

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

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

## ▶ To cite this version:

Adrien Cambier, Matthieu Chardy, Rosa Figueiredo, Adam Ouorou, Michael Poss. Optimizing the investments in mobile networks and subscriber migrations for a telecommunication operator. 2019. hal-02019853v2

# HAL Id: hal-02019853 https://hal.science/hal-02019853v2

Preprint submitted on 19 Jul 2019 (v2), 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.

#### Manuscript Draft

Manuscript Number:

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

Article Type: Research Article

Keywords: OR in Telecommunications; Mixed Integer Linear Programming;

Capacity Expansion; Bass model.

Corresponding Author: Mr. Adrien CAMBIER,

Corresponding Author's Institution: Orange Labs

First Author: Adrien CAMBIER

Order of Authors: Adrien CAMBIER; Matthieu CHARDY; Rosa FIGUEIREDO; Adam OUOROU; Michael POSS

Abstract: 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 guidelines requirements are demanded at the end of the time horizon. Following classical models, we consider that the willing of customers for shifting to a new technology follows an S-shape piecewise constant function. We propose a Mixed-Integer Linear Programming formulation and reinforce it through several valid inequalities. We assess the formulation numerically on real instances.

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

Adrien Cambier<sup>a,b</sup>, Matthieu Chardy<sup>a</sup>, Rosa Figueiredo<sup>b</sup>, Adam Ouorou<sup>a</sup>, Michael Poss<sup>c</sup>

<sup>a</sup>Orange Labs, 44 avenue de la République, 92320 Châtillon, France. <sup>b</sup>LIA, Avignon Université, Avignon, France. <sup>c</sup>UMR CNRS 5506 LIRMM, Université de Montpellier, Montpellier, France.

#### Abstract

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 guidelines requirements are demanded at the end of the time horizon. Following classical models, we consider that the willing of customers for shifting to a new technology follows an S-shape piecewise constant function. We propose a Mixed-Integer Linear Programming formulation and reinforce it through several valid inequalities. We assess the formulation numerically on real instances.

Keywords: OR in Telecommunications, Mixed Integer Linear Programming, Capacity Expansion, Bass model.

#### 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 (CISCO, 2017), traffic will globally reach 49 Exabytes per month in 2020 with a compound annual

 $Email\ address:\ {\tt adrien.cambier@orange.com}\ ({\rm Adrien\ Cambier})$ 

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 DigitalInformationWorld (2018))

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 Orange (2018)). 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 willing 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	$2.4~\mathrm{KB/s}$	64 KB/s	2  MB/s	$1~\mathrm{GB/s}$	> 1 GB/s

Table 2: Evolution of speed through mobile generations (from Sharma (2013)).

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 GSM<sup>1</sup> association (see GSMA (2018) 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 network (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

<sup>&</sup>lt;sup>1</sup>GSM association is an originally-European trade body that represents the interests of mobile network operators worldwide.

in 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 (ARCEP, 2018), strategic guidelines are decided by the operator at the end of 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, investments 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 Knippel and Lardeux (2007); Gollowitzer et al. (2013) among many others, including the case of multi-period planning (Garcia et al., 1998; Gendreau et al., 2006; Kubat and Smith, 2001).

Closer to the context of mobile capacity expansion, Chardy et al. (2016) present an exact mixed-integer formulation and a heuristic method to solve 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 Chardy et al. (2016) 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 users upgrades is hence not taken into account. Lim and Kim (2017) 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 move 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 Rogers (2003)), 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).

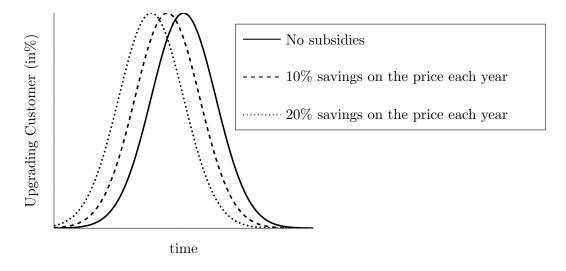


Figure 1: Curve of the diffusion of innovation and influence of subsidies

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 Bass (1995)) and generations (see Norton and Bass (1987)). 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 Bass (2004) for a discussion on 2G/3G upgrade; Lim and Kim (2017) for Bass model applied to the forecasting of the 5G upgrade; and Michalakelis et al. (2008) 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  $\sigma$ . 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 the strengthening of the MILP with several families of valid inequalities.

Our case study assesses in details the performance and solutions provided by the model on real instances. Our results show, in particular, the efficiency of the valid inequalities since the latter enable us to solve to near-optimality large real-life instances.

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, linearized and reinforced, is provided in Section 3. This model is numerically assessed in Section 4 on real-life instances. Concluding remarks are given in Section 5.

#### 2. Problem Description

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

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. For each site and for each deployed technology, adding new modules is possible with a unitary cost of  $CM_q$ ,  $g \in \mathcal{G}$ . With the deployment of the newest technology on a site comes at least the installation of one module simultaneously. 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 by site for each technology, which we note  $\overline{M}_g$ ,  $g \in \mathcal{G}$ . 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  $U_{s,o}^0, s \in \mathcal{S}, o \in \mathcal{G}$ . 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  $\sigma$ . The (finite) set of possible values taken by  $\sigma$ 

will be denoted by  $\mathcal{K}$ . 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 at the end of a given time period will be measured as the range of the proportion of sites on which NG technology is deployed, denoted  $\alpha$  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  $\mathcal{C} = \{1, \ldots, C\}$ . The function modeling the upgrade mechanism, denoted by  $f: \mathcal{K} \times \mathcal{C} \to [0,1]$ , 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  $c \in \mathcal{C}$  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 traffic demand depends on the network serving the subscribers rather than their current subscription. More precisely, a CG subscriber will have the demand  $D_{CG}^t$  at time period t and a t0 subscriber will have the demand t0 subscriber will have the proportion of sites covered by t0 at the end of the time horizon, which is denoted by t0 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 t0 subscribers associated with t0 sites. These subscribers benefit indeed from the new performances 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 t0 and t0.

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  $s \in \mathcal{S}$ ,
- $Z_{s,NG}^0$  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  $o \in \mathcal{G}$ ,
- ullet  $D_g^t$  is the demand of subscribers served by technology g at time period t,
- $f(\sigma, c)$  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  $c \in \mathcal{C}$ ,
- $\alpha^0$  stands for the sites coverage at the beginning of the time horizon,
- $\alpha^t$  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  $\alpha^t$ ,
- $\alpha$  and 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. In Section 3.3 we linearize this formulation. We present several valid inequalities for our model in Section 3.4. Endly, an extension considering more than two technologies is given in Section 3.5.

#### 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  $u^t_{s,o}$  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  $\overline{U}^t_{s,o}$ ),
- 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  $\sigma^t$ ,  $c^t$ , and  $\alpha^t$  introduced in the previous section become optimization variables:

- For each  $t \in \mathcal{T}$ , let  $\sigma^t$  be the value of the subsidy, in  $k \in$ , 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  $\alpha^t = \frac{\sum\limits_{s \in S} z^t_{s,NG}}{N_S}$  be the redundant variable that denotes the NG sites coverage at the beginning of the time period t, i.e. the fraction of sites where NG technology is deployed at the end of time period t-1,
- For each  $t \in \mathcal{T}$ , let  $c^t$  denote the interval of  $\mathcal{C}$  to which belongs  $\alpha^t$ .

The upgrade function, representing the percentage of users reacting positively to a subsidy  $\sigma \in \mathcal{K}$  for a given coverage  $c \in \mathcal{C}$ , 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:

$$\min \sum_{t \in \mathcal{T}} \sigma^{t} \times f\left(\sigma^{t}, c^{t-1}\right) \sum_{s \in S} u_{s,CG}^{t-1} + \sum_{s \in S} \sum_{g \in \mathcal{G}} CM_{g}(m_{s,g}^{\bar{t}} - M_{s,g}^{0})$$

$$+ \sum_{s \in S} CA_{NG}(z_{s,NG}^{\bar{t}} - Z_{s,NG}^{0})$$

$$(1)$$

s.t. 
$$m_{s,CG}^t \leq \overline{M}_{CG}$$
  $\forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T},$  (2) 
$$m_{s,NG}^t \leq \overline{M}_{NG} z_{s,NG}^t \qquad \forall \ s \in \mathcal{S} \ \forall \ t \in \mathcal{T},$$
 (3)

$$z_{s,NG}^t \le m_{s,NG}^t \qquad \forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T},$$
 (4)

$$m_{s,g}^{t-1} \le m_{s,g}^t$$
  $\forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T}, \ \forall \ g \in \mathcal{G},$  (5)

$$u_{s,CG}^t = u_{s,CG,CG}^t \qquad \forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T}, \tag{6}$$

$$u_{s,NG}^t = u_{s,NG,CG}^t + u_{s,NG,NG}^t \qquad \forall s \in \mathcal{S} \ \forall \ t \in \mathcal{T}$$
 (7)

$$u_{s,NG,CG}^t \le \overline{U}_{s,NG}^t (1 - z_{s,NG}^t) \qquad \forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T},$$
(8)

$$D_g^t \sum_{o \in \mathcal{G}} u_{s,o,g}^t \le CAP_g m_{s,g}^t \qquad \forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T}, \ \forall \ g \in \mathcal{G},$$

$$(9)$$

$$u_{s,CG}^{t} = u_{s,CG}^{t-1} - f\left(\sigma^{t}, c^{t-1}\right) \times u_{s,CG}^{t-1} \qquad \forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T},$$

$$(10)$$

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

$$(11)$$

$$\sum_{s \in \mathcal{S}} u_{s,NG,NG}^{\bar{t}} \ge \underline{QoE} (\sum_{s \in \mathcal{S}} U_{s,NG}^0 + U_{s,CG}^0), \tag{12}$$

$$\alpha^{\bar{t}} \ge \underline{\alpha},$$
 (13)

$$\alpha^t \in [L_{c^t}, U_{c^t}] \qquad \forall \ t \in \mathcal{T}, \tag{14}$$

$$u_{s,o}^{0} = U_{s,o}^{0} \qquad \forall s \in \mathcal{S}, \ \forall \ o \in \mathcal{G}, \tag{15}$$

$$m_{s,g}^0 = M_{s,g}^0 \qquad \forall s \in \mathcal{S}, \ \forall \ g \in \mathcal{G}, \tag{16}$$

$$z_{s,NG}^0 = Z_{s,NG}^0 \qquad \forall \ s \in \mathcal{S}, \tag{17}$$

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

$$z_{s,NG}^t \in \{0,1\} \qquad \forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T} \ \cup \ \{0\}, \tag{19}$$

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

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

$$\sigma^t \in \mathcal{K} \tag{22}$$

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

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)–(5) 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) impose that, when the newest technology is deployed on a site, at least one module of this technology is added. Constraints (5) prevent from decommissioning by imposing the number of modules of each technology to be non-decreasing during the time horizon.

Constraints (6)–(9) are the network dimensioning constraints, in charge of making the link between the network dynamic and the subscriber dynamic. Constraints (6) ensure the technical incompatibility stating that CG subscribers cannot be served by NG technology. Constraints (7)and (8) ensure the load-balancing rule. Constraints (9) 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 (10)–(11) 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 (12)–(13) stand for the model strategic guidelines and refer to the end of time horizon. Constraint (12) 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 (13) imposes that the threshold on the number of sites on which NG is deployed is met. Constraints (14) make the link between the sites coverage and the level of coverage used in the subsidy function. Constraints (15)–(17) refer to the initial conditions. Finally, constraints (18)–(23) define the domain of the variables. The next section details function f and linear modeling of constraints (14).

#### 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  $\sigma$  and the range of coverage c. Figure 2 provides an example with four ranges.

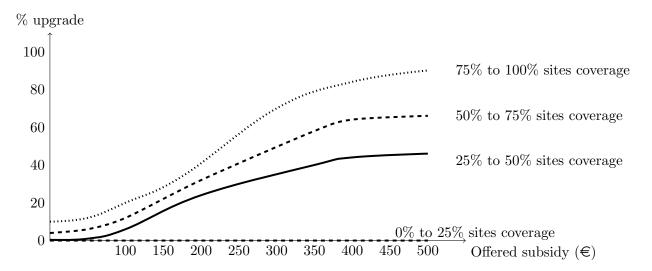


Figure 2: Example of upgrade function

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  $\sigma^t$  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 (10) and (11) can be written as:

$$u_{s,CG}^{t} = u_{s,CG}^{t-1} - \sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} f_{\sigma,c} \delta_{\sigma,c}^{t} u_{s,CG}^{t-1} \quad \forall \ s \in \mathcal{S} \ \forall \ t \in \mathcal{T},$$
 (24)

$$u_{s,NG}^{t} = u_{s,NG}^{t-1} + \sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} f_{\sigma,c} \delta_{\sigma,c}^{t} u_{s,CG}^{t-1} \quad \forall \ s \in \mathcal{S} \ \forall \ t \in \mathcal{T}.$$
 (25)

We linearize the products of binary variables  $\delta_{\sigma,c}^t$  and continuous variables  $u_{s,CG}^{t-1}$  using a classical method (Fortet, 1960). Consequently, the MMMP can be formulated as the following MILP.

$$\min \sum_{t \in \mathcal{T}} \sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} \sum_{s \in S} \sigma f_{\sigma,c} \pi_{\sigma,c,s,CG}^t + \sum_{s \in S} \sum_{g \in \mathcal{G}} CM_g(m_{s,g}^{\bar{t}} - M_{s,g}^0)$$

$$+ \sum_{s \in S} CA_{NG}(z_{s,NG}^{\bar{t}} - Z_{s,NG}^0)$$

$$(26)$$

s.t. 
$$(2) - (9), (12), (13), (15) - (17)$$

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

$$(27)$$

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

$$(28)$$

$$\sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} \delta_{\sigma,c}^t = 1 \qquad \forall t \in \mathcal{T}, \tag{29}$$

$$\sum_{\sigma \in \mathcal{K}} \delta_{\sigma,c}^t \le 1 + U_c - \alpha^{t-1} \qquad \forall \ t \in \mathcal{T}, \ \forall \ c \in \mathcal{C}, \tag{30}$$

$$\sum_{\sigma \in \mathcal{K}} \delta_{\sigma,c}^t \le 1 + \alpha^{t-1} - L_c \qquad \forall \ t \in \mathcal{T}, \ \forall \ c \in \mathcal{C}, \tag{31}$$

$$\pi_{\sigma,c,s,CG}^{t} \leq \delta_{\sigma,c}^{t} \overline{U}_{s,CG}^{t-1} \qquad \forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T}, \ \forall \ \sigma \in \mathcal{K}, \ \forall \ c \in \mathcal{C},$$
 (32)

$$\pi_{\sigma,c,s,CG}^{t} \leq u_{s,CG}^{t-1} \qquad \forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T}, \ \forall \ \sigma \in \mathcal{K}, \ \forall \ c \in \mathcal{C},$$
 (33)

$$\pi_{\sigma,c,s,CG}^{t} \ge u_{s,CG}^{t-1} - (1 - \delta_{\sigma,c}^{t})\overline{U}_{s,CG}^{t-1} \qquad \forall \ s \in \mathcal{S}, \ \forall \ t \in \mathcal{T}, \ \forall \ \sigma \in \mathcal{K}, \ \forall \ c \in \mathcal{C},$$
 (34)

$$m_{s,q}^t \in \mathbb{Z}$$
  $\forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T} \cup \{0\}, \ \forall \ g \in \mathcal{G},$  (35)

$$z_{s,NG}^t \in \{0,1\} \qquad \forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T} \ \cup \ \{0\}, \tag{36}$$

$$u_{s,o}^t \ge 0$$
  $\forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T} \ \cup \ \{0\}, \ \forall \ o \in \mathcal{G},$  (37)

$$u_{s,o,g}^t \ge 0$$
  $\forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T}, \ \forall \ o, g \in \mathcal{G}^2,$  (38)

$$\delta_{\sigma,c}^t \in \{0,1\} \qquad \forall t \in \mathcal{T}, \ \forall \ \sigma \in \mathcal{K}, \ \forall \ c \in \mathcal{C}, \tag{39}$$

$$\pi_{\sigma,c,s,CG}^{t} \ge 0 \qquad \forall t \in \mathcal{T}, \ \forall \ \sigma \in \mathcal{K}, \ \forall \ c \in \mathcal{C}, \ \forall \ s \in \mathcal{S}.$$
 (40)

We denote this formulation by (MMMP). Constraints (27) and (28) are the linearizations respectively of constraints (24) and (25). Constraints (29) 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 (30) and (31) ensure that, for each time period, variables  $\delta_{\sigma,c}^t$  are set according to the coverage at the previous time period. Constraints (30) (respectively (31)) set all  $\delta$  related to a range at 0 if the coverage is greater (resp. smaller) than the upper (resp. lower) bound of the range. Constraints (32)–(34) are the typical linearizations of the products of a binary variable with a continuous one. Constraints (35)–(40) define the domain of all variables in the formulation.

## 3.4. Valid inequalities

Preliminary computational experiments on small instances showed that the solutions of linear relaxation present variables z and  $\delta$  fractionary. Consequently, we propose several valid inequalities in this subsection to reinforce the model. The strength of these inequalities is assessed numerically in Section 4.2.

**Proposition 3.1.** Considering a time period  $t \in \mathcal{T} \cup \{0\}$  and a site  $s \in \mathcal{S}$ , inequality

$$z_{s,NG}^t \le z_{s,NG}^{t+1} \tag{41}$$

is valid for (MMMP).

*Proof.* This result is induced by constraints (3),(4) and (5).

**Proposition 3.2.** 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

$$\sum_{\sigma \in \mathcal{K}} \sum_{c' < c} \delta_{\sigma, c'}^{t'} \le 1 - \sum_{\sigma \in \mathcal{K}} \sum_{c' \ge c} \delta_{\sigma, c'}^{t}$$

$$\tag{42}$$

is valid for (MMMP).

*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 can not be smaller. As defined in Section 2, the NG sites coverage is indeed non-decreasing over the time horizon.

**Proposition 3.3.** Considering a time period  $t \in \mathcal{T}$  and a site  $s \in \mathcal{S}$ , equality

$$\sum_{c \in \mathcal{C}} \sum_{\sigma \in \mathcal{K}} \pi_{\sigma,c,s,CG}^t = u_{s,CG}^{t-1} \tag{43}$$

is valid for (MMMP).

Proof. Following the Reformulation Linearization Techniques (see Sherali and Adams (1998) for more details), we obtained these constraints by multiplying each constraint from set (29) 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 3.4.** Considering a time period  $t \in \mathcal{T}$  and a range of coverage  $c \in \mathcal{C}$ , inequality

$$\sum_{\sigma \in \mathcal{K}} \lceil N_S L_c \rceil \, \delta_{\sigma,c}^t \le \sum_{s \in \mathcal{S}} z_{s,NG}^t \tag{44}$$

is valid for (MMMP).

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

**Proposition 3.5.** 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  $s \in \mathcal{S}$ , inequality

$$\left\lceil \frac{D_{NG}^t \underline{U}_{s,NG}^t}{CAP_{NG}} \right\rceil z_{s,NG}^t \le m_{s,NG}^t \tag{45}$$

is valid for (MMMP).

*Proof.* If NG technology is deployed on a site  $s \in \mathcal{S}$  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 (9).

**Proposition 3.6.** Let  $\overline{U}_{s,NG}^t = U_{s,NG}^0 + U_{s,CG}^0[1 - (1 - \max_{\sigma} f_{\sigma,C})^t]$  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  $s \in \mathcal{S}$ , every optimal solution of (MMMP) verifies the following inequality:

$$m_{s,NG}^t \le \max(M_{s,NG}^0, \left\lceil \frac{D_{NG}^t \overline{U}_{s,NG}^t}{CAP_{NG}} \right\rceil) z_{s,NG}^t . \tag{46}$$

Proof. If NG technology is not deployed on a site  $s \in \mathcal{S}$  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  $s \in \mathcal{S}$  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 (9). Installing more than this bound costs  $CA_{NG}$  by additional module without any impact of the feasibility, and such a solution will hence be eliminated by the objective function minimization.

#### 3.5. 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, proposed 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 indifference between 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 indifference between 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).

A mixed-integer formulation for the Mobile Master Plan problem within this framework is provided in the appendix.

#### 4. 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 and the impact of the proposed valid inequalities. Second, we observe the characteristics of the solutions in terms of costs, considering several businesses scenario, including when the investment expenses are smoothed along the time horizon.

#### 4.1. 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 includes the number of modules and subscribers for each site. Others 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  $k \in$  each and a maximum of five 4G modules of 10 MHZ (bandwidth) with a capacity of 25 Mbps and a cost of 16  $k \in$  each. The cost for adding the 4G technology on a site is 75  $k \in$ . As for subsidies hypothesis, we take 10 discrete values for the discretization of the amount

of subsidy offered,  $\sigma \in [0, 100, 150, 200, 250, 300, 350, 400, 450, 500] \in$ . We also use 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 Table 3, which constitutes our reference upgrade function. Two other upgrades functions are also

Coverage level \Subsidies (in €)	0	100	150	200	250	300	350	400	450	500
low		0	0	0	0	0	0	0	0	0
medium low		5	12	21	30	40	42	43	44	45
medium high		10	20	30	40	50	60	62	64	65
high	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.

considered 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  $k \in \mathbb{R}$ .

The computations have been made on a server of 16 processors Intel Xeon of CPU 5110 and clocked at 1.6 GHz each. The code has been written in Julia 0.5.0, and the solver used is CPLEX 12.6 (default branch-and-bound algorithm). The time limit for MILP solving is set to 1800 seconds for sections 4.2 and 4.3.

## 4.2. Results for algorithmic tests

Our objectives in this section are to assess the scalability of our formulation (and the impact of proposed valid inequalities) and to test its sensitivity to the type of areas. We will hence 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 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, surburban S or urban U). The solutions of the linear relaxation and of the MILP are computed with and without the valid inequalities from section 3.4. We test formulations (MMMP), (MMMP + each valid inequality) and (MMMP + all 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 a '\*' if the branch-and-bound procedure has 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 density) tested with each family of valid inequalities

	stance	gap varues for		,	- ,	ot gap			•
$N_S$	Density	(MMMP)	+ (41)	+ (42)	+ (43)	+ (44)	+ (45)	+ (46)	+ (41)-(46)
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 (43). Improvements on the root gaps can also be seen in urban instances by adding the sets (41) or (46). Indeed, when the new technology is already deployed on most sites, the non-decreasing of z reduces the search space. For this reason, these instances are the ones for which adding all inequalities rather than only inequalities (43) significantly improves the root gap. Moreover, we observe that inequalities (43) also have the best impact among the different valid inequalities for reducing the final gap (see Table 6). The solution found with inequalities (43) 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,

Table 5: Best solution for 12 instances (4 sizes, 3 density) tested with each family of valid inequalities

In	stance		Best solution found by CPLEX								
111				DC30 k		<u> </u>			I		
$N_S$	Density	(MMMP)	+ (41)	+ (42)	+ (43)	+ (44)	+ (45)	+ (46)	+ (41)-(46)		
50	R	4173	4103*	4103*	4103*	4103*	4103*	4103*	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*	3861*	3861*	3861*	3861*	3861*	3861*		
150	R	12783	12308	12247	12242	12344	12247	12263	12242		
	S	11179	10049	10027	9990	10049	10146	10030	9990		
	U	5692	5522*	5522*	5522*	5522*	5522*	5522*	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		

with (for most instances) a slightly lower final gap.

For the following tests and observations, we will hence focus on the formulation with all the valid inequalities added. 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 is easier 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 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

Table 6: Final gaps for 12 instances (4 sizes, 3 density) tested with each family of valid inequalities

In	stance	Final gap							
$N_S$	Density	(MMMP)	+ (41)	+ (42)	+ (43)	+ (44)	+ (45)	+ (46)	+ (41)-(46)
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

optimal solution for the urban instance.

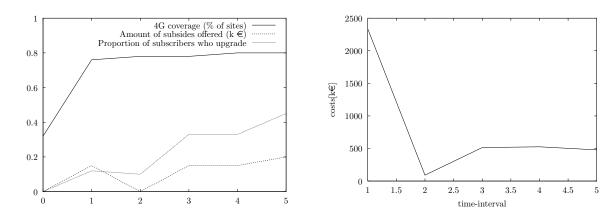


Figure 3: Evolution of the coverage, subsidies decisions Figure 4: Evolution of the costs over the time horizon and reactions over the time horizon (optimal solution of (optimal solution of 50 sites suburban instance).

50 sites suburban instance).

Moreover, we aim to analyze the characteristics of an optimal solution over the time horizon. We thus draw the features of the solution for the 50 sites suburban instance 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. This observation leads us to consider the case where the costs are smoothed over the time horizon (see Section 4.3).

#### 4.3. 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. In order to smooth the costs, we introduce a cost equilibrium set of constraints which bounds the budget fluctuations by a percentage p. Let the budget spent in year  $t \in \mathcal{T}$  be denoted by

$$B_t = \sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} \sum_{s \in \mathcal{S}} \sigma f_{\sigma,c} \pi_{\sigma,c,s,CG}^t + \sum_{s \in \mathcal{S}} \sum_{g \in \mathcal{G}} CM_g(m_{s,g}^t - m_{s,g}^{t-1}) + \sum_{s \in \mathcal{S}} CA_{N_G}(z_{s,NG}^t - z_{s,NG}^{t-1}) \forall t \in \mathcal{T}. \tag{47}$$

The objective function becomes equal to  $\sum_{t \in \mathcal{T}} B_t$ . The cost equilibrium set of constraints can be written as follows:

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

$$(48)$$

The 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. Note that the valid inequalities still hold with this constraint, except for inequalities (46), which is hence removed. Results for both formulations in Table 7 are displayed respectively below "(MMMP)" and "(MMMP + cost equilibrium)". The indicators provided are the same as in Section 4.2. 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.

From a financial point of view, we consider the 50 sites instances in order to discuss optimal solutions. For the 50 sites suburban instance with the upgrade for the reference markets, we plot in Figure 5 the counterpart of Figure 3 when requiring the costs 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  $250 \in$  per user instead of 200 are proposed in the last year). Besides, in the second year, a subsidy of  $150 \in$  is proposed 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

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

			final gap and root gap for co			1	Ctions				
	Instance			(MMMP)			(MMMP) + cost equilibrium				
$N_S$	density	upgrade	sol	f-gap	r-gap	sol	f-gap	r-gap			
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		

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.

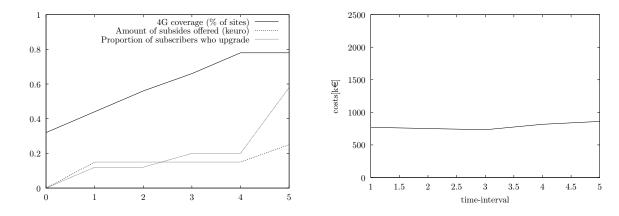


Figure 5: Evolution of the coverage, subsidies decisions Figure 6: Evolution of the costs over the time horizon and reactions over the time horizon (solution of 50 sites (solution of 50 sites suburban instance) when cost equilibrium constraints are librium constraints are imposed imposed

#### 4.4. Computational tests on large instances

In this section, we want to assess our model on 10 instances corresponding to different French territorial divisions (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)). As the planification is made for 5 years, the computational time is not

what matters the most for the operator, so we want to test longer 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 " $\alpha^0$ ". 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.

Table 8: Solution and final gap for large instances

Instanc		MILP (l	nalf an hour)	MILP (t	two hours)	MILP (five hours)		
Ter. Div.	$N_S$	$\alpha^0$	sol	f-gap	sol	f-gap	sol	f-gap
Finistère	210	36	13675	5.44	13405	3.41	13405	3.38
Côtes d'Armor	149	29	10444	2.85	10420	1.99	10420	1.00
Morbihan	168	38	11178	2.27	11178	2.22	11178	2.12
Ile et Vilaine	214	43	12400	3.49	12115	1.94	12115	1.94
Mayenne	73	31	4879	0.81	4879	0.00	-	-
Sarthe	116	33	7752	1.87	7728	0.78	7728	0.00
Maine et Loire	145	28	9876	1.78	9876	1.78	9876	1.64
Bretagne	741	37	47433	4.22	47406	3.94	47105	3.32
Pays de la Loire	334	30	22464	4.64	22464	4.00	22464	4.00
Full instance	1075	35	169968	93.5	169968	92.8	69497	3.52

We observe that, with a five-hour time limit, the final gap obtained is below 5% for all large real-life instances. The convergence of the branch-and-bound procedure is obtained for the two smallest instances. Increasing the time limit from two hours to five hours improves the solution found only for the two largest instances. Additional longer tests (12 hours) performed on these two instances show no further improvement of the solution value nor the gap.

### 5. 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. An extension for tackling a more-than-two generations framework has been provided. 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

ARCEP, 2018. Arcep rankings. Accessed 10/30/2018.

URL https://www.arcep.fr/actualites/les-communiques-de-presse/detail/n/qualite-des-services-mobiles-1.html

Bass, F. M., 1995. Empirical generalizations and marketing science: A personal view. Marketing Science 14 (3\_supplement), G6–G19.

URL https://doi.org/10.1287/mksc.14.3.G6

Bass, F. M., 2004. Comments on a new product growth for model consumer durables the bass model. Management Science 50 (12\_supplement), 1833–1840.

URL https://doi.org/10.1287/mnsc.1040.0300

Chardy, M., Yahia, M. B., Bao, Y., Sept 2016. 3g/4g load-balancing optimization for mobile network planning. In: 2016 17th International Telecommunications Network Strategy and Planning Symposium (Networks). pp. 7–12.

- CISCO, 2017. Cisco visual networking index: Global mobile data traffic forecast update, 20162021 white paper. Accessed 10/30/2018.
  - URL https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html
- DigitalInformationWorld, 2018. Accessed 10/30/2018.
  - URL https://www.digitalinformationworld.com/2018/05/infographic-internet-minute-2018.html
- Fortet, 1960. Boole algebra and its application to operation research. Trabajos de Estadistica, 111–118.
- Garcia, B.-L., Mahey, P., LeBlanc, L. J., 1998. Iterative improvement methods for a multiperiod network design problem. European Journal of Operational Research 110 (1), 150–165.
- Gendreau, M., Potvin, J.-Y., Smires, A., Soriano, P., 2006. Multi-period capacity expansion for a local access telecommunications network. European Journal of Operational Research 172 (3), 1051–1066.
- Gollowitzer, S., Gouveia, L., Ljubic, I., 2013. Enhanced formulations and branch-and-cut for the two level network design problem with transition facilities. European Journal of Operational Research 225 (2), 211–222.
- GSMA, 2018. The mobile economy 2018 (gsm association annual report).
  - URL https://www.gsma.com/mobileeconomy/wp-content/uploads/2018/05/The-Mobile-Economy-2018.pdf
- Knippel, A., Lardeux, B., 2007. The multi-layered network design problem. European Journal of Operational Research 183 (1), 87–99.
- Kubat, P., Smith, J. M., 2001. A multi-period network design problem for cellular telecommunication systems. European Journal of Operational Research 134 (2), 439–456.
- Lim, D.-E., Kim, T., 2017. An application of a multi-generation diffusion model to forecast 5g mobile telecommunication service subscribers in south korea. International Journal of Pure and Applied Mathematics 116 (23), 809–817.

Michalakelis, C., Varoutas, D., Sphicopoulos, T., 2008. Diffusion models of mobile telephony in greece. Telecommunications Policy 32 (3), 234 – 245.

URL http://www.sciencedirect.com/science/article/pii/S0308596108000116

Norton, J. A., Bass, F. M., 1987. A diffusion theory model of adoption and substitution for successive generations of high-technology products. Management Science 33 (9), 1069–1086.

URL https://doi.org/10.1287/mnsc.33.9.1069

Orange, 2018. Mobile network worlwide activity (orange group webpage). Accessed 10/30/2018. URL https://www.orange.com/en/Group/Activities/Networks/Folder/Mobile-network

Rogers, E., 2003. Diffusion of Innovations, 5th Edition. Simon and Schuster.

Sharma, P., 2013. Evolution of mobile wireless communication networks-1g to 5g as well as future prospective of next generation communication network. International Journal of Computer Science and Mobile Computing 2 (8), 47–53.

Sherali, H., Adams, W., 1998. A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems. Nonconvex Optimization and Its Applications. Springer US. URL https://books.google.fr/books?id=sPzzL4VvWqsC

#### 6. Appendix

Considering a general set of generations  $\mathcal{G} = \{cg_1, \ldots, cg_N, NG\}$  of N current generations ordered by efficiency and the newest one NG, the notation proposed in Section 3.1 still hold. The (MMMP) formulation proposed in Section 3.2 can be extended as follows while considering assumption given in section 3.5.

$$\min \sum_{t \in \mathcal{T}} \sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} \sum_{s \in S} \sum_{o \in \mathcal{G} \setminus \{NG\}} \sigma f_{\sigma,c} \pi_{\sigma,c,s,o}^t + \sum_{s \in S} \sum_{g \in \mathcal{G}} CM_g(m_{s,g}^{\bar{t}} - M_{s,g}^0)$$

$$+ \sum_{s \in S} CA_{NG}(z_{s,NG}^{\bar{t}} - Z_{s,NG}^0)$$

$$(49)$$

s.t. 
$$m_{s,q}^t \le Z_{s,q}^0 \overline{M}_q$$
  $\forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T}, \ \forall \ g \in \mathcal{G} \setminus \{NG\},$  (50)

$$m_{s,NG}^t \le \overline{M}_{NG} z_{s,NG}^t \qquad \forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T},$$
 (51)

$$z_{s,NG}^{t} \le m_{s,NG}^{t} \qquad \forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T}, \tag{52}$$

$$m_{s,g}^{t-1} \le m_{s,g}^t$$
  $\forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T}, \ \forall \ g \in \mathcal{G},$  (53)

$$u_{s,o}^{t} = \sum_{q \in \mathcal{G}: q < o} u_{s,o,g}^{t} \qquad \forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T}, \ \forall \ o \in \mathcal{G},$$
 (54)

$$Z_{s,k}^{0} \sum_{o \in \mathcal{G}: o > k} \sum_{g \in \mathcal{G}: o < k} u_{s,o,g}^{t} = 0 \qquad \forall s \in \mathcal{S}, \ \forall \ t \in \mathcal{T}, \ \forall \ k \in \mathcal{G} \setminus \{NG\}$$
 (55)

$$\sum_{g \in \mathcal{G} \setminus \{NG\}} u_{s,NG,g}^t \le \overline{U}_{s,NG}^t (1 - z_{s,NG}^t) \qquad \forall \ s \in \mathcal{S}, \ \forall \ t \in \mathcal{T},$$
(56)

$$D_g^t \sum_{o \in \mathcal{G}} u_{s,o,g}^t \le CAP_g m_{s,g}^t \qquad \forall \ s \in \mathcal{S}, \ \forall \ t \in \mathcal{T}, \ \forall \ g \in \mathcal{G},$$
 (57)

$$u_{s,o}^{t} = u_{s,o}^{t-1} - \sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} f_{\sigma,c} \ \pi_{\sigma,c,s,o}^{t} \qquad \forall \ s \in \mathcal{S}, \ \forall \ t \in \mathcal{T}, \ \forall \ o \in \mathcal{G} \setminus \{NG\},$$
 (58)

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

$$(59)$$

$$\sum_{s \in \mathcal{S}} u_{s,NG,NG}^{\bar{t}} \ge \underline{QoE} (\sum_{s \in \mathcal{S}} \sum_{o \in \mathcal{G}} U_{s,o}^{0})$$

$$\tag{60}$$

$$\alpha^{\bar{t}} \ge \underline{\alpha} \tag{61}$$

$$\sum_{\sigma \in \mathcal{K}} \sum_{c \in \mathcal{C}} \delta_{\sigma,c}^t = 1 \qquad \forall t \in \mathcal{T}, \tag{62}$$

$$\sum_{\sigma \in \mathcal{K}} \delta_{\sigma,c}^t \le 1 + U_c - \alpha^{t-1} \qquad \forall \ t \in \mathcal{T}, \ \forall \ c \in \mathcal{C},$$
 (63)

$$\begin{split} \sum_{\sigma \in \mathcal{K}} \delta^t_{\sigma,c} & \leq 1 + \alpha^{t-1} - L_c \\ \pi^t_{\sigma,c,s,o} & \leq \delta^t_{\sigma,c} \overline{U}^{t-1}_{s,o} \\ \pi^t_{\sigma,c,s,o} & \leq \delta^t_{\sigma,c} \overline{U}^{t-1}_{s,o} \\ \pi^t_{\sigma,c,s,o} & \leq u^{t-1}_{s,o} \\ \pi^t_{\sigma,c,s,o} & \geq u^{t-1}_{s,o} \\ \pi^t_{s,o} & \in \mathbb{Z} \\ \pi^t_{s,o} & \in \mathbb{Z} \\ \pi^t_{s,o} & \in \mathbb{Z} \\ \pi^t_{s,o} & \leq 0 \\ \pi^t_{s,o} & \geq 0 \\ \pi^t_{s,o,o} & \geq 0 \\ \pi^t_{s,o,o$$

We denote this formulation by (MMMP-multigen). The objective function (49) minimizes both subscribers migration costs and network investments. The first term stands for the subsidies of-fered (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 (50)–(53) are the network dynamic constraints. Constraints (50)–(51) 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 (52) impose that, when the newest technology is deployed on a site, at least one module of this technology is added. Constraints (53) prevent from decommissioning by imposing the number of modules of each technology to be non-decreasing during the time horizon.

Constraints (54)–(57) are the network dimensioning constraints, in charge of making the link between the network and the subscriber dynamic. Constraints (54) ensure the technical incompatibility stating that subscribers cannot be served by a more recent technology than the one they have subscribed to. Constraints (55) and (56) ensure the load-balancing rule. Constraints (57) 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 (58)–(59) 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 (60)–(61) stand for the model strategic guidelines and refer to the end of time horizon. Constraint (60) 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 (61) imposes that the threshold on the number of sites on which NG is deployed is met.

Constraints (62) 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 (63) and (64) ensure that, for each time period, variables  $\delta_{\sigma,c}^t$  are set according to the coverage at the previous time period. Constraints (63) (respectively (64)) set all  $\delta$  related to a range at 0 if the coverage is greater (resp. smaller) than the upper (resp. lower) bound of the range. Constraints (65)–(67) are the typical linearizations of the products of a binary variable with a continuous one. Constraints (68)–(73) define the domain of all variables in the formulation. Constraints (74)–(76) refer to the initial conditions. Finally, constraints (77)–(80) define the domain of the variables.

## **Cover Letter**

Dear Editor

Please find attached the manuscript entitled "Optimizing the investments in mobile networks and subscriber migrations for a telecommunication operator" submitted for possible publication in Computers and OR.

With best regards

Adrien CAMBIER (Orange Labs)