at least one installed BS. On the other hand, (4.20) insure that *Traffic* TPs are only assigned to the BS they are covered by.

Capacity Constraints

$$\sum_{i \in I_t} x_{ijt} d_{it} \leq \sum_{k \in K_j} \sum_{l \in L} c_{jkl} y_{jklt} \qquad \forall j \in S, t \in T \tag{4.21}$$

The capacity constraints (4.21) insure that each active BS can satisfy the traffic demand of the covered Traffic TPs.

Assignment Constraints

$$\sum_{j \in S} x_{ijt} = 1 \qquad \forall i \in I_t, t \in T \tag{4.22}$$

The assignment constraints given by (4.22) impose that every traffic TP is assigned to only one Base Station.

Linking Constraints

$$\sum_{l \in L} y_{jklt} = z_{jk} \qquad \forall j \in S, k \in K_j, t \in T \tag{4.23}$$

Constraints (4.23) above are linking constraints between variables y and z.

Configuration Constraints

$$\sum_{k \in K_j} z_{jk} \leq 1 \qquad \forall j \in S \tag{4.24}$$

Configuration constraints (4.24) impose that at most one configuration is chosen for every CS.

Binary Constraints

$$z_{jk} \in \{0,1\} \qquad \forall j \in S, k \in K_j \tag{4.25}$$

$$x_{ijt} \in \{0,1\} \qquad \forall i \in I_t, j \in S, t \in T \tag{4.26}$$

$$y_{jklt} \in \{0,1\} \qquad \forall j \in S, k \in K_j, l \in L, t \in T \tag{4.27}$$

Finally, (4.25), (4.27) and (4.26) impose the binary values for the decision variables.

Chapter 5

Optimization Results

5.1 Resolution Approach

The mathematical model we propose has been programmed using the AMPL programming language and CPLEX solver. AMPL requires two different input files. The first one has the *.mod* extension and contains the declaration of all the parameters and variables, the description of the objective function and the list of constraints the problem is subject to; the second file, with the extension *.dat*, includes the values of all the parameters declared in the *.mod*.

In order to test the effectiveness of the model, we needed to generate realistic cellular network instances (that is to say, a truthful *.dat* file) so that the numbers of CSs and TPs are similar to the ones that can be found in real networks. Therefore, an *Instance Generator* (IG) was designed, implemented in C++ and used as an input for the CPLEX solver. The main features of the IG are explained in Section 5.1.1.

Once obtained an optimization result from CPLEX, in order to better understand it we also needed a tool which allowed us to display the solution. We choose R, a free software for statistical computing and graphics able to produce vector plots. To provide R with the positions and the characterization of CSs and TPs, in addition to the *.dat* file required by AMPL we let the IG generate some other text files containing the cartesian coordinates of all the network devices and the coverage radius typical of every couple configuration - power level. Then, to let R know which CSs, configurations and power levels have been chosen by CPLEX, we realized another C++ program fully discussed in Section 5.1.3.

5.1.1 Instance Generator

Instance Generator Input

The following is a list of the parameters that are used as an input to start the generation of instances. The Instance Generator takes the following entry parameters:

- dimensions of the covered area (length × height),
- number of CSs $|S|$,
- number of Coverage TPs $|I_c|$,
- number of Traffic TPs $|I_t|$,
- range of random cost for CSs,
- range of random traffic for Traffic TPs,
- number and values of possible configurations for the BSs,
- cost of each configuration,
- number and values of possible power levels,
- percentage of transmitted power for each power level,
- consumed power for each couple configuration - power level,
- minimum power needed for a TP to be covered by a BS (receiver sensitivity),
- BS capacity for each couple configuration - power level,
- number of time periods, with starting and ending time,
- percentages corresponding to the daily traffic profile,
- value of the trade-off parameters β and ϑ.

Now that we have stated all the elements that must be taken into account to produce the instance generator, we can present the algorithm.

Instance Generator Algorithm

- *Reading and assignment*
  - Read all the input parameters
  - Create coordinates for all Candidate Sites and Test Points
  - Assign, within the chosen range, a random cost parameter to each CS
  - Assign, within the chosen range, a random traffic value to each Traffic TP (corresponding to the maximum emitted traffic for that TP)

- *Computation of values*
  - *For every Test Point*
    - *Coverage or Traffic TP :* Verify if it is in the coverage range of al least one Candidate Site (if not, new random coordinates are computed for all CSs)
    - *Traffic TP :* Compute the traffic emitted in each time period by multiplying the maximum traffic value by p_t
  - *For every pair of Test Point and Candidate Site (TP, CS)*
    - Compute the mutual distance and the related channel attenuation based on the relation (4.12)
  - *For every possible configuration and power level*
    - Compute the value of the effective transmitted power and the corresponding coverage radius
    - Determine wether the pair (TP, CS) is feasible depending on the value of the received power (4.13).

Thus, the outcome of the Instance Generator is then a feasible set of Traffic and Coverage TPs as well as Candidate Sites that guarantee the feasibility of the instance. In other words, the instance algorithm allows us to find the appropriate a_{ijkl} defined in (4.14).

5.1.2 Test Scenarios

In order to assess the performance of the proposed approach, we considered different test scenarios generated using the IG. The scenarios features are described in Table 5.1. For each scenario, the first entry represents the area size (expressed in squared meters), the second entry is the number of Candidate Sites and the next are the numbers of the Coverage and Traffic Test Points placed in the area; finally, the Table presents the values of the configurations and power levels allowed in each scenario.

To guarantee a homogeneous basic wireless coverage for the whole area, the Coverage Test Points are arranged along a regular grid; instead, the Traffic Test Points and the Candidate Sites are randomly spread on the considered area. Figures 5.1, 5.2 and 5.3, prepared by using the R software, represent the disposition

Table 5.1: Parameters used for generating the test scenarios.

	Scenario nr.1	Scenario nr.2	Scenario nr.3
Area Size (m^2)	5000×5000	10000×10000	20000×20000
CSs Number	20	80	150
Coverage TPs Number	36	121	1681
Traffic TPs Number	30	100	180
Configurations Number	3	3	4
Chosen Configurations	C2, C3, C4	C2, C3, C4	C1, C2, C3, C4
Power Levels Number	4	4	5
Chosen Power Levels	P1, P2, P4, P5	P1, P2, P4, P5	P1, P2, P3, P4, P5

of Candidate Sites and Test Points in scenario nr.1, scenario nr.2 and scenario nr.3 respectively: the triangles stand for the Candidate Sites, the white circles symbolize the Coverage Test Points while the black ones denote the Traffic Test Points.

5.1.3 Solutions Display

In many cases, especially when handling huge instances, the understanding of CPLEX optimizations may not be so easy and immediate. So as to help their comprehension, we wrote a few lines of code using C++ which enable R to display CPLEX solutions. The program takes as input the following parameters:

- number of possible configurations for the BSs $|K_j|$,
- number of possible power levels $|L|$,
- values of the coverage radius of every pair configuration - power level, computed by the IG,
- CSs coordinates, listed by the IG,
- Coverage and Traffic TPs coordinates, also listed by the IG,
- assignment variables x set to 1, collected in a CPLEX output file,
- installed BSs characterization variables y set to 1, also collected in a CPLEX output file.

Using these elements, the program picks out the coordinates of the CSs where a BS is installed and identifies which BSs each Traffic TP is assigned to. Then, it creates two types of text files for each time period:

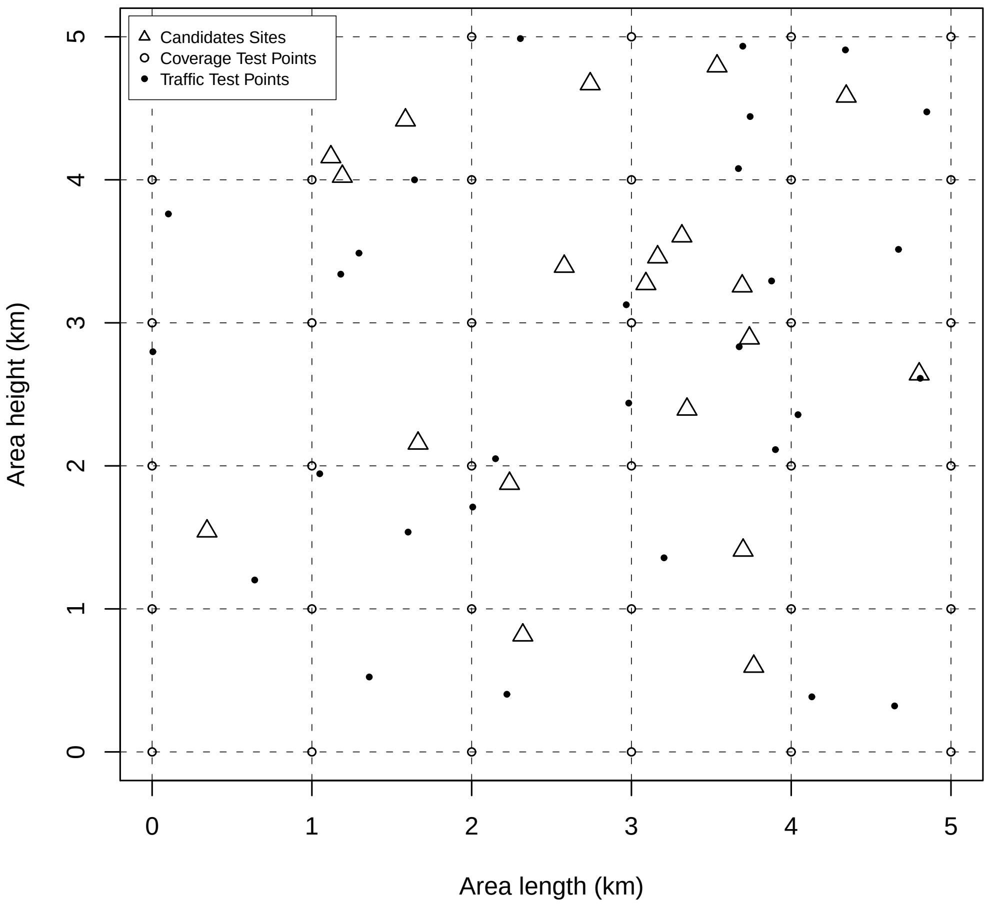


Figure 5.1: Test scenario nr.1.

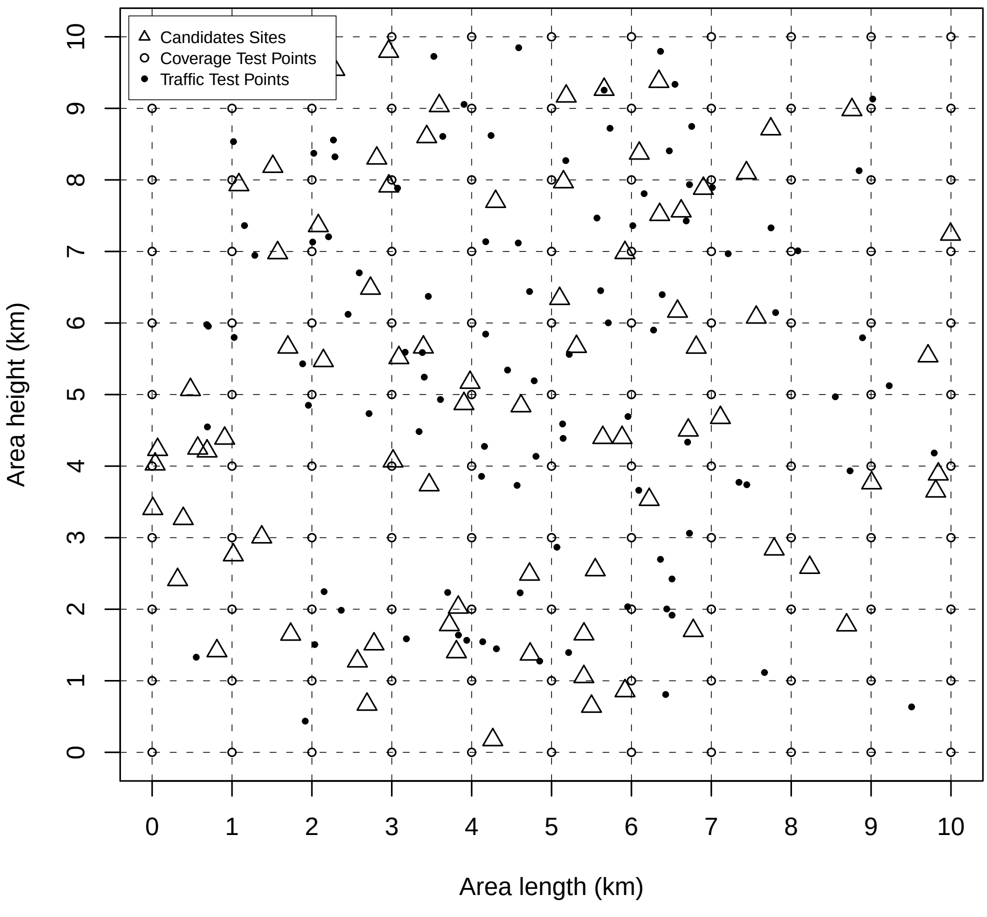


Figure 5.2: Test scenario nr.2.

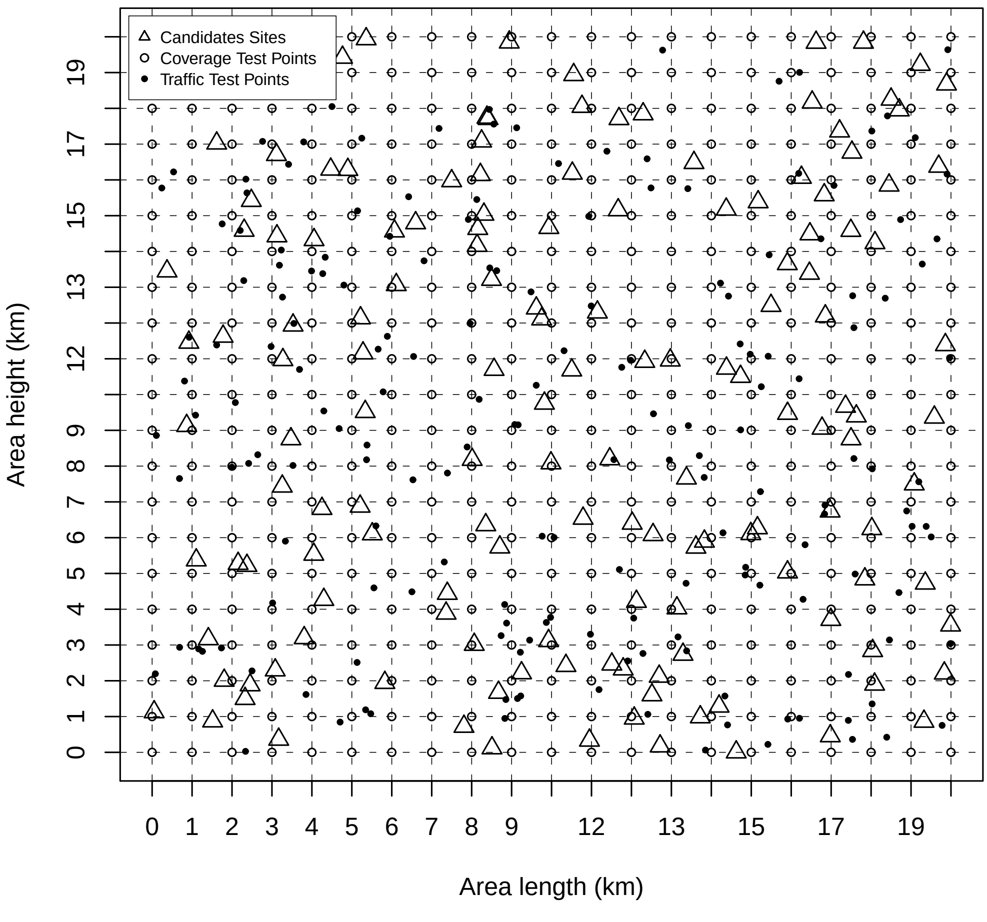


Figure 5.3: Test scenario nr.3.

- *Chosen Candidate Sites files*, which list the cartesian coordinates of the installed Base Stations only,
- *Assignment files*, which list the cartesian coordinates of all the Traffic Test Points and of the Base Stations they are assigned to.

At this point, R has all it needs to generate the required representations. By giving R the following entry text files:

- CSs coordinates file,
- Coverage and Traffic TPs coordinates file,
- Chosen CSs file,
- Assignment file,

we obtained six different displays for every scenario, one for each time period. In these pictures both selected and discarded CSs are depicted, the first symbolized by black triangles and the latter by white ones; black triangles are differently sized in order to give an idea of the type of Base Station (i.e., the configuration) installed in a selected site. The coverage area of every deployed Base Station is delimited by dotted circumferences, whose radius depends on the BS power level; however, when a powerful configuration is chosen, the coverage area may be greater than the analyzed one and the circumference may not be visible in the picture. Moreover, Traffic TPs are linked to the assigned BSs by dotted segments.

5.2 Experimental Results

The optimization of energy consumption, installation and operational costs was performed using the mathematical model presented in Section 4.3. For every scenario, we tested different values of the weight parameters β and ϑ; after some tests, we resolved to set the ratio β/ϑ to 10000 or 1000 in order to obtain well balanced results. By doing so, we strove to highlight the benefits achieved by jointly minimizing costs and power expenditures in the design and management phases, instead of limiting the optimization at the network planning stage.

Differences in Network Design

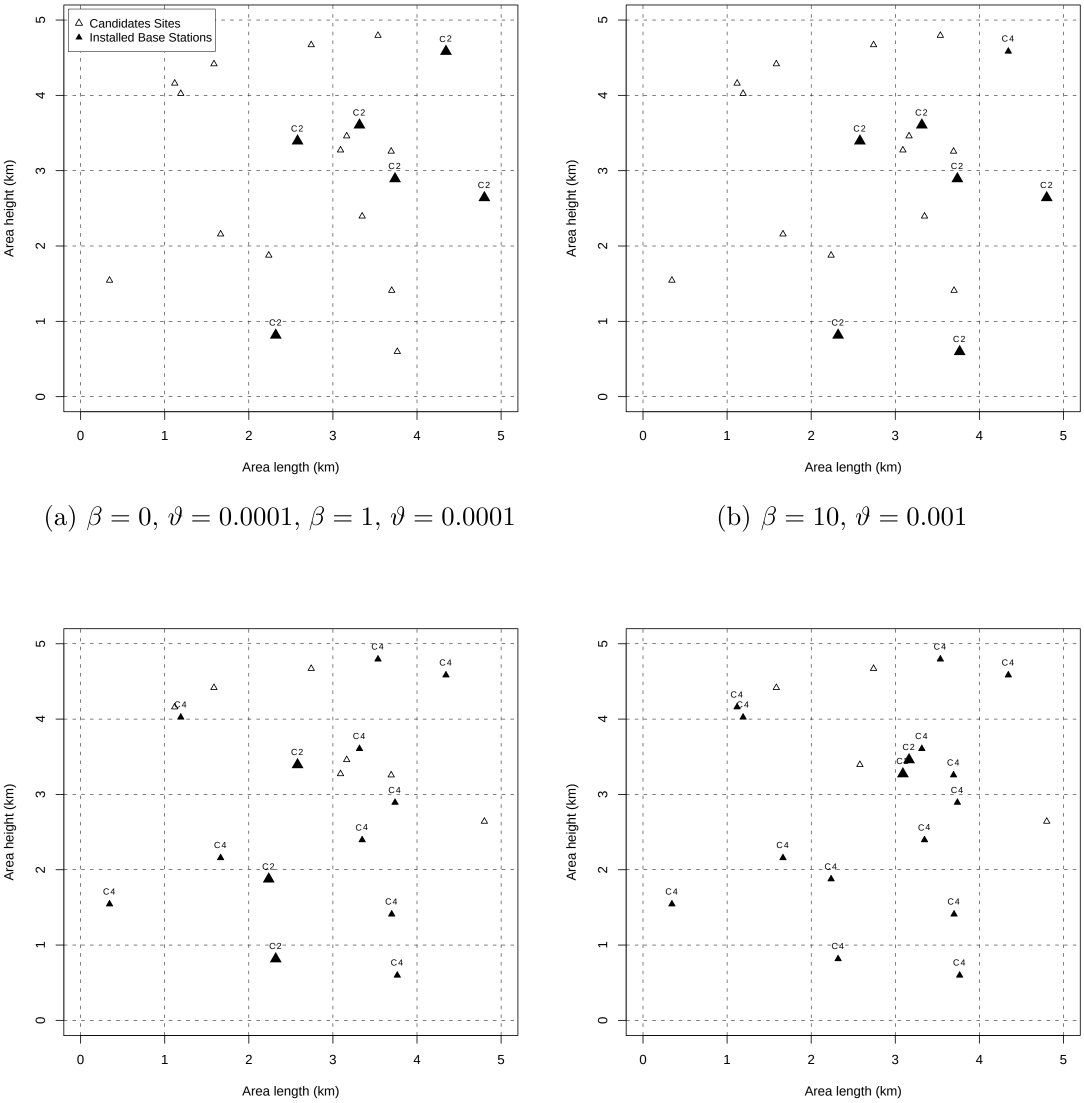
First of all, some brief observation about the influence of our approach on the mere network design can be done. We use examples by test scenario nr.1, but the same reflections apply for the other two scenarios.

In Figure 5.4, some BSs deployment obtained using different values of β and ϑ are represented. To symbolize the installed Base Stations we used black triangles , having different size depending on the chosen configuration; this last is also indicated above each triangle. When only Capex costs are considered (that is to say $\beta = 0$, Figure 5.4a), the model tries to minimize the installation costs by selecting the minimum number of Base Stations that is able to satisfy the traffic needs of the whole area. Obviously, since operational costs are neglected, each installed BS is of type C2 (the powerful one, in scenario nr.1). This is because, as reported in Section 4.1.3, covering the area with big cells is more convenient from an installation cost point of view; in our experiments, for example, the costs per covered square kilometer are 0.16, 4.14, and 7.46 for BS types C2, C3, C4, respectively. We can notice the same behavior when $\beta = 1$; in this case, even if we take into account the energetic costs by letting $\beta > 0$, the Capex component in the objective function is still leading and the previous design turns out to be still the best.

The network design changes when increasing the importance of the Opex part of the objective function. The first slight topology adjustment is displayed in Figure 5.4b, where β is equal to 10. While one more type C2 BS is powered on, we note that another one is turned into a type C4. Figure 5.4c, in which the value of β is raised to 100, clearly shows that a widespread BSs deployment is now the preferred option. As we can see, the optimization process selects a higher number of Candidate Sites but only a small part of the installed BSs is of type C2; by choosing the less powerful configuration C4 for the greater part of the BSs, the model tends to reach the topology which implies the minimum energy expenses.

Again, as we expected, the increase of β to 100000 has the effect of incrementing the number of deployed BSs while diminishing the ones of type C2.

Table 5.2 gathers the most important characteristic values of the topologies described for scenario nr.1. In addition to the number of installed BSs and the percentage of use of each configuration, for every couple of β and ϑ the table shows the minimum and mean power received by a TP and the maximum and



(a) $\beta = 0$, $\vartheta = 0.0001$, $\beta = 1$, $\vartheta = 0.0001$

(b) $\beta = 10$, $\vartheta = 0.001$

(c) $\beta = 100$, $\vartheta = 0.01$

(d) $\beta = 100000$, $\vartheta = 100$

Figure 5.4: Scenario nr.1: network design for different values of β.

Table 5.2: Scenario nr.1: important values in different network topologies.

	$\beta = 0$ $\vartheta = 0.0001$	$\beta = 1$ $\vartheta = 0.0001$	$\beta = 10$ $\vartheta = 0.001$	$\beta = 100$ $\vartheta = 0.01$	$\beta = 100000$ $\vartheta = 100$
Ratio β to ϑ	-	10000	10000	10000	1000
Installed BSs	6	6	7	13	16
Chosen Configurations	C2 - 100% C3 - 0% C4 - 0%	C2 - 100% C3 - 0% C4 - 0%	C2 - 86% C3 - 0% C4 - 14%	C2 - 18% C3 - 0% C4 - 82%	C2 - 13% C3 - 0% C4 - 87%
Mean Received Power (dBm)	-75.85	-79.82	-83.29	-89.56	-91.76
Min Received Power (dBm)	-89.34	-91.14	-98.91	-101.39	-101.77
Mean Distance BS - TP (m)	1143.3	1404.8	1505.5	1435.5	1370.4
Max Distance BS - TP (m)	2670.8	2886.5	3185.9	3294.3	3375.7

mean distance between a TP and an installed BS. The presented mean values refer to the entire day and they are computed by averaging out the mean values typical of every time period, each one weighted using the corresponding period duration $h(t)$, $t \in T$. Similarly, the maximum distance and the minimum received power are obtained by the weighted averages of the maximum distances and the minimum received powers in every time period.

Looking at the table, all distance and power values (with only one exception) grow worse when increasing the weight of the Opex costs. Although a greater number of CSs is selected, due to the high value of β the energy management is pushed to turn off many Base Stations during time periods when the traffic is low, that is to say during nocturnal hours (t_6). So, in good part of the day ($h(t_6) = 10$ hours, as reported in Table 4.1) the Test Points are served by few Base Stations having powerful configurations (in case of scenario nr.1, C2) but located far from the most of them. For this same reason, if β raises, TPs receive generally lower power from the switched on BSs.

Tables 5.3 and 5.4 present the main characteristics of the topologies found for scenarios nr.2 and nr.3 respectively. Even if the behavior of the displayed values cannot be simply interpreted, the two tables highlight that the key issue is to find out the pair of β and ϑ providing the network topology which is nearer to network operators and users requirements.

Table 5.3: Scenario nr.2: important values in different network topologies.

	$\beta = 0$ $\vartheta = 0.0001$	$\beta = 1$ $\vartheta = 0.0001$	$\beta = 10$ $\vartheta = 0.001$	$\beta = 100$ $\vartheta = 0.01$	$\beta = 100000$ $\vartheta = 100$
Ratio β to ϑ	-	10000	10000	10000	1000
Installed BSs	20	20	22	49	52
Chosen Configurations	C2 - 100% C3 - 0% C4 - 0%	C2 - 100% C3 - 0% C4 - 0%	C2 - 86% C3 - 0% C4 - 14%	C2 - 20% C3 - 0% C4 - 80%	C2 - 15% C3 - 0% C4 - 85%
Mean Received Power (dBm)	-75.48	-80.65	-82.90	-91.17	-89.34
Min Received Power (dBm)	-95.58	-98.04	-101.35	-101.98	-101.91
Mean Distance BS - TP (m)	1241.0	1473.5	1508.4	1681.9	1394.5
Max Distance BS - TP (m)	5440.5	5201.2	5045.4	5876.9	5743.7

Table 5.4: Scenario nr.3: important values in different network topologies.

	$\beta = 0$ $\vartheta = 0.0001$	$\beta = 1$ $\vartheta = 0.0001$	$\beta = 10$ $\vartheta = 0.001$	$\beta = 100$ $\vartheta = 0.01$	$\beta = 100000$ $\vartheta = 100$
Ratio β to ϑ	-	10000	10000	10000	1000
Installed BSs	21	21	32	58	75
Chosen Configurations	C1 - 90% C2 - 10% C3 - 0% C4 - 0%	C1 - 95% C2 - 5% C3 - 0% C4 - 0%	C1 - 19% C2 - 78% C3 - 0% C4 - 3%	C1 - 0% C2 - 48% C3 - 0% C4 - 52%	C1 - 0% C2 - 37% C3 - 0% C4 - 63%
Mean Received Power (dBm)	-72.05	-78.30	-87.1	-89.68	-89.67
Min Received Power (dBm)	-89.35	-97.28	-101.94	-101.62	-101.97
Mean Distance BS - TP (m)	2231.6	3065.9	2604.8	2437.9	2275.0
Max Distance BS - TP (m)	7842.7	11442.8	9662.3	6418.6	6569.1

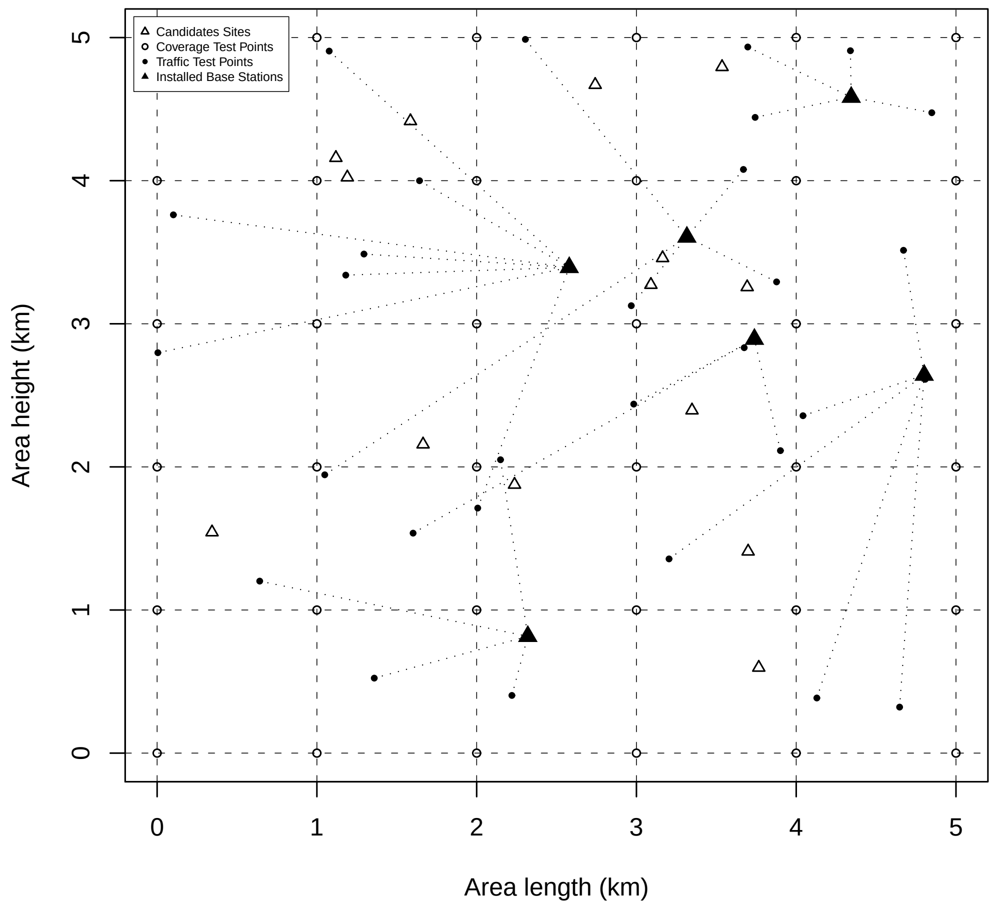


Figure 5.5: Scenario nr.1, $\beta = 0$ and $\beta = 1$ - time period 2.

Joint Approach Advantages

After the preliminary digression on the network topology changes, in this paragraph the real advantages of the joint design and management energy-aware model are pointed out. In order to give some straightforward examples of the type of results obtained, scenario nr.1 is analyzed first. Figure 5.5 displays the network provided by the optimization model with $\beta = 0$ (for the cases of $\vartheta = 0.001$ and $\vartheta = 0.0001$, which produce the same topology). When $\beta = 0$ only Capex costs are considered and operation costs due to energy expenditures are neglected. As indicated in the Figure's caption, the same network is also obtained with $\beta = 1$

($\vartheta = 0.001$ and $\vartheta = 0.0001$) since the Capex component largely prevails in the objective function. As observed in the previous paragraph, the optimal solution of the design problem includes 6 Base Stations, all of type C2.

Given that with $\beta = 0$ the energy cost is not relevant, the energy management part of the model simply keeps Base Stations on in all time periods. With $\beta = 1$, even if the energy-cost component of the objective function is much smaller than the installation cost and it does not impact on the selection of the number of Base Stations to be deployed, the energy management mechanism is encouraged to modify the topology according to the traffic level by switching on and off base stations and adjusting their power. The topology reported in Figure 5.5 shows the network topology in time period 2, when traffic is high and all 6 base stations are powered on, while Figure 5.6 shows the topology during time period 6 when traffic is low and only 2 base stations are kept on (the only two in the figure that have connected TPs).

Figure 5.7 reports the network designed using $\beta = 100000$ ($\vartheta = 100$). With such a high value of β, the Opex component in the objective function is much higher than the Capex and the optimization process is pushed to select the most energy efficient network topology, regardless of its installation cost. We can see that now more Base Stations (16) are installed and that several of them are small ones with short coverage range (types C3 and C4). While during the most traffic-intense period all the selected BSs are turned on, in time period 6, which is the one with the lowest traffic load, only 3 out of 16 Base Stations are used to serve TPs and the others are switched off, as shown in Figure 5.8. Notably, one of the on cells is a big one, while the other two are small ones. By letting $\vartheta = 10$ instead of 100, one more Base Station is required to cover the entire area; however, even if 17 BS are deployed and time period 2 corresponds to the peak traffic scenario, only 16 BSs are powered on, which shows the importance of energy management.

In Table 5.5 we show a summary of the results obtained solving scenario nr.1 with different values of β. The table reports the numerical values of the first two components of the objective function (corresponding to Capex and Opex), the total energy (given by Opex/β), the number of Base Stations installed, and the number of Base Stations powered on during the six time periods with the corresponding average Base Station utilization. The BSs utilization is calculated

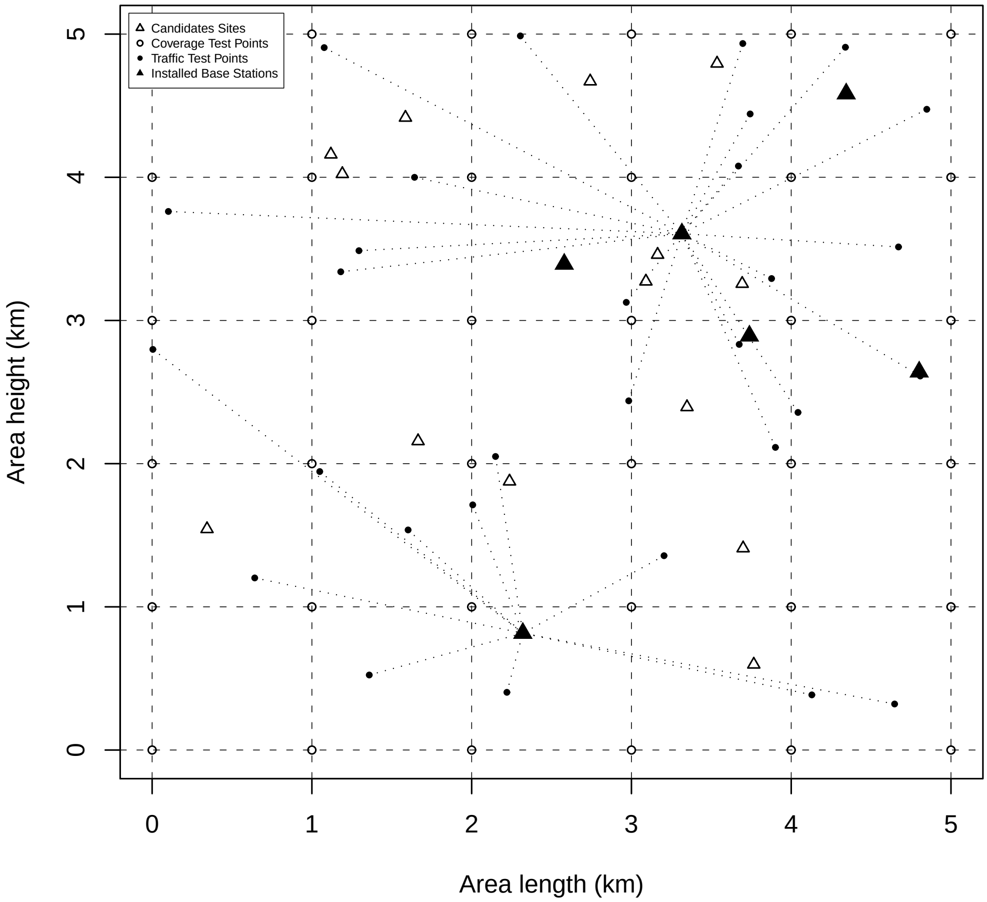


Figure 5.6: Scenario nr.1, $\beta = 1$ - time period 6.

as

$$\sum_{j \in S} \frac{\sum_{i \in I_t} \sum_{k \in K_j} \sum_{l \in L} \frac{x_{ijt} \cdot d_{it} \cdot 100}{c_{jkl} \cdot y_{jklt}}}{\sum_{k \in K_j} \sum_{l \in L} y_{jklt}}, \tag{5.1}$$

and, in other words, it symbolize the amount of the total power that is actually utilized for serving users. For this scenario, computation time to solve the problem varies from a few minutes to one hour approximately.

We observe that, by only taking into account the operational expenditures with the parameter β equal to 1, we can reach a considerable energy saving (54% compared to the case of $\beta = 0$) thanks to the power management mechanism, even if the solver brings no changes to the network topology. Augmenting the value

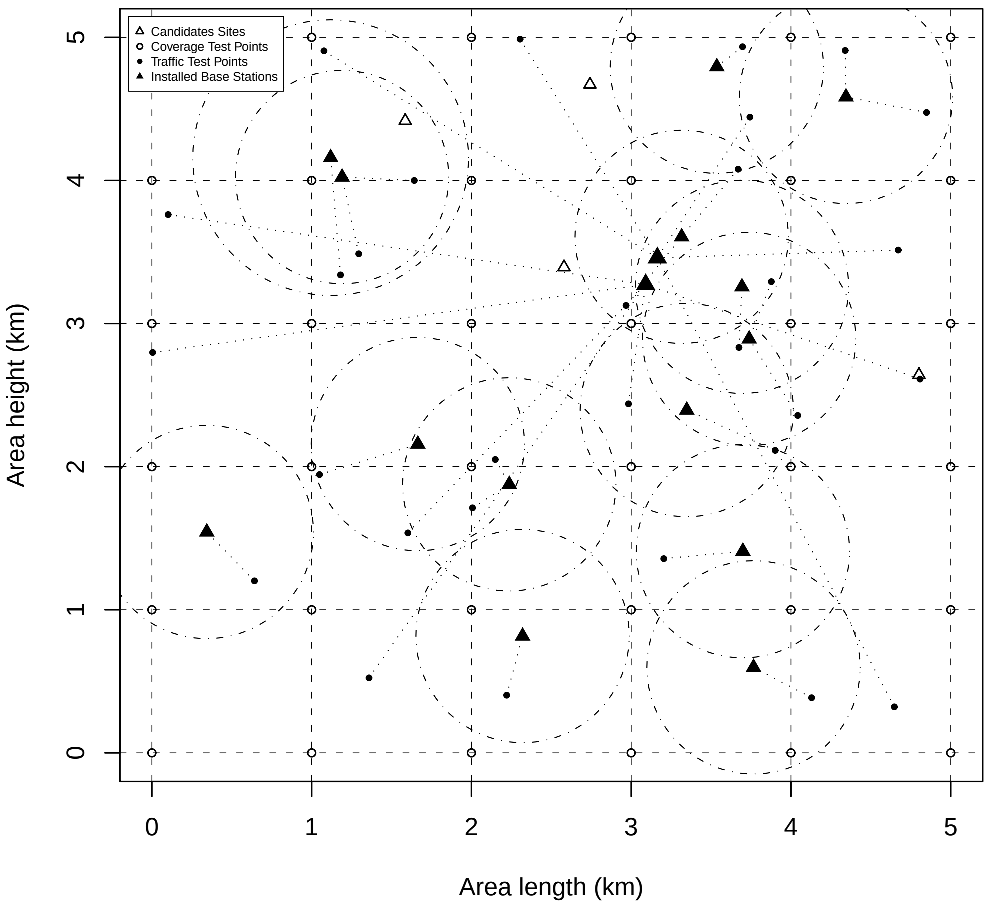


Figure 5.7: Scenario nr.1, $\beta = 100000$ - time period 2.

of β, the installation cost as well as the number of Base Stations increase, while the total energy consumed decreases. With respect to the case of $\beta = 0$ where only installation cost is considered and Base Stations are always powered on, the Capex increase is 79% with $\beta = 100$ and 118% with $\beta = 100000$, while energy decreases 70% with $\beta = 100$ and 71% with $\beta = 100000$. In this scenario, the value of $\beta = 100$ appears a reasonable compromise between installation cost and energy consumption. Increasing β further provides negligible improvements in energy saving at the price of a remarkable additional installation cost. Obviously, in general the most appropriate value of β depends on many issues including the characteristics of the problem instance and BS types. Consider however that β

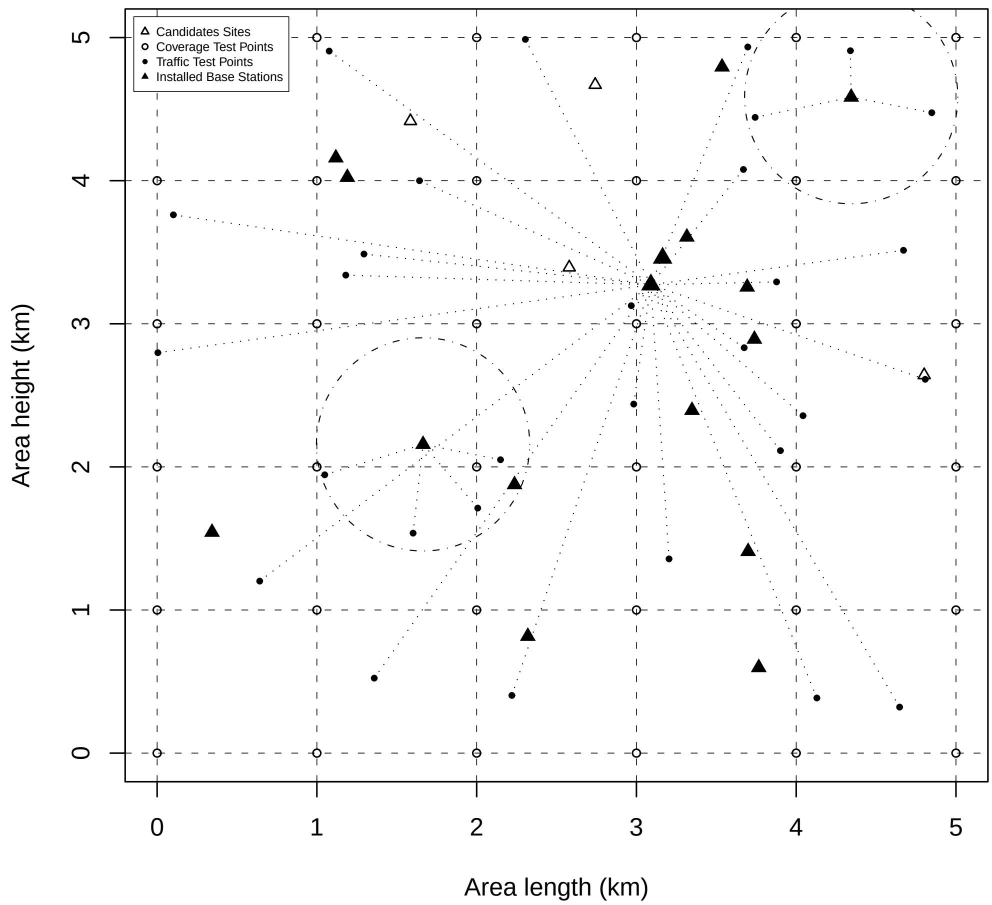


Figure 5.8: Scenario nr.1, $\beta = 100000$ - time period 6.

basically incorporates two important parameters: i) the energy cost and, ii) the time period over which the network designer wants to compute the operational expenditures in order to compare them with the installation one. Therefore, an analysis of the Pareto optimal solutions with different values of β allows the network designers to select the best option according to the network development policies of the mobile operator.

Tables 5.6 and 5.7 display the results calculated by CPLEX for the second and third scenario. In these cases, the computational time ranges from a few minutes (for $\beta = 0$ or $\beta = 1$) to several days (when $\beta \geq 10$).

Table 5.5: Scenario nr.1: Summary of the results with different values of β.

		$\beta = 0$ $\vartheta = 0.0001$	$\beta = 1$ $\vartheta = 0.0001$	$\beta = 10$ $\vartheta = 0.001$	$\beta = 100$ $\vartheta = 0.01$
Capex		384723.7	384723.7 *(+0%)*	429682.9 *(+12%)*	690577.8 *(+79%)*
Opex		0	66850.3	552482.8	4428686.5
Energy		146479.2	66850.3 *(-54%)*	55248.3 *(-62%)*	44286.9 *(-70%)*
Installed BSs		6	6	7	13
Powered On BSs	t_1	6 - 88%	6 - 88%	6 - 97%	11 - 88%
	t_2	6 - 97%	6 - 97%	7 - 83%	13 - 87%
	t_3	6 - 68%	5 - 82%	5 - 91%	11 - 74%
	t_4	6 - 78%	5 - 93%	5 - 93%	12 - 83%
	t_5	6 - 54%	4 - 80%	4 - 91%	9 - 82%
	t_6	6 - 19%	2 - 58%	2 - 76%	2 - 75%

Table 5.6: Scenario nr.2: Summary of the results with different values of β.

		$\beta = 0$ $\vartheta = 0.0001$	$\beta = 1$ $\vartheta = 0.0001$	$\beta = 10$ $\vartheta = 0.001$	$\beta = 100$ $\vartheta = 0.01$
Capex		1287673.2	1287673.2 *(+0%)*	1351488.7 *(+5%)*	2610669.3 *(+103%)*
Opex		0	205881.98	1882404.1	14449851.04
Energy		668029.5	205882.0 *(-69%)*	188240.4 *(-72%)*	144498.5 *(-78%)*
Installed BSs		20	20	22	49
Powered On BSs	t_1	20 - 87%	19 - 91%	21 - 91%	46 - 82%
	t_2	20 - 97%	20 - 97%	22 - 96%	47 - 85%
	t_3	20 - 68%	15 - 90%	17 - 89%	38 - 79%
	t_4	20 - 77%	17 - 91%	19 - 90%	42 - 80%
	t_5	20 - 53%	11 - 97%	13 - 96%	31 - 81%
	t_6	20 - 19%	6 - 64%	6 - 78%	9 - 64%

Table 5.7: Scenario nr.3: Summary of the results with different values of β.

		$\beta = 0$ $\vartheta = 0.0001$	$\beta = 1$ $\vartheta = 0.0001$	$\beta = 10$ $\vartheta = 0.001$	$\beta = 100$ $\vartheta = 0.01$
Capex		1651579.1	1666813.1 *(+1%)*	2102440.4 *(+27%)*	3280231.1 *(+99%)*
Opex		0	490453.87	3573296.9	30717028.6
Energy		1458408.1	490453.9 *(-66%)*	357329.7 *(-75%)*	307170.3 *(-79%)*
Installed BSs		21	21	32	58
Powered On BSs	t_1	21 - 87%	19 - 94%	31 - 91%	53 - 87%
	t_2	21 - 96%	21 - 94%	32 - 94%	55 - 89%
	t_3	21 - 67%	15 - 93%	26 - 92%	43 - 82%
	t_4	21 - 77%	18 - 88%	28 - 93%	49 - 84%
	t_5	21 - 53%	12 - 92%	22 - 87%	37 - 79%
	t_6	21 - 19%	5 - 86%	7 - 96%	8 - 88%

Table 5.8: Summary of the results with $\beta = 100000$ for the three scenarios.

		Scenario nr.1	Scenario nr.2	Scenario nr.3
		$\beta = 100000$ $\vartheta = 100$	$\beta = 100000$ $\vartheta = 10$	$\beta = 100000$ $\vartheta = 100$
Capex		839177.1 *(+118%)*	2858678.5 *(+122%)*	4401055.0 *(+166%)*
Opex		4299217106.0	13966808718.0	33753498898.0
Energy		42992.2 *(-71%)*	139668.1 *(-79%)*	337535.1 *(-77%)*
Installed BSs		16	54	75
Powered On BSs	t_1	14 - 82%	51 - 80%	68 - 76%
	t_2	16 - 83%	50 - 86%	69 - 78%
	t_3	13 - 77%	41 - 78%	59 - 68%
	t_4	13 - 90%	48 - 78%	65 - 71%
	t_5	9 - 82%	31 - 81%	44 - 71%
	t_6	4 - 43%	7 - 84%	12 - 63%

Scenario nr.3 in particular, which is also the most realistic one, provides very interesting results. In order to get clearer pictures, Candidate Sites and Coverage Test Points are not displayed in the figures concerning the third scenario; however, a Coverage TP is positioned in each intersection of the grid. A view with all the CSs have been presented in Figure 5.3.

As already pointed out for scenario nr.1, as β increases the energy expenses diminish while the cost for network devices installation becomes heavier. Figure 5.9 shows the topology for $\beta = 0$ ($\vartheta = 0.0001$). Since no energy expenditure is considered, all the 21 installed BSs are almost always kept on at the maximum power level; consequently, during low traffic periods, the percentage of BSs utilization drops until 19%, with respect to 96% in t_2. Even if some variations in the network topology can be observed, the same number of Base Stations (21) is deployed by letting $\beta = 1$ and $\vartheta = 0.0001$. As reported in Table 5.4, the solution found by CPLEX presents more type C1 Base Stations; besides, as expected because of the power management action, the mean connection quality indicators - i.e., distance between stations and users, power received by users - worsen of about 37% and 9% respectively. However, as shown in Table 5.7, considering the Opex in the objective function allows energy savings of about 66% with respect to the case of $\beta = 0$, with a negligible increase in Capex costs.

The values progress described above is in line with the behavior observed also for scenarios nr.1 and nr.2 (see the differences from $\beta = 0$ to $\beta = 1$ in Tables 5.3 and 5.4). Consider now the intermediate value $\beta = 10$. In that case, and by letting $\vartheta = 0.001$, 32 BSs are selected, most of them of type C2. The important

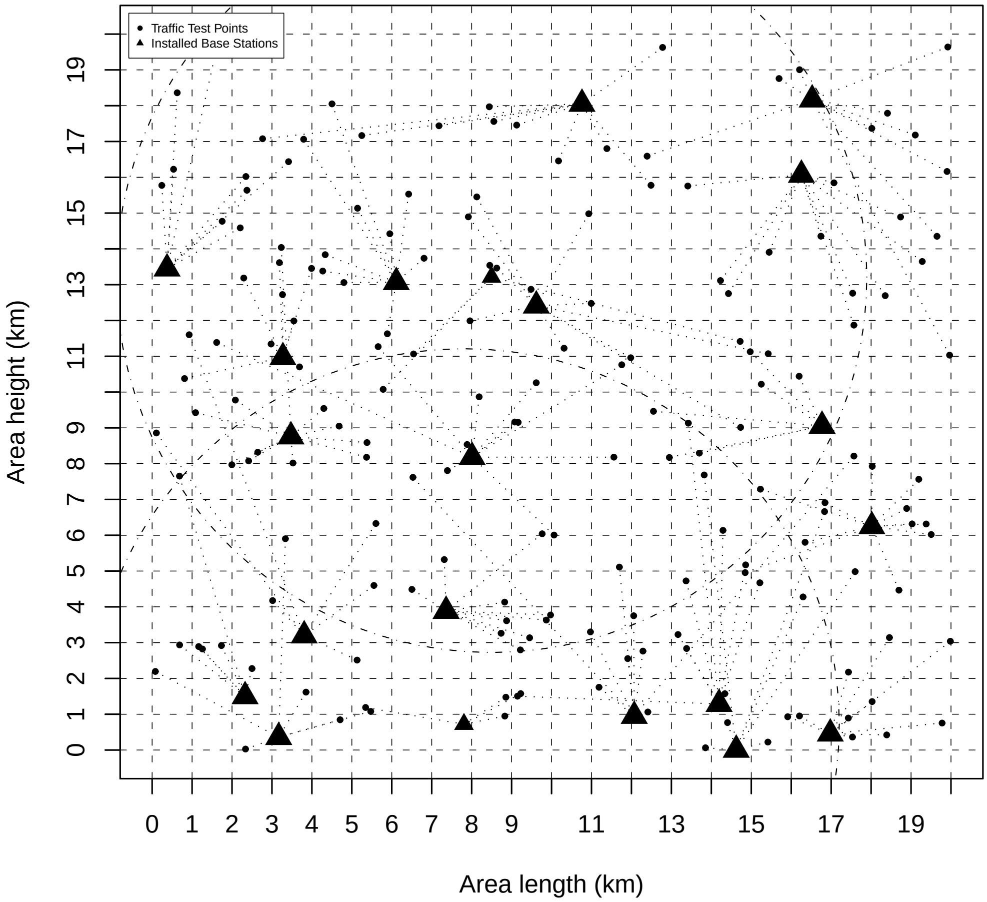


Figure 5.9: Scenario nr.3, $\beta = 0$ - time period 2.

outcome is that, at the cost of an increase in installation costs of about 27% with respect to the case of $\beta = 0$ (which can be considered really modest), a great saving of 75% in energy consumption is achieved. The topology for $\beta = 10$ is displayed in Figure 5.10, which shows also the effect of the energy management mechanism in nightly hours: thanks to that, the BS utilization percentage keeps around 90% during the whole day.

It's worth observing that further increase of β leads to excessive growth in Capex expenses: as an example, the parameter $\beta = 100000$ implies an increase of about 109% with respect to $\beta = 10$ (166% with respect to $\beta = 0$), while the energy expenditures decrease by only 9% compared with the case of $\beta =$

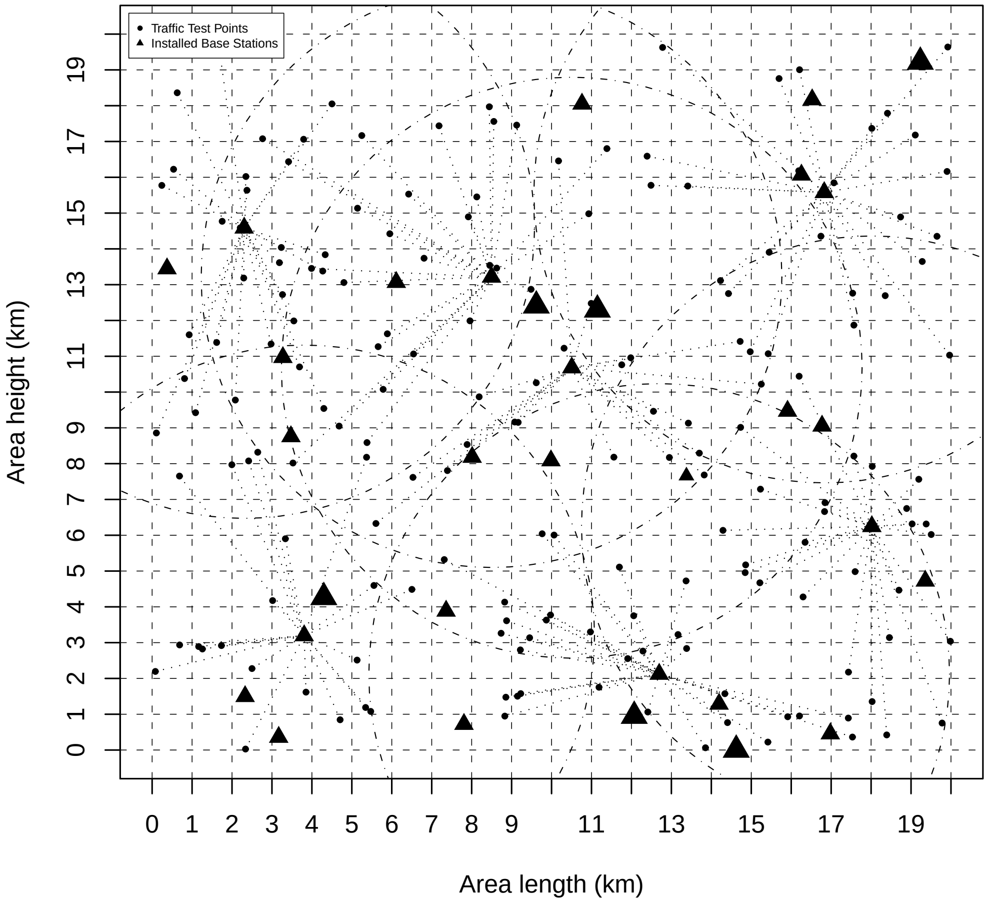


Figure 5.10: Scenario nr.3, $\beta = 10$ - time period 6.

10. The same response is noticeable for both scenario nr.1 and scenario nr.2 in the summarizing Table 5.8. Also, we can look at Figures 5.11, 5.12 and 5.13, which graphically resume the obtained results (for scenario nr.1, nr.2 and nr.3 respectively) by displaying how operators expenses change varying the value of β. The black lines link the increasing percentage values of Capex costs, while the dotted ones link the decreasing values of energy expenditures. Here, it is evident that the model behaves similarly for the three scenarios, giving good results until we reach a certain value of the trade-off parameter: the issue is that, beyond these values of β, the solutions found by the proposed model get worse, so becoming

disadvantageous. Note that, one again, the key point is selecting the best couple of β and ϑ which is able to meet network operators needs.

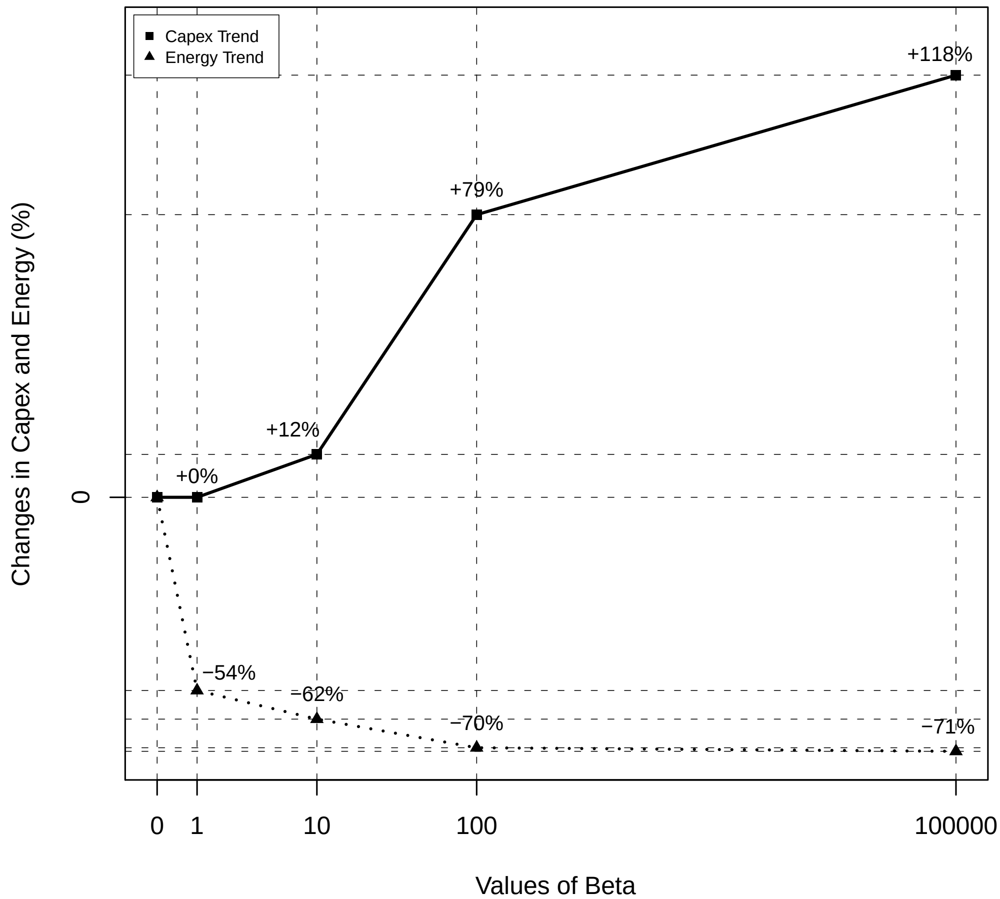


Figure 5.11: Scenario nr.1, changes of Capex and energy expenditures varying β.

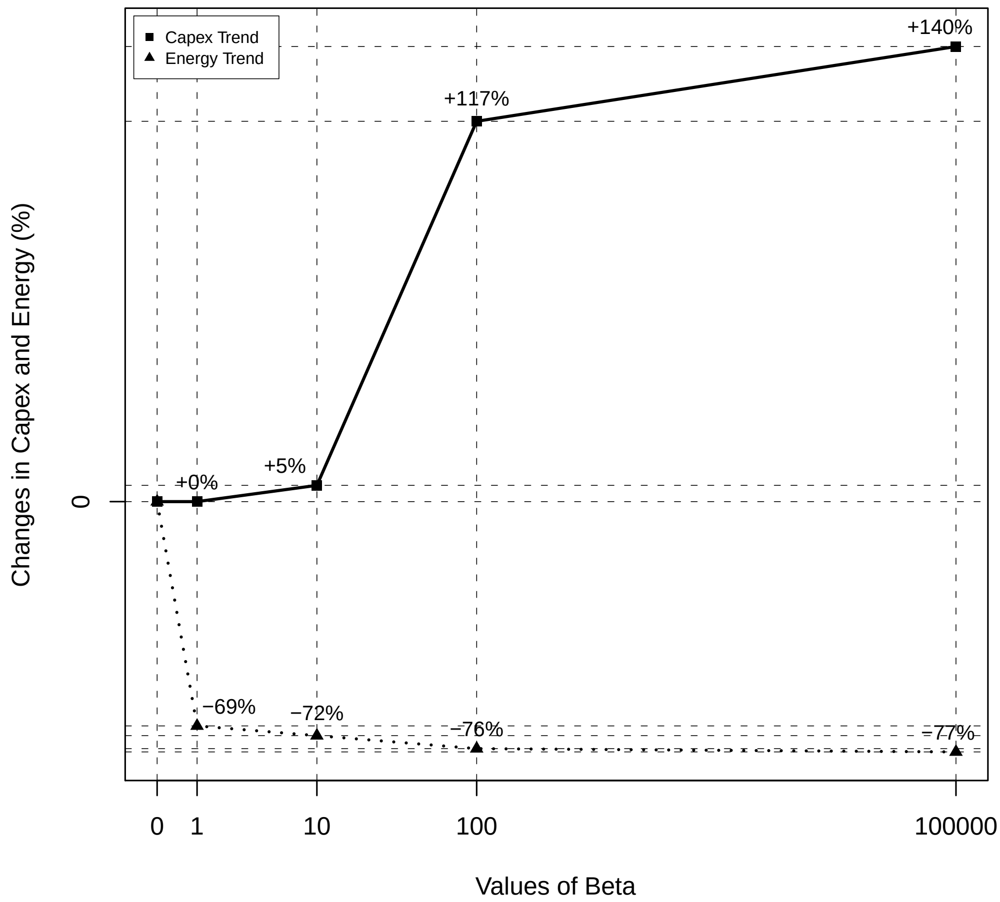


Figure 5.12: Scenario nr.2, changes of Capex and energy expenditures varying β.

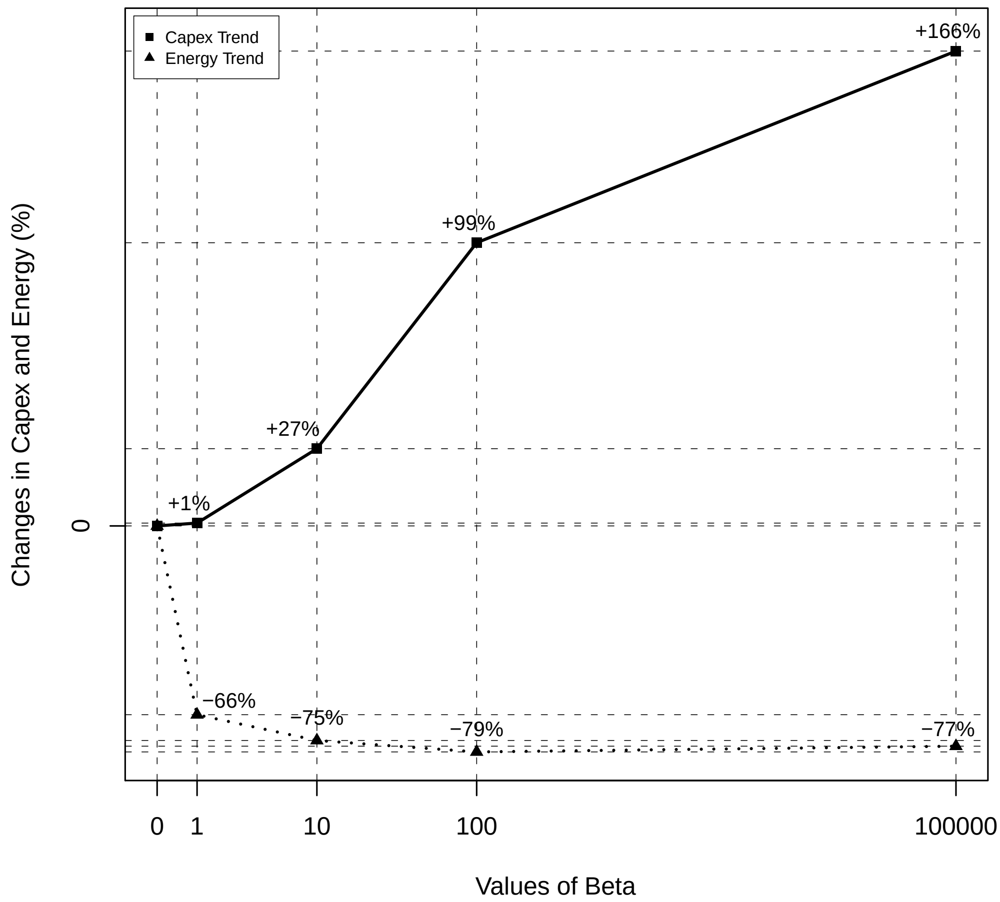


Figure 5.13: Scenario nr.3, changes of Capex and energy expenditures varying β.

Chapter 6

Conclusions

As the demand for telecommunication devices grows - due not only to increases in number of mobile accounts in emerging economies, but also to peer-to-peer content exchange - the supporting infrastructure needs to be widened and empowered. Because of this uninterrupted spread, the telecommunications sector's energetic footprint can no longer be ignored; then again, network operators have to cope with more and more expenses owing to installation and management of new devices.

In this context, green networking has become one of the most actual topics concerning the Information and Communication Technology sector. Referring in particular to cellular wireless systems, we tried to tackle the problem from an innovative point of view. Up to now, researchers have mainly focused on management aspects such as turning off some Base Stations when the traffic load is lower, ignoring the influence of the network design on energy expenditures.

Here we demonstrate this connection, claiming the importance of jointly considering radio planning and energy management to obtain the network design that minimizes installation and operational costs. For this reason, we started from a common radio planning and coverage formulation and we made some fundamental modifications: not only we refined variables and constraints to allow the power management mechanism to turn off some stations when not needed, but we also included in the objective function Opex costs. Operational expenses are weighted by a convenient parameter β which, in our opinion, besides energy cost, symbolizes the time period over which the designer wants to compute the operational expenditures in order to compare them with the installation ones. When set to zero, this parameter lets the objective function ignore Opex costs; this way, we could compare our model to the traditional one, which only considers installation expenses.

We defined a daily traffic pattern based on real measurings, which reflects mobile users'habits and states the active users percentages in every time period. Then we described four Base Stations categories. In our opinion, using different cell sizes for covering the service area is capital for energy savings. In fact, even if short-radius cells are usually considered a more efficient solution from the energetic point of view, only by using small and big cells at the same time we can dynamically switch on and off some Base Stations following the traffic fluctuations during the day. As we asserted the importance of having various sized cells, each selected BS type is characterized by a different coverage radius. We also allowed the Base Stations to regulate their emission power by introducing five power levels; this way, each BS can work at its maximum power (level 1) or can transmit only a percentage of it (level 2, 3 and 4). Level 5 is the main feature of the energy management mechanism and it indicates a BS in stand-by mode.

We showed that varying the tradeoff parameter β between installation and operational expenditures, the network topologies that we get through the proposed models have quite different characteristics. First of all we noticed that, when only Capex costs are considered, the minimum number of BSs able to cover the entire area is selected; all the installed stations are of the most powerful type and, since energy expenditures are ignored, no power management is adopted during low traffic periods. As they tend to use big cells, network topologies with a low installation cost are not very efficient from an energy consumption point of view. In terms of power waste, we calculated the percentages of BSs utilization in each time period. When Opex are neglected we observed that only in the most traffic intense periods the installed BSs are fully exploited, while during nightly hours the utilization decreases to 19%. The same results were found for all the three scenarios.

The situation totally changed when introducing Opex costs in the objective function. In that case, by letting β grow, we observed that more and more Base Stations are deployed, the most part of low power types, meaning that our model was correctly trying to reach the topology which implied the minimum energy expenditures, regardless of fixed costs. During daily hours, when traffic is most intense, all the selected Base Stations (or most of them) are turned on. On the other hand, in nocturnal hours the model pushes a great part of the stations to the stand-by mode while only a few ones are kept on. Now, the BSs utilization always keeps over 60% as a minimum, reaching a percentage between 86 and 96% during the night for scenario nr.3.

Generally, we noted that, during the day, small cells are preferred to bigger ones, especially when β assumes high values. This is because the energy consumption, which in case of high β weighs much more than deployment costs, usually decreases faster than the growth of the number of BSs required to cover the area. On the contrary, during the night, the trend is to power on only a few, high power Base Stations, eventually supported by some other smaller ones. In fact, we observed that the most energy efficient networks include not only small cells with low energy consumption, but also big cells in order to provide to the energy management mechanism enough flexibility to adjust network capacity according to traffic load.

In order to give an idea of the advantages reachable using our enhanced model, besides numerical results we provided saving percentages with regard to the traditional network design problem. For example we computed that, by simply including Opex costs in the objective function, without any additional installation expense, we can save between 54 and 69% in energetic costs. The percentage raises when increasing the value of β, with a modest growth of the Capex depending on the test scenario.

Concerning future works, several refinements of the presented results are possible. First, we can compare the performance obtained using energy-aware GSM networks to other kind of cellular networks deployed on the same test scenarios. As an example, we can modify propagation and Base Stations data we selected for our tests in order to cover the service area using UMTS or even Long Term Evolution (LTE) networks. The same way, our approach can be applyed to other wireless networks like WLANs and mesh architectures. Also, since here only traffic variations in a generic day are taken into account, the difference between working and weekend days with respect to traffic load can be considered in order to obtain further energy savings .

However, the most interesting topic we are going to treat concerns the *coverage on-demand* option. The main idea is that of allowing independent network configurations for data and signaling so as to satisfy the radio coverage constraint with minimal energy consumption. If we remove the complete coverage constraints (and so the Coverage TPs), coverage gaps in the data network are permitted; anyway the signaling network, which is static and fully covers all the service area, will be able to detect active users and provide them network capacity by dynamically switching on the required netowork elements.

Bibliography

[1] Energy efficient wireless internet access with cooperative cellular networks. *Computer Networks*, In Press, Uncorrected Proof, 2010".

[2] I.F. Akyildiz and Xudong Wang. A survey on wireless mesh networks. *Communications Magazine, IEEE*, 43(9):S23 – S30, sep. 2005.

[3] E. Amaldi, A. Capone, and F. Malucelli. Planning UMTS base station location: Optimization models with power control and algorithms. *IEEE Transaction on Wireless Communications*, 2(5):932–952, September 2003.

[4] E. Amaldi, A. Capone, F. Malucelli, and C. Mannino. *Optimization for planning cellular networks*, chapter 32, pages 917–940. 2006.

[5] S. Bosio, A. Capone, and M. Cesana. Radio planning of wireless local area networks. *IEEE/ACM Transactions on Networking*, 15(6):1414–1427, December 2007.

[6] C.D. Charalambous and N. Menemenlis. Dynamical spatial log-normal shadowing models for mobile communications. *Proceedings of the 27th General Assembly of the International Union of Radio Science*, August 2002.

[7] L. Chiaraviglio, M. Mellia, and F. Neri. Energy-aware networks: Reducing power consumption by switching off network elements. In *FEDERICA-Phosphorus tutorial and workshop (TNC2008)*, 2008.

[8] A. Eisenblätter and H.F. Geerd. Wireless network design: solution-oriented modeling and mathematical optimization. *Wireless Communications, IEEE*, 13(6):8–14, December 2006.

[9] TS ETSI. 145 005 v9.3.0 (3gpp ts 45.005 version 9.0 release 9). July 2010.

[10] A.J. Fehske, F. Richter, and G.P. Fettweis. Energy efficiency improvements through micro sites in cellular mobile radio networks. In *GLOBECOM Workshops, 2009 IEEE*, pages 1–5, November 2009.

[11] J. Gong, S. Zhou, Z. Yang, D. Cao, C. Zhang, Z. Niu, and P. Yang. Green mobile access network with dynamic base station energy saving. In *IEICE Tech. Rep., IA 2009*, volume 109, pages 25–29, October 2009.

[12] The Climate Group. SMART 2020: Enabling the low carbon economy in the information age. In *2010 State of Green Business*, June 2008.

[13] M. Gupta and S. Singh. Greening of the internet. In *Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communications*, pages 19–26, 2003.

[14] P. Heegaard. Empirical observations of traffic patterns in mobile and ip telephony. *Next Generation Teletraffic and Wired/Wireless Advanced Networking*, pages 26–37, 2007.

[15] H.Karl. An overview of energy-efficiency techniques for mobile communication systems. Technical report, Report of the Working Group 7 "Low-power broadband wireless communication" of the Arbeitsgruppe Mobikom, DLR/BMBF, October 2003.

[16] A. P. Jardosh, G. Iannaccone, K. Papagiannaki, and B. Vinnakota. Towards an energy-star WLAN infrastructure. In *Mobile Computing Systems and Applications, 2007. HotMobile 2007. Eighth IEEE Workshop on*, pages 85–90, March 2007.

[17] A. P. Jardosh, K. Papagiannaki, E. M. Belding, K. C. Almeroth, G. Iannaccone, and B. Vinnakota. Green WLANs: on-demand WLAN infrastructures. *Mobile Networks and Applications*, 14(6):798–814, December 2009.

[18] K. Johansson, A. Furuskar, P. Karlsson, and J. Zander. Relation between base station characteristics and cost structure in cellular systems. In *Personal, Indoor and Mobile Radio Communications, 2004. PIMRC 2004. 15th IEEE International Symposium on*, volume 4, pages 2627–2631, May 2004.

[19] J. Lorincz, A. Capone, and M. Bogarelli. Energy savings in wireless access networks through optimized network management. In *Wireless Pervasive Computing (ISWPC), 2010 5th IEEE International Symposium on*, pages 449–454, May 2010.

[20] J. T. Louhi. Energy efficiency of modern cellular base stations. In *Telecommunications Energy Conference, 2007. INTELEC 2007. 29th International*, pages 475–476, October 2007.

[21] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo. Optimal energy savings in cellular access networks. In *Proc. of GreenComm'09*, June 2009.

[22] M. A. Marsan, L. Chiariviglio, D. Ciullo, and M. Meo. A simple analytical model for the energy-efficient activation of access points in dense WLANs. In *e-Energy 2010 - First International Conference on Energy-Efficient Computing and Networking*, April 2010.

[23] M. A. Marsan and M. Meo. Energy efficient management of two cellular access networks. In *Proc. of GreenMetrics 2009*, June 2009.

[24] R. Mathar and T. Niessen. Optimum positioning of base stations for cellular radio networks. *Wireless Networks*, 6(4):421–428, 2000.

[25] H. Mellah and B. Sansò. Review of facts, data and proposals for a greener Internet. In *Proceedings of Broadnets09*, September 2009.

[26] F. Richter, A.J. Fehske, and G.P. Fettweis. Energy efficiency aspects of base station deployment strategies for cellular networks. In *VTC '09*, 2009.

[27] S.Y. Seidel and T.S. Rappaport. 914 mhz path loss prediction models for indoor wireless communications in multifloored buildings. *Antennas and Propagation, IEEE Transactions on*, 40(2):207–217, February 1992.

[28] M.L. Sichitiu. Wireless mesh networks: opportunities and challenges. In *Proceedings of World Wireless Congress*, 2005.

[29] T. D. Todd, A. A. Sayegh, M. N. Smadi, and D. Zhao. The need for access point power saving in solar powered wlan mesh networks. *IEEE Network Magazine*, 22(3), May/June 2008.

[30] K. Tutschku. Demand-based radio network planning of cellular mobile communication systems. In *IEEE INFOCOM*, volume 3, pages 1054–1061, 1998.