



HAL
open science

Optimizing the investments in mobile networks and subscriber migrations for a telecommunication operator

Adam Ouorou, Michael Poss, Adrien Cambier, Matthieu Chardy, Rosa Figueiredo, | Adam Ouorou, | Michael Poss

► To cite this version:

Adam Ouorou, Michael Poss, Adrien Cambier, Matthieu Chardy, Rosa Figueiredo, et al.. Optimizing the investments in mobile networks and subscriber migrations for a telecommunication operator. 2019. hal-02019853v3

HAL Id: hal-02019853

<https://hal.science/hal-02019853v3>

Preprint submitted on 25 Oct 2019 (v3), last revised 8 Jul 2020 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimizing the investments in mobile networks and subscriber migrations for a telecommunication operator

Adrien Cambier^{1,2} | Matthieu Chardy¹ | Rosa Figueiredo² | Adam Ouorou¹ | Michael Poss³

¹Orange Labs, 92320 Châtillon, France

²LIA, Avignon Université, Avignon, France.

³UMR CNRS 5506 LIRMM, Université de Montpellier, Montpellier, France.

Correspondence

Adrien CAMBIER, Orange Labs, 44 avenue de la République, 92320 Châtillon, France
Email: adrien.cambier@orange.com

Funding information

We consider the context of a telecommunication company that is at the same time an infrastructure operator and a service provider. When planning its network expansion, the company can leverage over its knowledge of the subscriber dynamic to better optimize the network dimensioning, therefore avoiding unnecessary costs. In this work, the network expansion represents the deployment and/or reinforcement of several technologies (e.g., 2G, 3G, 4G), assuming that subscribers to a given technology can be served by this technology or older ones. The operator can influence subscriber dynamic by subsidies. The planning is made over a discretized time horizon while some strategic guideline requirements are required at the end of the time horizon. Following classical models, we consider that the willingness of customers for shifting to a new technology follows an *S*-shape piecewise

constant function. We propose a Mixed-Integer Linear Programming formulation, improved through several valid inequalities and a heuristic algorithm. We assess the formulation numerically on real instances.

KEYWORDS

OR in telecommunications, Capacity Expansion, Investments Optimization, Subscriber Dynamic, Bass model, Mixed-Integer Linear Programming

1 | INTRODUCTION

Over the last few years, new bandwidth-consuming usages such as video streaming (see Table 1) have appeared, increasing the average monthly consumption by user, known as Average Usage per User. This phenomenon, correlated with an increase in the number of users, induces natural traffic growth. According to the Visual Networking Index of the IT and network company CISCO [5], traffic will globally reach 49 Exabytes per month in 2020 with a compound annual growth rate of 47%, this growth being particularly important in Africa (65%). Network expansion is necessary to support such traffic growth.

Service	2015	2018
Youtube videos viewed	2.78 M	4.3 M
Netflix hours watched	69444	266000
Instagram scrolling	38000	174000

TABLE 1 Number of usages of some services happening worldwide on the internet in 60 seconds (from [6])

Whenever possible, telecommunication companies must hence satisfy the request of subscribers in speed and volume to remain competitive, which requires network investments (several billion € to improve the mobile network in the last six years, see [17]). Facing both needs of offering a satisfying service and of limiting the investments, the operator does not want to under-/over- dimension its network.

In many countries, telecommunication companies are both infrastructure operators and service providers. As infrastructure operators, these companies are responsible for planning their network expansion. As service providers, they design the offers for users and have an influence on network traffic. In this manuscript, we consider a model where each offer can be characterized by the technology that its subscribers can reach. This model will simplify the notation used throughout since

offers and technologies are in bijection from the subscriber viewpoint. The network expansion decisions can benefit from taking into account the subscriber dynamic, and vice-versa. Thus, operators wish to understand the willingness of subscribers to shift to a new technology in order to optimally plan the investments in new mobile generations.

Having a deep and rigorous analysis of the demand evolution can be an advantage for an operator compared to others. This enables the operator to plan financial subsidies, e.g., cost reduction on a phone having access to the newest generation, in order to manage the network expansion and its market share.

	1G	2G	3G	4G	5G
Date	1980s	1990	2003	2009	2020
Speed	2.4 KB/s	64 KB/s	2 MB/s	1 GB/s	> 1 GB/s

TABLE 2 Evolution of speed through mobile generations (from [19]).

Planning the network expansion is a process that is inherently multi-period since investments must be distributed along a couple of years. As often in such problems, it is more efficient to use a strategy that considers multiple years simultaneously. This is even more important in mobile networks because of the quick progress of mobile technology. This is illustrated in Table 2 that shows the speed increase through mobile generations. According to the forecasts of GSM¹ association (see [11] for the detailed report), 4G will become the leading mobile network technology worldwide by number of connections (more than 3 billion) in 2019 while early commercial launches will start for 5G. This fast roll-out of mobile generations leads to a cyclic dependency between the subscriber and the network dynamics as investments in the network promote new subscriptions which in turn lead to new investments.

As we could expect, the network dynamic adds important constraints that must be considered when planning investments related to new mobile generations. For example, dismantling one generation of a mobile network is not an easy option since operational teams are reluctant to abandon well-functioning (and robust) technologies for new ones without back-up. Moreover, several services may need old(er) technologies (machine-2-machine, roaming, ...). Hence, different technologies have to co-exist and operators have to maintain simultaneously several generations.

1.1 | Mobile Master Plan

The points raised above motivate us to study the design of *multi-period master plans for mobile networks* (Mobile Master Plans), which consists in deciding, for a given set of time points and in a given area served by telecommunication sites, how to invest in the evolution of network technologies regarding three aspects: densification, sites coverage extension and subscription upgrades. An *investment in densification* means adding new pieces of equipment (modules) of a given technology in

¹GSM association is an originally-European trade body that represents the interests of mobile network operators worldwide.

order to increase the capacity of a site already covered by this technology. An *investment in coverage extension* means the deployment of a given technology on telecommunication sites not covered so far by this technology. Finally, an *investment in user upgrades* means the offer of financial subsidies to the subscribers in order to promote upgrades from their current subscription to the newest technology available. As already mentioned, these three kinds of investment decisions must be synchronized.

In order to be well-placed in operator rankings [1], strategic guidelines are decided by the operator at the end of the time horizon and drive the investments in order to guarantee competitiveness (for instance: sites coverage, user coverage, experienced throughput). In practice, Mobile Master Plans are designed for a 5-year time horizon with decisions taken for each year in this period. Mobile Master Plans are driven by cost minimization while ensuring strategic targets over the whole time horizon.

1.2 | Related literature

In this work, we are interested in optimizing, over a time horizon, investment decisions related to network expansion and subscriber dynamics assuming the arrival of a new technology. Such problems have been treated in the literature but with other targets in mind and a focus on network investments.

1.2.1 | Network expansion

Capacity expansion problems in telecommunication networks have been studied for a long time in the integer programming community, see [12, 10] among many others, including the case of multi-period planning [8, 9, 13].

Closer to the context of mobile capacity expansion, the authors of [4] present an exact mixed-integer formulation and a heuristic method to compute a Mobile Master Plan in a restrictive framework. These models integrate the upgrade of subscribers thanks to subsidies. However, a limitation of the models studied in [4] is that the amount of subsidy offered to users is fixed and the number of users which upgrade is set as a variable, constrained only to be positive and upper bounded by the total number of users. The possibility of offering subsidies in order to increase user upgrades is hence not taken into account. In [14], the authors consider the transition of a generation to another, but from a subscriber migration point of view only, applied for a South Korean network.

1.2.2 | The Bass model

Our Mobile Master Plan shall offer financial subsidies to the customers to encourage them to shift to the newest technology. To model the effect of these subsidies, we rely on the well-known Bass model from the marketing literature.

First studies on diffusion of innovation and new products appeared in the 60's, in a period of high economic growth and important innovations (television, etc). Everett Rogers published the

diffusion of innovation theory in 1962 (see [18]), based on the adoption curve of Figure 1.

This curve presents the percentage of subscribers which adopt a new product during the time horizon. The curve assumes that the timing of a consumer's initial purchase is related to the number of previous buyers (imitation part) and enlightens different types of subscribers. The innovators are the easiest to convince and the laggards are those who adopt the last. External factors (marketing and attractiveness) due to decisions taken in the time horizon are not taken into account, as well as the generation effect (new generation replacing an older one).

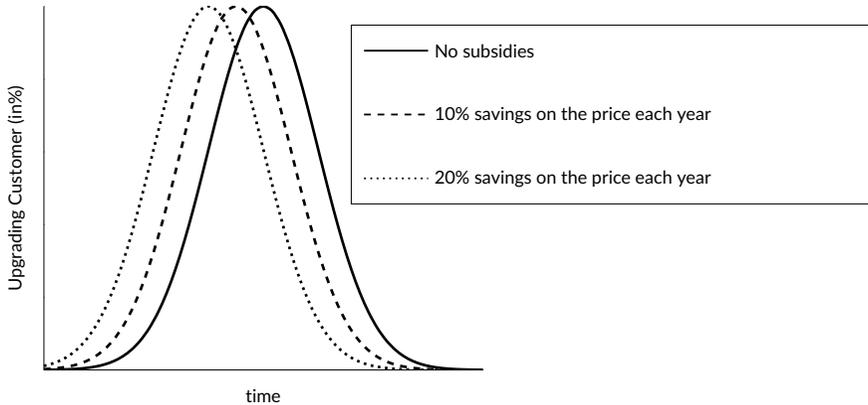


FIGURE 1 Curve of the diffusion of innovation and influence of subsidies

In 1969, Bass formalized Roger's model by using differential equations and later developed it to tackle some of the issues with external variables (see [2]) and generations (see [16]). This work considers marketing aspects by showing a left shift on the shape of adoptions curves when regular savings (equivalent to constant subsidies) on the price of a product are made, as illustrated in Figure 1. These models help the understanding of how subscribers react in a telecommunication market: see Section 5 in [3] for a discussion on 2G/3G upgrade; [14] for Bass model applied to the forecasting of the 5G upgrade; and [15] for an application to the Greek mobile market.

In order to adapt this formalism to the current telecommunication context, we consider two important factors for modeling the adoption of a new technology, described hereafter by function f . First, the percentage of subscribers shifting from older technologies is very sensitive to the price gap between the new technology and their current one, which will be referred to as σ . The second factor is the influence of network deployment (denoted c). Indeed, subscribers upgrade more easily when they are sure to benefit from the new service, i.e. if the newest technology is deployed.

1.3 | Contributions and structure of the paper

To the best of our knowledge, jointly optimizing the two dynamics (network and subscribers) under capacity and targeting constraints has not yet been studied. Hence, the first main contribution of

the manuscript is to provide a realistic model for the problem. Building on practical considerations, we come up with a non-linear non-convex Mixed-Integer Programming (MIP) formulation, the non-linearity of which comes from the aforementioned function f . Practical considerations will imply that the domain of f is finite, making f a discrete function. Hence, the linearization of the model will naturally follow using classical techniques.

The second main contribution of the manuscript lies in improving the performance of the model through (i) the strengthening of the MILP with several families of valid inequalities, and (ii) a heuristic algorithm that assigns fixed values of the decision variables (subsidies and coverage) and solves the resulting problem as a knapsack problem.

Our third contribution lies in assessing in details the performance and solutions provided by the model on real instances. Our results show, in particular, the efficiency of the valid inequalities and the heuristic algorithm, enabling us to solve large real-life instances to near-optimality. We also assess scalability performances and economic costs when is introduced a smoothing constraint stating that the operator does not want high budget discrepancy over the years.

The remainder of this article is organized as follows. Section 2 introduces our Mobile Master Plan problem for two technologies, for which a mixed integer formulation is provided and linearized in Section 3. Section 4 introduces the aforementioned valid inequalities and heuristic algorithm. These models are numerically assessed in Section 5 on real-life instances. Concluding remarks are given in Section 6.

2 | PROBLEM DESCRIPTION

The time horizon is taken as multi-period with equally-sized time periods denoted by $t \in \mathcal{T} = \{1, \dots, \bar{t}\}$ (typically 5 periods of one year each). We add "0" for denoting the beginning of the time horizon.

The whole area is served by existing telecommunication sites potentially equipped with at most two mobile network technologies. We thus consider a set $\mathcal{G} = \{CG, NG\}$ of mobile network generations (the current one and the newest one being deployed) and a set $\mathcal{S} = \{1, \dots, N_S\}$ of telecommunication sites. As we have already mentioned, the network capacity on a site can be increased in two ways: by deploying a technology on a site or by adding new modules of an already deployed technology. In this work, we assume that the current technology CG is deployed on all sites at the beginning of the time horizon and that only the newest technology NG can be deployed during the time horizon with a cost per site of CA_{NG} . The binary parameter $Z_{s,NG}^0$, $s \in \mathcal{S}$, is equal to 1 iff the newest network technology NG is deployed on site s at the beginning of the time horizon. The initial number of modules on each site is denoted by $M_{s,g}^0$, $s \in \mathcal{S}$, $g \in \mathcal{G}$. For each site and for each deployed technology, adding new modules is possible with a unitary cost of CM_g , $g \in \mathcal{G}$. The network decommissioning (possibility of removing modules) is linked to spectral considerations and hence is outside of the scope of this article. Technical constraints impose an upper bound on the number of modules that can be added to a site for each technology, which we note \bar{M}_g , $g \in \mathcal{G}$. Let us introduce CAP_g , $g \in \mathcal{G}$, the unitary capacity of a module of each network technology.

The initial number of subscribers to each technology associated with each site is denoted by $U_{s,o}^0, s \in \mathcal{S}, o \in \mathcal{G}$. These values are taken as equivalent value for the dimensioning. For technical reasons, subscribers cannot be served by a more recent technology than the one they subscribe to. Hence, CG subscribers have to be served by CG technology. For quality of experience motivations, we introduce a load-balancing rule stating that NG subscribers associated to a site $s \in \mathcal{S}$ are served by NG technology if deployed on s and CG technology otherwise.

As mentioned in the introduction, we assume in this work that the whole investments in user upgrades are made towards the newest technology NG . We also assume that the upgrade mechanism modeling subscribers willing to shift to NG technology depends only on two parameters. The first one is the value of the subsidy denoted by σ . The set of possible values taken by σ will be denoted by \mathcal{K} . This set is finite due to practical considerations (modeling traditionally used marketing offers : 20%, 30% savings on the new phone etc). The second one is an indicator of the level of NG technology deployment. This indicator will be taken as a range of coverage c (low, medium low, medium high and high coverages for instance). The range of coverage of a given time period will be measured as the range of the proportion of sites on which NG technology is deployed, denoted α and referred as sites coverage, in what follows. For modeling coverage ranges, we partition the interval $[0, 1]$ into C smaller intervals $[L_c, U_c[$, and define $C = \{1, \dots, C\}$. The function modeling the upgrade mechanism, denoted by $f : \mathcal{K} \times C \rightarrow [0, 1]$, provides the proportion of subscribers willing to shift to NG technology if they receive the given subsidy $\sigma \in \mathcal{K}$ under a given range of coverage $c \in C$ at the beginning of the time period $t \in \mathcal{T}$. This function is assumed non-decreasing in both arguments.

As already pointed out, network and subscriber dynamics are linked. First, each network generation has to be dimensioned to handle the traffic demand per subscriber $D_g^t, t \in \mathcal{T}, g \in \mathcal{G}$, defined as the 95% quantile of the possible demands occurring over the time period. Hence, the capacity has to be sufficient to handle the resulting traffic 95% of the time. We assume that this unitary traffic demand depends on the network serving the subscribers rather than their current subscription. More precisely, at each time period t , a CG subscriber will have the demand D_{CG}^t and a NG subscriber will have the demand D_{NG}^t if served by NG and D_{CG}^t otherwise. Second, telecommunication operators are ranked according to their performance. Therefore, we decide to focus on requiring satisfying levels for two key performance indicators: the proportion of sites covered by NG at the end of the time horizon, which is denoted by $\alpha^{\bar{t}}$ and the averaged quality of experience to the corresponding subscribers. The averaged quality of experience is guaranteed by asking for a minimal proportion of the total number of subscribers being NG subscribers associated with NG sites. These subscribers benefit indeed from the new performance and have the maximum throughput. The thresholds required at the end of the time horizon associated with these two targeting indicators are respectively denoted by $\underline{\alpha}$ and \underline{QoE} .

Decisions are taken over the time horizon. These decisions are the deployment of NG technology, the number of modules added (for all technologies), and the subsidies given to the subscribers from older technologies for upgrading towards NG technology. The problem defined in this work, denoted as the Mobile Master Plan Problem (MMPP), consists in finding the decisions which minimize network and subscribers investments over the time horizon while satisfying capacity and

targeting constraints.

Parameters introduced in this section are summed up below:

- CA_{NG} is the cost of adding NG technology,
- CM_g is the cost of adding a module of a technology $g \in \mathcal{G}$,
- $M_{s,g}^0$ stands for the initial number of modules of technology $g \in \mathcal{G}$ on site $s \in \mathcal{S}$,
- \overline{M}_g stands a technical upper bound on the number of modules of technology $g \in \mathcal{G}$,
- $Z_{s,NG}^0$ stands for the initial presence (yes/no) of NG technology on site $s \in \mathcal{S}$,
- $U_{s,o}^0$ is the initial number of subscribers on site $s \in \mathcal{S}$ to technology $o \in \mathcal{G}$,
- D_g^t is the unitary demand of a subscriber served by technology $g \in \mathcal{G}$ at time period $t \in \mathcal{T}$,
- CAP_g is the capacity of adding a module of a technology $g \in \mathcal{G}$,
- $f(\sigma, c)$ is the reaction to the subsidy offered $\sigma \in \mathcal{K}$ under range of coverage interval $c \in \mathcal{C}$,
- $\sigma^t \in \mathcal{K}$ is the value of subsidy offered at time period $t \in \mathcal{T}$,
- L_c stands for the lower bound of coverage range $c \in \mathcal{C}$,
- U_c stands for the upper bound of coverage range $c \in \mathcal{C}$,
- α^0 stands for the sites coverage at the beginning of the time horizon,
- α^t stands for the sites coverage at the end of time period $t \in \mathcal{T}$,
- $c^t \in \mathcal{C}$ is the range of coverage of α^{t-1} for each time period $t \in \mathcal{T}$,
- $\underline{\alpha}$ and \underline{QoE} are the thresholds fixed as strategic guidelines.

3 | MATHEMATICAL MODELING

We provide in this section a mixed-integer formulation for the problem described in Section 2. We define the set of decision variables used in our formulation in Section 3.1 and present a non-linear mixed-integer formulation in Section 3.2. Finally, in Section 3.3 we linearize this formulation.

3.1 | Decision variables

For modeling the network investment, we use the following variables:

- For $t \in \mathcal{T} \cup \{0\}$, $s \in \mathcal{S}$, let us introduce the binary variable

$$z_{s,NG}^t = \begin{cases} 1, & \text{if the newest technology is deployed at site } s \text{ at the end of time period } t, \\ 0, & \text{otherwise.} \end{cases}$$

- For $t \in \mathcal{T} \cup \{0\}$, $s \in \mathcal{S}, g \in \mathcal{G}$, the integer variable $m_{s,g}^t$ represents the total number of modules of technology g deployed on site s at the end of time period t .

As for modeling the number of users on each site, we use the following continuous variables:

- For each $t \in \mathcal{T} \cup \{0\}$, $s \in \mathcal{S}$, $o \in \mathcal{G}$, let $u_{s,o}^t$ denote the total number of subscribers to technology o in site s at the end of time period t (we denote an upper bound on this quantity by $\bar{U}_{s,o}^t$),
- For each $t \in \mathcal{T}$, $s \in \mathcal{S}$, $o, g \in \mathcal{G}^2$, let $u_{s,o,g}^t$ denote the total number of subscribers to technology o served by technology g in site s at the end of time period t .

In addition, the notation σ^t , c^t , and α^t introduced in the previous section become optimization variables:

- For each $t \in \mathcal{T}$, let σ^t be the value of the subsidy, in k€ , offered to subscribers to former technologies for upgrading to technology NG at the beginning of time period t ,
- For each $t \in \mathcal{T}$, let $\alpha^t = \frac{\sum_{s \in \mathcal{S}} z_{s,NG}^t}{N_S}$ be the redundant variable that denotes the NG sites coverage (fraction of sites where NG technology is deployed) at the end of the time period t
- For each $t \in \mathcal{T}$, let c^t denote the interval of C to which α^{t-1} belongs.

The upgrade function, representing the percentage of users reacting positively to a subsidy $\sigma \in \mathcal{K}$ for a given coverage $c \in C$, is denoted by $f(\sigma, c)$ and will be modeled explicitly in Section 3.3.

3.2 | General Formulation

The MMPP can be modeled as follows:

$$\begin{aligned} \min \quad & \sum_{t \in \mathcal{T}} \sigma^t f(\sigma^t, c^t) \sum_{s \in \mathcal{S}} u_{s,CG}^{t-1} + \sum_{s \in \mathcal{S}} \sum_{g \in \mathcal{G}} CM_g(m_{s,g}^{\bar{i}} - M_{s,g}^0) \\ & + \sum_{s \in \mathcal{S}} CA_{NG}(z_{s,NG}^{\bar{i}} - Z_{s,NG}^0) \end{aligned} \quad (1)$$

$$\text{s.t. } m_{s,CG}^t \leq \bar{M}_{CG} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \quad (2)$$

$$m_{s,NG}^t \leq \bar{M}_{NG} z_{s,NG}^t \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \quad (3)$$

$$m_{s,g}^{t-1} \leq m_{s,g}^t \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \forall g \in \mathcal{G}, \quad (4)$$

$$u_{s,NG}^t = u_{s,NG,CG}^t + u_{s,NG,NG}^t \quad \forall s \in \mathcal{S} \forall t \in \mathcal{T}, \quad (5)$$

$$u_{s,NG,CG}^t \leq \bar{U}_{s,NG}^t (1 - z_{s,NG}^t) \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \quad (6)$$

$$D_{CG}^t (u_{s,CG}^t + u_{s,NG,CG}^t) \leq CAP_{CG} m_{s,CG}^t \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \forall g \in \mathcal{G}, \quad (7)$$

$$D_{NG}^t u_{s,NG,NG}^t \leq CAP_{NG} m_{s,NG}^t \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \forall g \in \mathcal{G}, \quad (8)$$

$$u_{s,CG}^t = u_{s,CG}^{t-1} - f(\sigma^t, c^t) u_{s,CG}^{t-1} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \quad (9)$$

$$u_{s,NG}^t = u_{s,NG}^{t-1} + f(\sigma^t, c^t) u_{s,CG}^{t-1} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \quad (10)$$

$$\sum_{s \in \mathcal{S}} u_{s,NG,NG}^{\bar{i}} \geq \underline{QoE} \left(\sum_{s \in \mathcal{S}} U_{s,NG}^0 + U_{s,CG}^0 \right), \quad (11)$$

$$\alpha^t \geq \underline{\alpha}, \quad (12)$$

$$\alpha^t N_S = \sum_{s \in \mathcal{S}} z_{s,NG}^t \quad \forall t \in \mathcal{T} \cup \{0\}, \quad (13)$$

$$\alpha^{t-1} \in [L_{c^t}, U_{c^t}] \quad \forall t \in \mathcal{T}, \quad (14)$$

$$u_{s,o}^0 = U_{s,o}^0 \quad \forall s \in \mathcal{S}, \forall o \in \mathcal{G}, \quad (15)$$

$$m_{s,g}^0 = M_{s,g}^0 \quad \forall s \in \mathcal{S}, \forall g \in \mathcal{G}, \quad (16)$$

$$z_{s,NG}^0 = Z_{s,NG}^0 \quad \forall s \in \mathcal{S}, \quad (17)$$

$$m_{s,g}^t \in \mathbb{Z} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \cup \{0\}, \forall g \in \mathcal{G}, \quad (18)$$

$$z_{s,NG}^t \in \{0, 1\} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \cup \{0\}, \quad (19)$$

$$u_{s,o}^t \geq 0 \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \cup \{0\}, \forall o \in \mathcal{G}, \quad (20)$$

$$u_{s,o,g}^t \geq 0 \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \forall o, g \in \mathcal{G}^2, \quad (21)$$

$$\sigma^t \in \mathcal{K} \quad \forall t \in \mathcal{T}, \quad (22)$$

$$c^t \in \mathcal{C} \quad \forall t \in \mathcal{T}. \quad (23)$$

We denote this formulation by \mathcal{M}^{NL} . The objective function (1) minimizes both subscribers migration costs and network investments. The first term stands for the offered subsidies (user upgrades), the second term for the adding of new modules for increasing the capacity (densification), and the third term for the deployment of the newest technology NG (coverage extension). Constraints (2)–(4) are the network dynamic constraints. Constraints (2)–(3) define the upper bounds on the numbers of modules for each technology deployed on each site. These constraints also ensure that if a technology is not deployed, no corresponding modules can be added. Constraints (4) prevent from decommissioning by imposing the number of modules of each technology to be non-decreasing during the time horizon.

Constraints (5)–(7) are the network dimensioning constraints, in charge of making the link between the network dynamic and the subscriber dynamic. Constraints (5) and (6) ensure the load-balancing rule. Constraints (7) are the capacity constraints: the installed capacities of each technology on each site have to be sufficient for providing services for all users located at this site and having to be served by this technology. They also ensured the technical incompatibility stating that CG subscribers cannot be served by NG technology.

Constraints (9)–(10) are the subscriber dynamic constraints. They define the total number of subscribers to CG and NG technologies at each site and each time period, taking into account former CG subscribers who decide to shift to NG technology, thanks to subsidies and coverage improvements. Constraints (11)–(12) stand for the model strategic guidelines and refer to the end of the time horizon. Constraint (11) ensures the threshold of subscribers covered by the newest technology is met. The indicator is proportional to the quality of experience which measures the percentage of users having access to the new technology throughput. Constraint (12) imposes that the threshold on the number of sites on which NG is deployed is met. Constraints (13) stand for defining variables α^t . Constraints (14) define the range of coverage used in the subsidy function as the range of coverage to which belongs α^{t-1} . Constraints (15)–(17) refer to the initial conditions.

Finally, constraints (18)–(23) define the domain of the variables. Section 3.3 details function f and linear modeling of constraints (14).

Remark 1 The size of formulation \mathcal{M}^{NL} can be reduced by replacing variables $m_{s,g}^t$ for each $t \in \mathcal{T} \setminus \{\bar{t}\}$ by $m_{s,g}^{\bar{t}}$ and removing constraints (4). Indeed variables m are only required to be lower and upper bounded and to be non-decreasing over the time horizon while the objective function only depends on $m_{s,g}^{\bar{t}}$. However, we choose to present a model with variables $m_{s,g}^t$ as a generic basis for businesses applications where budget has to be controlled over time, thus requiring the temporal dynamic of the number of modules. We will assess numerically in Section 5 a set of constraints smoothing the costs for the operator by bounding budget fluctuations by a percentage p . Let the budget spent in each year be denoted by

$$B_t = \sigma^t f(\sigma^t, c^t) \sum_{s \in S} u_{s,CG}^{t-1} + \sum_{s \in S} \sum_{g \in \mathcal{G}} CM_g(m_{s,g}^t - m_{s,g}^{t-1}) + \sum_{s \in S} CA_{NG}(z_{s,NG}^t - z_{s,NG}^{t-1}) \quad \forall t \in \mathcal{T}. \quad (24)$$

so the objective function (1) is actually equal to $\sum_{t \in \mathcal{T}} B_t$. The cost equilibrium set of constraints can be written as follows:

$$(1-p) \times \frac{\sum_{t' \in \mathcal{T}} B_{t'}}{\bar{t}} \leq B_t \leq (1+p) \times \frac{\sum_{t' \in \mathcal{T}} B_{t'}}{\bar{t}} \quad \forall t \in \mathcal{T}. \quad (25)$$

With constraints (25) added to formulation \mathcal{M}^{NL} , variables $m_{s,g}^t$ are no longer redundant.

Remark 2 The time dependency of the technology installation variable z cannot be removed. Indeed, the upgrade reaction at period $t \in \mathcal{T}$ depends on variables z^{t-1} (constraints (13) and (14)). Hence, modifying the period of installation impacts the total upgrade reaction and the objective value. Note that the coverage will not always be set to the highest value due to the subsidies costs.

3.3 | Upgrade function modeling

As we have mentioned, the upgrade function (function characterizing the proposition of CG subscribers that shift to NG technology) is non-decreasing in both the subsidy amount σ and the range of coverage c . Figure 2 provides an example with four ranges.

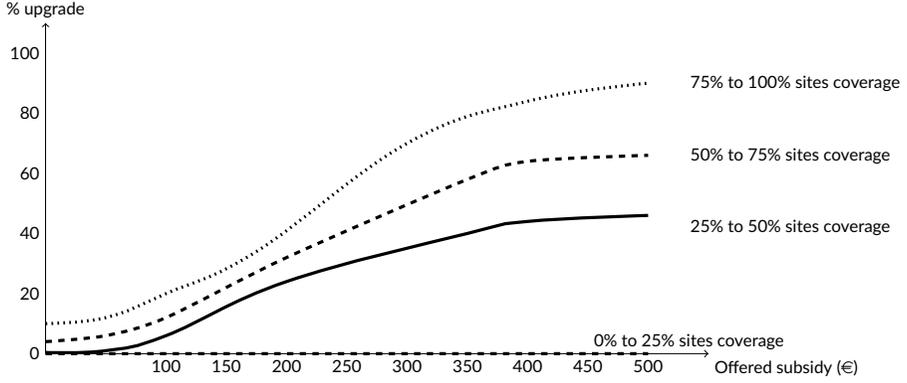


FIGURE 2 Example of upgrade function

To shorten notation, we denote by $f_{\sigma,c}$ the percentage of subscribers that react positively when subsidy $\sigma \in \mathcal{K}$ is offered and the NG sites coverage belongs to the range $[L_c, U_c]$, formally defined as $f(\sigma, c)$. Aiming to incorporate this in our MILP, we introduce a binary variable $\delta_{\sigma,c}^t$ for each $t \in \mathcal{T}, \sigma \in \mathcal{K}, c \in \mathcal{C}$, taking value equal to 1 iff σ^t is offered and $\alpha^{t-1} \in [L_c, U_c]$. The first term of objective function (1) from Section 3.2 can be rewritten as follows:

$$\sum_{t \in \mathcal{T}} \sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} \sigma f_{\sigma,c} \delta_{\sigma,c}^t \sum_{s \in \mathcal{S}} u_{s,CG}^{t-1}.$$

Also, constraints (9) and (10) can be written as:

$$u_{s,CG}^t = u_{s,CG}^{t-1} - \sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} f_{\sigma,c} \delta_{\sigma,c}^t u_{s,CG}^{t-1} \quad \forall s \in \mathcal{S} \forall t \in \mathcal{T}, \quad (26)$$

$$u_{s,NG}^t = u_{s,NG}^{t-1} + \sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} f_{\sigma,c} \delta_{\sigma,c}^t u_{s,CG}^{t-1} \quad \forall s \in \mathcal{S} \forall t \in \mathcal{T}. \quad (27)$$

We linearize the products of binary variables $\delta_{\sigma,c}^t$ and continuous variables $u_{s,CG}^{t-1}$ using a classical method [7]. Consequently, the MMPP can be formulated as the following MILP.

$$\begin{aligned} \min \quad & \sum_{t \in \mathcal{T}} \sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} \sum_{s \in \mathcal{S}} \sigma f_{\sigma,c} \pi_{\sigma,c,s,CG}^t + \sum_{s \in \mathcal{S}} \sum_{g \in \mathcal{G}} CM_g (m_{s,g}^{\bar{}} - M_{s,g}^0) \\ & + \sum_{s \in \mathcal{S}} CANG (z_{s,NG}^{\bar{}} - Z_{s,NG}^0) \end{aligned} \quad (28)$$

s.t. (2) – (7), (11) – (13), (15) – (17)

$$u_{s,CG}^t = u_{s,CG}^{t-1} - \sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} f_{\sigma,c} \pi_{\sigma,c,s,CG}^t \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \quad (29)$$

$$u_{s,NG}^t = u_{s,NG}^{t-1} + \sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} f_{\sigma,c} \pi_{\sigma,c,s,CG}^t \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \quad (30)$$

$$\sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} \delta_{\sigma,c}^t = 1 \quad \forall t \in \mathcal{T}, \quad (31)$$

$$\sum_{\sigma \in \mathcal{K}} \delta_{\sigma,c}^t \leq 1 + U_c - \alpha^{t-1} \quad \forall t \in \mathcal{T}, \forall c \in \mathcal{C}, \quad (32)$$

$$\sum_{\sigma \in \mathcal{K}} \delta_{\sigma,c}^t \leq 1 + \alpha^{t-1} - L_c \quad \forall t \in \mathcal{T}, \forall c \in \mathcal{C}, \quad (33)$$

$$\pi_{\sigma,c,s,CG}^t \leq \delta_{\sigma,c}^t \bar{U}_{s,CG}^{t-1} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \forall \sigma \in \mathcal{K}, \forall c \in \mathcal{C}, \quad (34)$$

$$\pi_{\sigma,c,s,CG}^t \leq u_{s,CG}^{t-1} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \forall \sigma \in \mathcal{K}, \forall c \in \mathcal{C}, \quad (35)$$

$$\pi_{\sigma,c,s,CG}^t \geq u_{s,CG}^{t-1} - (1 - \delta_{\sigma,c}^t) \bar{U}_{s,CG}^{t-1} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \forall \sigma \in \mathcal{K}, \forall c \in \mathcal{C}, \quad (36)$$

$$m_{s,g}^t \in \mathbb{Z} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \cup \{0\}, \forall g \in \mathcal{G}, \quad (37)$$

$$z_{s,NG}^t \in \{0, 1\} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \cup \{0\}, \quad (38)$$

$$u_{s,o}^t \geq 0 \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \cup \{0\}, \forall o \in \mathcal{G}, \quad (39)$$

$$u_{s,o,g}^t \geq 0 \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \forall o, g \in \mathcal{G}^2, \quad (40)$$

$$\delta_{\sigma,c}^t \in \{0, 1\} \quad \forall t \in \mathcal{T}, \forall \sigma \in \mathcal{K}, \forall c \in \mathcal{C}, \quad (41)$$

$$\pi_{\sigma,c,s,CG}^t \geq 0 \quad \forall t \in \mathcal{T}, \forall \sigma \in \mathcal{K}, \forall c \in \mathcal{C}, \forall s \in \mathcal{S}. \quad (42)$$

We denote this formulation by \mathcal{M} . Constraints (29) and (30) are the linearizations respectively of constraints (26) and (27). Constraints (31) ensure that one and only one subsidy from the set \mathcal{K} is offered at each time period, the case when no subsidy is given being represented by $\sigma = 0$. Constraints (32) and (33) ensure that, for each time period, variables $\delta_{\sigma,c}^t$ are set according to the coverage. Constraints (32) (respectively (33)) set all δ related to a range to 0 if the coverage is greater (resp. smaller) than the upper (resp. lower) bound of the range. Constraints (34)–(36) are the typical linearizations of the products of a binary variable with a continuous one. Constraints (37)–(42) define the domain of all variables in the formulation.

Remark 3 *The range coverage being an increasing variable, and the initial range of coverage c_{init} being determined by the value of parameter $Z_{s,NG}^0$ for each $s \in \mathcal{S}$, we can use set $C' = \{c_{init}, \dots, C\}$ instead of C , hence reducing the formulation size.*

We show in the Appendix how this model can be extended to tackle frameworks where several generations (more than three generations) co-exist.

4 | STRENGTHENING AND UPPER BOUND

We strengthen our formulation with several valid inequalities in Section 4.1 and provide in Section 4.2 a heuristic algorithm based on the properties of the problem where the offered subsidies and the coverage are fixed.

4.1 | Valid inequalities

Preliminary computational experiments on small instances showed that the solutions of linear relaxation present variables z and δ fractional. Consequently, we propose several valid inequalities in this subsection to reinforce the model. The strength of these inequalities is assessed numerically in Section 5.2.

Proposition 4 *Considering a time period $t \in \mathcal{T} \cup \{0\}$ and a site $s \in \mathcal{S}$, inequality*

$$z_{s,NG}^t \leq z_{s,NG}^{t+1} \quad (43)$$

is valid for formulation \mathcal{M} .

Proof This result is implied by constraints (3) and (4).

Proposition 5 *Considering a time period $t \in \mathcal{T}$ and a range of coverage $c \in \mathcal{C}$, for all time periods t' posterior to t , inequality*

$$\sum_{\sigma \in \mathcal{K}} \sum_{c' < c} \delta_{\sigma,c'}^{t'} \leq 1 - \sum_{\sigma \in \mathcal{K}} \sum_{c' \geq c} \delta_{\sigma,c'}^t \quad (44)$$

is valid for formulation \mathcal{M} .

Proof This set of constraints states that if at a time period $t \in \mathcal{T}$, the range of coverage is greater or equal to $c \in \mathcal{C}$, then the range of coverage for posterior time periods cannot be smaller. As defined in Section 2, the NG sites coverage is indeed non-decreasing over the time horizon.

Proposition 6 *Considering a time period $t \in \mathcal{T}$ and a site $s \in \mathcal{S}$, equality*

$$\sum_{c \in \mathcal{C}} \sum_{\sigma \in \mathcal{K}} \pi_{\sigma,c,s,CG}^t = u_{s,CG}^{t-1} \quad (45)$$

is valid for formulation \mathcal{M} .

Proof Following the Reformulation Linearization Techniques (see [20] for more details), we obtained these constraints by multiplying each constraint from set (31) by variables $u_{s,CG}^{t-1}$ for each $s \in \mathcal{S}$. The product obtained in the left member is then replaced by the corresponding linearization variable.

Proposition 7 Considering a time period $t \in \mathcal{T}$ and a range of coverage $c \in C$, inequality

$$\lceil N_S L_c \rceil \sum_{\sigma \in \mathcal{K}} \delta_{\sigma,c}^t \leq \sum_{s \in S} z_{s,NG}^t \quad (46)$$

is valid for formulation \mathcal{M} .

Proof The network is in a range of coverage $c \in C$ only if technology NG is deployed on at least $\lceil N_S L_c \rceil$ sites (remember that N_S is the total number of sites and is constant over the time horizon).

Proposition 8 Let $\underline{U}_{s,NG}^t = U_{s,NG}^0 + U_{s,CG}^0 [1 - (1 - \min_{\sigma} f_{\sigma,c^0})^t]$ denote a lower bound on the number of NG subscribers on site s at time period t . Considering a time period $t \in \mathcal{T}$ and a site $s \in S$, inequality

$$\left\lceil \frac{D_{NG}^t \underline{U}_{s,NG}^t}{CAP_{NG}} \right\rceil z_{s,NG}^t \leq m_{s,NG}^t \quad (47)$$

is valid for formulation \mathcal{M} .

Proof If NG technology is deployed on a site $s \in S$ at a time period $t \in \mathcal{T}$, we know that NG subscribers have to be served by NG technology. By computing a lower bound on the quantity of NG subscribers at this site and on this time period, we can hence compute a corresponding lower bound on the number of modules required for satisfying the capacity constraints (7).

Proposition 9 Let $\bar{U}_{s,NG}^t = U_{s,NG}^0 + U_{s,CG}^0 [1 - (1 - \max_{\sigma} f_{\sigma,c})^t]$ denote an upper bound on the number of NG subscribers on site s at time period t . Considering a time period $t \in \mathcal{T}$ and a site $s \in S$, every optimal solution of the MMPP satisfies the following inequality:

$$m_{s,NG}^t \leq \max(M_{s,NG}^0, \left\lceil \frac{D_{NG}^t \bar{U}_{s,NG}^t}{CAP_{NG}} \right\rceil) z_{s,NG}^t \quad (48)$$

Proof If NG technology is not deployed on a site $s \in S$ at a time period t , the number of modules for this technology on this site at this time period is 0. If NG technology is deployed on a site $s \in S$ at a time period t , we know that only NG subscribers on this site can be served by NG technology. By computing an upper bound of the quantity of NG subscribers on this site at this time period, we can hence compute a corresponding upper bound of the number of modules needed to satisfy the capacity constraints (7). Installing more than this bound costs CAP_{NG} by additional module without any impact of the feasibility, and such a solution will hence be eliminated by the objective function minimization.

4.2 | Heuristic approach: fixing the upgrade reaction values

In this section, we discuss a heuristic algorithm that can also help to solve the MMPP. The main difficulties of formulation \mathcal{M}^{NL} come from the non-linear constraints (9), (10) and (14). Hence, a

simple heuristic approach to the problem would be to fix variables σ^t (the amount of the subsidy offered to CG subscribers) and c^t (the coverage range) to specific values $\bar{\sigma}^t \in C$ and $\bar{c}^t \in \mathcal{K}$ for each period $t \in \mathcal{T}$ and solve the resulting problem optimally. Note that \bar{c}^1 is already fixed to the initial coverage c_{init} .

Let us now denote the problem where $\sigma^t = \bar{\sigma}^t$ and $c^t = \bar{c}^t$ as the MMPP($\bar{\sigma}, \bar{c}$). Applying constraints (9) recursively over the time-horizon, we obtain

$$U_{s,CG}^t = \prod_{i=1}^t (1 - f_{\bar{\sigma}^i, \bar{c}^i}) U_{s,CG}^0 \quad \forall t \in \mathcal{T}, \forall s \in S. \quad (49)$$

Similarly, summing up constraints (9) and (10) gives $u_{s,CG}^t + u_{s,NG}^t = u_{s,CG}^{t-1} + u_{s,NG}^{t-1}$ and hence recursively that the total number of subscribers on a site $s \in S$ (denoted by $U_s^{TOT} = U_{s,CG}^0 + U_{s,NG}^0$) is constant over the time horizon. We hence have

$$U_{s,NG}^t = U_s^{TOT} - U_{s,CG}^t \quad \forall t \in \mathcal{T}, \forall s \in S. \quad (50)$$

We can also compute the amount of money spent in subsidies offered to the subscribers, which is equal to the constant

$$upgradecost = \sum_{t \in \mathcal{T}} \bar{\sigma}^t f_{\bar{\sigma}^t, \bar{c}^t} U_{s,CG}^{t-1}. \quad (51)$$

We are now able to reformulate capacity and load-balancing constraints as a set of constraint depending only on the fixed constants $U_{s,CG}^t$ and $U_{s,NG}^t$ and of variables m and z . Consequently, the MMPP($\bar{\sigma}, \bar{c}$) can be formulated as follows

$$\begin{aligned} \min \quad & upgradecost + \sum_{s \in S} \sum_{g \in \mathcal{G}} CM_g (m_{s,g}^{\bar{c}} - M_{s,g}^0) \\ & + \sum_{s \in S} C_{ANG} (z_{s,NG}^{\bar{c}} - Z_{s,NG}^0) \end{aligned} \quad (52)$$

$$\text{s.t. } m_{s,CG}^t \leq \bar{M}_{CG} \quad \forall s \in S, \forall t \in \mathcal{T}, \quad (53)$$

$$m_{s,NG}^t \leq \bar{M}_{NG} z_{s,NG}^t \quad \forall s \in S \forall t \in \mathcal{T}, \quad (54)$$

$$m_{s,g}^{t-1} \leq m_{s,g}^t \quad \forall s \in S, \forall t \in \mathcal{T}, \quad (55)$$

$$D_{CG}^t U_s^{TOT} (1 - z_{s,NG}^t) \leq CAP_{CG} m_{s,CG}^t \quad \forall s \in S, \forall t \in \mathcal{T}, \quad (56)$$

$$D_{CG}^t U_{s,CG}^t \leq CAP_{CG} m_{s,CG}^t \quad \forall s \in S, \forall t \in \mathcal{T}, \quad (57)$$

$$D_{NG}^t U_{s,NG}^t z_{s,NG}^t \leq CAP_{NG} m_{s,NG}^t \quad \forall s \in S, \forall t \in \mathcal{T}, \quad (58)$$

$$\sum_{s \in S} U_{s,NG}^{\bar{c}} z_{s,NG}^{\bar{c}} \geq \underline{QoE} \sum_{s \in S} U_s^{TOT}, \quad (59)$$

$$\sum_{s \in S} z_{s,NG}^{\bar{c}} \geq \bar{\alpha} \quad (60)$$

$$\sum_{s \in \mathcal{S}} z_{s,NG}^{t-1} \geq \lfloor L_{\bar{c}^t} N_S \rfloor \quad \forall t \in \mathcal{T}, \quad (61)$$

$$\sum_{s \in \mathcal{S}} z_{s,NG}^{t-1} \leq \lfloor U_{\bar{c}^t} N_S \rfloor \quad \forall t \in \mathcal{T}, \quad (62)$$

$$m_{s,g}^0 = M_{s,g}^0 \quad \forall s \in \mathcal{S}, \forall g \in \mathcal{G}, \quad (63)$$

$$z_{s,NG}^0 = Z_{s,NG}^0 \quad \forall s \in \mathcal{S}, \quad (64)$$

$$m_{s,g}^t \in \mathbb{Z} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \cup \{0\}, \forall g \in \mathcal{G}, \quad (65)$$

$$z_{s,NG}^t \in \{0, 1\} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \cup \{0\}. \quad (66)$$

The objective (1) can be reformulated into (52) where the term standing for subsidies cost is now a constant. Constraints (56)–(58) ensure the load-balancing rules (seen from a network point of view rather than from a subscriber point of view). Constraints (56) ensure that CG technology has to handle the consumption of both CG and NG subscribers when NG technology is not installed. Constraints (57) ensure that CG technology has to handle the consumption of CG subscribers (note that these constraints hold for all sites and for all time period, but are only active where and when NG technology is installed). Constraints (58) ensure that when installed, NG technology has to handle the consumption of NG subscribers. The threshold constraint can be reformulated as (59) since NG subscribers are served by NG if and only if NG is installed. Constraints (60)–(62) are obtained by applying the definition of α .

Remark 10 *This formulation for the MMPP($\bar{\sigma}, \bar{c}$) still holds when smoothing constraints (25) are added.*

For the remainder of the section, we will focus on simplified formulations for the MMPP($\bar{\sigma}, \bar{c}$). First, we remove time dependency on variables m (replacing $m_{s,g}^t$ by $m_{s,g}$ the number of modules installed at the end of the time horizon for each site s and generation g), which means that smoothing constraints (25) cannot be added to the formulations provided (see Remark 1). Moreover, let us recall that $D_g^{t-1} \leq D_g^t$ for each $g \in \mathcal{G}$ and for each $t \in \mathcal{T}$. Notice that for each $t \in \mathcal{T}$, $U_{s,CG}^{t-1} \geq U_{s,CG}^t$ while $U_{s,NG}^{t-1} \leq U_{s,NG}^t$. Consequently, the above formulation can be simplified significantly. Namely, one readily verifies that constraints (56)–(58) are dominated by the following constraints:

$$D_{CG}^t U_s^{TOT} (1 - z_{s,NG}^t) \leq CAP_{CG} m_{s,CG} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \quad (67)$$

$$\max_{t \in \mathcal{T}} D_{CG}^t U_{s,CG}^t \leq CAP_{CG} m_{s,CG} \quad \forall s \in \mathcal{S}, \quad (68)$$

$$D_{NG}^{\bar{t}} U_{s,NG}^{\bar{t}} z_{s,NG}^{\bar{t}} \leq CAP_{NG} m_{s,NG} \quad \forall s \in \mathcal{S}. \quad (69)$$

We observe that we can use these constraints to compute a closed form for the optimal value taken by variables $m_{s,g}$ which depends on the values taken by variables $z_{s,NG}^t$. We know indeed that for each site $s \in \mathcal{S}$ the number of modules installed at the end of the time horizon is:

- for CG technology:

- if NG is already installed ($Z_{s,NG}^0 = 1$): $\tilde{m}_{s,CG}^{AI} = \max\left(\left[\frac{\max_{i \in \mathcal{T}} D_{CG}^i U_{s,CG}^i}{CAP_{CG}}\right], M_{s,CG}^0\right)$ (only CG subscribers are served by CG technology),
- if NG is not installed over the time horizon ($z_{s,NG}^{\bar{i}} = 0$): $\tilde{m}_{s,CG}^{NI} = \max\left(\left[\frac{D_{CG}^{\bar{i}} U_{s,CG}^{TOT}}{CAP_{CG}}\right], M_{s,CG}^0\right)$ (all subscribers are served by CG technology),
- if NG is installed at time period $t \in \mathcal{T}$ ($z_{s,NG}^t - z_{s,NG}^{t-1} = 1$):

$$\tilde{m}_{s,CG}^t = \max\left(\left[\frac{D_{CG}^{t-1} U_{s,CG}^{TOT}}{CAP_{CG}}\right], \left[\frac{\max_{i \geq t} D_{CG}^i U_{s,CG}^i}{CAP_{CG}}\right], M_{s,CG}^0\right),$$
- for NG technology when it is installed ($z_{s,NG}^{\bar{i}} = 1$): $\tilde{m}_{s,NG} = \max\left(\left[\frac{D_{NG}^{\bar{i}} U_{s,NG}^{\bar{i}}}{CAP_{NG}}\right], M_{s,NG}^0\right)$.

Note that if, on a site $s \in \mathcal{S}$, $\tilde{m}_{s,CG}^{AI} > \bar{M}_{CG}$ then $z_{s,NG}^{\bar{i}} = 1$ and if $\tilde{m}_{s,NG} > \bar{M}_{NG}$ then $z_{s,NG}^{\bar{i}} = 0$. If both happen, the instance is unfeasible.

Consequently, we can provide a formulation for the MMPP($\bar{\sigma}$, \bar{c}) using only binary variables z .

$$\begin{aligned} \min \sum_{s \in \mathcal{S}} \text{upgradecost} + CA_{NG} (z_{s,NG}^{\bar{i}} - Z_{s,NG}^0) \\ + (\tilde{m}_{s,NG} - M_{s,NG}^0) CM_{NG} z_{s,NG}^{\bar{i}} + (\tilde{m}_{s,CG}^{NI} - M_{s,CG}^0) CM_{CG} (1 - z_{s,NG}^{\bar{i}}) \\ + \sum_{t \in \mathcal{T}} (\tilde{m}_{s,CG}^t - M_{s,CG}^0) CM_{CG} (z_{s,NG}^t - z_{s,NG}^{t-1}) + (\tilde{m}_{s,CG}^{AI} - M_{s,CG}^0) CM_{CG} Z_{s,NG}^0 \end{aligned} \quad (70)$$

$$s.t. \sum_{s \in \mathcal{S}} U_{s,NG}^{\bar{i}} z_{s,NG}^{\bar{i}} \geq QoE \sum_{s \in \mathcal{S}} U_s^{TOT}, \quad (71)$$

$$z_{s,NG}^{t-1} \leq z_{s,NG}^t \quad \forall t \in \mathcal{T}, \quad (72)$$

$$\sum_{s \in \mathcal{S}} z_{s,NG}^{\bar{i}} \geq \bar{\alpha}, \quad (73)$$

$$\sum_{s \in \mathcal{S}} z_{s,NG}^{t-1} \geq \lceil L_{\bar{c}^t} N_S \rceil \quad \forall t \in \mathcal{T}, \quad (74)$$

$$\sum_{s \in \mathcal{S}} z_{s,NG}^{t-1} \leq \lfloor U_{\bar{c}^t} N_S \rfloor \quad \forall t \in \mathcal{T}, \quad (75)$$

$$Z_{s,NG}^0 = Z_{s,NG}^0 \quad \forall s \in \mathcal{S}, \quad (76)$$

$$z_{s,NG}^t \in \{0, 1\} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}. \quad (77)$$

We define by $S_{CG} \subset \mathcal{S}$ the subset of the sites where NG is not installed at the beginning of the time horizon. We also remove from set S_{CG} the sites for which we already know if we will install NG technology or not due to unfeasibilities. We notice that we only have to solve the problem on sites of set S_{CG} .

Our heuristic algorithm cannot afford to enumerate for each time period $t \in \mathcal{T}$ all possible values of $(\bar{\sigma}^t, \bar{c}^t) \in \mathcal{K} \times \mathcal{C}$ since that would result in solving $(K \times C)^{\bar{i}}$ problems. This approach would actually solve the MMPP exactly. Instead, we propose to test a restricted subset $\mathcal{W} \subset (\mathcal{K} \times \mathcal{C})^{\bar{i}}$ of different couples of values for each time period, see Algorithm 1.

Algorithm 1: Heuristic algorithm

INPUT : $\mathcal{W} \subset (\mathcal{K} \times C)^{\bar{t}}$;
for $(\bar{\sigma}, \bar{c}) \in \mathcal{W}$ do $cost(\bar{\sigma}, \bar{c}) \leftarrow$ optimal solution cost of $MMPP(\bar{\sigma}, \bar{c})$;
return $\min_{(\bar{\sigma}, \bar{c}) \in \mathcal{W}} cost(\bar{\sigma}, \bar{c})$

As explained next, for tractability reasons we assume that

$$\bar{c} = (c_{init}, C, \dots, C) \text{ for each } (\bar{\sigma}, \bar{c}) \in \mathcal{W}. \quad (78)$$

Assumption (78) implies that all network investments are performed in the first time period (see objective (70)). Hence, we can replace variables $z_{s,NG}^t$ for each $t \in \mathcal{T}$ by variables $z_{s,NG}^{\bar{t}}$.

All modules costs for sites of set $S \setminus S_{CG}$ are labeled into constant $netcost$. Constant N_{inst} labelled the number of sites where we know NG is installed at the end of the time horizon. For each site $s \in S_{CG}$, $C1_s = CA_{NG} + \bar{m}_{s,NG} CM_{NG}$ denotes the cost implied by deciding to install NG technology and $C2_s = (\bar{m}_{s,CG}^{NI} - \bar{m}_{s,CG}^{AI}) CM_{CG}$ the cost implied by deciding not to. Therefore, the $MMPP(\bar{\sigma}, \bar{c})$ can be further reformulated as the following bidimensional knapsack problem:

$$\min \sum_{s \in S_{CG}} C1_s z_{s,NG}^{\bar{t}} + C2_s (1 - z_{s,NG}^{\bar{t}}) + upgradecost + netcost \quad (79)$$

$$st \sum_{s \in S_{CG}} U_{s,NG}^{\bar{t}} z_{s,NG}^{\bar{t}} \geq QoE \sum_{s \in S} U_s^{TOT} - \sum_{s \in S \setminus S_{CG}} U_{s,NG}^{\bar{t}}, \quad (80)$$

$$\sum_{s \in S_{CG}} z_{s,NG}^{\bar{t}} \geq \max(\bar{\alpha}, LCNS) - N_{inst}, \quad (81)$$

$$z_{s,NG}^{\bar{t}} \in \{0, 1\} \quad \forall s \in S_{CG}. \quad (82)$$

Proposition 11 *The $MMPP(\bar{\sigma}, \bar{c})$ can be solved in $O(NS + |S_{CG}|^2(CA_{NG} + \bar{M}_{NG} CM_{NG}))$ when $\bar{c}^t = C, \forall t \in \{2, \dots, \bar{t}\}$.*

Proof From the arguments given above, this result follows from the fact that the problem becomes a knapsack problem which can be solved by a dynamic programming algorithm.

5 | CASE STUDY FOR 3G AND 4G

The purpose of this case study is two-fold. First, we assess the scalability of the exact MILP model, the impact of the proposed valid inequalities and the quality of our heuristic algorithm. Second, we observe the characteristics of the solutions in terms of costs, considering several businesses scenarios, including when the investment expenses are smoothed along the time horizon.

5.1 | Instances and platform

Numerical tests are performed on instances of French telecommunication operator Orange in the French areas of Brittany and Pays de la Loire representing a case study with two network generations: 3G and 4G. The full area contains 1075 sites: 700 equipped only with 3G technology and 375 equipped with both technologies. We create a set of smaller instances, out of this large instance, in order to have different scenarios characterized by the number of sites and the 4G initial coverage: rural scenarios where the 4G technology is initially deployed on 17% of the sites, suburban scenarios where this proportion is equal to 34% and urban scenarios where it is equal to 68%.

The real data include the number of modules and subscribers for each site. Others parameter values of this case study are realistic values taken from telecommunication equipment sellers. Each site can carry a maximum of four 3G modules (carrier) of 5MHZ with a capacity of 3 Mbps and a cost of 3 k€ each and a maximum of five 4G modules of 10 MHZ (bandwidth) with a capacity of 25 Mbps and a cost of 16 k€ each. The cost for adding the 4G technology on a site is 75 k€. As for subsidies, we considered 10 values for the of subsidy offered, $\sigma \in \{0, 100, 150, 200, 250, 300, 350, 400, 450, 500\}$ €. We also rely on four levels of 4G sites coverage: *low*, *medium low*, *medium high* and *high* respectively refer to ranges (in%) $[0, 25]$, $[25, 50]$, $[50, 75]$ and $[75, 100]$. The curves from Figure 2 are hence discretized into the table of values shown in Table 3, which constitutes our reference upgrade function. Two other upgrades functions are also consid-

Coverage level \ Subsidies (in €)	0	100	150	200	250	300	350	400	450	500
<i>low</i>	0	0	0	0	0	0	0	0	0	0
<i>medium low</i>	0.5	5	12	21	30	40	42	43	44	45
<i>medium high</i>	5	10	20	30	40	50	60	62	64	65
<i>high</i>	10	20	33	45	58	70	80	83	87	90

TABLE 3 Reaction of the subscribers (in%) on reference markets for given subsidies and coverage levels.

ered for adapting to very technology-reluctant markets and technology-friendly markets. We ask for final target objectives of $\underline{QoE} = 80\%$ for the quality of experience and $\underline{\alpha} = 70\%$ for the 4G sites coverage. We optimize on the typical time horizon of five years discretized in five time periods of one year. The unit of the objective value is k€.

The computations have been made on a server of 32 processors Intel Xeon of CPU 5110 clocked at 1.6 GHz each. The code has been written in Julia 1.1.0, and the solver used is CPLEX 12.8 (default branch-and-bound algorithm). The time limit for MILP solving is set to 1800 seconds for Sections 5.2 and 5.3. We have observed in our experiments that removing the dependency of variables m on time period t does not help improving the solution found nor the gap.

5.2 | Results for algorithmic tests

Our objectives in this section are to assess the scalability of our formulation (including the impact of the proposed valid inequalities) and to test its sensitivity to the type of areas. Hence we focus on twelve instances considering four sizes between 50 and 200 sites and the three types of areas mentioned above. The upgrade function used here is the reference one (see Table 3). Instances features are displayed in Tables 4, 5 and 6, column " N_S " standing for the number of sites and column "density" standing for the density scenario (rural R , suburban S or urban U). The solutions of the linear relaxation and of the MILP are computed with and without the valid inequalities from Section 4.1. more precisely, we test formulations (\mathcal{M}), (\mathcal{M} + each family of valid inequality) and (\mathcal{M} + all families of valid inequalities). The obtained root gap for each tested formulation is displayed in Table 4. The best root gap among the formulations with a single valid inequality is in bold. For each formulation, the value of the best solution found by CPLEX within the time limit is displayed in Table 5. We label the value with the character "*" if the branch-and-bound procedure converged. The best solution found is in bold. The corresponding final gap is displayed in Table 6. The best value for the final gap is in bold and the second best is in italic.

TABLE 4 Root gap values for 12 instances (4 sizes, 3 densities) tested with each family of valid inequalities

Instance		Root gap							
N_S	density	\mathcal{M}	+ (43)	+ (44)	+ (45)	+ (46)	+ (47)	+ (48)	+ (43)-(48)
50	R	25	25	25	19	25	25	25	19
	S	32	32	32	24	32	32	32	24
	U	61	58	61	45	61	61	55	27
100	R	26	26	26	20	26	26	26	20
	S	31	32	31	24	31	31	31	23
	U	62	59	62	47	62	62	57	28
150	R	28	25	25	20	26	25	25	20
	S	38	31	31	24	31	32	31	24
	U	63	59	62	47	62	62	57	29
200	R	28	25	24	18	24	24	25	18
	S	36	31	31	23	30	32	30	23
	U	63	59	62	46	62	62	56	29

First, we provide insight on the relative efficiency of each family of valid inequalities. We observe, on Table 4, tighter relaxations when using the valid inequalities. More precisely, a significant improvement on the relaxation is enabled by the RLT set of inequalities (45). Improvements on the root gaps can also be seen in urban instances by adding the sets (43) or (48). Indeed, when the new technology is already deployed on most sites, the non-decreasing of z reduces the search space.

TABLE 5 Best solution for 12 instances (4 sizes, 3 densities) tested with each family of valid inequalities

Instance		Best solution found by CPLEX							
N_S	density	M	+ (43)	+ (44)	+ (45)	+ (46)	+ (47)	+ (48)	+ (43)-(48)
50	R	4173	4103*						
	S	3458	3458*	3458*	3458*	3458*	3458*	3458*	3458*
	U	2021*	2021*	2021*	2021*	2021*	2021*	2021*	2021*
100	R	8347	8347	8347	8347	8401	8347	8347	8347
	S	7036	6902	6861	6861	6861	6861	6861	6861
	U	3864	3861*						
150	R	12783	12308	12247	12242	12344	12247	12263	12242
	S	11179	10049	10027	9990	10049	10146	10030	9990
	U	5692	5522*						
200	R	17021	16309	16167	16036	16127	16036	16318	16036
	S	14266	13305	13305	13094	13094	13567	13221	13094
	U	7828	7616	7616	7616	7619	7616	7619	7616

For this reason, these instances are the ones for which adding all inequalities rather than only inequalities (45) significantly improves the root gap. Moreover, we observe that inequalities (45) also have the best impact among the different valid inequalities for reducing the final gap (see Table 6). The solution found with inequalities (45) is always the best one found, as it can be observed in Table 5, but using the inequalities all together enables us to find the same solutions, with (for most instances) a slightly lower final gap.

For the following tests and observations, we will hence focus on the formulation including all families of valid inequalities. First, referring to scalability, we observe that the branch-and-bound procedure converges to optimality for the three instances of 50 sites and the urban instances of 100 and 150 sites (and nearly converges for the urban instance of 200 sites). Besides, the final gap remain under 4 % for all instances. Second, we focus on the sensitivity to the type of area. We notice that the problem proves easier to solve in urban areas, which can be explained by the decisions on coverage extension needed to satisfy the strategic targets: having more sites already covered by the newest technology results indeed in fewer decisions to take. However, the relaxation is weaker (see Table 4) due to the shape of the subsidy function (beginning with an high coverage means higher reactions but also higher gaps between continuous reactions and discrete reactions). This enlightens that the problem practical difficulty is strongly correlated with the question of coverage extension. We notice that these resulting coverage extension investments also have a significant financial impact. For instance, the optimal solution for the rural instance of 50 sites is around two times more expensive than the optimal solution for the urban instance.

TABLE 6 Final gaps for 12 instances (4 sizes, 3 densities) tested with each family of valid inequalities

Instance		Final gap							
N_S	density	M	+ (43)	+ (44)	+ (45)	+ (46)	+ (47)	+ (48)	+ (43)-(48)
50	R	6.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	S	5.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	U	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	R	10.62	4.52	3.13	3.59	6.33	1.90	6.46	1.14
	S	14.92	4.81	3.17	2.55	3.60	3.67	4.37	2.50
	U	7.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00
150	R	16.71	9.65	5.62	4.43	7.91	6.87	6.87	4.12
	S	20.72	10.49	4.34	3.91	4.13	10.45	7.71	3.47
	U	7.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00
200	R	16.73	10.53	7.99	3.59	10.55	8.84	9.80	2.48
	S	18.85	10.06	12.22	2.77	11.03	13.59	12.28	3.06
	U	7.73	1.83	1.01	0.49	1.96	1.86	2.27	0.07

5.3 | Results for business-oriented tests

In this section, we assess the impact of several business-oriented scenarios from both algorithmic and financial perspectives:

- smoothing the costs over the time horizon,
- considering three upgrade functions: the reference upgrade used in the algorithmic tests, the technology-reluctant and technology-friendly upgrades.

For these purposes, and in order to have a sufficient expected number of optimal solutions (according to the algorithmic tests) for assessing financial aspects, we consider a set of 18 instances made of

- for the 50 sites instances: the three densities and the three types of market (9 instances)
- for the 100 to 200 sites instances: suburban density and the three types of markets (9 instances).

These 18 instances are presented in Table 7. The type of market is displayed in column “upgrade”: “T.F markets”, “Ref. markets” and “T.R. markets” stand, respectively for technology-friendly, reference and technology-reluctant markets.

We want to assess on these instances the impact of cost smoothing, i-e when we add con-

straints (25). These constraints enforce all period expenses to lie between $(1 - p)$ and $(1 + p)$ times the quotient of the total expenses over the time horizon by the number of time periods, where p is a parameter setting the maximal allowed budget fluctuation - set at 10% for the following tests.

Note that the valid inequalities still hold, except for inequalities (48), which is hence removed. Results for both formulations “ \mathcal{M} ” and “ \mathcal{M} + cost equilibrium” are displayed in Table 7. The indicators provided are the same as in Section 5.2. The last column “overcost” gives the resulting overcost (the relative gap in % between the values of the solutions without and with the cost equilibrium set of constraints).

From a computational point of view, adding the cost equilibrium constraints hardens the problem. We see indeed in Table 7 that the proof of optimality is obtained only for the urban instances of 50 sites.

TABLE 7 Best solution found, final gap and root gap for cost equilibrium and other upgrade functions

Instance			\mathcal{M}			\mathcal{M} + cost equilibrium			overcost
50	R	T.F. markets	3622	0.00	23	4410	5.19	32	22
		Ref. markets	4103	0.00	19	4589	0.90	26	12
		T.R. markets	4417	0.00	19	4962	2.05	26	12
S		T.F. markets	2890	0.00	30	3600	1.73	40	25
		Ref. markets	3458	0.00	23	3877	1.70	31	12
		T.R. markets	3847	0.00	23	4192	0.04	28	9
U		T.F. markets	1264	0.00	43	1742	0.00	77	38
		Ref. markets	2021	0.00	27	2043	0.00	11	1
		T.R. markets	2443	0.00	22	2470	0.00	34	1
100	S	T.F. markets	5729	0.00	29	7142	2.54	40	25
		Ref. markets	6861	0.25	23	7767	7.95	31	13
		T.R. markets	7625	2.83	22	8389	7.64	28	10
150	S	T.F. markets	8459	1.71	30	13650	26.22	54	61
		Ref. markets	9990	0.98	23	11642	11.12	34	17
		T.R. markets	11150	3.59	23	12130	6.66	29	9
200	S	T.F. markets	11075	3.27	28	17984	26.15	53	62
		Ref. markets	13094	3.06	23	35768	52.99	71	173
		T.R. markets	14517	3.43	22	15949	6.49	28	10

From a financial point of view, we consider the 50 sites instances in order to analyze the char-

acteristics of an optimal solution. We thus draw the features of the solution for the 50 sites suburban instance with the upgrade for the reference markets, when we do not require the cost to be smoothed, in Figure 3. The 4G sites coverage, the amount of subsidies given and the reaction of the subscribers are plotted. On the 4G sites coverage curve, we can notice that the 4G sites coverage at the end of the time horizon is 80% and that this value is not reached progressively throughout the time horizon. Indeed, the 4G sites coverage at the end of the first year is already nearly equal to this final value. This fast deployment is made to benefit from more upgrade thanks to coverage improvements (for instance the natural effect of coverage improvement can be observed at the second time period, over which the switch from coverage range *medium low* to range *high* enables the model not to offer any subsidy). This enlightens the financial interest for the operator in quickly having a network of good quality. However, it results also in large budget variations, with the first year costing more than four times the second most expensive year, and almost nothing spent over the second year, as we can see in Figure 4. These important variations do not match with the financial context of a telecommunication operator as investments should be distributed along the whole time horizon.

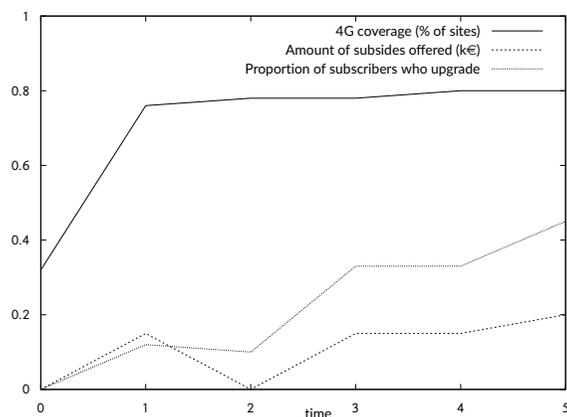


FIGURE 3 Evolution of the coverage, subsidies decisions and reactions over the time horizon (optimal solution of 50 sites suburban instance).

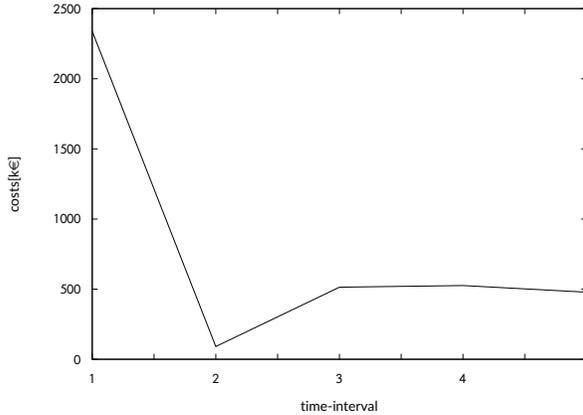


FIGURE 4 Evolution of the costs over the time horizon (optimal solution of 50 sites suburban instance).

We plot in Figure 5 the counterpart of Figure 3 when cost are required to be smoothed. We see in these curves that the range of coverage *high* is reached in four years in the solution with cost equilibrium instead of one without cost equilibrium (see Figure 3). The effect of the subsidies is hence considerably weakened, which can be observed on the reaction curve of Figure 5. This has an impact on the upgrade investments, which become higher (for instance 250€ per user instead of 200 are offered in the last year). Besides, in the second year, a subsidy of 150€ is offered while in the solution without imposing cost equilibrium the coverage improvements enabled the model not to offer any subsidy. These have important effects on the costs of the corresponding time periods, as can be observed in Figures 4 and 6. By comparing the previous instance with the other 50 sites instances for reference markets in rural and urban areas, we see the influence of the initial density on the overcost resulting from the cost smoothing. This effect can be seen in Table 7, column “overcost” and rows *Ref. markets*: the overcost is around 12% for rural and suburban instances while it is only 1% for urban ones. The needed investments for reaching the upper range of coverages are indeed lower when starting from higher initial 4G sites coverages, reducing gap between solutions with and without cost equilibrium. We hence are able to quantify the overcost to get business-fit solutions and see that this overcost is particularly reduced for urban instances.

Finally, with regards to the type of markets, we notice that there seems to be no algorithmic sensitivity to the choice of the upgrade function. In what concerns the financial sensitivity, we notice that, as expected, since more upgrade investments are needed, the cost is higher on reluctant markets. For instance, the cost for the suburban instance of 50 sites with the upgrade for reluctant markets is 33% more expensive than the instance with the upgrade for technology-friendly markets.

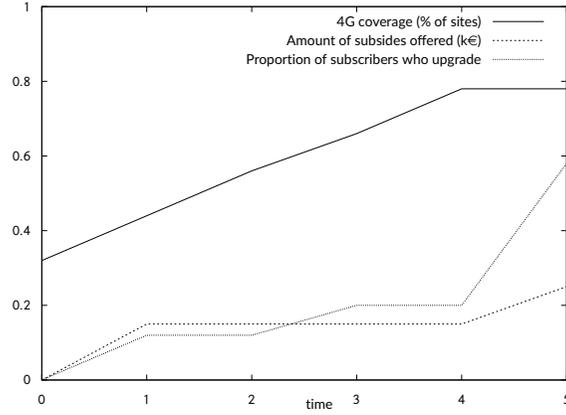


FIGURE 5 Evolution of the coverage, subsidies decisions and reactions over the time horizon (solution of 50 sites suburban instance) when cost equilibrium constraints are imposed.

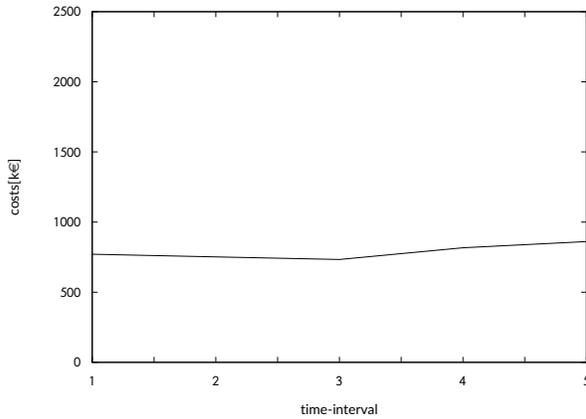


FIGURE 6 Evolution of the costs over the time horizon (solution of 50 sites suburban instance) when cost equilibrium constraints are imposed.

5.4 | Computational tests on large instances

In this section, we assess our exact solution method on 10 instances corresponding to different French territorial divisions and compare it with the heuristic of Section 4.2. Two regions: Bretagne (divided into 4 departments: Finistère, Côtes d'Armor, Morbihan and Ile et Vilaine) and part of Pays de la Loire (divided into 3 departments: Mayenne, Sarthe, Maine et Loire) are hence considered. As the planification is made for 5 years, the computational time is not what matters most for the operator, so we test larger time limits in order to see if it enables us to find better solutions and to

reduce the gaps.

Instances are displayed in Table 8. The name of the territorial division (department/region), its number of sites and its initial 4G sites coverage in % are respectively stored under “Ter. Div.,” “ N_S ” and “ α^0 ”. The best solution found, the final gap and root gap are labeled in the same way as before, and indicated for three different time limits: half-an hour, two hours and five hours.

We observe that, with a five-hour time limit, the final gap obtained is below 5% for 8 large real-life instances. However, the model struggles to find a good quality solution for the two largest instances.

TABLE 8 Solution and final gap for large instances

Instance			MILP (half an hour)		MILP (two hours)		MILP (five hours)	
Ter. Div.	N_S	α^0	sol	f-gap	sol	f-gap	sol	f-gap
Finistère	210	36	13885	7.00	13406	4.91	13406	4.61
Côtes d’Armor	149	29	10420	3.38	10420	1.94	10420	1.48
Morbihan	168	38	11178	4.08	11178	3.32	11178	2.75
Ile et Vilaine	214	43	12400	3.76	12115	2.73	12115	2.15
Mayenne	73	31	4879	1.62	4879	0.92	4879	0.50
Sarthe	116	33	7729	3.23	7729	2.39	7728	1.64
Maine et Loire	145	28	9877	4.68	9877	4.05	9877	3.65
Bretagne	741	37	∞	∞	128109	100.00	128109	57.03
Pays de la Loire	334	30	∞	∞	22470	4.26	22464	4.00
Full instance	1075	35	∞	∞	169968	92.80	169968	92.80

We hence assess the interest of the heuristic of Section 4.2 for finding feasible solutions. As mentioned previously, we will look for a solution where the range (“high”) is reached over the first period and we enumerate the amount of subsidies $\bar{\sigma} \in \mathcal{W}$ (ten possibilities if we do not restrict) at each time period (five) so as to solve each resulting problem $\text{MMPP}(\bar{\sigma}, \bar{c})$ with the pseudo-polynomial model provided in Section 4.2. This means that we have to solve 10^5 $\text{MMPP}(\bar{\sigma}, \bar{c})$ problems, which we cannot afford. In addition, we know that when the subsidies reaction is not sufficient to reach the threshold of 80% of subscribers being 4G subscribers, the corresponding problem is unfeasible. We also observe that with 250 euros of subsidies at each period, we have less than 1% of the initial 3G subscribers remaining at the end of the time horizon, and that this situation is feasible for our instances. Hence, in our heuristic, we enumerate all $\bar{\sigma} \in \{0, 100, 150, 200, 250\}$ and we solve the problem only if the reaction is sufficient to reach the threshold. This give at most $5^5 = 3125$ $\text{MMPP}(\bar{\sigma}, \bar{c})$ problems to solve.

In a second step, the solution found by the heuristic is used as initial solution (MIPstart) for the solver. The time limit given to the solver is 7200 seconds minus the time of the heuristic in order to compare with the MIP solving in 7200 seconds without initial solution provided.

Results are presented in Table 9. The column “heuristic” stands for the algorithm described above, the column “MILP” for the MILP without initial solution provided and the column “MIPstart” for the MILP with the heuristic solution provided as MIPstart. The column “gapMILP” reports the gap between the heuristic value and the MILP solution value.

TABLE 9 Solution and final gap for large instances

Instance			heuristic		gapMIP	MILP		MIPstart	
Ter. Div.	N_S	α^0	sol	time		sol	f-gap	sol	f-gap
Finistère	210	36	13406	505	0	13406	4.91	13406	3.14
Côtes d'Armor	149	29	10420	617	0	10420	1.94	10420	1.64
Morbihan	168	38	11178	551	0	11178	3.32	11178	2.06
Ile et Vilaine	214	43	12115	776	0	12115	2.73	12115	2.32
Mayenne	73	31	4879	127	0	4879	0.92	4879	0.00
Sarthe	116	33	7729	186	0	7729	2.38	7229	0.00
Maine et Loire	145	28	9877	221	0	9877	4.06	9877	0.72
Bretagne	741	37	47106	3197	-63.41	128109	100.00	47106	3.51
Pays de la Loire	334	30	22467	4113	-0.01	22470	4.26	22464	3.01
Full instance	1075	35	69497	5997	-59.00	169968	92.80	69497	5.42

We observe that the heuristic finds very good quality solutions for all instances in two hours of total computation time (heuristic + MIPstart). For the two largest instances, these solutions are far better (around 60% savings) than the best solution found without heuristic by the MIP in two hours. These solutions are not improved afterwards by CPLEX but using the heuristic as MIPstart enables to obtain the proof of convergence for the two smallest instances and to have all final gaps below 6%.

6 | CONCLUSION

In this article, we introduced a problem of multi-year investments planning for a telecommunication operator. Encompassing several real aspects faced by operators, our problem consists in optimizing network and subscriber dynamics under capacity and strategic constraints. In particular, we have modeled the fraction of subscribers adopting a new technology as depending on the coverage of that technology. In addition, the operator can provide subsidies to encourage the subscribers to shift faster to that technology. We have provided a non-linear MIP formulation for this problem, which we linearize and reinforce with several sets of valid inequalities. Computational tests have been made for a real 3G/4G case-study. The efficiency of the valid inequalities in improving the performances has been underlined, as well as the relevance of the branch-and-bound procedure performed on the tightened MILP for solving scaled real-life instances. For the largest instance,

the solver struggles to find a feasible solution and so we have proposed a heuristic to find a good-quality primal feasible solution. This heuristic is based on fixing the upgrade parameters (subsidies amount and state of coverage) and enables us to find very good quality solutions while running much faster than the branch-and-bound procedure on the exact MILP. Our results also illustrate the effect of imposing additional business-oriented constraints needed from the operational viewpoint. For instance, we have modeled the possibility for the operator of smoothing its investments along the time horizon, which enables us to quantify the overcost due to smoothing policies. This overcost is significant, especially in suburban and rural areas. This could push operators to reconsider their investment policies and decide to perform an important initial investment in order to make savings on the whole time horizon.

References

- [1] ARCEP, 2018. Arcep rankings. <https://www.arcep.fr/actualites/les-communiques-de-presse/detail/n/qualite-des-services-mobiles-1.html> (Accessed 10/30/2018).
- [2] F.M. Bass, *Empirical generalizations and marketing science: A personal view*, *Mark. Sci.* **14** (1995), G6–G19.
- [3] F.M. Bass, *Comments on "a new product growth for model consumer durables the bass model"*, *Manage. Sci.* **50** (2004), 1833–1840.
- [4] M. Chardy, M.B. Yahia, and Y. Bao, *3g/4g load-balancing optimization for mobile network planning*, 2016 17th International Telecommunications Network Strategy and Planning Symposium (Networks), Sept 2016, pp. 7–12.
- [5] CISCO, 2017. Cisco visual networking index: Global mobile data traffic forecast update, 2016–2021 white paper. <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html> (Accessed 10/30/2018).
- [6] DigitalInformationWorld, 2018.. <https://www.digitalinformationworld.com/2018/05/infographic-internet-minute-2018.html> (Accessed 10/30/2018).
- [7] Fortet, *Boole algebra and its application to operation research.*, *Trabajos de Estadística* (1960), 111–118.
- [8] B.L. García, P. Mahey, and L.J. LeBlanc, *Iterative improvement methods for a multiperiod network design problem*, *Eur. J. Oper. Res.* **110** (1998), 150–165.
- [9] M. Gendreau, J.Y. Potvin, A. Spires, and P. Soriano, *Multi-period capacity expansion for a local access telecommunications network*, *Eur. J. Oper. Res.* **172** (2006), 1051–1066.
- [10] S. Gollowitzer, L. Gouveia, and I. Ljubic, *Enhanced formulations and branch-and-cut for the two level network design problem with transition facilities*, *Eur. J. Oper. Res.* **225** (2013), 211–222.
- [11] GSMA, 2018. The mobile economy 2018 (gsm association annual report). <https://www.gsma.com/mobileeconomy/wp-content/uploads/2018/05/The-Mobile-Economy-2018.pdf> (Accessed 10/30/2018).

- [12] A. Knippel and B. Lardeux, *The multi-layered network design problem*, Eur. J. Oper. Res. **183** (2007), 87–99.
- [13] P. Kubat and J.M. Smith, *A multi-period network design problem for cellular telecommunication systems*, Eur. J. Oper. Res. **134** (2001), 439–456.
- [14] D.E. Lim and T. Kim, *An application of a multi-generation diffusion model to forecast 5g mobile telecommunication service subscribers in south korea*, Int. J. Pure Appl. Math. **116** (2017), 809–817.
- [15] C. Michalakelis, D. Varoutas, and T. Sphicopoulos, *Diffusion models of mobile telephony in greece*, Telecommunications Pol. **32** (2008), 234 – 245.
- [16] J.A. Norton and F.M. Bass, *A diffusion theory model of adoption and substitution for successive generations of high-technology products*, Manage. Sci. **33** (1987), 1069–1086.
- [17] Orange, 2018. Mobile network worldwide activity (orange group webpage). <https://www.orange.com/en/Group/Activities/Networks/Folder/Mobile-network> (Accessed 10/30/2018).
- [18] E. Rogers, *Diffusion of Innovations*, Simon and Schuster, 5th edition 2003.
- [19] P. Sharma, *Evolution of mobile wireless communication networks-1g to 5g as well as future prospective of next generation communication network*, Int. J. Comput. Sci. Mobile Comput. **2** (2013), 47–53.
- [20] H. Sherali and W. Adams, *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*, Nonconvex Optimization and Its Applications, Springer US, 1998.

Appendix: Extension to more-than-two generations frameworks

This model can be extended to tackle cases where several generations ($N \geq 3$ generations) co-exist. Three important questions arise for the operator when extending to more than two generations.

1. It has to decide its network investments policy, i.e. which technology(ies) can be deployed (or not) over the time horizon and for which technologies new modules can be added.
2. It has to define its subsidies policy, i.e. to which current subscriptions and for shifting to which technology(ies) the subsidies are offered. Note that, if allowed by the regulatory context, offered subsidies could have different values according to the current and/or targeted subscription. Finally, we stress the fact that reactions (and thus the modeled upgrade function) could also be different with respect to current and/or targeted technologies.
3. It has to define its load-balancing policy, i.e. which network technology(ies) is/are preferred to serve subscribers of different subscriptions, while respecting technical incompatibilities and deployment of the technology on the associated site. Note that this rule could be no preference among compatible and deployed technologies, a strict priority order (served by the most efficient compatible and deployed technology) or a mix of them (for instance, served by the most efficient and compatible technology if deployed, and no preference among less efficient technologies otherwise).

Let us define the framework where:

1. The operator allows module investments for all technologies, but the deployment of a technology on a site not yet equipped is limited to the newest technology. Besides, we suppose that the oldest generation is deployed everywhere at the beginning of the time horizon.
2. The subsidies are offered to subscribers from current subscriptions (all except the newest one) for shifting to the newest technology only. The subsidies offered to the subscribers and their reaction are considered independent of their initial subscription.
3. Subscribers cannot be served by a more recent technology than the one they have subscribed to (technical incompatibility). In addition, among the compatible technologies, we assume that they have to be served by the most recent and deployed technology (in particular the one they have subscribed to if it has been deployed).

Considering a general set of generations $\mathcal{G} = \{cg_1, \dots, cg_N, NG\}$ of N current generations ordered by efficiency and the newest one NG , the notation proposed in Section 3.1 still hold. The formulation proposed in Section 3.2 can be extended as follows for the framework we have defined:

$$\begin{aligned}
 \min \quad & \sum_{t \in T} \sum_{\sigma \in K} \sum_{c \in C} \sum_{s \in S} \sum_{o \in \mathcal{G} \setminus \{NG\}} \sigma f_{\sigma,c} \pi_{\sigma,c,s,o}^t & + \sum_{s \in S} \sum_{g \in \mathcal{G}} CM_g (m_{s,g}^i - M_{s,g}^0) \\
 & + \sum_{s \in S} CANG (z_{s,NG}^i - Z_{s,NG}^0) & \quad (83)
 \end{aligned}$$

$$\text{s.t. } m_{s,g}^t \leq Z_{s,g}^0 \overline{M}_g \quad \forall s \in S, \forall t \in \mathcal{T}, \forall g \in \mathcal{G} \setminus \{NG\}, \quad (84)$$

$$m_{s,NG}^t \leq \overline{M}_{NG} z_{s,NG}^t \quad \forall s \in S, \forall t \in \mathcal{T}, \quad (85)$$

$$m_{s,g}^{t-1} \leq m_{s,g}^t \quad \forall s \in S, \forall t \in \mathcal{T}, \forall g \in \mathcal{G}, \quad (86)$$

$$u_{s,o}^t = \sum_{g \in \mathcal{G}: g \leq o} u_{s,o,g}^t \quad \forall s \in S, \forall t \in \mathcal{T}, \forall o \in \mathcal{G}, \quad (87)$$

$$Z_{s,k}^0 \sum_{o \in \mathcal{G}: o \geq k} \sum_{g \in \mathcal{G}: g < k} u_{s,o,g}^t = 0 \quad \forall s \in S, \forall t \in \mathcal{T}, \forall k \in \mathcal{G} \setminus \{NG\} \quad (88)$$

$$\sum_{g \in \mathcal{G} \setminus \{NG\}} u_{s,NG,g}^t \leq \overline{U}_{s,NG}^t (1 - z_{s,NG}^t) \quad \forall s \in S, \forall t \in \mathcal{T}, \quad (89)$$

$$D_g^t \sum_{o \in \mathcal{G}} u_{s,o,g}^t \leq CAP_g m_{s,g}^t \quad \forall s \in S, \forall t \in \mathcal{T}, \forall g \in \mathcal{G}, \quad (90)$$

$$u_{s,o}^t = u_{s,o}^{t-1} - \sum_{\sigma \in \mathcal{K}} \sum_{c \in C} f_{\sigma,c} \pi_{\sigma,c,s,o}^t \quad \forall s \in S, \forall t \in \mathcal{T}, \forall o \in \mathcal{G} \setminus \{NG\}, \quad (91)$$

$$u_{s,NG}^t = u_{s,NG}^{t-1} + \sum_{o \in \mathcal{G} \setminus \{NG\}} \sum_{\sigma \in \mathcal{K}} \sum_{c \in C} f_{\sigma,c} \pi_{\sigma,c,s,o}^t \quad \forall s \in S, \forall t \in \mathcal{T}, \quad (92)$$

$$\sum_{s \in S} u_{s,NG,NG}^i \geq \underline{QoE} \left(\sum_{s \in S} \sum_{o \in \mathcal{G}} U_{s,o}^0 \right) \quad (93)$$

$$\alpha^i \geq \underline{\alpha} \quad (94)$$

$$\sum_{\sigma \in \mathcal{K}} \sum_{c \in C} \delta_{\sigma,c}^t = 1 \quad \forall t \in \mathcal{T}, \quad (95)$$

$$\sum_{\sigma \in \mathcal{K}} \delta_{\sigma,c}^t \leq 1 + U_c - \alpha^{t-1} \quad \forall t \in \mathcal{T}, \forall c \in C, \quad (96)$$

$$\sum_{\sigma \in \mathcal{K}} \delta_{\sigma,c}^t \leq 1 + \alpha^{t-1} - L_c \quad \forall t \in \mathcal{T}, \forall c \in C, \quad (97)$$

$$\pi_{\sigma,c,s,o}^t \leq \delta_{\sigma,c}^t \overline{U}_{s,o}^{t-1} \quad \forall s \in S, \forall t \in \mathcal{T}, \forall \sigma \in \mathcal{K}, \forall c \in C, \forall o \in \mathcal{G} \quad (98)$$

$$\pi_{\sigma,c,s,o}^t \leq u_{s,o}^{t-1} \quad \forall s \in S, \forall t \in \mathcal{T}, \forall \sigma \in \mathcal{K}, \forall c \in C, \forall o \in \mathcal{G}, \quad (99)$$

$$\pi_{\sigma,c,s,o}^t \geq u_{s,o}^{t-1} - (1 - \delta_{\sigma,c}^t) \overline{U}_{s,o}^{t-1} \quad \forall s \in S, \forall t \in \mathcal{T}, \forall \sigma \in \mathcal{K}, \forall c \in C, \forall o \in \mathcal{G}, \quad (100)$$

$$m_{s,g}^t \in \mathbb{Z} \quad \forall s \in S, \forall t \in \mathcal{T} \cup \{0\}, \forall g \in \mathcal{G}, \quad (101)$$

$$z_{s,NG}^t \in \{0, 1\} \quad \forall s \in S, \forall t \in \mathcal{T} \cup \{0\}, \quad (102)$$

$$u_{s,o}^t \geq 0 \quad \forall s \in S, \forall t \in \mathcal{T} \cup \{0\}, \forall o \in \mathcal{G}, \quad (103)$$

$$u_{s,o,g}^t \geq 0 \quad \forall s \in S, \forall t \in \mathcal{T}, \forall o, g \in \mathcal{G}^2, \quad (104)$$

$$\delta_{\sigma,c}^t \in \{0, 1\} \quad \forall t \in \mathcal{T}, \forall \sigma \in \mathcal{K}, \forall c \in C, \quad (105)$$

$$\pi_{\sigma,c,s,o}^t \in \{0, 1\} \quad \forall t \in \mathcal{T}, \forall \sigma \in \mathcal{K}, \forall c \in C, \forall s \in S. \quad (106)$$

$$u_{s,o}^0 = U_{s,o}^0 \quad \forall s \in S, \forall o \in \mathcal{G}, \quad (107)$$

$$m_{s,g}^0 = M_{s,g}^0 \quad \forall s \in S, \forall g \in \mathcal{G}, \quad (108)$$

$$z_{s,NG}^0 = Z_{s,NG}^0 \quad \forall s \in \mathcal{S}, \quad (109)$$

$$m_{s,g}^t \in \mathbb{Z} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \cup \{0\}, \forall g \in \mathcal{G}, \quad (110)$$

$$z_{s,NG}^t \in \{0, 1\} \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \cup \{0\}, \quad (111)$$

$$u_{s,o}^t \geq 0 \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \cup \{0\}, \forall o \in \mathcal{G}, \quad (112)$$

$$u_{s,o,g}^t \geq 0 \quad \forall s \in \mathcal{S}, \forall t \in \mathcal{T}, \forall o, g \in \mathcal{G}^2, \quad (113)$$

The objective function (83) minimizes both subscribers migration costs and network investments. The first term stands for the subsidies offered (user upgrades); the second term for the adding of new modules for increasing the capacity (densification); and the third term for the deployment of the newest technology NG (coverage extension).

Constraints (84)–(86) are the network dynamic constraints. Constraints (84)–(85) define the upper bounds on the numbers of modules for each technology added on each site. These constraints also ensure that if a technology is not deployed, no corresponding modules can be added. Constraints (86) prevent from decommissioning by imposing the number of modules of each technology to be non-decreasing during the time horizon.

Constraints (87)–(90) are the network dimensioning constraints, in charge of making the link between the network and the subscriber dynamic. Constraints (87) ensure the technical incompatibility stating that subscribers cannot be served by a more recent technology than the one they have subscribed to. Constraints (88) and (89) ensure the load-balancing rule. Constraints (90) are the capacity constraints: the installed capacities of each technology on each site have to be sufficient for providing services for all users located at this site and having to be served by this technology.

Constraints (91)–(92) are the subscriber dynamic constraints. They define the total number of subscribers to each technology at each site and each time period, taking into account subscribers to current generations who decide to upgrade to NG technology, thanks to subsidies and coverage improvements. Constraints (93)–(94) stand for the model strategic guidelines and refer to the end of time horizon. Constraint (93) ensures the threshold of subscribers covered by the newest technology is met. The indicator is proportional to the quality of experience which measures the percentage of users having access to the new technology throughput. Constraint (94) imposes that the threshold on the number of sites on which NG is deployed is met.

Constraints (95) ensure that one and only one subsidy from the set \mathcal{K} is offered at each time period, the case when no subsidy is given being represented by $\sigma = 0$. Constraints (96) and (97) ensure that, for each time period, variables $\delta_{\sigma,c}^t$ are set according to the coverage at the previous time period. Constraints (96) (respectively (97)) set all δ related to a range at 0 if the coverage is greater (resp. smaller) than the upper (resp. lower) bound of the range. Constraints (98)–(100) are the typical linearizations of the products of a binary variable with a continuous one. Constraints (101)–(106) define the domain of all variables in the formulation. Constraints (107)–(109) refer to the initial conditions. Finally, constraints (110)–(113) define the domain of the variables.

Remark 12 *The remark 1 on variables m still holds.*