Networks

WILEY

RESEARCH ARTICLE

Speeding Up the Solution of the Site and Power Assignment Problem in Wireless Networks

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

[1]DING, Università del Sannio, Benevento, Italy | [2]DIAG, Sapienza University of Rome, Rome, Italy

Correspondence: Alice Calamita (alice.calamita@uniroma1.it)

Received: 26 November 2023 | **Revised:** 22 January 2025 | **Accepted:** 30 January 2025

Funding: This author has been partially supported by Fondazione Ugo Bordoni, Rome, Italy.

Keywords: 0-1 linear programming | base station deployment | fixing heuristic | power assignment | reduced cost fixing | wireless network design

ABSTRACT

This paper addresses the optimal design of wireless networks through the site and power assignment problem. Given a set of candidate transmitters, this problem involves choosing optimal transmitter locations and powers to provide service coverage over a target area. In the modern context of increasing traffic, establishing suitable locations and power emissions for the transmitters in wireless networks is a relevant and challenging task due to heavy radio spectrum congestion. Traditional network design formulations are very ill-conditioned and suffer from numerical inaccuracies and limited applicability to large-scale practical scenarios. Our contribution consists of speeding up the solution of the problem under consideration by addressing its drawbacks from a modeling point of view. We propose valid cutting planes and various presolve operations to reduce the problem size and strengthen existing formulations, along with a reduction scheme based on reduced cost fixing to reduce the sources of numerical inaccuracies. Our proposals prove effective, allowing us to achieve optimality on large-scale instances obtained from a real 4G LTE network in solution times aligning well with planning windows.

1 | Introduction

A wireless network is a telecommunications network that uses radio waves, or other wireless communication technologies, to transmit and receive data without the need for physical cables, allowing for the wireless connectivity of devices. From a design point of view, the basic elements of wireless networks are transmitters and receivers. Hence, wireless network design (WND) consists of identifying the proper locations for the transmitters and setting their operational parameters—like frequency and/or power emission—in such a way as to cover the receivers located in the area of interest for providing service. This paper addresses the design of wireless networks for 4G LTE technology and considers the frequency as fixed, thus tackling a WND problem known as the *site and power assignment problem*.

Even if wireless networks rely on different technologies and standards based on the service they are meant to provide, they all share a common feature: The need to reach users scattered over a vast area with a radio signal that must be strong enough to prevail against other unwanted interfering signals. The quality of service indeed depends on the interplay of numerous signals, wanted and unwanted, generated from a large number of transmitting devices. The increasing traffic and the densification of the base stations along the service area have led to an increase in interfering signals. Consequently, establishing suitable locations and power emissions for all the transmitters, coexisting within a heavily congested radio spectrum, has become a challenging and relevant task. Indeed, in the current era of pervasive connectivity, the design of wireless networks plays a pivotal role in shaping modern societal infrastructure. The importance of studying wireless network design lies in its profound implications for enhancing communication and enabling technological advancements, and its influence on the evolving dynamics of global connectivity in daily lives.

1.1 | Literature Overview

As wireless networks are becoming denser, due to technological advancements and increased traffic [1], practitioners' traditional design approach based on trial-and-error supported by simulation has exhibited many limitations. The inefficiency of this approach led to the need for optimization approaches, that became critical for lowering costs and meeting user-demanded service quality standards (see, e.g., [2, 3]). Many optimization models for WND have been investigated over the years. However, the natural formulation on which most models are based presents severe limitations since it involves numerical problems in the problem-solving phase, which emerge even in small instances. Indeed, the constraint matrices of these models contain coefficients that range in a huge interval, as well as large big-M leading to weak bounds. In this paper we review the exact approaches proposed for WND problems, pointing out the contributions that highlight these numerical issues. We recommend [4] for a complete literature review on WND (also including heuristic approaches) and [5] for a thorough overview of the optimization challenges in modern WND.

The exact approaches proposed in the literature are mainly oriented toward non-compact formulations and row-generation methods. In [6], a mixed-integer linear programming (MILP) formulation is introduced, and the proposed exact solution method combines combinatorial and classical Benders decomposition and valid cuts. In [7–10], mixed-integer formulations are used to solve randomly generated instances through standard MIP solvers. In [2], a non-compact 0-1 formulation is investigated. The solution algorithm is based on a row-generation method. The same authors of [2] present a 0-1 model for a WND variant linked to the feasible server assignment problem in [11]. In [12], a non-compact formulation is proposed for the maximum link activation problem; the formulation uses cover inequalities to replace the source of numerical instability. In [13], the source of numerical issues in WND is deeply investigated, and the use of numerically safe LP solvers is suggested to make the solutions reliable. In [14], two non-compact 0-1 LP formulations are proposed for cellular network design under spectral efficiency constraints, and a separation algorithm is also provided. In [4], a compact reformulation for the siting problem has been proposed that allows for the exact solution of large instances. The papers explicitly addressing numerical issues are [2, 4, 11–14].

1.2 | Main Motivation and Contributions

Traditional solution methods, employing (mixed-)integer linear programs with (very) ill-conditioned coefficient matrices, suffer from numerical inaccuracies and limited applicability to large-scale practical scenarios. Indeed, the traditional modeling choice typically includes big-M coefficients to model service coverage conditions, leading to very weak linear relaxations and solutions—returned by state-of-the-art MIP solvers—typically far from the optimum [15]. Our contribution consists of speeding up the solution of the problem under consideration, by addressing its drawbacks, from a modeling point of view. Specifically, we discuss how to improve the natural formulation of the WND problem proposed in the literature by

1. introducing several presolve operations to reduce the number of problem variables and overall problem size;
2. providing valid cliques and variable upper bounds to accelerate solution times;
3. proposing an aggressive reduction scheme based on a reduced cost fixing procedure that strengthens the formulation by reducing the big-M values;
4. providing a resulting model that is compact and a straightforward resolution scheme easily manageable by its users (e.g., telecommunications practitioners).

Although reduced cost fixing is a well-known technique, its application to this problem has never been investigated before (as far as we know) and shows significant potential thanks to the positive impact on the reduction of numerical problems. To improve this technique, we provide a heuristic, producing near-optimal solutions very quickly. Our proposals to reduce memory and numerical issues will allow a rapid solution of large WND instances in line with the times required in the planning phase.

The remainder of this paper is organized as follows. Section 2 presents the problem statement and its mathematical formulation. In Section 3, our contribution to accelerating the problem solution is given. In Section 4, computational results are discussed. Conclusions are provided in Section 5.

2 | Problem Formulation

A wireless network consists of radio transmitters distributing service (i.e., wireless connection) to a target area. The target area is usually partitioned into elementary areas, called testpoints, in line with the recommendations of the telecommunications regulatory bodies. Each testpoint is considered a representative receiver of all the users in the elementary area. Testpoints receive signals from all the transmitters. The power received is classified as serving power if it relates to the signal emitted by the transmitter serving the testpoint; otherwise, it is classified as interfering power (see 4G LTE standard [16]). A testpoint is regarded as served by a base station if the ratio of the serving power to the sum of the interfering powers and noise power (Signal-to-Interference-plus-Noise Ratio or SINR) is above a threshold [17], whose value depends on the desired quality of service.

Since frequency planning is beyond the scope of this study, we assume the frequency channel is given and equal for all the transmitters. We also assume that the power emissions of the activated transmitters can be represented by a finite set of power values, which fits with the standard network planning practice of considering a small number of discrete power values. This practice of power discretization for modeling purposes has been introduced in [2]. Lastly, we assume, without loss of generality, that at least two transmitters will be deployed as serving the target area with a single transmitter is trivial and would not require the use of optimization to be solved.

Let $\mathcal{B}$ be the finite set of potential transmitters and $\mathcal{T}$ be the finite set of receivers located at the testpoints. Let $\mathcal{P} = \{P_1, \dots, P_{|\mathcal{P}|}\}$ be the finite set of feasible power values assumed by the activated transmitters, with $P_1 > 0$ and $P_{|\mathcal{P}|} = P_{\max}$. Hence $\mathcal{L} = \{1, \dots, |\mathcal{P}|\}$ is the finite set of power value indices (or simply power levels). We introduce the variables

$$z_{bl} = \begin{cases} 1 & \text{if transmitter } b \text{ is emitting at power } P_l \\ 0 & \text{otherwise} \end{cases} \qquad b \in \mathcal{B},\ l \in \mathcal{L}$$

and

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

To enforce the choice of only one (strictly positive) power level for each activated transmitter we use

$$\sum_{l \in \mathcal{L}} z_{bl} \leq 1 \qquad b \in \mathcal{B} \tag{1}$$

The mathematical formulation of the WND problem contains the so-called SINR inequalities used to assess service coverage conditions. We recall that a testpoint $t \in \mathcal{T}$ is regarded as served by a base station $\beta \in \mathcal{B}$ if the $\text{SINR}_{t\beta}$ is above a threshold. The $\text{SINR}_{t\beta}$ is given by the ratio of the serving power coming from β and the sum of the interfering powers and noise power measured in t. The serving power is given by the product of the power emitted by the serving base station β and the path loss coefficient, modeling the reduction of power in the signal from β to t. The interfering power is given by the sum of all interfering contributions measured at the testpoint t. These interfering contributions are modeled as the serving power, with the unique difference that the signal is emitted by all the activated base stations, except for the serving one (i.e., except for β). Let $\tilde{a}_{tb} > 0$ be the path loss coefficient applied to the signal received in $t \in \mathcal{T}$ and emitted by $b \in \mathcal{B}$. Let $\mu > 0$ be the system noise. Then a receiver t is served by a base station $\beta \in \mathcal{B}$ if the $\text{SINR}_{t\beta}$ is above a given SINR threshold $\delta > 0$, namely

$$\frac{\tilde{a}_{t\beta} \sum_{l \in \mathcal{L}} P_l z_{\beta l}}{\mu + \sum_{b \in \mathcal{B} \setminus \{\beta\}} \tilde{a}_{tb} \sum_{l \in \mathcal{L}} P_l z_{bl}} \geq \delta \qquad t \in \mathcal{T}, \beta \in \mathcal{B}:\ x_{t\beta} = 1 \tag{2}$$

where the numerator represents the serving signal in t (coming from β), and the denominator is the sum of the noise and the interfering signals in t (coming from all $b \neq \beta$). Following [2], we can rewrite the SINR condition (2) through the big-M constraints

$$\tilde{a}_{t\beta} \sum_{l \in \mathcal{L}} P_l z_{\beta l} - \delta \sum_{b \in \mathcal{B} \setminus \{\beta\}} \tilde{a}_{tb} \sum_{l \in \mathcal{L}} P_l z_{bl} \geq \delta\mu - M_{t\beta}(1 - x_{t\beta})$$
$$t \in \mathcal{T}, \beta \in \mathcal{B} \tag{3}$$

where $M_{t\beta}$ is a large (strictly) positive constant. When $x_{t\beta} = 1$, (3) reduces to (2); when $x_{t\beta} = 0$ and $M_{t\beta}$ is sufficiently large, (3) becomes redundant. We can set, for example,

$$M_{t\beta} = \delta\mu + \delta P_{\max} \sum_{b \in \mathcal{B} \setminus \{\beta\}} \tilde{a}_{tb} \tag{4}$$

Note that we can claim that a testpoint $t \in \mathcal{T}$ is served if and only if there exists at least one (t, β) with $\beta \in \mathcal{B}$ that can satisfy (3) with $x_{t\beta} = 1$.

A constraint to express a minimum service coverage of the service area is included. We assume that each testpoint weights $r_t \in \mathbb{R}$ to account for the fact that testpoints can represent a different number of users or be more or less crucial in service coverage. A minimum service coverage level $r \in \mathbb{R}$ of the testpoint is enforced by the constraint:

$$\sum_{t \in \mathcal{T}} r_t \sum_{b \in \mathcal{B}} x_{tb} \geq r \tag{5}$$

When r_t are all equal to 1, r represents the minimum number of testpoint to be served, and the constraint above corresponds to a territorial service coverage. When $r_t \in [0,1]$ can be interpreted as the fractions of the population living in the elementary area $t \in \mathcal{T}$, then $r \in [0,1]$ represents the fraction of the population to be covered by the service and the constraint represents a population service coverage.

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

$$\sum_{b \in \mathcal{B}} x_{tb} \leq 1 \qquad t \in \mathcal{T} \tag{6}$$

We can observe that when full-service coverage is required (i.e., $r = \sum_{t \in \mathcal{T}} r_t$), all constraints of type (6) are satisfied at equality and become redundant.

The objective functions proposed in the literature of WND are several, going from the maximization of the service coverage to the maximization of the quality of service. A considerable goal to pursue nowadays is the citizens' welfare; therefore, the model we refer to aims at identifying solutions with low environmental impact in terms of electromagnetic pollution and/or power consumption. Reducing electromagnetic pollution indeed involves reducing the power emitted by the transmitters [18]. Hence, we aim to minimize the total number of activated base stations with a penalization on the use of stronger power levels: The cost associated with the use of a power level equal to $l \in \mathcal{L}$, namely c_l, is greater the greater is P_l. For example, we can use $c_l = P_l$ for each $l \in \mathcal{L}$.

Thus, a natural formulation of the WND is the following 0-1 LP model:

$$\min_{x,z} \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} c_l z_{bl}$$
$$(x, z) \in S \tag{7}$$

where the feasible region S is defined as

$$S = \left\{ (x, z) \in \{0,1\}^{n+m} :\ \text{Satisfying (1), (3), (5), and(6)} \right\}$$

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

Sets, parameters, and variables used in the model are summarized in Table 1.

TABLE 1 | Sets, parameters, and variables.

Symbol	Notation
$\mathcal{B}$	Set of potential transmitters
$\mathcal{T}$	Set of receivers
$\mathcal{P}$	Set of feasible power values
$\mathcal{L}$	Set of feasible power indices
$\tilde{a}_{tb}$	Path loss coefficient applied to the signal from transmitter $b \in \mathcal{B}$ to testpoint $t \in \mathcal{T}$
c_l	Cost of using power level $l \in \mathcal{L}$
r_t	Weight of testpoint $t \in \mathcal{T}$
r	Service coverage level
δ	SINR threshold
μ	System noise
z_{bl}	0-1 Variable representing if transmitter $b \in \mathcal{B}$ is emitting at power level $l \in \mathcal{L}$
x_{tb}	0-1 Variable representing if testpoint $t \in \mathcal{T}$ is served by transmitter $b \in \mathcal{B}$

In principle, MIP solvers can solve the model (7). However, it is well-known that the following issues may arise, as widely described, for example, in [2, 13]:

- the power received in each testpoint lies in a large interval, from very small values (10^{-7}) to huge (10^5), which makes the range of the coefficients $\tilde{a}_{tb}$ in the constraint 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 the effectiveness of standard solution procedures;
- real-world problems lead to models with a large number of variables and constraints.

These issues make the solution of real-life instances of this problem very challenging. Therefore, our study aims at making the solution of this problem more efficient and fast, hence practicable.

3 | Our Contribution

We mentioned that practical WND problems are hard to solve using optimal procedures as numerical and memory issues arise even in small instances of the problem. In this section, we discuss how to strengthen natural formulations of the problem by means of presolve operations, valid cutting planes, and a coefficient tightening procedure. All these operations will speed up the solution of the problem and reduce its size, as shown in the computational results section.

3.1 | Presolve Operations

Reducing the model size is crucial as real-life instances typically involve many variables and constraints. In this section, we describe several operations that allow us to reduce the size of the problem by eliminating some x_{tb} and z_{bl} variables a priori. Note that the elimination of the x_{tb} variables also leads to the a priori elimination of the corresponding SINR_{tb} constraint (3). We remark that the preprocessing operations we propose for this problem are variations of similar problem-size reductions proposed in the WND literature (e.g., in [5]).

3.1.1 | Reducing the Number of Servers

Usually, for reasons related to signal quality, only a certain number of transmitters (the ones emitting the strongest signals received from the testpoint) are generally considered as possible servers for a given testpoint. Therefore, we establish a number of possible servers for each testpoint a priori. Namely, for each testpoint t, we select a subset of servers S_t, corresponding to the transmitters emitting the strongest signals received in t.

Then, to further reduce the size of the problem, we delete all the variables modeling transmitter-receiver pairs which any feasible solution would exclude. Let us consider the minimum number of transmitters to be deployed, one server and one interferer. Hence, for each testpoint t, we fix $x_{ts} = 0$ for all those transmitters $s \in S_t$ such that the SINR_{ts} is below the threshold δ for each possible single interferer $b \in \mathcal{B} : b \neq s$. Namely, we eliminate all x_{ts} such that

$$\text{SINR}_{ts} = \frac{\tilde{a}_{ts} P_i}{\mu + \tilde{a}_{tb} P_j} < \delta \qquad \forall b \in \mathcal{B} \setminus \{s\}, \forall P_i, P_j \in \mathcal{P}$$

which can be easily verified directly with

$$\text{SINR}_{ts}^{\max} = \frac{\tilde{a}_{ts} P_{\max}}{\mu + \tilde{a}_{th} P_{\min}} < \delta \qquad \text{with} \quad \tilde{a}_{th} := \min_{b \in \mathcal{B} \setminus \{s\}} \{\tilde{a}_{tb}\}$$

3.1.2 | Reducing Power Levels

To reduce the size of the problem, we also select a subset of transmitter power levels that any feasible solution would exclude. Let us consider again the minimum number of transmitters to be deployed, one server and one interferer. Hence, for each transmitter $b \in \mathcal{B}$ and level of power $l \in \mathcal{L}$, we fix $z_{bl} = 0$ for all those (b, l) such that the SINR is below the threshold δ for each possible interferer $\beta \in \mathcal{B} : \beta \neq b$ and for each testpoint $t \in \mathcal{T}$. Namely, we eliminate all z_{bl} such that

$$\text{SINR}_{tb} = \frac{\tilde{a}_{tb} P_l}{\mu + \tilde{a}_{t\beta} P_j} < \delta \qquad \forall t \in \mathcal{T}, \forall \beta \in \mathcal{B} \setminus \{b\}, \forall P_j \in \mathcal{P}$$

which can be easily verified directly with

$$\text{SINR}_{tb}^{\max} = \frac{\tilde{a}_{tb} P_l}{\mu + \tilde{a}_{th} P_{\min}} < \delta \qquad \forall t \in \mathcal{T}, \text{ with}$$
$$\tilde{a}_{th} := \min_{\beta \in \mathcal{B} \setminus \{b\}} \{\tilde{a}_{t\beta}\}$$

3.1.3 | Heuristic Sparsification

We perform a heuristic sparsification to deal with the numerical issues arising from the coefficients of the SINR inequalities. Specifically, we set a minimum threshold ε on the received power below which the received power can be considered null. Namely, for each $t \in \mathcal{T},\ b \in \mathcal{B},\ P_i \in \mathcal{P}$ we set

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

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

3.2 | Cutting Planes

A standard procedure for solving 0-1 LPs is the branch-and-bound algorithm, which can significantly be improved by cutting planes, that is, inequalities that are valid for all integer solutions but not for some solutions of linear relaxation. By means of such inequalities, fractional linear relaxation solutions can be cut off. Valid inequalities are internally generated by state-of-the-art MIP solvers. However, MIP solvers cannot take advantage of the particular problem structure known to the user. For the problem at hand, we identify some problem-specific cutting planes, including variable upper bounds and families of clique inequalities, that we provide in this section.

3.2.1 | Families of Clique Inequalities

Let us consider again the minimum number of transmitters to be deployed, one server and one interferer. We observe that we can exclude (i) potential levels of power for a certain transmitter $b \in \mathcal{B}$ and (ii) potential serving signals for a certain testpoint $t \in \mathcal{T}$ simply considering the minimum SINR required in t.

Hence, for each testpoint $t \in \mathcal{T}$, interferer $b \in \mathcal{B}$ with power $P_l \in \mathcal{P}$ and server $\beta \in \mathcal{B}$ such that the SINR measured in t and that considers as the only interferer b is always below the threshold δ, that is,

$$\frac{\tilde{a}_{t\beta} P_i}{\mu + \tilde{a}_{tb} P_l} < \delta \quad \forall P_i \in \mathcal{P}$$

which can be easily verified directly with

$$\frac{\tilde{a}_{t\beta} P_{\max}}{\mu + \tilde{a}_{tb} P_l} < \delta \tag{8}$$

we can exclude the possibility that b is activated at a power level equal to l and simultaneously β serves t using

$$z_{bl} + x_{t\beta} \leq 1 \tag{9}$$

Using (1) and (6), we can strengthen the cliques (9).

Theorem 1. *Given $(t, b, \beta, l) \in \{\mathcal{T}, \mathcal{B}, \mathcal{B}, \mathcal{L}\}$ such that (8) is satisfied, with l denoting the minimum power level satisfying (8) and $b \neq \beta$, the following cliques are valid inequalities*

$$\sum_{\lambda \in \mathcal{L}:\ \lambda \geq l} z_{b\lambda} + x_{t\beta} \leq 1 \qquad b \in \mathcal{B}, t \in \mathcal{T}, \beta \in \mathcal{B} \setminus \{b\} \tag{10}$$

Proof. If $x_{t\beta} = 1$, then for each $\lambda \geq l$ we have $z_{b\lambda} = 0$ otherwise the $\text{SINR}_{t\beta}$ constraint (3) is violated. If instead $z_{b\lambda} = 1$ for one $\lambda \geq l$, then $x_{t\beta} = 0$ $(\beta \neq b)$ since the $\text{SINR}_{t\beta}$ constraint (3) is violated. Hence, we cannot have simultaneously that $\sum_{\lambda \in \mathcal{L}:\ \lambda \geq l} z_{b\lambda} = 1$ and $x_{t\beta} = 1$.

Moreover, inequalities (9) are implied by (10) as

$$z_{bl} + x_{t\beta} \leq \sum_{\lambda \in \mathcal{L}:\ \lambda \geq l} z_{b\lambda} + x_{t\beta} \leq 1$$

□

Theorem 2. *Given $(t, b, \beta, l) \in \{\mathcal{T}, \mathcal{B}, \mathcal{B}, \mathcal{L}\}$ such that (8) is satisfied for all $\beta \neq b$, the following cliques are valid inequalities*

$$z_{bl} + \sum_{\beta \in \mathcal{B} \setminus \{b\}} x_{t\beta} \leq 1 \qquad b \in \mathcal{B}, t \in \mathcal{T}, l \in \mathcal{L} \tag{11}$$

Proof. If $z_{bl} = 1$, then each $x_{t\beta} = 0\, (\beta \neq b)$ since the $\text{SINR}_{t\beta}$ constraint (3) is violated. If instead $x_{t\beta} = 1$ for one $\beta \neq b$, then $z_{bl} = 0$ otherwise the $\text{SINR}_{t\beta}$ constraint (3) is violated. Hence, we cannot have simultaneously that $z_{bl} = 1$ and $\sum_{\beta \in \mathcal{B} \setminus \{b\}} x_{t\beta} = 1$.

Moreover, inequalities (9) are implied by (11) as

$$z_{bl} + x_{t\beta} \leq z_{bl} + \sum_{\beta \in \mathcal{B} \setminus \{b\}} x_{t\beta} \leq 1$$

□

Moreover, given the testpoint $t \in \mathcal{T}$ and the server $b \in \mathcal{B}$ with power $P_l \in \mathcal{P}$, if it holds

$$\frac{\tilde{a}_{tb} P_l}{\mu + \tilde{a}_{t\beta} P_j} < \delta \quad \forall \beta \in \mathcal{B} \setminus \{b\}, \forall P_j \in \mathcal{P}$$

then b cannot serve t. This condition is verified for all $\beta \in \mathcal{B} \setminus \{b\}$ (whatever is their power level $P_j \in \mathcal{P}$) if it holds for the minimum possible interferer at the minimum power level $P_{\min}$. Hence, for all (t, b, l) such that

$$\frac{\tilde{a}_{tb} P_l}{\mu + \tilde{a}_{th} P_{\min}} < \delta, \qquad \tilde{a}_{th} := \min_{\beta \in \mathcal{B} \setminus \{b\}} \{\tilde{a}_{t\beta}\} \tag{12}$$

we can exclude that b is activated at a power level equal to l and simultaneously b serves t, namely we can impose

$$z_{bl} + x_{tb} \leq 1 \qquad \text{for all } (t, b, l) \text{ satisfying (12)} \tag{13}$$

Using (1), we can strengthen the cliques (13).

Theorem 3. *Given $(t, b, l) \in \{\mathcal{T}, \mathcal{B}, \mathcal{L}\}$ such that (12) is satisfied and l corresponds to the maximum power level such that (12) is satisfied, the following cliques are valid inequalities*

$$\sum_{\lambda \in \mathcal{L}:\ \lambda \leq l} z_{b\lambda} + x_{tb} \leq 1 \qquad b \in \mathcal{B}, t \in \mathcal{T} \tag{14}$$

Proof. If $x_{tb} = 1$, then for each $\lambda \leq l$ we have $z_{b\lambda} = 0$, otherwise the SINR_{tb} constraint (3) is violated. If instead $z_{b\lambda} = 1$ for one $\lambda \leq l$, then $x_{tb} = 0$ since the SINR_{tb} constraint (3) is violated. Hence, we cannot have simultaneously that $\sum_{\lambda \in \mathcal{L}:\ \lambda \leq l} z_{b\lambda} = 1$ and $x_{tb} = 1$.

Moreover, inequalities (13) are implied by (14) as

$$z_{bl} + x_{tb} \leq \sum_{\lambda \in \mathcal{L}:\ \lambda \leq l} z_{b\lambda} + x_{tb} \leq 1$$

□

3.2.2 | Variable Upper Bounds

Variable upper bound constraints (VUBs)

$$x_{tb} \leq \sum_{l\in\mathcal{L}} z_{bl} \qquad t\in\mathcal{T}, b\in\mathcal{B} \tag{15}$$

enforce that a testpoint $t \in \mathcal{T}$ can be assigned to a transmitter $b \in \mathcal{B}$ only if b is activated, that is, only if b is using one strictly positive power level. They are known to strengthen the quality of linear relaxation significantly.

Let us denote by $\mathcal{L}_{tb} \subseteq \mathcal{L}$ the subset of power levels l that satisfy (12) for a given transmitter-receiver pair $(b,t) \in \{\mathcal{B},\mathcal{T}\}$, meaning that we can exclude that b is activated at a power level $l \in \mathcal{L}_{tb}$ and simultaneously b serves t. The VUBs (15) can be tightened to

$$x_{tb} \leq \sum_{l\in\mathcal{L}\setminus\mathcal{L}_{tb}} z_{bl} \qquad t\in\mathcal{T}, b\in\mathcal{B} \tag{16}$$

3.3 | Tightening Procedure for the Big-*M*: A Reduced Cost Fixing Approach

To further reduce the model size, we propose a reduced cost fixing method, in short RCF (see [19] for a survey on presolve techniques). Although this procedure is well-known and widespread, no one has ever tried to see its effects on this type of problem (based on our knowledge).

By solving the linear relaxation of the problem, we can get the lower bound LBD and the corresponding reduced costs $\overline{c}_{bl}$ associated with the sole variables z_{bl} in the optimal solution of the linear relaxation. Then, given an upper bound UBD > LBD, if $\overline{c}_{bl} \geq \text{UBD} - \text{LBD}$ for some $b \in \mathcal{B}$, $l \in \mathcal{L}$, the corresponding z_{bl} must be at its lower bound in every optimal solution; hence we can fix $z_{bl} = 0$.

Whenever the fixing of a variable z_{bl} occurs at a given $l \in \mathcal{L}$ such that $P_l = P_{\max}$, we can recompute and reduce the big-M, resulting in a tightening of the formulation. Indeed, after the RCF, the transmitter b may no longer be able to emit at the maximum power level l such that $P_l = P_{\max}$, since the corresponding z_{bl} variable has been fixed to zero. In such cases, the value by which a_{tb} is weighted in the big-M (see (4)) can be reduced to the highest power value that b can assume, which is strictly less than $P_{\max}$.

To formalize it, let us define the set $\mathcal{B}^R \subseteq \mathcal{B}$ of base stations affected by RCF, that is, such that $b \in \mathcal{B}^R$ if the variable z_{bl} has been fixed to zero for at least one $l \in \mathcal{L}$. Then, for a given $b \in \mathcal{B}^R$, we can define the set $\mathcal{L}_b^R \subset \mathcal{L}$ of power levels that b can assume after the RCF. We denote by $P_{b,\max}^R$ the power value corresponding to the maximum power level that $b \in \mathcal{B}^R$ can assume. Since $\mathcal{L}_b^R \subset \mathcal{L}$, we have that $P_{b,max}^R \leq P_{\max}$. Using this notation, we can write down the new value of the big-M

$$\begin{aligned} M'_{t\beta} &= \delta\mu + \delta\left(P_{\max}\sum_{b\in\mathcal{B}\setminus\{\beta,\mathcal{B}^R\}} \tilde{a}_{tb} + \sum_{b\in\mathcal{B}^R} P_{b,\max}^R \tilde{a}_{tb}\right) \\ &= \delta\mu + \delta\sum_{b\in\mathcal{B}\setminus\{\beta\}} \tilde{P}_b \tilde{a}_{tb} \end{aligned}$$

which satisfies $M_{t\beta} \geq M'_{t\beta}$ since $P_{\max} \geq \tilde{P}_b = \begin{cases} P_{\max} & \text{if } b\in\mathcal{B}\setminus\mathcal{B}^R \\ P_{b,\max}^R & \text{if } b\in\mathcal{B}^R. \end{cases}$

The smaller the optimality gap given by the estimated lower and upper bounds, the greater the number of z_{bl} variables that can be fixed to zero, and the smaller the big-M coefficients. Hence, applying a standard algorithm—as implemented in MIP commercial solvers—to the tightened formulation (i.e., the formulation got after the RCF) produces stronger bounds and a faster solution.

We can apply a further reduction. Given an upper bound UBD, the largest number of activated transmitters γ can be obtained by considering all transmitters at the minimum power level $l = 1$, which means all transmitters with a cost equal to c_1 and hence that $\text{UBD} \leq \gamma c_1$, from which we have $\gamma = \lceil \frac{\text{UBD}}{c_1} \rceil$. Given the information on the maximum number of transmitters that can be installed, we can further reduce the value of the big-M by replacing the sum of all the interfering signals in the testpoint t with the sum of the γ *strongest* interfering signals in t. Given γ, the big-M can be computed as

$$M''_{t\beta} = \delta\mu + \delta\sum_{b\in\mathcal{A}_t\setminus\{\beta\}} \tilde{P}_b\tilde{a}_{tb} \leq M'_{t\beta} \leq M_{t\beta} \tag{17}$$

where $\mathcal{A}_t \subset \mathcal{B}$ is the set of the γ base stations emitting the strongest signals received in t, that is, $|\mathcal{A}_t| = \gamma$. The smaller is γ, and the smaller is the big-M; therefore, the estimate of γ should be as accurate as possible.

We observe that getting a good lower bound is straightforward after the inclusion of all the proposed VUBs and cliques. Conversely, finding a good upper bound is a more daunting task. Although commercial MIP solvers can be used to derive a feasible solution, it may be time-consuming. Consequently, we derived a fixing heuristic based on the observation that the fractional values of the variables in the LP relaxation are often good predictors of zero/non-zero variables in an optimal ILP solution. This especially occurs when the LP relaxation is extremely tight. Our fixing heuristic is made up of the following steps:

1. Solve the LP relaxation of the problem and take the fractional solution
2. Identify the set of fractional variables that are likely to be zero (i.e., whose value is less than a very low threshold in the fractional solution)
3. Fix the selected variables to zero and perform bound strengthening to propagate implications
4. Solve the resulting ILP problem to get a near-optimal feasible solution.

3.4 | The Final Formulation

The final formulation we propose differs from the initial formulation (7) since:

TABLE 2 | Characteristics of the instances: SINR threshold (δ) and fraction of population to be served (r).

Instance	BO1	BO2	BO3	BO4	BO5	BO6	BO7	BO8	BO9	BO10	BO11	BO12	BO13	BO14	BO15	BO16
δ (dBW)	−10	−9.5	−9	−8.5	−8	−7.5	−7	−46.5	−6	−5.5	−5	−3	0	+3	+5	+7
r	1	1	1	1	1	1	1	1	1	0.99	0.99	0.95	0.85	0.75	0.70	0.65

1. the number of servers and the number of levels for the power has been reduced according to the three operations described in Section 3.1;
2. it includes the addition of the VUBs (16) and the cliques (10), (11), and (14);
3. it is the result of an RCF procedure;
4. thanks to RCF, we can use the smaller big-M in (17) to formulate the SINR constraints.

4 | Computational Experience

The code has been implemented in Python and the experiments have been carried out on a Ubuntu server with an Intel(R) Xeon(R) Gold 5218 CPU running at 2.30 GHz, with 96 GB of RAM and 8 cores. Gurobi Optimizer 10.0.1 [20] with default settings has been employed as a MIP solver. We set a time limit of 4 h for computation time.

4.1 | Testbed

The testbed is made of instances obtained from an existing 4G LTE 800 MHz network in the Municipality of Bologna (Italy) provided by the Fondazione Ugo Bordoni (FUB) [21]. The system noise and the power received in each receiver by each possible transmitter have been simulated by the FUB using the *Cost Hata* model [16, 22, 23]. The power values are in W and scaled by a factor of 10^{10} to avoid numerical issues and obtain better accuracy on optimal solutions, as suggested in [13]. Accordingly, the threshold on the quality of service δ and the system noise μ are expressed in W; μ is also scaled by 10^{10}. The emitted power values considered in this study are three (hence three power levels): 20, 40, and 80 W, corresponding to the typical values considered in planning situations (i.e., the minimum, medium, and maximum power). However, to investigate the impact of the number of power levels on solution times, we conducted additional tests, reported in the sensitivity analysis section, where the number of power levels is doubled. The threshold ϵ has been set around −110 dBmW [24, 25]. In (5), the parameter r_t has been set to the fraction of the population living in the elementary area $t \in \mathcal{T}$, and r to the minimum fraction of the population to serve. We used $|S_t| = 10$, namely, we selected the strongest ten signals as possible servers for each testpoint t. Indeed, this value proved reasonable in our instances, as larger values, which may lead to more difficult problems, offer no improvement in the solution (weak signals are either eliminated during presolve operations or fail to meet SINR requirements).

From this network, we derived 16 instances (BO1–BO16), each of them having $|\mathcal{B}| = 135$ and $|\mathcal{T}| = 4693$. Each instance differs in the fraction of the population to be served (r) and the quality of service required (given by δ, chosen in the typical range of LTE services), as reported in Table 2. The estimate of the upper bound has then been obtained on each instance using the fixing heuristic.

TABLE 3 | Characteristics of the tested formulations for the variable-power case.

Formulation	Characteristics
B	Basic formulation (7)
B+CPs	Formulation (7) plus the addition of cutting planes (i.e., VUBs (16) and cliques (10), (11), and (14))
P	Presolve operations reported in Section 3.1 applied to the basic formulation (7)
P+CPs+RCF	Final setting reported in Section 3.4, including RCF steps 3 and 4
P+CPs	Final setting reported in Section 3.4, excluding RCF steps 3 and 4

4.2 | Results

In this section, we show the impact of the operations discussed in Section 3. To this end, we compare the results obtained using formulation (7)—denoted as the basic formulation B—and the final setting described in Section 3.4. Table 3 provides an explicit description of the tested formulations. We use three evaluation criteria, namely the size, the sparsity, and the quality of the gaps at the root node and the end of the optimization.

We observe that we differentiate between two distinct final settings in our analysis: P+CPs+RCF, where the RCF procedure is applied along with a corresponding reduction of the big-M value, and P+CPs, where these procedures are not included. This distinction is necessary as RCF can only be applied to certain instances. Specifically, instances requiring low to medium service values (BO1– to BO11) can successfully be solved using RCF. For instances with high service demands (BO12–BO16, that is, $\delta > -5$ dBW), RCF is not a viable procedure, as both Gurobi and our heuristic fail to find good feasible solutions to be used in the RCF within a reasonable time. Hence, we apply to set P+CPs+RCF when possible, otherwise, we apply to set P+CPs.

We report Table 4 showing the average number of variables, constraints, and non-zeros of each formulation. Figure 1 instead shows four box plots, one for each formulation, displaying the distribution of the data concerning the relative gap at the root node and the end of the optimization, based on a summary of five numbers (minimum value, first quartile, median value, third quartile, and maximum value).

From Table 4 and Figure 1 we can claim that:

- formulation B+CPs, leads to a significant improvement in the root relaxation value, affecting the gaps, revealing the decisive effect of adding the cutting planes;
- formulation P, leads to a significant reduction in the number of variables, constraints, and non-zeros, however, this is not enough to get an improvement in the solution times (both root and final gaps are still very high)
- the aggressive reduction scheme in P+CPs+RCF, combining all the operations proposed, leads to a definitely reduced and sparser formulation;
- the lighter reduction scheme in P+CPs, not including the RCF procedure, leads to a reduced and sparser formulation compared to B+CPs, but slightly bigger and less sparse than P+CPs+RCF;
- the smallest gaps are achieved using P+CPs+RCF or P+CPs; however, good final gaps can be reached already using B+CPs.

The optimization results are reported in Table 5. The evaluation metrics considered are: number of variables (NumVars), constraints (NumConstrs) and non-zeros (NumNZs); relative optimality gap at the root node (RootGap), final relative optimality gap (Gap); number of explored nodes (Nodes); total solution time (Time). Denoting the optimal value with Opt, the value of the lower bound at the root node with RootLB, and the value of the best lower bound with BestLB, the RootGap is computed as 100(Opt-RootLB)/Opt, and the Gap as 100(Opt-BestLB)/Opt. We reported in bold the gap of instances solved to optimality within the time limit.

TABLE 4 | Average number of variables, constraints, and non-zeros in the tested formulations.

Formulation	NumVars	NumConstrs	NumNZs
B	6.34E+05	6.38E+05	2.56E+08
B+CPs	6.34E+05	9.27E+05	2.57E+08
P	4.73E+04	5.17E+04	1.24E+07
P+CPs+RCF	4.72E+04	1.93E+05	8.05E+06
P+CPs	4.73E+04	6.62E+05	1.48E+07

Results show that formulations P+CPs+RCF and P+CPs are definitely reduced in size and sparser than B. This is mainly due to the presolve operations and the reduced cost-fixing procedure (when viable). The quality of the bounds at the root node is better in P+CPs+RCF and P+CPs, mainly due to the addition of cutting planes. Thanks to a good root bound, the number of the explored nodes is heavily reduced in P+CPs+RCF and P+CPs, and consequently, the time spent on the branch-and-cut tree search. Overall, solution times are significantly reduced if we compare P+CPs+RCF or P+CPs to B. Indeed, none of the basic formulations have been successfully solved within the time limit of 4 h, whereas all final formulations but one have been solved to optimality in less than 45 min.

Regarding the comparison between the final settings and B+CPs, we observe that the final settings lead to smaller and sparser formulations, and also to faster solution times. Indeed, with B+CPs we can solve only one instance to optimality within the time limit. As for the comparison between the final settings and P, even if the sparsity and the number of variables are similar, the lack of cutting planes in P leads to very poor relaxation bounds, making it impossible to solve the instances within the time limit.

In the end, setting P+CPs+RCF, or P+CPs in case RCF is not viable, turns out to be much more competitive than B or B+CPs or P.

In Table 6 we also report the computational time of every phase of the setting P+CPs+RCF.

In particular, by LBTime we denote the time needed to get the lower bound given by the linear relaxation of the problem, by

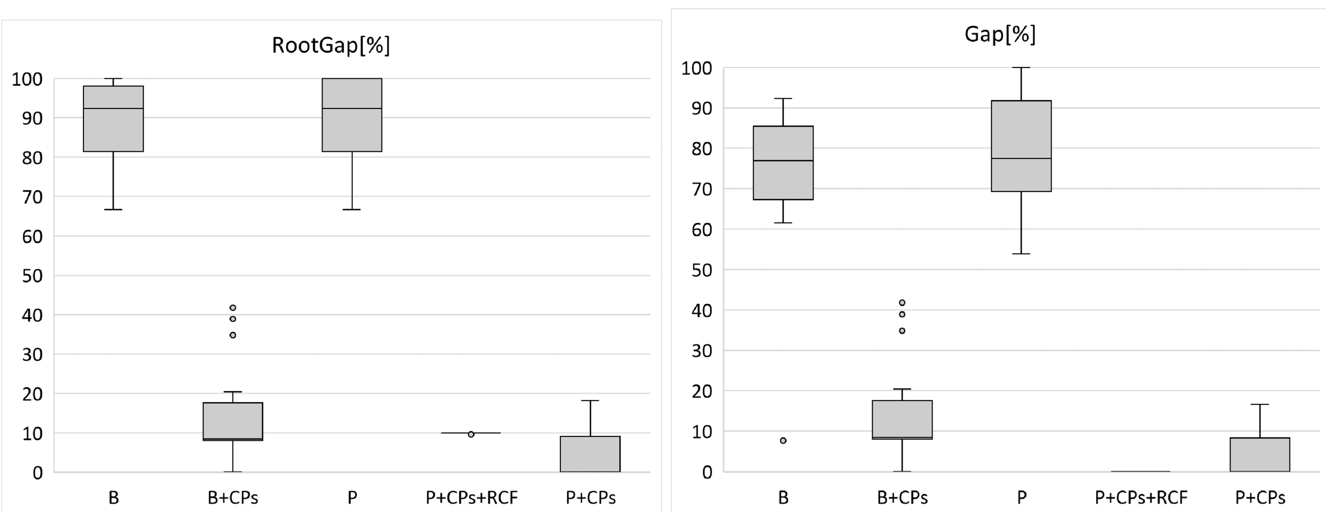

FIGURE 1 | Box plots of the relative gap at the root node (on the left) and the end of the optimization phase (on the right) in the tested formulations (see Table 3).

TABLE 5 | Optimization results for the variable-power case.

Instance	Formulation	NumVars	NumConstrs	NumNZs	RootGap (%)	Gap (%)	Nodes	Time (s)
BO1	B	633 960	638 383	2.56E+08	100.00	92.31	1484	TL
	B+CPs	633 960	750 990	2.57E+08	7.68	**0.00**	74	14 354.51
	P	47 335	51 758	1.24E+07	100.00	100.00	1442	TL
	P+CPs+RCF	47 201	164 365	8.17E+06	9.97	**0.00**	83	429.96
BO2	B	633 960	638 383	2.56E+08	100.00	7.69	1503	TL
	B+CPs	633 960	755 652	2.57E+08	8.06	8.06	1	TL
	P	47 335	51 758	1.24E+07	100.00	69.23	14 557	TL
	P+CPs+RCF	47 201	169 027	8.19E+06	9.97	**0.00**	110	534.67
BO3	B	633 960	638 383	2.56E+08	100.00	69.23	14 472	TL
	B+CPs	633 960	760 640	2.57E+08	8.32	8.32	1	TL
	P	47 335	51 758	1.24E+07	100.00	69.23	14 766	TL
	P+CPs+RCF	47 201	174 015	8.20E+06	9.97	**0.00**	122	518.33
BO4	B	633 960	638 383	2.56E+08	92.31	76.92	14 505	TL
	B+CPs	633 960	765 973	2.57E+08	8.17	8.17	1	TL
	P	47 335	51 758	1.24E+07	100.00	92.31	14 752	TL
	P+CPs+RCF	47 201	179 348	8.22E+06	9.97	**0.00**	129	471.71
BO5	B	633 960	638 383	2.56E+08	100.00	76.92	12 182	TL
	B+CPs	633 960	771 677	2.57E+08	8.32	8.32	1	TL
	P	47 335	51 758	1.24E+07	100.00	92.31	14 657	TL
	P+CPs+RCF	47 201	185 052	8.27E+06	9.97	**0.00**	105	541.17
BO6	B	633 960	638 383	2.56E+08	92.31	76.92	15 032	TL
	B+CPs	633 960	777 637	2.57E+08	8.45	8.45	1	TL
	P	47 335	51 758	1.24E+07	92.31	69.23	14 540	TL
	P+CPs+RCF	47 200	190 259	8.24E+06	9.97	**0.00**	151	680.76
BO7	B	633 960	638 383	2.56E+08	92.31	61.54	14 708	TL
	B+CPs	633 960	784 138	2.57E+08	8.49	8.49	1	TL
	P	47 335	51 758	1.24E+07	92.31	69.23	15 156	TL
	P+CPs+RCF	47 200	197 513	8.27E+06	9.97	**0.00**	293	586.75
BO8	B	633 960	638 383	2.56E+08	92.31	84.62	14 353	TL
	B+CPs	633 960	791 167	2.57E+08	8.41	8.41	1	TL
	P	47 335	51 758	1.24E+07	92.31	92.31	14 439	TL
	P+CPs+RCF	47 201	204 542	8.32E+06	9.97	**0.00**	127	598.99
BO9	B	633 960	638 383	2.56E+08	92.31	69.23	14 793	TL
	B+CPs	633 960	798 671	2.57E+08	8.53	8.53	1	TL
	P	47 335	51 758	1.24E+07	92.31	53.85	14 743	TL
	P+CPs+RCF	47 202	212 046	8.39E+06	9.97	**0.00**	86	537.43
BO10	B	633 960	638 384	2.57E+08	90.00	90.00	14 756	TL
	B+CPs	633 960	806 851	2.57E+08	9.01	9.01	1	TL
	P	47 335	51 759	1.24E+07	90.00	90.00	29 536	TL
	P+CPs+RCF	47 158	220 226	7.11E+06	9.63	**0.00**	37	1922.65
BO11	B	633 960	638 384	2.57E+08	90.00	90.00	14 753	TL
	B+CPs	633 960	815 638	2.57E+08	4.54	4.54	1	TL
	P	47 335	51 759	12 423 085	90.00	90.00	30 179	TL
	P+CPs+RCF	47 157	229 013	7.16E+06	9.63	**0.00**	37	2311.70

(Continues)

TABLE 5 | (Continued)

Instance	Formulation	NumVars	NumConstrs	NumNZs	RootGap (%)	Gap (%)	Nodes	Time (s)
BO12	B	633 960	638 384	2.57E+08	85.71	85.71	24 276	TL
	B+CPs	633 960	860 727	2.57E+08	0.00	**0.00**	1	10 587.45
	P	47 335	51 759	1.24E+07	85.71	85.71	29 363	TL
	P+CPs	47 335	274 102	1.33E+07	0.00	**0.00**	1	181.99
BO13	B	633 960	638 384	2.57E+08	80.00	80.00	14 964	TL
	B+CPs	633 960	981 892	2.58E+08	20.42	20.42	1	TL
	P	47 335	51 759	1.24E+07	80.00	80.00	14 684	TL
	P+CPs	47 335	395 267	1.38E+07	18.20	16.60	3659	TL
BO14	B	633 960	638 384	2.57E+08	75.00	75.00	14 641	TL
	B+CPs	633 960	1 227 228	2.59E+08	38.90	38.90	1	TL
	P	47 335	51 759	1.24E+07	75.00	75.00	29 705	TL
	P+CPs	47 335	640 603	1.46E+07	0.00	**0.00**	1	2128.19
BO15	B	633 960	638 384	2.57E+08	66.67	66.67	14 800	TL
	B+CPs	633 960	1 455 551	2.60E+08	34.83	34.83	1	TL
	P	47 327	51 751	1.24E+07	66.67	66.67	30 451	TL
	P+CPs	47 327	867 412	1.55E+07	0.00	**0.00**	1	1503.43
BO16	B	633 960	638 384	2.57E+08	66.67	66.67	15 068	TL
	B+CPs	633 960	1 729 581	2.61E+08	41.77	41.77	1	TL
	P	47 281	51 705	1.24E+07	66.67	66.67	29 797	TL
	P+CPs	47 281	1 132 474	1.67E+07	0.00	**0.00**	1	2556.58

Note: Time is expressed in seconds.
Abbreviation: TL, time limit reached.

TABLE 6 | Detailed solution times on the final formulation.

Instance	LBTime (s)	UBTime (s)	SolTime (s)	Time (s)
BO1	34.85	31.53	363.58	429.96
BO2	33.61	34.56	466.50	534.67
BO3	32.79	34.88	450.66	518.33
BO4	37.57	41.40	392.74	471.71
BO5	33.25	42.21	465.71	541.17
BO6	39.08	47.49	594.19	680.76
BO7	34.14	52.14	500.47	586.75
BO8	33.75	54.54	510.70	598.99
BO9	36.43	64.36	436.64	537.43
BO10	47.49	28.61	1846.55	1922.65
BO11	52.87	29.39	2229.44	2311.70

Note: Time is expressed in seconds.

UBTime the time needed to get an upper bound using our fixing heuristic, by SolTime the time needed to solve the problem after the RCF procedure, and by Time the overall solution time is given by the sum of the previous components. The efficiency in solving the instances using the scheme based on RCF is attributed to a rapid computation of the lower and upper bounds. Both depend on the good quality of the linear relaxation of the problem, achieved thanks to the inclusion of the cuts we introduced. Note that for higher-quality values, unfortunately, the lower bound gets slightly worse and it is no longer possible to exploit our fixing heuristic. We also tried to exploit Gurobi as a heuristic to obtain an upper bound on high service instances, but unfortunately, the bounds produced by Gurobi in times compatible with the use of this procedure were not good enough to actually apply the RCF.

TABLE 7 | Maximum big-M values on the instances BO1–BO11 where we can apply RCF.

Instance	Formulation	MaxBigM
BO1	B	91 958.73
	P+CPs+RCF	45 979.28
BO2	B	103 179.39
	P+CPs+RCF	103 179.20
BO3	B	115 769.18
	P+CPs+RCF	115 768.96
BO4	B	129 895.16
	P+CPs+RCF	64 947.46
BO5	B	145 744.77
	P+CPs+RCF	145 744.49
BO6	B	161 284.63
	P+CPs+RCF	80 642.16
BO7	B	183 481.79
	P+CPs+RCF	183 481.44
BO8	B	205 869.96
	P+CPs+RCF	205 869.56
BO9	B	230 989.89
	P+CPs+RCF	115 494.79
BO10	B	259 174.92
	P+CPs+RCF	259 174.08
BO11	B	290 799.04
	P+CPs+RCF	142 654.51

Finally, Table 7 compares the maximum values of the big-M in B and P+CPs+RCF. We observe that the value of the largest big-M used in P+CPs+RCF is half that used in B in half the instances.

TABLE 8 | Optimization results with six power levels. Time is expressed in seconds.

Instance	Formulation	NumVars	NumConstrs	NumNZs	RootGap (%)	Gap (%)	Nodes	Time (s)
BO1	B	634 365	638 383	5.11E+08	100.00	100.00	1931	TL
	P+CPs+RCF	47 584	183 138	2.05E+07	10.00	**0.00**	140	1022.79
BO4	B	634 365	638 383	5.11E+08	100.00	100.00	2012	TL
	P+CPs+RCF	47 584	202 538	2.06E+07	10.00	**0.00**	91	1203.81
BO7	B	634 365	638 383	5.11E+08	100.00	100.00	2233	TL
	P+CPs+RCF	47 586	225 923	2.08E+07	10.00	**0.00**	75	1174.86
BO10	B	634 365	638 384	5.11E+08	100.00	100.00	7476	TL
	P+CPs+RCF	47 294	255 200	1.13E+07	9.60	**0.00**	46	672.64
BO13	B	634 365	638 384	5.11E+08	100.00	100.00	7262	TL
	P+CPs	47 740	478 053	2.73E+07	18.20	10.00	4666	TL
BO16	B	634 365	638 384	5.11E+08	100.00	100.00	7330	TL
	P+CPs	47 686	1 426 379	3.28E+07	0.00	**0.00**	1	5708.71

Note: TL, time limit reached.

TABLE 9 | Comparison of the (#3) vs. (#6) power levels instances: % changes on the KPIs obtained as (KPI(#6)- KPI(#3))/min{KPI(#6), KPI(#3)}.

Instance	Formulation	NumVars (%)	NumConstrs (%)	NumNZs (%)	Time (s)
BO1	B	0.06	0.00	99.51	TL
	P+CPs+RCF	0.81	11.42	150.85	137.88%
BO4	B	0.06	0.00	99.51	TL
	P+CPs+RCF	0.81	12.93	150.51	155.20%
BO7	B	0.06	0.00	99.51	TL
	P+CPs+RCF	0.82	14.38	151.39	100.23%
BO10	B	0.06	0.00	99.31	TL
	P+CPs+RCF	0.29	15.88	59.47	−185.84%
BO13	B	0.06	0.00	99.31	TL
	P+CPs	0.86	20.94	98.75	TL
BO16	B	0.06	0.00	99.31	TL
	P+CPs	0.86	2595	96.72	123.29%

Note: TL, time limit reached.

4.3 | Sensitivity Analysis

In the previous section, we presented results for problems with three power levels. To assess the impact of increasing the number of power levels on solution times, we conducted additional experiments on a subset of testbed instances (one in three) using six power levels: 20, 30, 40, 50, 60, and 80 W. We compared the characteristics and key performance indicators (KPIs) of the basic formulation B with those of the final settings P+CPs+RCF or P+CPs when RCF cannot be applied. The numerical results are summarized in Table 8, while Table 9 shows the percentage changes in the number of variables, constraints, non-zeros, and solution times between the three (#3) and six (#6) power level instances. The results in Table 9 indicate that doubling the number of power levels increases problem size and solution times. However, our proposed approach remains effective (see Table 8): The final settings produce smaller and sparser problems that significantly reduce solution times compared to the basic formulation. Furthermore, with the final settings, all the tested instances except one are solved within the time limit, aligning well with planning time windows.

5 | Conclusions

In this paper, we worked on improving the natural formulation proposed in the literature on wireless network design for the site and power assignment problem. Specifically, we provided several presolve operations to reduce the number of problem variables and overall problem size, along with valid cliques and variable upper bounds to reduce solution times. We also proposed an aggressive reduction scheme based on a reduced cost-fixing procedure that reduces the big-M values, strengthening the formulation and reducing the problem size. Computational tests were conducted on authentic 4G LTE data from the Municipality of Bologna in Italy. The results confirmed the efficacy of our proposals, which, compared to the standard solution of traditional problem formulation, resulted in faster solution times. Indeed, only with our proposals, we could efficiently solve large instances of the problem to optimality, in solution times consistent with planning operations. However, the reduced cost fixing procedure can be used only when proper upper bounds can be computed quickly: This is not trivial in all cases, particularly when high-quality service is required.

Acknowledgments

This research was carried out in collaboration with Fondazione Ugo Bordoni, a higher education and research institution under the supervision of the Italian Ministry of Enterprises and Made in Italy that operates in the telecommunication field, providing innovative services for government bodies. The authors gratefully acknowledge the support of Fondazione Ugo Bordoni and, in particular, Manuel Faccioli, Federica Mangiatordi, Luca Rea, Guido Riva, Pierpaolo Salvo, and Antonio Sassano.

Data Availability Statement

The data that support the findings of this study are available from Fondazione Ugo Bordoni. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from the author(s) with the permission of Fondazione Ugo Bordoni.

References

1. A. Israr, Q. Yang, W. Li, and A. Y. Zomaya, "Renewable Energy Powered Sustainable 5G Network Infrastructure: Opportunities, Challenges and Perspectives," *Journal of Network and Computer Applications* 175 (2021): 102910.

2. F. D'Andreagiovanni, C. Mannino, and A. Sassano, "GUB Covers and Power-Indexed Formulations for Wireless Network Design," *Management Science* 59, no. 1 (2013): 142–156.

3. S. Dehghan, "A New Approach," *3GSM Daily 2005* 1 (2005): 44.

4. P. Avella, A. Calamita, and L. Palagi, "A Compact Formulation for the Base Station Deployment Problem in Wireless Networks," *Networks* 82 (2023): 52–67.

5. J. Kennington, E. Olinick, and D. Rajan, *Wireless Network Design: Optimization Models and Solution Procedures*, vol. 158 (Springer Science & Business Media, 2010).

6. J. Naoum-Sawaya and S. Elhedhli, "A Nested Benders Decomposition Approach for Telecommunication Network Planning," *Naval Research Logistics (NRL)* 57, no. 6 (2010): 519–539.

7. D. Ageyev and A. Al-Ansari, "Lte Ran and Services Multi-Period Planning," in *2015 Second International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T)* (IEEE, 2015), 272–274.

8. D. Ageyev and A. Al-Anssari, "Optimization Model for Multi-Time Period Lte Network Planning," in *2014 First International Scientific-Practical Conference Problems of Infocommunications Science and Technology* (IEEE, 2014), 29–30.

9. O. Bondarenko, D. Ageyev, and O. Mohammed, "Optimization Model for 5g Network Planning," in *2019 IEEE 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM)* (IEEE, 2019), 1–4.

10. A. Dmytro, A.-A. Ali, and Q. Nameer, "Multi-Period Lte Ran and Services Planning for Operator Profit Maximization," in *The Experience of Designing and Application of CAD Systems in Microelectronics* (IEEE, 2015), 25–27.

11. F. D'andreagiovanni, C. Mannino, and A. Sassano, "Negative Cycle Separation in Wireless Network Design," in *International Conference on Network Optimization* (Springer, 2011), 51–56.

12. A. Capone, L. Chen, S. Gualandi, and D. Yuan, "A New Computational Approach for Maximum Link Activation in Wireless Networks Under the Sinr Model," *IEEE Transactions on Wireless Communications* 10, no. 5 (2011): 1368–1372.

13. F. D'andreagiovanni and A. M. Gleixner, "Towards an Accurate Solution of Wireless Network Design Problems," in *International Symposium on Combinatorial Optimization* (Springer, 2016), 135–147.

14. G. Ecker, D. Yuan, A. M. Koster, and A. Schmeink, "Accurate Optimization Models for Interference Constrained Bandwidth Allocation in Cellular Networks," *Computers & Operations Research* 101 (2019): 1–12.

15. F. D'andreagiovanni, "Pure 0-1 Programming Approaches to Wireless Network Design," *4OR* 10, no. 2 (2012): 211.

16. M. Rumney, *LTE and the Evolution to 4G Wireless: Design and Measurement Challenges* (John Wiley & Sons, 2013).

17. T. S. Rappaport, *Wireless Communications: Principles and Practice*, vol. 2 (Prentice Hall PTR, 1996).

18. L. Chiaraviglio, A. S. Cacciapuoti, G. Di Martino, et al., "Planning 5G Networks Under Emf Constraints: State of the Art and Vision," *IEEE Access* 6 (2018): 51021–51037.

19. T. Achterberg, R. E. Bixby, Z. Gu, E. Rothberg, and D. Weninger, "Presolve Reductions in Mixed Integer Programming," *INFORMS Journal on Computing* 32, no. 2 (2020): 473–506.

20. Gurobi, "Gurobi Optimizer," 2021, https://www.gurobi.com/products/gurobi-optimizer/ accessed: October 08, 2021.

21. FUB, "Fondazione Ugo Bordoni," 2021, https://www.fub.it/en/ accessed: October 4, 2021.

22. R. V. Akhpashev and A. V. Andreev, "Cost 231 Hata Adaptation Model for Urban Conditions in Lte Networks," in *2016 17th International Conference of Young Specialists on Micro/Nanotechnologies and Electron Devices (EDM)* (IEEE, 2016), 64–66.

23. J. D. Gadze, K. A. Agyekum, S. J. Nuagah, and E. Affum, "Improved Propagation Models for Lte Path Loss Prediction in Urban & Suburban Ghana," 2020. *arXiv preprint arXiv:2001.05227*.

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

25. Usat, "Understanding LTE Signal Strength," (2021), accessed: May 18, 2021, https://usatcorp.com/faqs/understanding-lte-signal-strength-values/.