ISSN 2281-4299

Dipartimento di Ingegneria informatica
automatica e gestionale Antonio Ruberti

Sapienza
Università di Roma

A Compact Formulation for the Base Station Deployment Problem in Wireless Networks

Pasquale Avella
Alice Calamita
Laura Palagi

Technical Report n. 01, 2022

A Compact Formulation for the Base Station Deployment Problem in Wireless Networks

Pasquale Avella[1], Alice Calamita[2], Laura Palagi[2]

[1] Dipartimento di Ingegneria, Università del Sannio, Benevento, 82100, Italy
[2] DIAG, Sapienza University of Rome, Rome, 00185, Italy

Abstract

In this paper, we consider the base station deployment problem in a wireless network. The natural formulation of this problem usually leads to numerical and memory issues, preventing users from dealing with real-world problems. We provide a compact reformulation that allows us to get beyond the limitations of the natural formulation. We test the new formulation on three instances derived from simulations of realistic LTE scenarios. The computational results demonstrate the efficiency of the proposed formulation, which enables to quickly identify the optimal solution with extreme accuracy.

Keywords: Wireless network planning, base station deployment, 0-1 linear programming, compact formulation, siting model, SINR inequalities

1 Introduction

This paper addresses the base station deployment optimization problem consisting of the selection of base station locations, among a set of possible ones, according to an objective and so that the selected base stations are able to guarantee coverage on a target area. The research was carried out in collaboration with the Fondazione Ugo Bordoni (FUB) in response to their request for the identification of a compact formulation for real-world problems of this kind that must be implemented using standard modeling languages and tractable by commercial solvers. FUB is a higher education and research institution under the supervision of the Ministry of Economic Development that operates in the telecommunication field, providing innovative services for government bodies [14]. Their interest in this topic comes from the fact that technological advancement and increase in traffic are leading to a densification of the number of base stations along the territory, with the consequent increase in interfering signals. Therefore, careful planning of the base station location becomes crucial to reduce the capital expenditures costs that operators will incur and to meet the service quality requirements demanded by users.

Optimization models for Wireless Network Design (WND) and for the specific problem of base station deployment have been proposed since decades. However, the available compact models present severe limitations when applied to realistic situations: the resolution of these models involves strong numerical issues that arise even in small instances. This is due to the fact that the constraints matrix contains coefficients that greatly vary in their order of magnitude. Furthermore, in real-world instances, the size of the model is huge. Due to the inability of dealing with practical problems, most contributors preferred to tackle large-size instances with a heuristic approach. Only a few tried to optimally solve real-life problems, using non-compact formulations and/or nonstandard optimization procedures.

In contrast to these works, we propose a compact formulation that can be used to optimally solve real-life cases directly using Mixed-Integer Programming (MIP) solvers in default settings. The formulation that we propose overcomes the numerical and size criticalities of the natural compact formulation and is obtained making an a priori reformulation of the critical constraints. Our formulation of the problem has three main advantages:

- it is compact, hence easily manageable by its users (e.g., telecommunications practitioners);
- it is tractable by commercial solvers using standard optimization procedures and this is an advantage since solvers usually employ excellent heuristics leading to very good and fast results (for a discussion of how heuristics of MIP solvers have a crucial impact on solving complex problems see [13]);
- it is reduced in size compared to the natural formulation hence we can deal with large instances, as in a real-world problem.

The remaining of this paper is organized as follows. Section 2 discusses the main contributions in the field of wireless network planning. Section 3 reports the statement of the problem and Section 4 presents a basic natural formulation for the base station deployment problem and explains its limitations. In Section 5 a compact reformulation that overcomes those limits is proposed. Section 6 compares the computational results obtained using the basic and our strengthened formulation on realistic instances provided by FUB. Conclusions are given in Section 7.

2 Literature Review

The base station deployment problem can be traced back to a WND problem, namely the problem of configuring a set of transmitters to provide service on a target area. The term configuring refers both to the optimal identification of the positions, hence to the base station deployment problem, and to the optimal identification of some parameters of the transmitters. The design of wireless networks is a widely studied topic with a large number of contributions, we can divide them into two main strands: those using exact methods, those using heuristic methods. A summary classification of the cited works can be found in Table 1.

As for the strand of exact methods, a non-compact binary formulation is proposed in [8], in which both locations and power emissions of the antennas are optimized, with the aim of maximizing the territorial coverage. The main contribution of this paper lies in the formulation, particularly suitable for a rapid numerical resolution since it does not make use of the big-M coefficients, which are normally employed to model logical relationships bringing to numerical problems in large instances. Nonetheless, the formulation contains an exponential number of constraints and cannot be represented in a compact way. The resolution algorithm is a combination of row-generation and branch-and-cut repeated on different combinations of power emission levels.

The authors of [22] tackle the problem of optimal base station locations and power emissions, maximizing the service provider's profit. They use a Mixed-Integer Linear Programming (MILP) formulation with big-M constraints to be solved with an exact solution method that combines combinatorial Benders decomposition, classical Benders decomposition, and valid cuts in a nested way.

As for the papers proposing formulations to be solved in standard exact fashion, we can find a recent work [6] focusing on 5G planning that provides a mixed-integer formulation with a big-M approach. It takes into account a wide range of decision variables, including the locations, and considers the presence of mutually exclusive transmitters. The goal of the authors is the maximization of mobile operator profits, which consists of both maximizing coverage and minimizing costs (i.e. activation, installation, transmission of bits costs). This paper takes the basis on the formulations proposed in previous works [1, 2, 9] related to LTE-RAN applications, in which a set of services (except for [2]) and multi-period planning are also considered in the decision process. The formulations proposed in the aforementioned papers [1, 2, 6, 9] are solved using standard solvers but tests have been done on randomly generated instances with very small size (when declared), hence the criticalities arising from practical instances are not faced.

Among the heuristic proposals, we can cite many works. One of the first that has identified the presence of numerical issues is [21]. The authors of this paper proposed a heuristic algorithm to solve very large instances of MILP network planning models in which the revenue associated with the coverage is maximized. The addressed decisions variables are power emissions and frequency channels. In [10] a MILP problem for the optimal allocation of power emissions and frequency channels is tackled by means of a genetic algorithm, which is suitable for large-scale problems. The model maximizes territorial coverage. In [20] the authors employ a MILP formulation for the power assignment problem in wireless networks, in which the signal orientation of the antenna is also considered. Their formulation uses the notorious big-M approach and its exact resolution is not viable for real-world problems. Therefore, a constructive heuristic followed by an improving local search is introduced to overcome this issue. Another formulation for the optimal allocation of power emissions, frequency channels and transmission scheme can be found in [12], in which a matheuristic is also proposed to solve the problem. The matheuristic combines a genetic algorithm exploiting a suitable linear 2 relaxation of the problem and an integer linear programming heuristic improving the solutions found with the relaxation.
Among the recent works addressing 5G network planning, there is [16], in which the minimization of the total number of deployed base stations is pursued. The proposed model uses 0-1 variables and is solved with a meta-heuristic approach. Another interesting contribution is [7], which considers the design of 5G networks using an optimization model that allocates locations and frequencies and contains also constraints on electromagnetic emissions and on the distance between sensitive areas and activated antennas. The goal is the maximization of coverage together with the minimization of installation costs. The authors propose a heuristic to solve practical problems.

In summary, the majority of the contributions focus on heuristics approaches and a very limited number of papers addressed the numerical and size issues that arise in real cases using exact methods.

3 Problem Definition

The objective of the base station deployment problem is to identify an optimal placement for base stations, given a range of possible sites. For our purpose, base stations can be considered as transmitters providing service (i.e. wireless connection) to a target area. The target area is usually partitioned into elementary areas of identical size in line with the recommendations of the telecommunications regulatory bodies. Each elementary area is called testpoint and it is assumed to be representative of all users within the corresponding elementary area, namely each testpoint is a representative receiver.

Each testpoint receives the signals coming from all the transmitters: the power received by each receiver is classified as serving power if it relates to the power emitted by the transmitter serving that receiver, otherwise is classified as interfering power. This is in line with the physical layer specifications of the LTE standard [24], which corresponds to the technology we are modeling. A receiver is regarded as served by a base station if the

Paper	Decision Variables	Method	Formulation	Technology
[22]	location, power emission	exact	MILP	UMTS, W-CDMA
[8]	location, power emission	exact	0-1 LP	WiMAX, DVB-T
[21]	power emission, frequency channel	heuristic	MILP	DVB-T
[10]	power emission, frequency channel	heuristic	MILP	WiMAX
[20]	power emission, horizontal orientation	heuristic	MILP	DVB
[12]	power emission, frequency channel, transmission scheme	heuristic	0-1 LP	DVB-T
[16]	location	heuristic	0-1 LP	5G
[7]	location, frequency channel	heuristic	MILP	5G
[6]	location, power emission, frequency channel	exact	MILP	5G
[1, 2, 9]	location, power emission, frequency channel, service (except for [2]), period	exact	MILP	LTE-RAN

Key: DVB-T, Digital Video Broadcasting-Terrestrial; WiMAX, Worldwide Interoperability for Microwave Access; UMTS, Universal Mobile Telecommunications System; W-CDMA, Wideband Code Division Multiple Access; LTE-RAN, Long Term Evolution-Radio Access Network; 5G, 5$^{\text{th}}$ Generation.

Table 1: Paper classification scheme.

ratio of the serving power to the sum of the interfering powers and noise power (Signal-to-Interference Ratio or SIR) is above a threshold [23], whose value depends on the desired quality of service. We also consider the presence of received noise.

In our model, power emission, frequency channel, and transmission scheme of the transmitters are given and equal for all of them without loss of generalization. Indeed, in the planning phase, there is no need to consider more than one power value since planning considers either the mean or the maximum, the value can then be re-evaluated using traffic data in real-time; for different frequency channels, the problem decomposes since there is no interference among non-co-channel signals; the transmission scheme affects the received signal that in our case is given.

4 A Basic Compact Formulation

Since we assume that both the frequency channel and the power emissions are given and are the same for all the transmitters, we only need to define which transmitters are activated and which receivers are served by a transmitter. Thus the basic formulation of the SINR inequalities is a special case of the discrete big-M formulation with a single power value reported e.g. in [8].

Let $\mathcal{B}$ the set of possible transmitters and $\mathcal{T}$ the set of receivers located at the testpoints. We introduce the variables

$$z_b = \begin{cases} 1 & \text{if transmitter } b \text{ is activated} \\ 0 & \text{otherwise} \end{cases} \qquad b \in \mathcal{B}$$

and

$$x_{tb} = \begin{cases} 1 & \text{if testpoint } t \text{ is served by transmitter } b \\ 0 & \text{otherwise.} \end{cases} \qquad b \in \mathcal{B},\ t \in \mathcal{T}$$

The quality of the received signal at a testpoint can be measured through several criteria, we adopt the Signal-to-Interference-plus-Noise Ratio (SINR). Let $a_{tb} > 0$ be the power value measured in $t \in \mathcal{T}$ of the signal emitted by $b \in \mathcal{B}$, then a receiver t is served by $\beta \in \mathcal{B}$, if the SINR of the serving power to the sum of the interfering powers and noise $\mu > 0$ is above a given SINR threshold $\delta \geq 0$, namely

$$\frac{a_{t\beta} z_\beta}{\mu + \sum_{b \in \mathcal{B} \setminus \{\beta\}} a_{tb} z_b} \geq \delta \qquad t : \ x_{t\beta} = 1. \tag{1}$$

The way serving and interfering power signals are evaluated is the same, what changes is the role of the transmitter emitting the signal: if a transmitter is the server of the receiver t, then the signal will be classified as serving, otherwise it will be classified as interfering. Following [8], we can rewrite the SINR condition by means of the following big-M constraints

$$a_{t\beta} z_\beta - \delta \sum_{b \in \mathcal{B} \setminus \{\beta\}} a_{tb} z_b \geq \delta\mu - M_{t\beta}(1 - x_{t\beta}) \qquad t \in \mathcal{T}, \beta \in \mathcal{B} \tag{2}$$

where $M_{t\beta}$ is a large positive constant. When $x_{t\beta} = 1$, (2) reduces to (1); when $x_{t\beta} = 0$ and $M_{t\beta}$ is sufficiently large, (2) becomes redundant. We can set, e.g., $M_{t\beta} = \delta\mu + \delta \sum_{b \in \mathcal{B} \setminus \{\beta\}} a_{tb}$.

A minimum territorial coverage, namely a minimum number c of served testpoints, is enforced by the following

$$\sum_{b \in \mathcal{B}} \sum_{t \in \mathcal{T}} x_{tb} \geq c \tag{3}$$

where $0 \leq c \leq |\mathcal{T}|$.

Each testpoint must be covered by at most one serving base station, namely

$$\sum_{b \in \mathcal{B}} x_{tb} \leq 1 \qquad t \in \mathcal{T}. \tag{4}$$

In accordance with FUB purposes, we aim at minimizing the total number of activated base stations. Thus the basic formulation is the following 0-1 Linear Programming model:

$$\begin{aligned} \min_{x,z} \quad & \sum_{b \in \mathcal{B}} z_b \\ & (x, z) \in S \end{aligned} \tag{5}$$

where S is the feasible region defined as

$$S = \Big\{(x, z) \in \{0,1\}^{n+m} : \text{ satisfying } (2), (3), (4)\Big\}$$

with $x = (x_{tb})_{t\in\mathcal{T}, b\in\mathcal{B}}, z = (z_b)_{b\in\mathcal{B}}$ and $n = |\mathcal{T}| \times |\mathcal{B}|, m = |\mathcal{B}|$.

Sets and parameters used in the model are summarized in Table 2.

Symbol	**Notation**
$\mathcal{B}$	Set of potential transmitters
$\mathcal{T}$	Set of receivers
a_{tb}	Power value measured in $t \in \mathcal{T}$ of the signal emitted by $b \in \mathcal{B}$
c	Minimum number of receivers that must be covered
δ	SINR threshold
μ	System noise

Table 2: Sets and parameters of the optimization model.

In principle, model (5) can be solved by commercial solvers, however, it is well-known that both numerical and memory issues arise when trying to solve practical instances, as widely described in [8, 11]. The main issues encountered are:

- the power received in each testpoint ranges in a large interval, from very small values (order of 10^{-7}) to huge (10^5), which makes the range of coefficients a_{tb} in the constraints matrix very large and the solution process numerically unstable and possibly affected by error;
- the big-M coefficients lead to poor quality bounds that impact on the effectiveness of standard solution procedures;
- practical problems lead to models with a very large number of variables and constraints.

The consequences are that real-world base station deployment problems are hard to solve using classic optimal procedures.

5 A Priori Reformulation

We present a reformulation tightening the basic formulation and reducing the number of constraints in order to overcome the numerical criticalities outlined in Section 3.

5.1 VUB-based tightening

First we tighten the basic formulation adding the following Variable Upper Bounds (VUBs):

$$x_{tb} \leq z_b \quad t \in \mathcal{T}, b \in \mathcal{B} \tag{6}$$

enforcing that a testpoint $t \in \mathcal{T}$ can be assigned to the transmitter $b \in \mathcal{B}$ only if b is activated. VUBs are trivially valid inequalities for S. They turned out to be computationally very effective as reported in Section 6.

Furthermore, we can also strengthen each SINR constraint (2) by replacing the z_β with the $x_{t\beta}$ as in the following:

$$a_{t\beta}x_{t\beta} - \delta \sum_{b\in\mathcal{B}\setminus\{\beta\}} a_{tb}z_b \geq \delta\mu - M_{t\beta}(1 - x_{t\beta}) \qquad t \in \mathcal{T}, \beta \in \mathcal{B}. \tag{7}$$

Theorem 1 *Inequalities* (7) *are valid for* S*.*

Proof When $x_{t\beta} = 1$, the corresponding inequality of (7) becomes $a_{t\beta} - \delta \sum_{b\in\mathcal{B}\setminus\{\beta\}} a_{tb}z_b \geq \delta\mu$ which is the required SINR inequality. When $x_{t\beta} = 0$ we get $-\delta \sum_{b\in\mathcal{B}\setminus\{\beta\}} a_{tb}z_b \geq \delta\mu - M_{t\beta}$ that is trivially satisfied.

Theorem 2 *Inequalities* (2) *are implied by inequalities* (6) *and* (7)*.*

Proof For each $t \in \mathcal{T}$ and $\beta \in \mathcal{B}$, inequalities (2) and (7) can be written respectively as $z_\beta \geq \frac{h}{a_{t\beta}}$ and $x_{t\beta} \geq \frac{h}{a_{t\beta}}$ where $h = \delta \sum_{b\in\mathcal{B}\setminus\{\beta\}} a_{tb}z_b + \delta\mu - M_{t\beta}(1 - x_{t\beta})$ and using $a_{t\beta} > 0$. Inequality (6) can be written as $z_\beta - x_{t\beta} \geq 0$. Hence we get that inequality (2) can be obtained summing up inequalities (6) and (7), meaning that it is implied by them.

5.2 Constraint aggregation

In order to reduce the dimension of the problem, we can aggregate the constraints (7) using (4), producing the following aggregate SINR constraints

$$(1 + \delta) \sum_{b\in\mathcal{B}} a_{tb}x_{tb} - \delta \sum_{b\in\mathcal{B}} a_{tb}z_b \geq \delta\mu - M_t(1 - \sum_{b\in\mathcal{B}} x_{tb}) \qquad t \in \mathcal{T} \tag{8}$$

where the big-M term depends only on the testpoint t and could be set to $M_t = \delta\mu + \delta \sum_{b\in\mathcal{B}} a_{tb} \geq M_{t\beta}$.

Theorem 3 *Inequalities* (8) *are valid for* S*.*

Proof Given $t \in \mathcal{T}$, we can have the following cases according to (4): t is covered by exactly one base station β, or t is not covered by any base station. In the former case, thanks to (4), the sum $\sum_{b\in\mathcal{B}} a_{tb}x_{tb}$ reduces at most to a single element $a_{t\beta}x_{t\beta}$, being β the unique transmitter serving testpoint t, and the same is for $\sum_{b\in\mathcal{B}} x_{tb} = x_{t\beta}$. Thus (8) reduced to

$$(1 + \delta)a_{t\beta}x_{t\beta} - \delta \sum_{b\in\mathcal{B}} a_{tb}z_b \geq \delta\mu - M_t(1 - x_{t\beta}).$$

Since $x_{t\beta} = 1$ implies $z_\beta = 1$, we get that (7) is satisfied . In the latter case, no transmitter is serving testpoint t and we get $\sum_{b\in\mathcal{B}} a_{tb}x_{tb} = 0$ and $\sum_{b\in\mathcal{B}} x_{tb} = 0$. Thus (8) reduced to

$$-\delta \sum_{b\in\mathcal{B}} a_{tb}z_b \geq \delta\mu - M_t$$

which satisfy (7) when $x_{t\beta} = 0$.

Observe that Theorems 1, 2 and 3 allow to use the aggregate constraints (8) for replacing all the SINR constraints (2) in problems (5). Since for each receiver t the SINR aggregate constraint is only one, whereas for each receiver t the SINR constraints are $m = |\mathcal{B}|$, we obtain a reduced formulation in terms of number of constraints.

5.3 Coefficient tightening

To act on the numerical issues arising from the coefficients of the SINR inequalities, we also performed a heuristic sparsification by setting a minimum threshold ϵ on the received power below which it is possible to consider the power received in a receiver by a transmitter to be null, namely we set

$$\tilde{a}_{tb} = \begin{cases} a_{tb} & \text{if } a_{tb} \geq \varepsilon \\ 0 & \text{otherwise.} \end{cases}$$

This allows to reduce the size of the problem eliminating some x_{tb} variables a priori.

Another way to deal with numerical problems is to use the smallest possible big-M. Consider the big-M suitable for the aggregate SINR constraints, namely

$$M_t = \delta\mu + \delta \sum_{b\in\mathcal{B}} a_{tb} \quad \text{with } t \in \mathcal{T}.$$

One way to decrease its value is to replace the term which makes the sum of the signal power received from all the transmitters in the testpoint t with the sum of the strongest interferers. In particular, only the strongest α interferers might be considered, where α is an upper bound of the optimal number of activated base stations, i.e. $\alpha \geq \sum_{b\in\mathcal{B}} z_b^*$. Therefore, the big-$M$ with the proposed reduction becomes

$$\widetilde{M}_t = \delta\mu + \delta \sum_{b\in\mathcal{A}_t} a_{tb} \quad \text{with } t \in \mathcal{T}$$

where $\mathcal{A}_t$ is the set of the α base stations emitting the strongest signals received in t, i.e. $|\mathcal{A}_t| = \alpha$. The smaller α, the smaller the big-M and the better the formulation, so the estimate of α must be as accurate as possible. A couple of procedures that can be used to identify α are an a priori estimate (that requires the knowledge of the dataset) or the resolution of the siting problem only through a heuristic that identifies a good feasible solution.

5.4 The final formulation

The new strengthened and reduced formulation differs from the basic formulation since it contains the aggregate SINR constraints (8) instead of the SINR constraints (2) and considers the addition of VUBs (6). It can be formulated as follows

$$\begin{aligned}
\min \quad & \sum_{b\in\mathcal{B}} z_b && \\
& (1+\delta)\sum_{b\in\mathcal{B}} \tilde{a}_{tb}x_{tb} - \delta\sum_{b\in\mathcal{B}} \tilde{a}_{tb}z_b \geq \delta\mu - \widetilde{M}_t(1-\sum_{b\in\mathcal{B}} x_{tb}) && t\in\mathcal{T} \\
& x_{tb} \leq z_b && t\in\mathcal{T}, b\in\mathcal{B} \\
& \sum_{b\in\mathcal{B}}\sum_{t\in\mathcal{T}} x_{tb} \geq c && \\
& \sum_{b\in\mathcal{B}} x_{tb} \leq 1 && t\in\mathcal{T} \\
& x_{tb}\in\{0,1\} && t\in\mathcal{T}, b\in\mathcal{B} \\
& z_b\in\{0,1\} && b\in\mathcal{B}
\end{aligned} \tag{9}$$

where the new big-M coefficient can be defined as $\widetilde{M}_t = \delta\mu + \delta\sum_{b\in\mathcal{A}_t} a_{tb}$, in which $\mathcal{A}_t$ is the set of the α transmitters emitting the strongest signals received in t and the coefficient α is an upper bound of the optimal value. We have also carried out a heuristic coefficient tightening with which it was possible to eliminate the smallest coefficients of the SINR aggregate constraints with the setting of a minimum threshold on the received signal.

6 Computational Results

The scope of this section is to compare the performance obtained using the basic and the final formulations to solve the base station deployment problem on the same instances. The code has been implemented in Python and the experiments have been carried out on a Virtual Machine server characterized by 8 GHz CPU and 96 GB RAM employing *Gurobi Optimizer 9* [17] as the solver. The setting used on Gurobi is the default one, with an exception related to the use of heuristics, which has been increased as the rapid identification of a feasible solution turned out to be crucial in closing the problem resolution.

6.1 The instances

The instances used for the comparison are obtained from simulations of realistic scenarios provided by FUB concerning the LTE signal in the Municipality of Bologna. Both the system noise and the power received in each receiver by each possible transmitter are calculated using the *Cost Hata* model [3, 15, 24].

The values of the parameter a_{tb} are expressed in W and are scaled by 10^{10} to obtain a better accuracy on the optimal solutions, as suggested in [11]. Accordingly, both the threshold on the quality of service δ and the system noise μ are expressed in W, and the latter is also

scaled by 10^{10}.
The value of the received power below which the signal can be considered neglectable is set around -110 dBm [25, 26].
Not all transmitters are considered possible serving base stations for each testpoint: a subset of them is selected for each testpoint based on the power values received at the testpoint by the base stations. In particular, if we are not in the simple case in which one transmitter is able to cover the entire target area, which is trivially testable, the best-case scenario is the one in which two base stations will be deployed. According to this case, we can certainly exclude that base station $b \in \mathcal{B}$ serves testpoint $t \in \mathcal{T}$ for all those b whose power received in t is not sufficient to satisfy the specific SINR threshold δ considering each possible combination with a single interferer. Namely, we can set $x_{tb} = 0$ for all those (t, b) such that

$$\frac{a_{tb}}{\mu + a_{th}} < \delta \qquad \forall\ h \in \mathcal{B} \setminus \{b\}.$$

Note that this is not the same as setting the coefficient $a_{tb} = 0$; in fact, in this case, a_{tb} still appears in the interfering signal, which depends only on the activation of the base station, namely on $z_b = 1$.
The total number of x_{tb} that we can be set to zero following this rule are 225 636, 341 121 and 446 626 in instance 1, 2, 3 respectively, with an average per testpoint of 48, 73, 95 transmitters, among the 135 available, for instance 1, 2, 3 respectively.

Since the technology we refer to is LTE, the elementary areas should have a square shape with a side of 100 m in accordance with the guidelines provided by the Italian Resilience Plan [18] which are based on the guidelines given by the Body of European Regulators for Electronic Communications [4, 5]. However, since the number of the testpoints, which is strictly related to the size of the elementary areas, has a huge impact on the size of the problem and since it has been tested that for the purpose of optimization such a refined subdivision of the territory is not necessary, we decided to select a subset of the testpoints and use only this subset in the instances. The territorial distribution of the testpoints selected to model the Municipality of Bologna is depicted in Figure 1. The black and blue

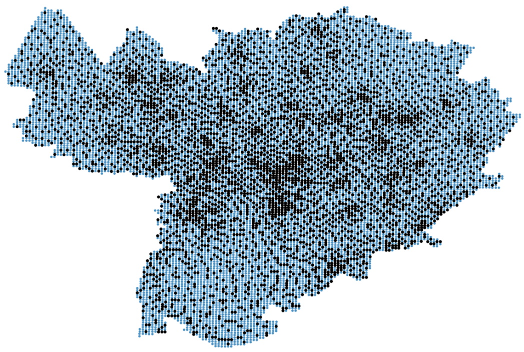

Figure 1: In black the testpoints of the Municipality of Bologna selected for the instances, in blue those discarded.

dots together form the 14 078 testpoints of the Municipality of Bologna identified according to [4, 5, 18]. Only the black ones are selected for optimization purposes.

The selection was made using a k-means clustering procedure with the aim of identifying the most representative testpoints. In fact, the use of a clustering procedure with clusters of different dimensions suggests intuitively the use of different sizes for elementary areas based on the characteristics of the specific area: e.g., larger size in rural areas, smaller size in urban areas. This procedure is widespread in many fields: for example, the famous coarsening and refinement approach of the METIS algorithm [19] suggests a first selection and a subsequent refinement in the parts of the graph where it is necessary. In our case, the areas where there is a greater variation of the received power values in adjacent testpoints, have smaller clusters and therefore a greater number of selected testpoints. Conversely, other areas do not need to be modeled as finely, since the variation of the received power values increases with increasing distance at a slower speed than the previously described areas. In other words, the densification of the selected testpoints reflects the morphology of the territory.

The values of the fixed parameters that are used in the optimization model are collected in Table 3 and are the same for all instances.

Symbol	**Value**	**Notation**
$\|\mathcal{B}\|$	135	Number of potential transmitters
$\|\mathcal{T}'\|$	14 078	Number of receivers according to [4, 5, 18]
$\|\mathcal{T}\|$	4 693	Number of receivers selected for modeling purposes
f	800 MHz	Frequency band
B_w	20 MHz	Channel bandwidth
μ	-130.97 dBW	System noise

Table 3: Values of the parameters.

The instances obtained from this case study are three and are indicated by numbers 1, 2, 3. Each instance differs in the quality of service required (increasing with the number) in the receivers and in the number of receivers that can be covered accordingly, like reported in Table 4. The upper bounds have been estimated a priori not to affect resolution times, therefore a more accurate estimate may be obtained using e.g. Gurobi heuristics before solving the problem entirely.

6.2 Results

The experiments reported in Table 6 aim at analyzing the effect of the reformulating operations discussed in the previous section on the resolution times. The table shows the results obtained when solving instance 1 using the five different formulations described in Table 5. In instance 1, all testpoints are required to be covered, hence for all formulations (B to F) instead of constraints (3) and (4), we used $\sum_{b\in\mathcal{B}} x_{tb} = 1$ for $t \in \mathcal{T}$ since each testpoint must be covered by exactly one base station. The evaluation metrics are model size and sparsity,

Instance	**Value**	**Notation**
1	- 7.56 dB	SINR threshold (δ)
	100%	Desired percentage of territory to be covered
	20	Upper bound of the optimal number of activated base stations (α)
2	0 dB	SINR threshold (δ)
	85%	Desired percentage of territory to be covered
	10	Upper bound of the optimal number of activated base stations (α)
3	+ 7 dB	SINR threshold (δ)
	65%	Desired percentage of territory to be covered
	10	Upper bound of the optimal number of activated base stations (α)

Table 4: Characteristics of the instances.

efficiency in the resolution expressed through the number of explored nodes and resolution times, quality of the root bounds, and accuracy of the optimal solution.

Formulation	**Characteristics**
B	Basic formulation in (5)
C	Formulation in (5) plus the addition of VUBs (6)
D	Formulation C plus the variable replacement in the SINRs like in (7)
E	Formulation C plus the replacement of the SINRs (2) with the aggregate SINRs (8)
F	Final setting in (9)

Table 5: Characteristics of the tested formulations. The letters B and F stand for basic and final, whereas the letters C, D, and E refer to intermediate formulations.

The results under column B report the performance of the basic formulation. First of all, its optimization process has been interrupted before finding the optimal solution as the time exceeded the limit of 36 hours, preventing us from evaluating the accuracy obtained. In fact, the resolution times required by formulation B are very long, confirming this formulation cannot be used in real-life cases. Furthermore, the reported number of activated transmitters, i.e. the value of the optimal solution, is an interval that goes from the best lower bound to the best upper bound found.

Formulation C, which with respect to B contains the VUBs, is reported in order to underline the decisive effect of adding these constraints. Indeed, their addition makes the problem resolution much more efficient: a definitely reduced number of nodes is explored, significantly reducing the Branch-and-Cut tree search time. This may depend on the goodness of the solution found at the end of the root relaxation: the first lower bound identified in B at the root node is 0, which is definitely worse than what found in C, i.e. 11.70, which is instead very close to the optimal solution (equal to 13).

For what concerns the comparison between formulations C and D, a little sparsification and

Instance 1

Parameters	**B**	**C**	**D**	**E**	**F**
Variables	633 690	633 690	633 690	633 690	409 434
Constraints	638 248	1 271 803	1 271 803	642 941	418 686
Non-zeros	86 210 410	87 477 520	87 430 590	3 167 775	2 046 630
Root Relax Sol Value	0.00	11.70	11.70	10.23	10.54
Root Best Lower Bound	1.00	11.99	12.00	11.97	11.99
Root Incumbent	26	18	19	24	17
Root Optimality Gap	96.15%	33.39%	36.84%	50.12%	29.47%
Final Optimality Gap	76.92%	0%	0%	0%	0%
Activated Transmitters	[3,13]	13	13	13	13
Explored Nodes	> 92 637	169	90	39	100
Presolve Time	14	12	11	3	1
Root Relax Sol Time	0	1	1	1	1
B&C Tree Search Time	> 2146	109	74	28	26
Total Resolution Time	> 2160	122	86	32	28
Inst Ideal Min Coverage	100%	100%	100%	100%	100%
Inst Real Coverage	N/A	100%	100%	100%	100%

Key: N/A, not available; Relax Sol, Relaxation Solution; B&C, Branch-and-Cut; Inst, Instance; Min, Minimum

Table 6: Optimization results on instance 1 of Bologna: a comparison between five different formulations (see Table 5 for descriptions). Time is expressed in minutes.

a good improvement in the times can be observed with the sole replacement of the variables in the SINR constraints.
The aggregation of the SINR constraints involves a slight worsening of the bounds but allows to obtain a definitely reduced and more sparse formulation: this is what can be observed by comparing the results under formulations D and E. The sparsity continues to increase after performing coefficient tightening operations: formulation F is characterized by fewer non-zeros.
As for the accuracy of the solutions, ideal and real coverage on the instance coincide in all the available results. This result is probably due to the scaling operation made in advance on the coefficients and to the effectiveness of Gurobi in dealing with numerical issues. Furthermore, the heuristic coefficient tightening seems not to affect the feasible set: ideal and real coverage coincide also in the results of formulation F.
In the end, our formulation F turns out to be much more competitive with respect to the other formulations, especially to the basic formulation B.

In Table 7, a comparison between the basic formulation in (5) and the final formulation in (9) is made for each instance of Bologna. The evaluation metrics are the same of Table

6.

	Instance 1		Instance 2		Instance 3	
Parameters	**B**	**F**	**B**	**F**	**B**	**F**
Variables	633 690	409 434	633 690	409 434	633 690	409 434
Constraints	638 248	418 686	638 249	418 687	638 249	418 687
Non-zeros	86 210 410	2 046 630	86 843 965	2 455 929	86 843 965	2 455 929
Root Relax Sol Value	0.00	10.54	0.00	2.09	0.00	1.30
Root Best Lower Bound	1.00	11.99	1.00	4.28	1.00	3.00
Root Incumbent	26	17	7	-	-	3
Root Optimality Gap	96.15%	29.47%	85.71%	-	-	0.00%
Final Optimality Gap	76.92%	0%	83.33%	0%	-	0.00%
Activated Transmitters	[3,13]	13	[1,6]	6	[1,-]	3
Explored Nodes	> 92 637	100	> 18 171	29 610	> 44 949	1
Presolve Time	14	1	14	3	15	4
Root Relax Sol Time	0	1	1	11	1	3
B&C Tree Search Time	> 2146	26	> 2145	983	> 2144	28
Total Resolution Time	> 2160	28	> 2160	997	> 2160	35
Inst Ideal Min Coverage	100%	100%	85%	85%	65%	65%
Inst Real Coverage	N/A	100%	N/A	94%	N/A	81%
Munic Real Coverage	N/A	100%	N/A	94%	N/A	81%

Key: N/A, not available; Relax Sol, Relaxation Solution; B&C, Branch-and-Cut; Inst, Instance; Min, Minimum; Munic, Municipality

Table 7: Optimization results on the three instances of Bologna: a comparison between the basic (indicated with letter B) and the final (indicated with letter F) formulations. Time is expressed in minutes.

From the results, it turns out that the final model is much more sparse in the coefficients and smaller in the size than the basic. And in fact, the quality of the bounds at the root node is better and each phase of its resolution is faster making the total resolution time heavily reduced. Furthermore, none of the basic formulations have been successfully solved within the time limit of 36 hours, therefore the reported number of activated transmitters is an interval and the evaluation of the coverage was not possible.

Figures 2, 3, 4 show the assignment of testpoints to each activated base station via a color code. The black dots represent the uncovered testpoints, whereas the grey dots represent the testpoints not included in the instance, i.e. the testpoints belonging to $\mathcal{T}' \setminus \mathcal{T}$. The solutions used to depict the figures are those obtained with the optimization of the final formulations. The distribution of testpoints to the serving base stations seems to be coherent according to the positions of the testpoints having the same color as the base station and to the direction in which the signal is emitted by the antenna (illustrated in a simplified fashion with an arrow).

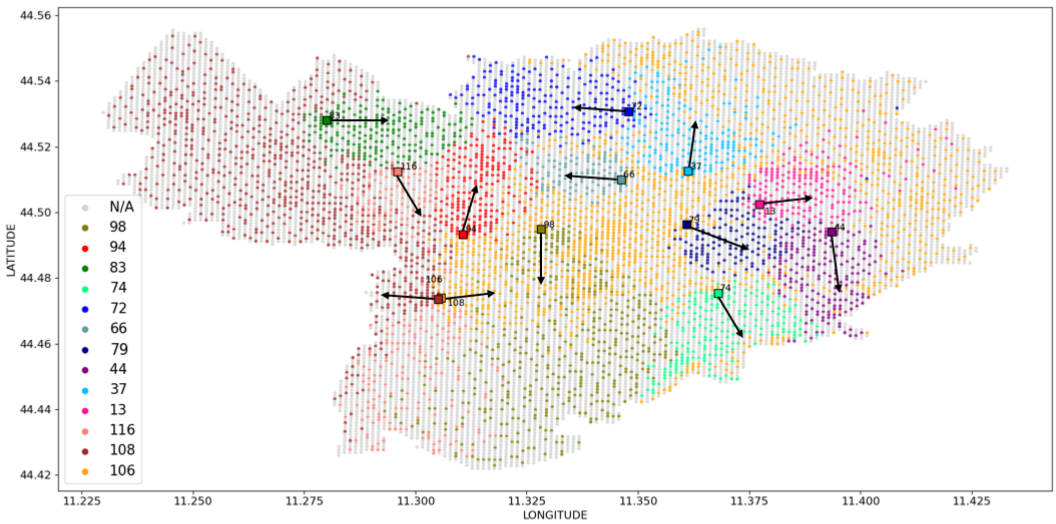

Figure 2: Optimal assignment of testpoints to the base stations activated according to the solution found on instance 1 using the final formulation.
N/A, not available since the testpoint is not in the instance

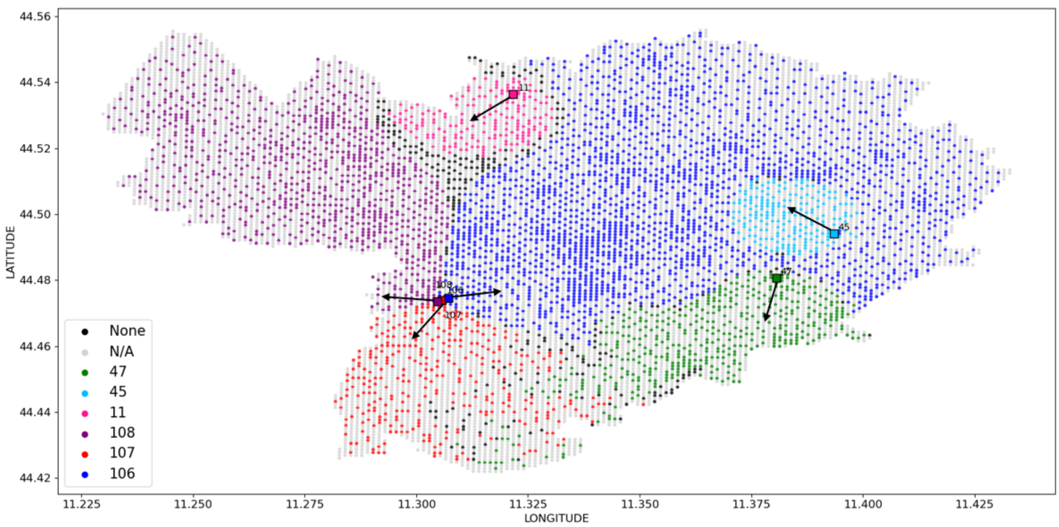

Figure 3: Optimal assignment of testpoints to the base stations activated according to the solution found on instance 2 using the final formulation.
None, the testpoint is covered by none of the activated antennas; N/A, not available since the testpoint is not in the instance

Let us now focus on the coverage obtained by the optimal solutions. The Instance Ideal Minimum Coverage is the minimum percentage of testpoints of the instance that must

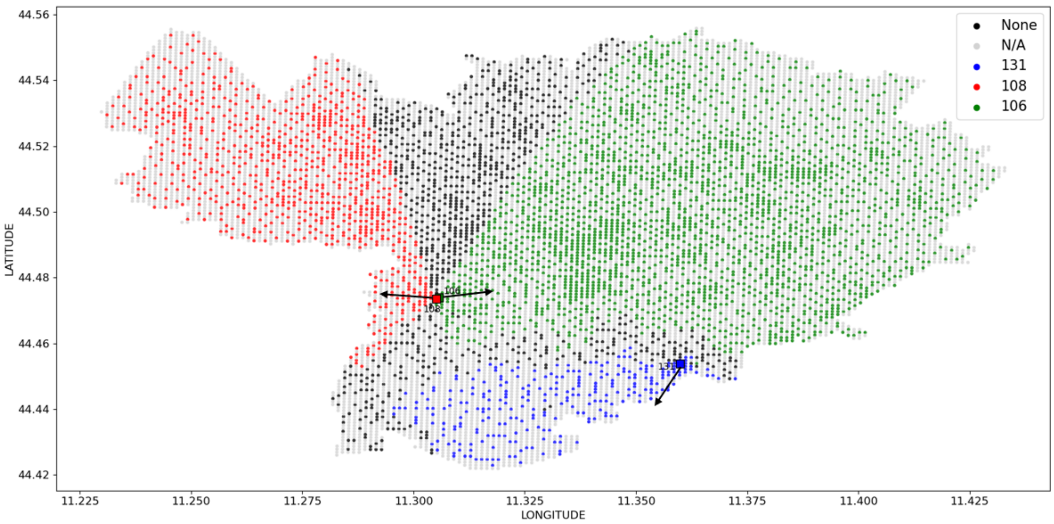

Figure 4: Optimal assignment of testpoints to the base stations activated according to the solution found on instance 3 using the final formulation.
None, the testpoint is covered by none of the activated antennas; N/A, not available since the testpoint is not in the instance

be covered by the optimal solution, whereas the real coverage is the exact percentage of testpoints that is covered by the optimal solution found. The real coverage can be evaluated on the testpoints included in the instance, i.e. included in $\mathcal{T}$, and in this case, is denoted by Instance Real Coverage. Otherwise, it can be evaluated on the testpoints of the entire municipality, i.e. those included in $\mathcal{T}'$, and in this case is denoted by Municipality Real Coverage. The results reported in Table 7 demonstrate that for optimization purposes, the selection of testpoints made through clustering is effective and sufficient, and in fact, the Instance Real Coverage and the Municipality Real Coverage coincide.

7 Conclusions

This study, conducted in cooperation with the Fondazione Ugo Bordoni, mainly focused on tackling numerical and memory issues encountered when solving base station deployment problems. The main scope was to provide a compact formulation suitable for real-life instances. Starting from a basic natural optimization model, we intervened on the formulation by carrying out both strengthening operations (i.e. additions of variable upper bounds, variable replacement, heuristic coefficient tightening, big-M reduction) and size reduction operations (i.e. constraints aggregation). The new formulation was tested to solve the problem in different cases concerning the LTE technology in the Municipality of Bologna. The proposed formulation allows us to efficiently solve the problem to a very high accuracy using commercial solvers.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the Fondazione Ugo Bordoni, and in particular of Manuel Faccioli, Federica Mangiatordi, Luca Rea, Guido Riva, Antonio Sassano.

References

[1] D. Ageyev and A. Al-Ansari, *Lte ran and services multi-period planning*, 2015 Second International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T), 2015, pp. 272–274.

[2] D. Ageyev and A. Al-Anssari, *Optimization model for multi-time period lte network planning*, 2014 First International Scientific-Practical Conference Problems of Infocommunications Science and Technology, 2014, pp. 29–30.

[3] R.V. Akhpashev and A.V. Andreev, *Cost 231 hata adaptation model for urban conditions in lte networks*, 2016 17th International Conference of Young Specialists on Micro/Nanotechnologies and Electron Devices (EDM), 2016, pp. 64–66.

[4] BEREC, 2020. *BEREC Guidelines on Very High Capacity Networks.* `https://berec.europa.eu/eng/document_register/subject_matter/berec/regulatory_best_practices/guidelines/9439-berec-guidelines-on-very-high-capacity-networks` (Accessed June 24, 2021).

[5] BEREC, 2020. *BEREC Guidelines to assist NRAs on the consistent application of Geographical surveys of network deployments.* `https://berec.europa.eu/eng/document_register/subject_matter/berec/regulatory_best_practices/guidelines/9027-berec-guidelines-to-assist-nras-on-the-consistent-application-of-geographical-su` (Accessed June 24, 2021).

[6] O. Bondarenko, D. Ageyev, and O. Mohammed, *Optimization model for 5g network planning*, 2019 IEEE 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM), 2019, pp. 1–4.

[7] L. Chiaraviglio, C. Di Paolo, and N.B. Melazzi, *5G network planning under service and EMF constraints: Formulation and solutions*, IEEE Trans. Mobile Comput. (2021).

[8] F. D'Andreagiovanni, C. Mannino, and A. Sassano, *GUB covers and power-indexed formulations for wireless network design*, Manage. Sci. **59** (2013), 142–156.

[9] A. Dmytro, A.A. Ali, and Q. Nameer, *Multi-period lte ran and services planning for operator profit maximization*, The Experience of Designing and Application of CAD Systems in Microelectronics, 2015, pp. 25–27.

[10] F. D'andreagiovanni, *On improving the capacity of solving large-scale wireless network design problems by genetic algorithms*, European Conference on the Applications of Evolutionary Computation, 2011, pp. 11–20.

[11] F. D'andreagiovanni and A.M. Gleixner, *Towards an accurate solution of wireless network design problems*, International Symposium on Combinatorial Optimization, 2016, pp. 135–147.

[12] F. D'andreagiovanni, H. Lakhlef, and A. Nardin, *A matheuristic for joint optimal power and scheduling assignment in dvb-t2 networks*, Algorithms **13** (2020), 27.

[13] M. Fischetti and A. Lodi, *Heuristics in mixed integer programming*, Wiley Encyclopedia Oper. Res. Manage. Sci. (2010).

[14] FUB, 2021. *Fondazione Ugo Bordoni.* `https://www.fub.it/en/` (Accessed October 4, 2021).

[15] J.D. Gadze, K.A. Agyekum, S.J. Nuagah, and E. Affum, *Improved propagation models for lte path loss prediction in urban & suburban ghana*, arXiv preprint arXiv:2001.05227 (2020).

[16] H. Ganame, L. Yingzhuang, H. Ghazzai, and D. Kamissoko, *5g base station deployment perspectives in millimeter wave frequencies using meta-heuristic algorithms*, Electronics **8** (2019), 1318.

[17] Gurobi, 2021. *Gurobi Optimizer.* `https://www.gurobi.com/products/gurobi-optimizer/` (Accessed October 08, 2021).

[18] Infratel Italia, 2021. *Mappatura 2021 reti fisse a banda ultralarga.* `https://www.infratelitalia.it/archivio-documenti/documenti/mappatura-2021-reti-fisse-a-banda-ultralarga-consultazione-degli-operatori` (Accessed June 24, 2021).

[19] G. Karypis and V. Kumar, *A fast and high quality multilevel scheme for partitioning irregular graphs*, SIAM J. scientific Comput. **20** (1998), 359–392.

[20] C. Mannino, S. Mattia, and A. Sassano, *Planning wireless networks by shortest path*, Comput. Optim. Appl. **48** (2011), 533–551.

[21] C. Mannino, F. Rossi, and S. Smriglio, *The network packing problem in terrestrial broadcasting*, Oper. Res. **54** (2006), 611–626.

[22] J. Naoum-Sawaya and S. Elhedhli, *A nested benders decomposition approach for telecommunication network planning*, Naval Res. Logist. (NRL) **57** (2010), 519–539.

[23] T.S. Rappaport et al., *Wireless communications: principles and practice* Vol. 2, Prentice hall PTR New Jersey, 1996.

[24] M. Rumney et al., *LTE and the evolution to 4G wireless: Design and measurement challenges*, John Wiley & Sons, 2013.

[25] Telco Antennas, 2021. *4G LTE Signal Strength Reference Guide.* https://www.telcoantennas.com.au/blog/guide-to-mobile-networks/4g-lte-signal-strength-reference-guide/ (Accessed May 18, 2021).

[26] Usat, 2021. *Understanding LTE Signal Strength.* https://usatcorp.com/faqs/understanding-lte-signal-strength-values/ (Accessed May 18, 2021).