



HAL
open science

Optimal design of Fiber To The Home networks

Vincent Angilella

► **To cite this version:**

Vincent Angilella. Optimal design of Fiber To The Home networks. Operations Research [math.OC]. Institut National des Télécommunications, 2018. English. NNT : 2018TELE0004 . tel-01892562

HAL Id: tel-01892562

<https://theses.hal.science/tel-01892562v1>

Submitted on 10 Oct 2018

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.



**THÈSE DE DOCTORAT DE
TÉLÉCOM SUDPARIS**

Spécialité

Recherche Opérationnelle

École doctorale Informatique, Télécommunications et Électronique (Paris)

Présentée par

Vincent ANGILELLA

Pour obtenir le grade de

DOCTEUR de TÉLÉCOM SUDPARIS

Sujet de la thèse :

Design optimal des réseaux Fiber To The Home

soutenue le 16 juin 2018

devant le jury composé de :

| | | |
|----------------------------|--------------------|-----------------------------------|
| M. Walid BENAMEUR | Directeur de thèse | Telecom SudParis |
| M. Matthieu CHARDY | Encadrant | Orange Labs |
| M. Michał PIÓRO | Rapporteur | Université de Varsovie |
| M. Dritan NACE | Rapporteur | Université Technique de Compiègne |
| Mme. Ivana LJUBIĆ | Examineur | ESSEC |
| M. Dimitri PAPANIMITRIOU | Examineur | Nokia Bell Labs |
| Mme. Marie-Christine COSTA | Examineur | ENSTA |
| M. Viet NGUYEN | Examineur | Université Paris 6 |

N° NNT : 2018TELE0004

Remerciements

J'aimerais avant tout remercier mes encadrants, Matthieu Chardy et Walid Benameur, grâce à qui tout ce travail a été possible. Ils se sont montrés très pédagogiques, patients, compréhensifs, et je pense que c'est en grande partie grâce à eux que cette thèse est un travail de qualité. Je tiens à les remercier aussi d'avoir été très disponibles malgré leurs responsabilités, ainsi que d'avoir fait preuve de très bonnes qualités humaines.

Par ailleurs, je voudrais remercier toute l'équipe MSA qui m'a accueillie en son sein. Je pense particulièrement aux chefs d'équipe que j'ai eues, Marie Françoise Colinas et Olfa Benhaddada. Mais aussi, pour leur compagnie et pour les discussions intéressantes que l'on a eues, à Matthieu, Stéphane, Rachel, Taoufik, Aurélien, Catherine, Elie, Mamdouh, Bruno, Stanislas, Yu, Manon, Alain, Amal, Eric, Adam, Nancy, ainsi que d'autres qui n'ont été que de passage.

Je voudrais aussi remercier mes collègues thésards pour leur solidarité, notamment Mariem, Mohamed, Yassine, Guanglei, Antoine, Wilfried, Iaad, Georges, Ayman, Quentin et Adrien.

Bien évidemment, ce travail n'aurait pas été possible sans ma famille, qui m'a supporté durant trois années.

Contents

| | | |
|----------|---|------------|
| 1 | Résumé en Français - French Abstract | iii |
| 1.1 | Présentation | iii |
| 1.2 | Problèmes abordés | x |
| 1.3 | Résultats | xi |
| 2 | Introduction | 1 |
| 2.1 | Problem Presentation | 2 |
| 2.1.1 | Context | 2 |
| 2.1.2 | General elements | 2 |
| 2.2 | Literature Review | 10 |
| 2.2.1 | FTTx network design | 10 |
| 2.2.2 | Related problems | 12 |
| 2.3 | Problems tackled | 14 |
| 3 | Fiber Cables Network Design Problem | 16 |
| 3.1 | Integer Programming | 17 |
| 3.1.1 | Notation | 17 |
| 3.1.2 | Formulation | 18 |
| 3.1.3 | Preprocessing | 23 |
| 3.1.4 | Valid inequalities | 23 |
| 3.2 | Complexity | 27 |
| 3.3 | Asymptotic behavior | 30 |
| 3.4 | Numerical results | 32 |
| 3.5 | Conclusion | 37 |
| 4 | Fiber Cables Network Design in an Arborescence Problem | 39 |
| 4.1 | Complexity | 40 |
| 4.1.1 | Proof of NP-completeness | 40 |
| 4.1.2 | Inapproximability | 42 |
| 4.1.3 | On the possible sets of decision variables | 42 |
| 4.2 | Path-based integer programming formulation | 44 |
| 4.2.1 | Notation | 44 |
| 4.2.2 | Formulation | 45 |
| 4.2.3 | Preprocessing | 48 |

| | | |
|----------|---|------------|
| 4.2.4 | Valid inequalities | 51 |
| 4.3 | Arc based integer programming formulation | 58 |
| 4.3.1 | Notation | 58 |
| 4.3.2 | Formulation | 59 |
| 4.3.3 | Preprocessing | 63 |
| 4.3.4 | Valid inequalities | 64 |
| 4.4 | Computational results | 67 |
| 4.4.1 | Test description and setting | 67 |
| 4.4.2 | Model comparison | 68 |
| 4.4.3 | Optimal design characteristics and operational recommen- dations | 76 |
| 4.4.4 | Restricted problem analysis | 78 |
| 4.5 | Conclusion | 80 |
| 5 | Backfeed in Fiber Cables Network Design | 81 |
| 5.1 | Integer Programming Formulation | 82 |
| 5.1.1 | Formulation | 82 |
| 5.1.2 | Valid Inequalities | 84 |
| 5.2 | Complexity | 85 |
| 5.3 | Numerical Results | 87 |
| 5.4 | Conclusion | 89 |
| 6 | Easy-Maintenance Fiber Cables Network Design | 90 |
| 6.1 | Integer Programming | 91 |
| 6.1.1 | SFCND | 91 |
| 6.1.2 | ESFCND | 94 |
| 6.2 | Dynamic Programming for ESFCND | 98 |
| 6.2.1 | Exact Algorithm | 98 |
| 6.2.2 | Approximation algorithm | 100 |
| 6.3 | Complexity | 102 |
| 6.3.1 | SFCND | 103 |
| 6.3.2 | ESFCND | 104 |
| 6.3.3 | Synthesis | 105 |
| 6.4 | Results | 106 |
| 6.4.1 | Models comparison | 106 |
| 6.4.2 | Sensitivity analysis | 107 |
| 6.4.3 | Operational considerations | 109 |
| 6.5 | Conclusion | 110 |
| 7 | Synthesis | 111 |

Chapter 1

Résumé en Français - French Abstract

1.1 Présentation

1.1.1 Contexte

Le déploiement des réseaux FTTH est actuellement l'un des défis les plus importants des opérateurs de télécommunications, pour des raisons financières (la couverture de la France est estimée à plusieurs dizaines de milliards d'euros) ainsi que pour l'image de la marque. La solution majoritairement choisie par les opérateurs est de déployer des réseaux optiques passifs (PON) (voir [45], [55]). Le répartiteur (splitter) optique passif est l'équipement clef d'un réseau PON. Cet appareil est connecté à une fibre optique d'un côté (fibre amont), et à plusieurs fibres de l'autre côté (fibres aval, voir Fig. 1.1). Il divise le signal venant de la fibre en amont sur les fibres aval, et regroupe le signal venant des fibres en aval sur la fibre en amont. Les architectures PON sont composées de N niveaux de répartition (voir [35], N étant généralement compris entre 1 et 3). On travaille ici avec des réseaux à un niveau de répartiteurs, mais le travail effectué s'applique à n'importe quelle architecture.

Cette thèse aborde un sous-problème du design des réseaux FTTH, qui consiste à relier un point où sont situés les équipements (centre de répartiteurs ou NRO) à plusieurs autres (clients ou répartiteurs), avec des câbles de fibres, à coût minimal. Dans un PON à plusieurs niveaux, le problème apparaît entre chaque niveau (voir Fig. 1.1). Par exemple, les câbles peuvent avoir à relier un centre de répartiteurs de haut niveau à plusieurs centres de plus bas niveau (voir Fig. 1.1b). L'architecture PON ainsi que la position des répartiteurs est supposée connue et fixée par des décisions faites auparavant (voir Section 2.2.1).

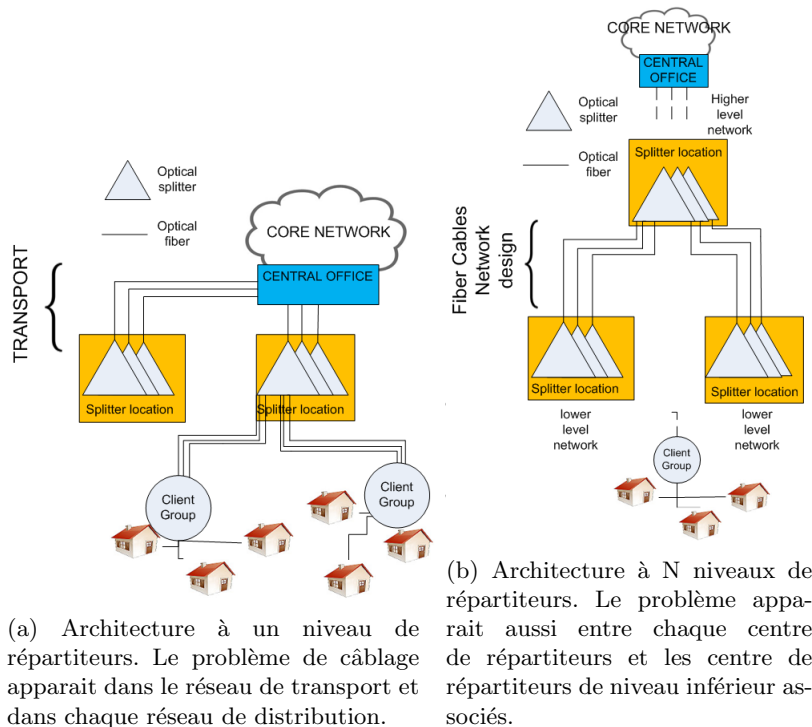


FIGURE 1.1: Architectures optiques

1.1.2 Eléments

Câbles et Génie Civil

Les câbles sont composés de trois niveaux : un câble contient plusieurs modules, et un module contient plusieurs fibres. Tous les modules d'un même câble sont identiques. Dans ce travail, on impose à tous les modules d'un réseau FTTH d'être identiques. Les demandes sont exprimées en nombre de modules, dont le but est de servir des groupes de clients. Cela nous permet de considérer uniquement deux niveaux, les câbles et les modules (voir Fig. 1.2a). Un module actif est un module connecté aux équipements à ses deux extrémités. Les autres sont inutilisés.

Les câbles sont déployés le long du génie civil (voir Fig. 1.2b), composé de conduites existantes (venant du réseau de cuivre, du réseau électrique, ...). Les câbles les plus larges ont un diamètre de 2,5 centimètres (voir les catalogues de constructeurs de câbles, [23], [33]), tandis que les conduites des réseaux de télécommunication sont généralement d'au moins 30 centimètres de diamètre. (voir [49]). Cela nous amène à supposer que l'on peut installer autant de câbles que l'on veut dans une conduite. Creuser de nouvelles conduites est en général exclu à cause de coûts trop importants. Le long des conduites se trouvent des

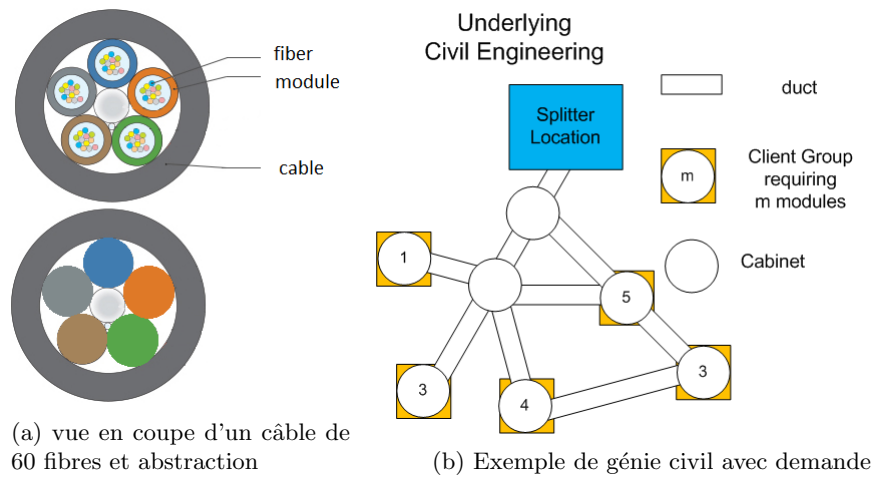


FIGURE 1.2: Structure des câbles et génie civil

cabinets (aussi appelés alvéoles ou chambres) dans lesquels les câbles peuvent subir des opérations. Certains de ces cabinets ont une demande en fibre.

Opérations

Dans les cabinets, les câbles peuvent subir deux opérations différentes, qui amènent à avoir trois configurations de base (voir Fig. 1.3) :

- Continuation : le câble traverse le cabinet sans subir aucun changement.
- Epanouissement (splicing) : cette opération consiste à couper un câble au niveau du cabinet, souder une partie de ses modules aux modules de nouveaux câbles, et cela dans une boîte de soudure. La taille de la boîte dépend de la taille du câble épanoui.
- Piquage (tapping) : cette opération consiste à ouvrir la gaine du câble, souder une partie de ses modules à ceux de nouveaux câbles, le tout dans une boîte de soudures. La taille de la boîte correspond à la taille du câble piqué. Celui-ci continue hors du cabinet avec plus de modules inactifs qu'il n'en avait. Comparé à un épanouissement, cette opération requiert moins de soudures, mais un coût de déploiement de câbles plus important.

En utilisant ces techniques, on peut construire un réseau de câbles satisfaisant la demande. Cependant, des règles d'ingénierie dépendant du contexte doivent être prises en compte. Voici celles suivies par Orange :

- Pour faciliter la maintenance, un câble créé ne peut jamais contenir des modules venant de deux câbles différents (voir Fig. 1.4).

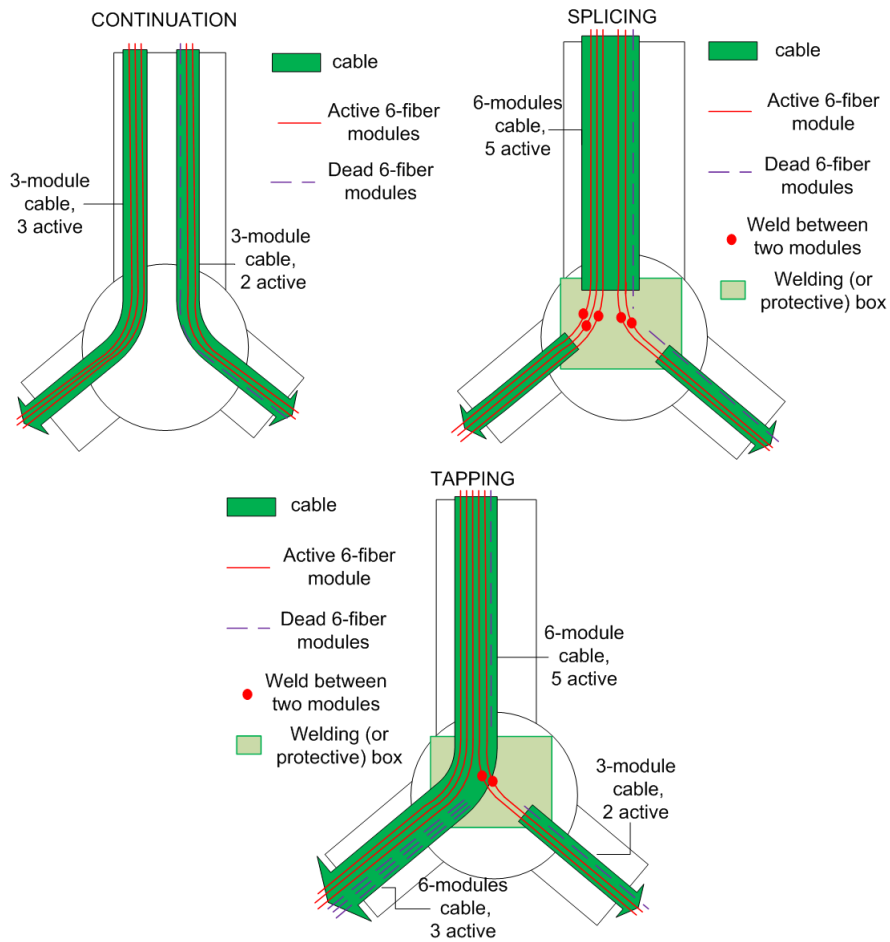


FIGURE 1.3: dans un cabinet, câbles continués, épanouis et piqués

- Au plus un câble peut être soit épanoui soit piqué dans une même cabinet. Cela est imposé par les autorités de régulation afin de laisser de l'espace physique à la concurrence (voir [5]).

Demande

La demande doit toujours être satisfaite avec des modules venant d'un seul et même câble. Cela peut être fait de deux façons différentes (voir Fig. 1.5)

Elle est soit servie par câble, lorsqu'un câble entier est utilisé, soit servie par modules, lorsque des modules venant d'un épanouissement d'un ou piquage sont utilisés. Afin d'éviter les différences de qualité de service perçues par des clients voisins, ces deux techniques ne peuvent pas être utilisées ensemble.

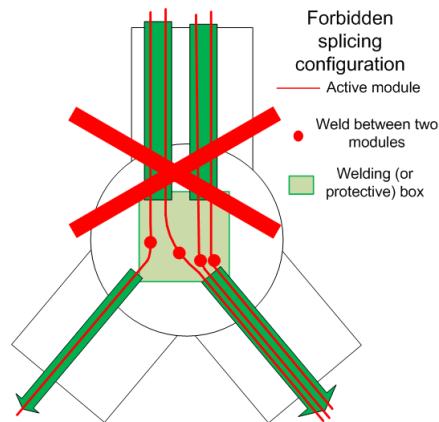


FIGURE 1.4: Un câble ne peut pas contenir des modules de deux câbles différents

Coûts

Les coûts considérés dans ce travail sont ceux payés par l'opérateur Orange (voir Fig. 1.6). Ils additionnent le prix du matériel et le prix de la main d'oeuvre :

- Le coût d'une boîte de soudures est une fonction en escalier croissante en fonction de la taille du boîtier. Cela vient du fait que l'on a un nombre limité de boîtiers différents proposés par les constructeurs de câbles.
- Le coût des soudures est une fonction croissante linéaire par morceaux en fonction du nombre de soudures à effectuer dans un cabinet.
- Le coût d'un câble est croissant et concave en fonction du type de câble (nombre de modules), et linéaire en fonction de sa longueur.

Le coût total de déploiement d'un réseau que l'on considère est la somme du prix des câbles, des soudures, et des boîtes. Un compromis apparaît entre le prix des câbles, qui dépend de leur longueur, et les prix des boîtiers et des soudures, qui n'en dépendent pas.

Le design d'un plan de déploiement de câbles (solution du problème) est complètement décrit par les éléments suivants :

- Une liste de conduites à utiliser
- Une liste de câbles, et pour chacun sa taille, son nombre de modules actifs, et le chemin sur lequel il doit être déployé.
- Une liste de points de séparation (épanouissement ou piquage) liant les câbles entre eux
- Pour chaque point de demande, le câble servant cette demande, et la façon dont elle est servie (par modules ou par câbles).

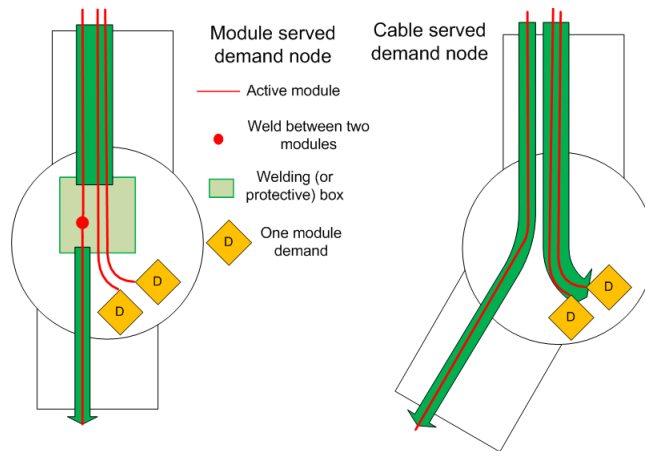


FIGURE 1.5: La demande peut être servie de deux façons différentes, qui ne peuvent pas être combinées

De plus, les règles d'ingénierie spécifiques au problème considéré doivent être respectées afin d'avoir une solution faisable.

Backfeed

Le long des chemins sur lesquels on peut déployer les câbles, certains peuvent aller vers l'amont du génie civil, c'est à dire depuis la demande de fibres vers la source (voir Fig. 1.7). On appelle cela « backfeed », et il se peut que cela fasse partie d'un déploiement de coût minimal.

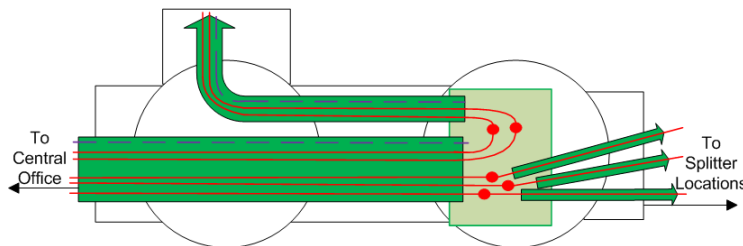


FIGURE 1.7: Illustration du backfeed

Différentes politiques de backfeed sont explorées au chapitre 5.

Maintenance

On introduit ici une restriction possible du problème qui peut être imposée à cause de considérations de maintenance. On oblige tous les câbles traversant

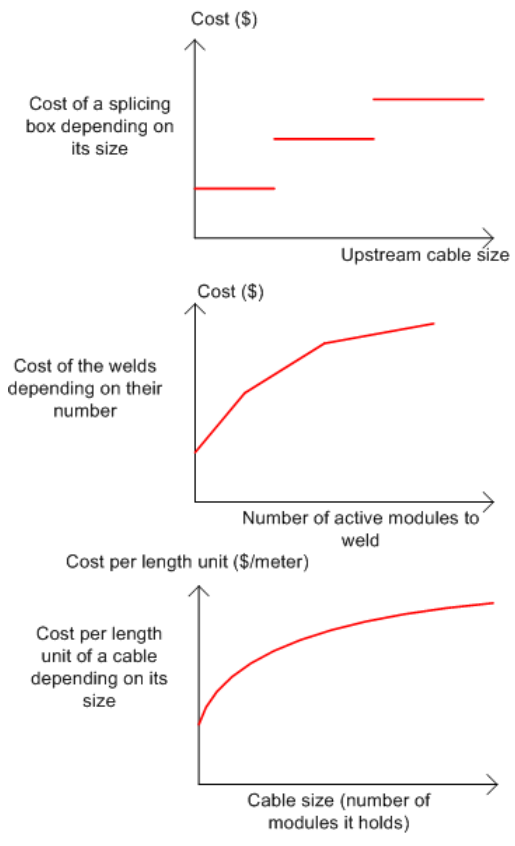


FIGURE 1.6: Coûts des boîtes, soudures et câbles

une même conduite à venir d'un même cabinet (éventuellement la source des fibres). Cette restriction est illustrée Fig. 1.8. En effet, en supposant que tous les câbles traversant une conduite sont abimés, une opération doit être faite dans tous les cabinets d'origine de chacun de ces câbles. Si la règle est respectée, une intervention est nécessaire dans seulement un cabinet.

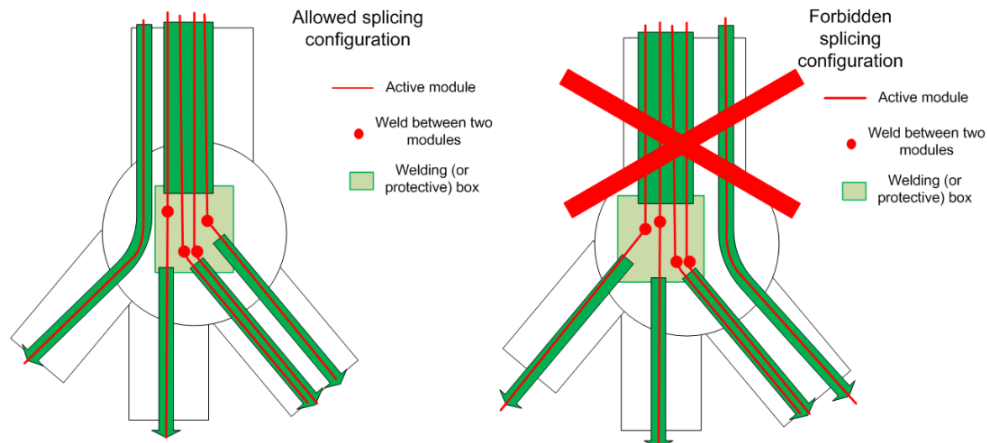


FIGURE 1.8: Gauche : épanouissement respectant la règle de maintenance. Sur chaque conduite, les câbles viennent d'un seul cabinet. Droite : épanouissement ne respectant pas la règle. Sur la conduite en bas à droite, deux câbles différents viennent de cabinets différents.

Ces considérations de maintenance sont traitées au chapitre 6.

1.2 Problèmes abordés

En introduisant ou pas des contraintes d'ingénierie, on peut obtenir une grande variété de problèmes. On résume ici les différents problèmes traités.

On définit Fiber Cables Network Design (FCND) de la façon suivante :

- Le génie civil disponible est un graphe général.
- L'épanouissement et le piquage sont tous deux autorisés.
- La demande peut être servie par câble et par modules.
- Le backfeed n'est pas autorisé.
- La règle de maintenance ne s'applique pas.

C'est le sujet du chapitre 3.

Fiber Cables Network Design in an Arborescence (FCNDA) est défini comme une restriction de FCND où le génie civil disponible a une structure d'arbre. On le traite au chapitre 4.

On définit Backfeed Fiber Cables Network Design (BFCND) comme suit :

- Le génie civil disponible a une structure d'arbre
- Seul l'épanouissement est autorisé.
- Seul le service de la demande par câble est autorisé.
- Le backfeed est permis.
- La règle de maintenance ne s'applique pas.

Ce problème est étudié au chapitre 5.

On définit SFCND (Splicing only Fiber Cables Network Design) comme une restriction de FCNDA où seul l'épanouissement est autorisé (pas de piquage). Enfin, on définit ESFCND (Easy-maintenance Splicing only Fiber Cables Network Design) comme une restriction de SFCND où la règle de maintenance décrite Section 1.1.2 s'applique. Ces deux problèmes sont étudiés au chapitre 6.

Une description succincte se trouve table 1.1.

TABLE 1.1: Propriétés des différents problèmes

| Problem | génie civil | piquage autorisé | service demande par modules autorisé | backfeed autorisé | règle de maintenance |
|---------|-------------|------------------|--------------------------------------|-------------------|----------------------|
| FCND | graphe | oui | oui | non | non |
| FCNDA | arbre | oui | oui | non | non |
| BFCND | arbre | non | non | oui | non |
| SFCND | arbre | non | oui | non | non |
| ESFCND | arbre | non | oui | non | oui |

1.3 Résultats

1.3.1 Chapitre 3

Ce chapitre se concentre sur le problème FCND. On propose de le résoudre par programmation en nombres entiers. En observant les solutions, on s'est aperçus que l'introduction des techniques de piquage et d'épanouissement dans le design de réseaux permet des économies significatives. Cela justifie le fait de se concentrer sur le design de réseaux de câbles uniquement.

Sur le plan théorique, le problème est prouvé être NP-difficile et dur à approximer. Sa relation avec deux autres problèmes classiques, l'arbre de Steiner et l'arborescence des plus courts chemins est montrée.

Sur le plan algorithmique, la solution de programmation entière proposée est capable de résoudre les plus petites instances. Les inégalités valides introduites réduisent les temps de calcul de façon significative. Cependant, les instances plus larges restent hors de portée. Cela justifie le fait d'étudier le problème indépendamment d'éléments comme le positionnement des répartiteurs, ce qui le rendrait encore plus dur à résoudre. Cela nous a aussi amenés à étudier FCNDA dans le chapitre 4.

1.3.2 Chapitre 4

Ce chapitre se concentre sur le problème Fiber Cables Network Design in an Arborescence, restriction de FCND où une arborescence du génie civil est déjà sélectionnée. Il introduit deux modèles pour la résolution, un basé sur des chemins et l'autre sur des graphes. Le problème est prouvé être NP-difficile et dur à approximer avec un ratio $\frac{3}{2}$. On montre aussi la difficulté qu'il y a à modéliser avec un certain jeu de variables, justifiant la complexité des modèles utilisés.

Les programmes en nombres entiers proposés ici sont assez performants pour résoudre des instances de taille réelle. Les deux modèles sont améliorés à l'aide d'inégalités valides qui améliorent de façon significative les performances.

Du point de vue financier, en introduisant la possibilité d'effectuer des opérations d'épanouissement et de piquage, on arrive à économiser 25 % des coûts totaux. Cela est comparable à ce que l'on avait avec FCND.

Le travail effectué dans ce chapitre sera la base pour l'implémentation d'un outil d'aide à la décision à Orange. En effet, il donne des économies significatives sans prendre trop de temps sur les instances réelles. De plus, des contraintes additionnelles peuvent être ajoutées sur les deux modèles, comme des contraintes de capacités physiques, notamment avec le modèle chemin.

Cependant, il reste des opérations effectuées par les techniciens qui ne sont pas couvertes par ce chapitre. C'est le cas du Backfeed, qui est le sujet du chapitre suivant.

1.3.3 Chapitre 5

Ce chapitre traite le problème Backfeed Fiber Cables Network Design (BFCND). L'introduction du backfeed dans le processus d'optimisation permet d'obtenir des gains significatifs, avec cependant des calculs plus longs. L'effet des inégalités valides introduites pour notre formulation suggère que ces temps peuvent être encore réduits.

Sur le plan opérationnel, le backfeed rend les opérations de maintenances plus difficiles. Le chapitre suivant introduit une contrainte ayant pour but de faciliter cette maintenance.

1.3.4 Chapitre 6

Ce chapitre traite deux problèmes, l'un étant contraint par une règle de maintenance : Easy maintenance Splicing Only Fiber Cables Network design (ESFCND) ;

l'autre ne l'étant pas : Splicing Only Fiber Cables Network design (SFCND). Pour le problème non contraint, une résolution par programmation en nombres entiers est proposée. Des inégalités valides la rendent plus utilisable en pratique. Pour le problème contraint, on a introduit deux méthodes de résolution : une basée sur de la programmation mixte incluant des inégalités valides, et l'autre par programmation dynamique. Elles sont complémentaires, la programmation dynamique étant plus efficace sur des instances où les noeuds du graphe de génie civil ont un degré faible.

Sur le plan théorique, le problème non contraint est plus complexe. Il reste NP-difficile avec certains paramètres fixés, alors que le problème contraint devient solvable en temps polynomial, ou approximable, suivant les paramètres choisis. Les expérimentations numériques ont confirmé cette tendance sur des instances réelles.

Au niveau opérationnel, la règle de maintenance peut être considérée comme un compromis raisonnable entre les coûts de déploiement et les coûts de maintenance. L'appliquer augmente les coûts de déploiement de 3,7 % sur nos instances test.

1.3.5 Implémentation

Ces travaux seront utilisés pour l'implémentation d'un outil informatique pour la planification du déploiement de réseaux FTTH par Orange. Un logiciel existant sélectionne tout d'abord un arbre du génie civil. Puis, les demandes sont groupées en « Points A » (PA). Sur cette arborescence, le problème FCNDA est résolu, en utilisant le modèle par chemins du chapitre 4. Enfin, les PA sont reliés aux points de demande, en utilisant éventuellement le backfeed.

Cela représente une amélioration considérable par rapport à l'ancienne version du programme qui prenait ces décisions de façon sous-optimale en utilisant une programmation dynamique gloutonne.

Un autre avantage de cette approche est la flexibilité de la programmation en nombres entiers. Elle permet en effet d'ajouter aisément des contraintes spécifiques à une instance, ce qui n'était pas le cas de la technique utilisée auparavant.

Chapter 2

Introduction

Chapter Abstract

This chapter introduces the topic of this thesis, FTTH network design. Section 2.1 presents the issues related to the design of fiber cables networks, such as cable separation techniques and associated technical constraints. Section 2.2 is a short summary of the literature available regarding the FTTH network design. It provides some insight about the classical problems on which the FTTH network design inherits, and describes some works tackling the FTTH network design.

2.1 Problem Presentation

2.1.1 Context

Deploying a FTTH network is one of the most important challenges currently faced by telecommunications operators for economical reasons (France covering is estimated to several dozen billion euros) and brand image. The technological solution overwhelmingly chosen by operators is to deploy Passive Optical Networks (PON) (see [45], [55]).

The passive optical splitter (also called “optical splitter” or “splitter”) is the key equipment of a PON. This equipment is connected to one optical fiber on one side (upstream fiber), and to several fibers on the other side (downstream fibers, see Fig. 2.1). It splits the signal coming from the upstream fiber on downstream fibers, and groups the signals coming from the downstream fibers on the upstream fiber. PON architectures are, in their standardized definition, composed of N-splitting stages/levels (see [35], N being usually between 1 and 3 in operational contexts due to service eligibility concerns). We consider here the single-splitting stage PON (see Fig. 2.1) since it is the one imposed by the French regulator for deployment in medium population density areas in France (see [6]), and thus currently the most widely deployed.

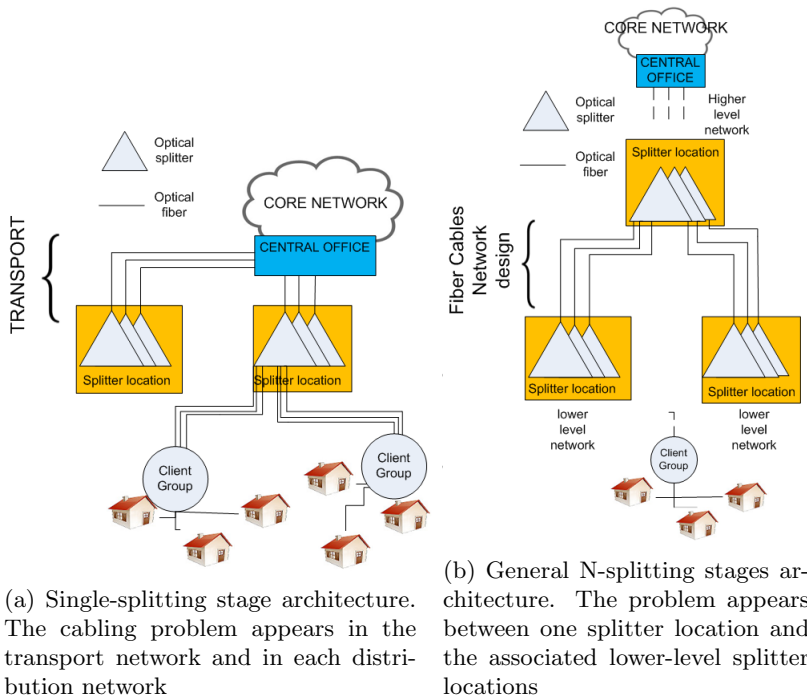
The set of cables linking a Central Office (CO) to the splitters is referred to as “transport network” (which can be referred to as “feeder network” or “feeder tree” depending on the context and optical architecture). Notice that several splitters are usually installed in the same location. The set of cables linking these co-located splitters to multiple households is referred to as “distribution network”. In the following, a fiber network will refer to a single transport or distribution network.

This thesis tackles a sub-problem of the FTTH network design problem, which consists in linking one equipment location (CO or splitters) to several others (respectively splitters or clients), using fiber cables, at minimal cost. This problem appears several times in the deployment of a single-splitting stage FTTH network (see Fig. 2.1). In a PON with several splitting stages, the same problem appears between each splitting stage. In this case, cables are to be deployed between a high-level splitter location and associated lower-level splitter locations (see Fig. 2.1b). The PON architecture and splitter locations are supposed to be fixed due to previous decision making (see Section 2.2.1).

2.1.2 General elements

Cable Structure

Fiber cables are made of three layers: one cable containing several fiber modules, and one fiber module containing several optical fibers. All fiber modules from a cable are identical (hold the same number of fibers) due to the cable manufacturer’s design. In this thesis, we assume that all modules from a fiber network are identical.



(a) Single-splitting stage architecture. The cabling problem appears in the transport network and in each distribution network

(b) General N-splitting stages architecture. The problem appears between one splitter location and the associated lower-level splitter locations

Figure 2.1: Optical architectures

Due to economies of scale and engineering rules, demands are expressed as a number of modules, meant to serve groups of clients. This allows us to consider only two layers, fiber cables and fiber modules (see Fig. 2.2a left). An “active module” is a module connected to the optical splitter at one extremity, and to the demand equipment at the other. The other ones, which remain unused, are referred to as “dead modules”.

Cables are laid across the civil engineering structure, made of existing ducts (copper line network, electric lines, ...). Largest cables used for FTTH deployment have a diameter between 0.2 and 2.5 centimeters (see cable manufacturer’s catalogs such as [23], [33]), while ducts for telecommunication networks are usually at least 30 centimeters wide (see [49]) and at least 60 centimeters deep (see [49] and local regulations on ducts). We therefore assume it is possible to install as many cables as desired in a duct. Digging new ducts is generally excluded due to its high cost. Along the ducts of this structure are located cabinets (also referred to as concrete rooms or man holes) where cables can be separated. Some of these cabinets have a demand.

Cable separation techniques

At a cabinet, the cables can endure two different separation operations leading to three basic scenarii (see Fig. 2.3):

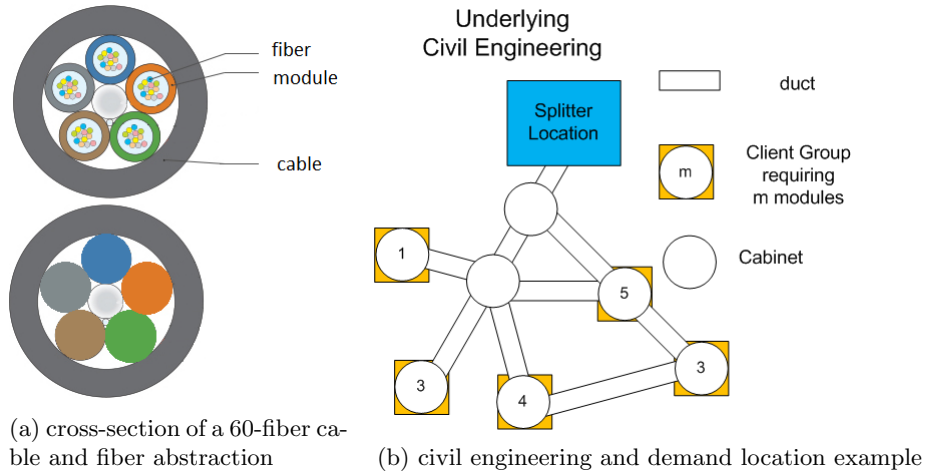


Figure 2.2: Cable structure and civil engineering

- Cable extending: a cable goes through a cabinet without undergoing any change.
- Cable splicing: this operation consists in cutting a cable at the cabinet, welding some of its modules to those of new cables, and putting a protective box in this cabinet. The size of the box matches the size of the spliced cable.
- Cable tapping: this operation consists in opening the sheath of the cable, welding some of its modules to those of new cables, and putting a protective box in this cabinet. The size of the box matches the size of the tapped cable. The tapped cable goes out of the cabinet with more dead modules than it had upstream. Note that compared to a splicing, this operation yields less welds, but may incur a higher cable cost.

In the following, spliced and tapped cables will be referred to altogether as “separated cables”, while splicing and tapping points will be grouped under the term “separation point”. After a splicing in a cabinet, the new cables starting at this cabinet are called “born cables”. As for a tapping, the continuation of the tapped cable is referred to as a “continued cable”, while the other cables starting at the same cabinet are also called born cables. Finally, cables starting at the root are referred to as born cables. Born cables and continued cables are grouped under the term “created cables”.

Using these separation techniques, one can build a fiber cable network satisfying demands. However, some context-dependent engineering rules have to be taken into account. We enumerate below the rules considered by Orange:

- In order to facilitate maintenance, a created cable can never contain active modules coming from two different cables (see Fig. 2.4 for illustration).

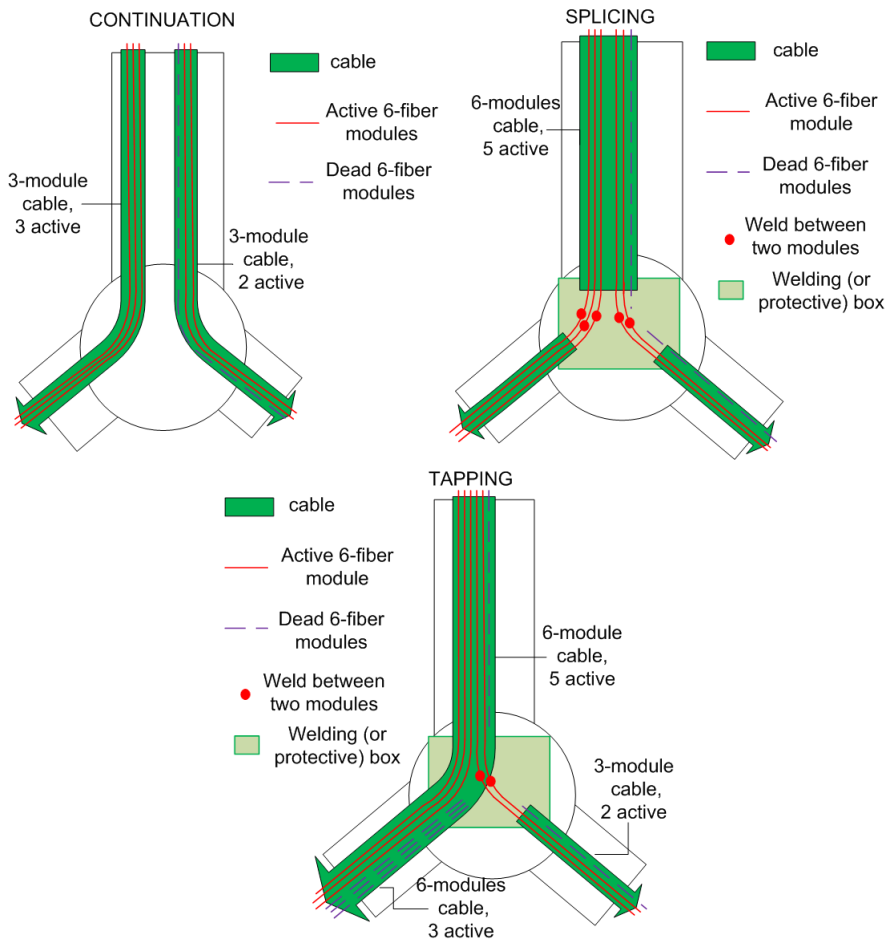


Figure 2.3: extended, spliced and tapped cables at a cabinet

- At most one cable can be either spliced or tapped at a given cabinet. This is imposed by the french regulation authority in order to leave some physical space for competitors (see [5]).

Demand supplying techniques

The demand should always be satisfied by modules from one cable, and can then be served in two different ways (see Fig. 2.5). It is either “cable-served” when only one full cable is used, or “module-served” when modules taken from a splicing or tapping point are used. In order to avoid differences in terms of quality of service perceived by neighboring customers, these two ways (cable-served and module-served) cannot be combined in a cabinet.

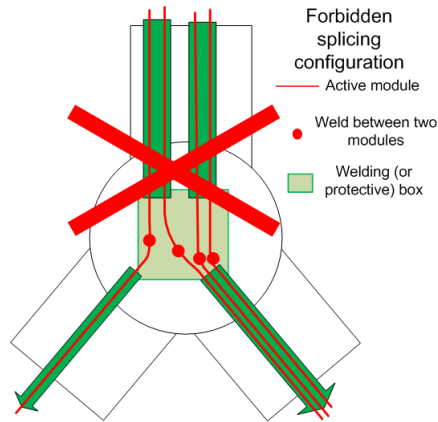


Figure 2.4: One cable cannot contain active modules coming from two different cables

Related cost items

The costs considered in this work are those currently incurred by the network operator (see Fig. 2.6). They include both material and manpower aspects:

- The cost of a welding box is an increasing scale function with respect to the size of the box. This comes from the limited number of available box sizes in manufacturer's catalogs.
- The manpower cost for welding fibers is a linearly piecewise concave increasing function with respect to the number of fibers in a given cabinet.
- The cost of a cable includes manpower and material costs. This cost is concave with respect to the cable type (number of modules), and linear with respect to its length.

The total cost is the sum of the cost of cables, the cost of boxes and the cost of welds. A trade-off appears between cable deployment costs, which depend on the cable length, and cable separation costs (boxes and welds). Intuitively, long distances will result in a lot of cable separations, since welds and boxes costs become negligible, while short distances will result in few cable separations.

The design of a cable deployment scheme (solution to the problem) is fully described by the following elements.

- A set of ducts to be used.
- A list of cables, each one including its size, its number of active modules, and the path on which it is deployed.
- A list of separation points linking the cables.

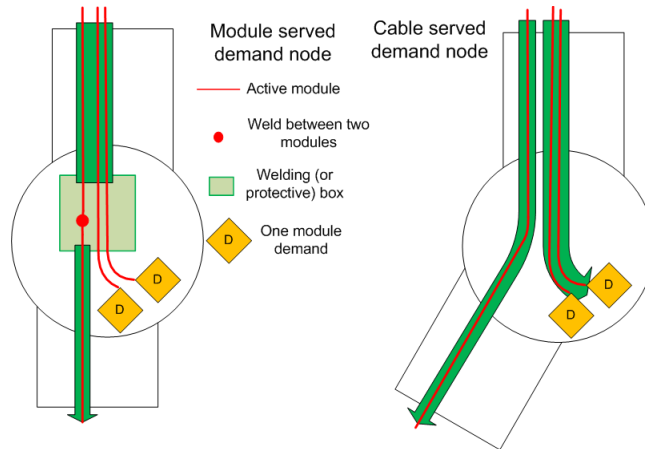


Figure 2.5: The demand can be served in two separate ways, which cannot be combined.

- For each demand point, the cable serving it, and the way the demand is served.

Additionally, the engineering rules specific to the problem tackled must be respected in order to have a feasible solution.

The informed reader can guess that the space of possible solutions will be explored by using integer variables. It is possible to use variables based on the description above (for instance, using one boolean variable per path, cable size, and number of active modules within this cable size). However, this full description is computationally quite expensive. The several sets of variables used were actually all different from a straightforward description. When needed, we will specify how to get a cable deployment scheme from a feasible set of variables.

Backfeed policies

Among the paths on which one can deploy cables, some may go "upstream" the civil engineering structure, that is on a direction from the fiber demand to the fiber source (see Fig. 2.7). This is referred to as "backfeed", and can be part of a minimum cost deployment.

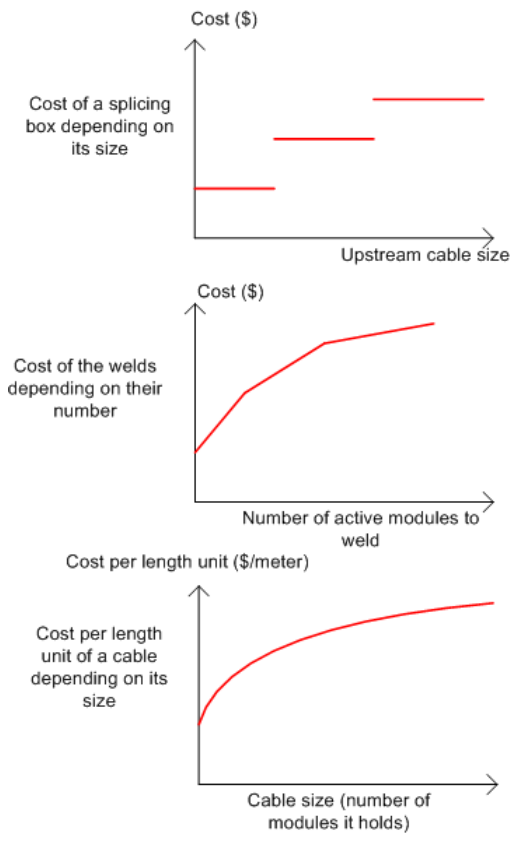


Figure 2.6: Cost of the boxes, welds and cables

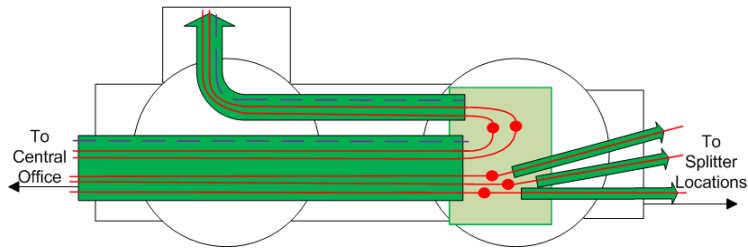


Figure 2.7: Illustration of backfeed

Backfeed policies are explored in Chapter 5

Operations Administration and Maintenance considerations

We introduce here a possible restriction of the problem that can be imposed due to maintenance considerations. We restrict the problem by imposing that all cables going through a given duct are born in the same cabinet (eventually the fiber source). This restriction is illustrated in Fig. 2.8. It is motivated by operations and maintenance considerations. Indeed, assuming all the cables of a given duct are damaged, then an intervention has to be done at the cabinets where each of these cables are born. If the rule is respected, an intervention is necessary in only one cabinet.

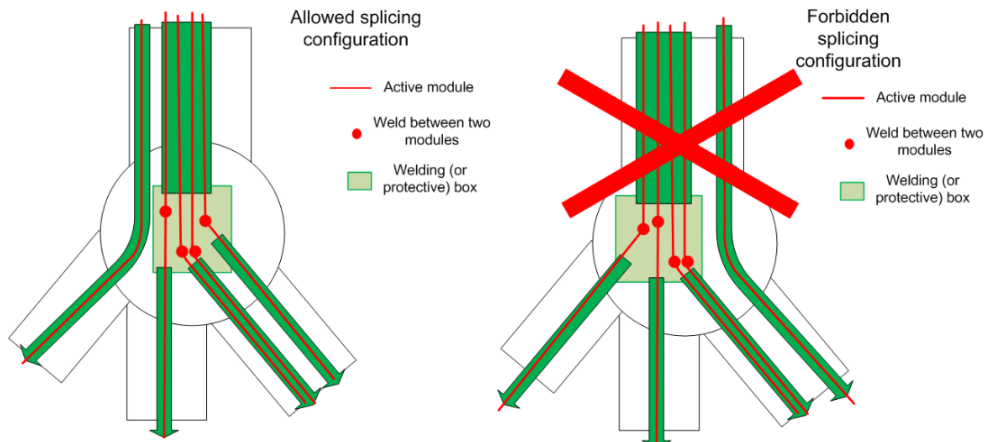


Figure 2.8: Left: Splicing configuration respecting the maintenance rule. On all edges, cables are born in the same cabinet; Right: Possible splicing configuration in conflict with the maintenance rule. On the bottom-right duct, two different cables are born in different cabinets.

Such OAM considerations are investigated in Chapter 6.

2.2 Literature Review

Some of the work presented in this thesis can also be found in [1], [2], [4] and [3].

2.2.1 FTTx network design

The fiber networks can be classified between the endpoint of the network. The option that offers the most bandwidth is the Fiber To The Home. However, one can also find Fiber To The Building (FTTB) networks, where the fiber stops at the base of a building, and is continued by legacy copper networks. Frequently encountered networks also include Fiber To The Cabinet (FTTC), where the fiber is stopped at a street cabinet, an Fiber To The District (FTTD), where the fiber is stopped in a commutation device located in the district. A slightly different case is the Fiber To The Antenna (FTTA), where the fibers connect antennas rather than households. All these networks are gathered by the generic term FTTx. The FCND problem applies to FTTH, FTTB and FTTC, since the operations after the cabinet are ignored.

In earlier works, we can find comparisons of these technologies, where the operator had to decide where to stop the fiber network, and where to use copper (see [53]). Nowadays, the lower cost of fiber optics, reduced by technological improvements, make the fiber to the home more and more widespread.

An abundant literature can be found on different FTTx optimization problems. The changes in local regulations as well as choices regarding the architecture make different problems appear. Moreover, the wide range of equipment to be installed rises an even wider range of decision problems. Indeed, the decisions of installing several different equipment are usually combined. This rises more elaborate problems from "classical" ones.

A worthy but never reached goal would be to devise a decision process able to decide all the different elements of a FTTH network at once. Some recent works, such as [44], introduce a MIP-based method for a global view of the FTTH optimization, including many levels of decision (OLT cards, optical architecture, splitter location, ducts, cables, splicing operations). The cable network design introduced (cables and splicing operations) is simpler than in FCND (no tapping considered for instance). Despite this simplification, the computations are done in a heuristic fashion. One can infer that a global approach including high level decisions (architecture, OLT cards and splitter location) as well as all the elements described in Section 2.1 would be too much to handle for the state of the art solvers and machines. It is also concluded in the survey [30].

This explains why in an operational context, decisions of different types are successively questioned and validated. Notice that the order of decisions is not fixed and depends on the operator. For instance, in the work [57], splitters and OLT cards are installed after the ducts have been selected and cables have been laid out. It is assumed that enough cables have been deployed so that the cable level of decision can be ignored. In this thesis, we will assume that the elements are selected in the following order.

- First, the optical architecture is decided.

- Then, for a given zone, the position of the Central Office and the Splitter Locations are decided.
- Finally, the civil engineering ducts are selected and the cables network design is done.

The optical architecture chosen is often one of the first elements to be decided. Note that a homogenous optical architecture is not always present, see for instance [56]. The work [45] compare the single-stage architecture (no splitting) to the single-splitting stage architecture. Its conclusions favor the single-splitting stage on the use-cases available.

Once the optical architecture is fixed, one can tackle the facility location/allocation and network design problem (see Fig. 2.1). It consists in deciding the splitter location and the routes the fiber cables will take. A significant amount of work considering this problem with different hypothesis is available in the literature. In the case of the 1-splitting stage FTTH (also called 2-level FTTH, see Fig. 2.1a), the reader can refer for instance to [29], [17], [28], [46], or [11]. The case of the 2-splitting stages is dealt with in [14], and a generalization for any number of splitting stages is proposed in [32] (the title points to a 2-splitting stage network, which is the main focus of the paper, but a generalization is proposed nonetheless). Heuristic approaches are also available for large-scale instances (see [54], [43], [39]).

Throughout this entire thesis, we will consider that the optical architecture and splitter locations have already been selected thanks to the works referred to above. In most works tackling splitter location problems, ducts are also selected. In Chapter 3, we will discard the ducts selected by this process of the decision making and include it in a new step. In Chapters 4, 5 and 6, we will consider the set of ducts in which the cables are to be deployed as already selected by a previous decision process.

From an operator’s point of view, the goal of a FTTH network is to connect and charge the clients. There is no guarantee that a client will ask for its connection, or to which operator it will ask. This appears in several papers, and eventually leads to specific marketing strategies (see [50]) or game theory modeling (see [12]). Other papers, such as [40], include a prize collecting aspect, in case the operator wants to decide to leave some areas to the competition (lost prize). The strategy chosen by Orange is to be able to connect almost every FTTH client in France (see [7]). When clients have a fiber connection, Orange is always present, eventually together with a competing operator. This decision is mostly due to brand image considerations. Under these conditions, the prize collecting aspect will not be considered (we have to be able to ”collect all prizes”).

Some works consider client uncertainty (such as [32], [31]). The client uncertainty is an interesting aspect worth considering, since it matches the client behavior. However, in our case, we will consider the uncertainty of each demand location individually which leads to straightforward simplifications. The demand to be served at each cabinet is considered fixed.

In many different networks, one has to ensure some redundancy between the nodes (equipment, facilities, ...). In case of failure, the network should be able to become quickly operational. This can be done by rerouting traffic on alternate routes, defined proactively. Resiliency is especially important for core networks, where a lot of traffic is at stake. A large literature is available regarding the design of such resilient networks (see [51], [58], [22]), including a wide range of failures (see [47]). However, in the case of FTTH networks, a redundant system is considered too expensive to be deployed. This is mostly due to the nature of optical splitters: being passive equipment, they are unable to reroute the traffic in case of a damaged link. The failures are to be repaired as fast as possible. This explains why the maintenance rule described in 2.1.2 has been introduced.

In the recent survey [30], the issue of cable separation techniques (splicing and tapping, see Section 2.1) is highlighted as an incomplete field of study, especially regarding manpower cost. We should note that some works such as [42], [44] and [37], do include the splicing technique, and cable separation costs depending only on the cable size. A heuristic algorithm presented in [19] includes cables weld cost and the splicing technique, but provides suboptimal solutions. Some papers consider the protective sheath surrounding the cables (see [36]), without exploring cable separation techniques. This thesis aims to tackle cables network design more accurately, by including for instance tapping, backfeed and module-served demand nodes (see Section 2.1). All these operations are currently practiced by the technicians in charge of laying down the network.

2.2.2 Related problems

Steiner Tree Problem

The Steiner Tree Problem (STP) is one of the earliest network design problems to be introduced. It has many variants from different application fields (electricity networks, telecommunications, ...). The STP has been extensively studied in the literature. It is one of the 21 NP-complete problems from the paper [34], some of the first ones to be introduced. As we will see in Chapter 3, a significant amount of the work that has been done for the STP can be reused for the FCND problem. This concerns especially the works done in polyhedral analysis.

The STP can be formulated as follows. Given a graph $G = (V, E)$, a set of terminal nodes $T \subseteq V$, find a connected subgraph $G' = (V', E')$ of G such that $|E'|$ is minimal and spanning the terminal nodes ($T \subseteq V'$). The minimality of E' ensures G' is a tree.

As we will see in Chapter 3, a significant amount of the work that has been done for the STP can be reused for the FCND problem. This concerns especially the works done in polyhedral analysis.

Among the papers that tackle it, we will refer to [26] for a study of the different possible MIP formulations of the STP. It compares several possible formulations and shows that they have different qualities. The formulation on which we base ourselves is the following.

Let us consider $G = (V, E)$ an oriented symmetrical graph, meaning for $(i, j) \in E$, $(j, i) \in E$. For $v \in V$, let us call $\Gamma^-(v)$ the set of incoming edges of v , and $\Gamma^+(v)$ the set of outgoing edges of v . Let us pick an artificial root node $r \in T$ from the set of terminal nodes. Let us define, for $e \in E$, x_e the boolean variable equal to 1 iff e is part of the minimal steiner tree. Let us define, for $e \in E$ and $t \in T$ a continuous flow variable $f_{e,t}$. The minimal Steiner Tree Problem can be formulated as follows.

$$\min \sum_{e \in E} x_e \quad (2.1)$$

$$\text{such that } \forall v \in V, \sum_{e \in \Gamma^-(v)} x_e \leq 1 \quad (2.2)$$

$$\forall t \in T, \sum_{e \in \Gamma^-(r)} f_{e,t} - \sum_{e \in \Gamma^+(r)} f_{e,t} = -1 \quad (2.3)$$

$$\forall t \in T, \forall t' \in T \setminus \{r\} \sum_{e \in \Gamma^-(t')} f_{e,t} - \sum_{e \in \Gamma^+(t')} f_{e,t} = 1 \quad (2.4)$$

$$\forall t \in T, \forall v \in V \setminus T \sum_{e \in \Gamma^-(v)} f_{e,t} - \sum_{e \in \Gamma^+(v)} f_{e,t} = 0 \quad (2.5)$$

$$\forall e \in E, \forall t \in T, f_{e,t} \leq x_e \quad (2.6)$$

(2.1) denotes the objective to be minimized. (2.2) ensure the solution does not contain any loop. Constraints (2.3), (2.4) and (2.5) make sure the solution is connected. A path denoted by the flow goes from the root node to each terminal. Finally, (2.6) ensures that if an edge e is on the selected path from the root node to a terminal, then e is part of the solution.

Remark 2.2.1 *In a practical application, one often sees the weighted STP, where weights w_e are affected to each edge $e \in E$, and the objective to be minimized is $\sum_{e \in E} w_e \cdot x_e$. For applications related to the FTTH, one can also consider a weight associated to the flow variables, $w_{e,t}^f$ which corresponds to the cost of fibers.*

Some advantages of this formulation are shown in [26]. It has a higher continuous relaxation than its non-oriented counterpart. Another advantage is its compacity (polynomial number of variables and constraints).

Regarding the polyhedral analysis, a significant amount of work is available in [10], [25], [15], [16] and [21]. Those papers introduce many types of facets and valid inequalities for the STP (called "odd-hole", "anti odd-hole", "wheel", partitions, ...). As we will see in chapter 3, any polyhedral study of the STP can be reused for the FCND.

Knapsack problem

The Knapsack problem was one of the earliest combinatorial decision problems to be introduced. It can be formulated as follows. Given a maximum weight

W , a set of n items, each with value v_i and weight w_i for $i \in \{1, \dots, n\}$, find a subset of items $S \subseteq \{1, \dots, n\}$ maximizing the total values $\sum_{i \in S} v_i$ such that the total weight of the items is smaller than W ($\sum_{i \in S} w_i \leq W$).

This problem is shown to be NP-complete in [34]. It can be solved optimally by dynamic programming, as shown in [52]. This dynamic programming solution can be modified into a Fully Polynomial Time Approximation Scheme (FPTAS), available in [27].

The Knapsack problem can be formulated as an integer program as follows. With the same notation, let us introduce for $i \in \{1, \dots, n\}$ the boolean variable X_i equal to 1 iff object i is selected in the Knapsack. Then, we can solve the following integer program:

$$\begin{aligned} & \max \sum_{i \in \{1, \dots, n\}} v_i \cdot X_i \\ \text{such that} \quad & \sum_{i \in \{1, \dots, n\}} w_i \cdot X_i \leq W \\ & \forall i \in \{1, \dots, n\}, X_i \in \{0, 1\} \end{aligned} \tag{2.7}$$

Constraint 2.7 expresses that the total weight cannot exceed the limit W .

The covering problem is the minimization version of this problem. It can be formulated as an integer program as follows:

$$\begin{aligned} & \min \sum_{i \in \{1, \dots, n\}} v_i \cdot X_i \\ \text{such that} \quad & \sum_{i \in \{1, \dots, n\}} w_i \cdot X_i \geq W \\ & \forall i \in \{1, \dots, n\}, X_i \in \{0, 1\} \end{aligned} \tag{2.8}$$

One can see that this formulation is quite simple, since it has only one constraint. While it is not a network design problem, its simple structure makes it relate to many other problems, including cables network design. The covering problem will be used in Chapter 6 for two purposes, on one hand to provide valid inequalities, and on the other hand to propose a FPTAS based on the FPTAS for Knapsack.

2.3 Problems tackled

The consideration (or lack of) of regulatory and engineering rules from Section 2.1 can lead to a large set of variants of fiber cables network design problems. Below is a Synthesis of the decision problems tackled in this thesis and the organization of the manuscript.

We sum up here how the different problems tackled in this thesis differ.

Fiber Cables Network Design (FCND) is defined as follows

- The civil engineering structure available is a general graph.
- Both tapping and splicing are allowed.
- Demand nodes can be both cable-served and module-served.
- No backfeed is allowed
- The maintenance rule does not apply

It will be tackled in Chapter 3.

Fiber Cables Network Design in an Arborescence (FCNDA) is defined as a restriction of FCND where the civil engineering structure available is an arborescence. It will be addressed in Chapter 4.

We define Backfeed Fiber Cables Network Design (BFCND) as follows

- The civil engineering structure available has a tree topology.
- Only splicing is allowed.
- Demand nodes can be only cable-served.
- Backfeed is allowed
- The maintenance rule does not apply

This problem is studied in Chapter 5.

We define SFCND (Splicing only Fiber Cables Network Design) as a restriction of FCNDA where only splicing operations are allowed. Finally, we define ESFCND (Easy-maintenance Splicing only Fiber Cables Network Design) as a restriction of SFCND where the maintenance rule described in Section 2.1.2 applies. Both SFCND and ESFCND will be considered in Chapter 6.

These problems are summed up in Table 2.1.

Table 2.1: Properties of the different problems

| Problem | given civil engineering | tapping allowed | module-served demand nodes allowed | backfeed allowed | maintenance rule applies |
|---------|-------------------------|-----------------|------------------------------------|------------------|--------------------------|
| FCND | graph | yes | yes | no | no |
| FCNDA | tree | yes | yes | no | no |
| BFCND | tree | no | no | yes | no |
| SFCND | tree | no | yes | no | no |
| ESFCND | tree | no | yes | no | yes |

Chapter 3

Fiber Cables Network Design Problem

Chapter Abstract

This chapter tackles the FCND problem. Recall that it can be stated as follows. Given a civil engineering infrastructure, the decision problem consists in selecting an arborescence from this infrastructure, a set of cables with enough active modules to serve the demand on each arc and associated separation points on the nodes. Separation points should conserve the number of active modules so that active modules are not interrupted from the fiber source to the demand. While minimizing the total cost (cable deployment, welding boxes and welds), each demand node should be either cable served or module served by a single cable, and at most one separation point per node is allowed.

An integer linear programming model is proposed in Section 3.1. The model is reinforced thanks to valid inequalities in Section 3.1.4, and its relation with the Steiner Tree Problem (STP) is discussed. Asymptotic behavior is studied in Section 3.3. A complexity study including inapproximability results is done in Section 3.2. Finally, computational experiments on real-life instances are presented in Section 3.4.

Hypothesis

An a graph describing the civil engineering structure is given, as well as demand points, and the number of fiber modules to be brought at each demand. The civil engineering structure selected should be an arborescence. In this arborescence, fiber cables have to be deployed in order to satisfy the demand.

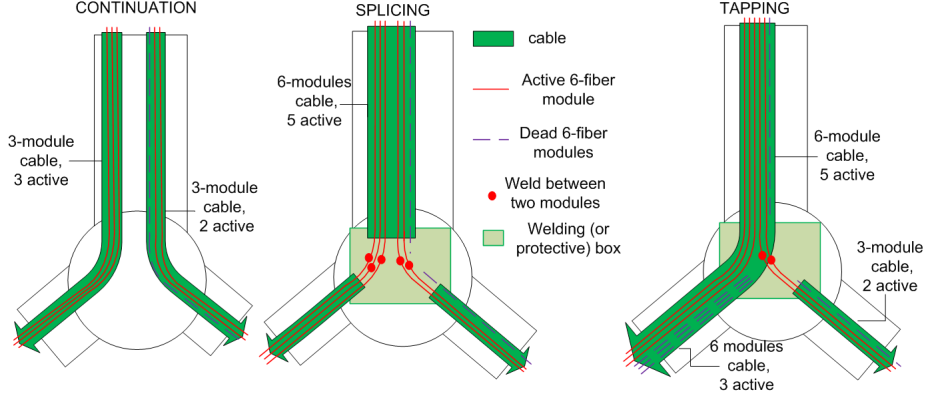


Figure 3.1: extended, spliced and tapped cables at a cabinet

In this problem, one has the possibility to:

- Use two different cable separation techniques: splicing and tapping (see Fig. 3.1). At most one of these operations can be done in a given node. The separation techniques conserve the number of active modules.
- Serve the demand in one of two different fashions: cable-served when one full cable is used, or module-served when modules from a cable are used.

Backfeed is not allowed, and the maintenance constraint does not apply.

3.1 Integer Programming

3.1.1 Notation

An symmetric directed graph $G = (V, E)$ describes the civil engineering structure, $V^* = V \setminus r$ denotes the set of nodes without the root r . Each node i has a demand of D_i modules. V_D denotes the set of nodes with a non-zero demand, and V_N the set of nodes with no demand.

For $i \in V$, we note $\Gamma^+(i) = \{j \in V \mid (i, j) \in E\}$ (since G is symmetric, we have $\Gamma^+(i) = \{j \in V \mid (j, i) \in E\}$). Each arc $(i, j) \in E$ has a length $\Delta_{(i,j)}$.

The cables available are from a discrete set $\mathcal{L} = \{1, \dots, L\}$, each cable having a number of modules M_l with $l \in \{1, \dots, L\}$. The set of possible active modules inside a cable of size l is denoted by $\mathcal{M}_l = \{1, \dots, M_l\}$. Each cable size has a cost

per length unit of C_l^{le} , the cost of a welding box of size l is PB_l , and the cost for welding m modules is PW_m .

The decision variables are the following:

- $\forall (i, j) \in E, X_{i,j} \in \{0, 1\}$: the binary variable which is equal to 1 iff some cables are installed in the duct described by arc (i, j) .
- $\forall (i, j) \in E, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, k_{i,j,l,m} \in \{0, \dots, |V_D|\}$: the number of cables with M_l fiber modules from which m are active going through arc (i, j) .
- $\forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, b_{i,l,m}^{spl} \in \{0, 1\}$: the binary variable equal to 1 if we splice in node i a cable with M_l modules of fibers from which m are active.
- $\forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, k_{i,l,m}^{born} \in \{0, \dots, M_L\}$: the number of cables born in node i having M_l modules of fibers from which m are active.
- $\forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, b_{i,l,m}^{tap} \in \{0, 1\}$ stands for the binary variable equal to 1 iff we tap a cable in node i having M_l modules of fibers from which m are active.
- $\forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, b_{i,l,m}^{ctn} \in \{0, 1\}$: the binary variable equal to 1 iff the continuation of a tapped cable in node i has M_l modules of fibers from which m are active modules.
- $\forall i \in V^*, \forall m \in \mathcal{M}_L, w_{i,m} \in \{0, 1\}$: the binary variable equal to 1 iff m modules are weld in either a tapping or a splicing in node i .
- $\forall i \in V_D, u_i$: the binary variable which equal to 1 iff the demand is served thanks to remaining modules of a spliced or tapped cable.

3.1.2 Formulation

A description of a cabling solution by the model is available in Fig. 3.2 (only variables not equal to 0 are noted).

The cables fiber network design problem can be cast into an integer linear program as follows:

$$\begin{aligned} \min \quad & \sum_{(i,j) \in E, l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} C_l^{le} \cdot \Delta_{(i,j)} \cdot k_{i,j,l,m} + \sum_{i \in V^*, l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} (b_{i,l,m}^{spl} + b_{i,l,m}^{tap}) \cdot PB_l \\ & + \sum_{i \in V^*, m \in \mathcal{M}_L} w_{i,m} \cdot PW_m \end{aligned}$$

such that

$$\sum_{j \in \Gamma^+(i), l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot k_{j,i,l,m} =$$

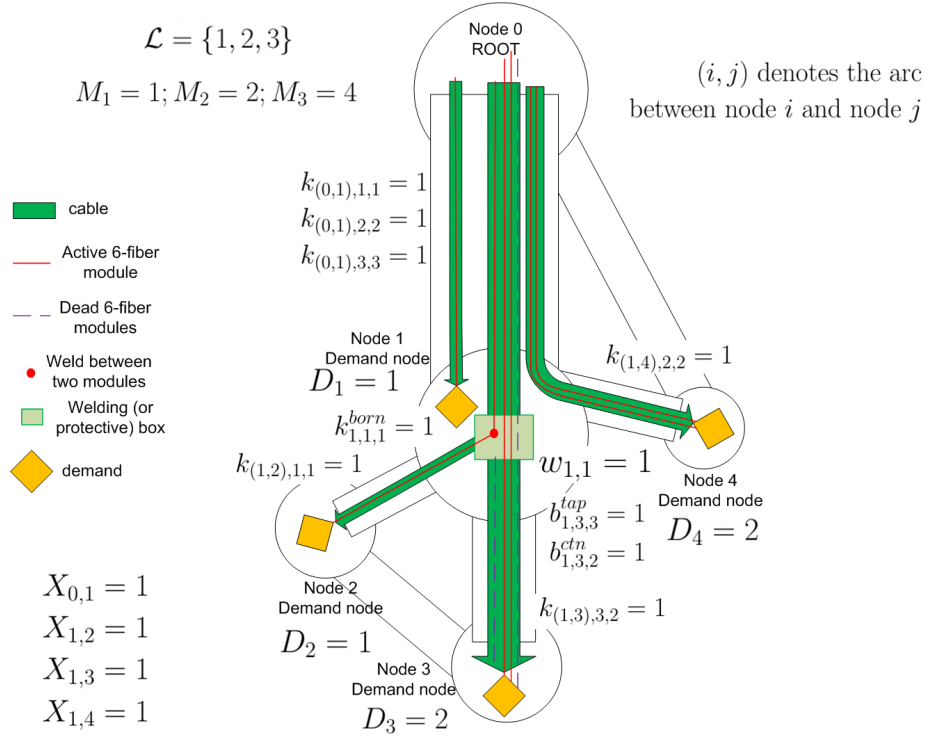


Figure 3.2: Model example

$$\sum_{j \in \Gamma^+(i)} \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot k_{i,j,l,m} + D_i \quad \forall i \in V^*, \quad (3.1)$$

$$\sum_{j \in \Gamma^+(i)} k_{j,i,l,m} - b_{i,l,m}^{spl} - b_{i,l,m}^{tap} = \sum_{j \in \Gamma^+(i)} k_{i,j,l,m} - k_{i,l,m}^{born} - b_{i,l,m}^{ctn} \quad \forall i \in V_N, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, \quad (3.2)$$

$$\sum_{j \in \Gamma^+(i)} k_{i,j,l,m} - k_{i,l,m}^{born} - b_{i,l,m}^{ctn} = \sum_{j \in \Gamma^+(i)} k_{j,i,l,m} - b_{i,l,m}^{spl} - b_{i,l,m}^{tap} \quad \forall i \in V_D, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l \setminus \{D_i\}, \quad (3.3)$$

$$\sum_{j \in \Gamma^+(i)} k_{i,j,l,m} - k_{i,l,m}^{born} - b_{i,l,m}^{ctn} \leq$$

$$\sum_{j \in \Gamma^+(i)} k_{j,i,l,m} - b_{i,l,m}^{spl} - b_{i,l,m}^{tap} \quad \forall i \in V_D, \forall l \in \mathcal{L}, m = D_i, \quad (3.4)$$

$$\begin{aligned} & \left(\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} \sum_{j \in \Gamma^+(i)} k_{i,j,l,m} - k_{i,l,m}^{born} \right. \\ & \left. - b_{i,l,m}^{ctn} \right) + 1 - u_i = \\ & \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} \sum_{j \in \Gamma^+(i)} k_{j,i,l,m} - b_{i,l,m}^{spl} - b_{i,l,m}^{tap} \quad \forall i \in V_D, \end{aligned} \quad (3.5)$$

$$0 \leq \sum_{j \in \Gamma^+(i)} k_{i,j,l,m} - k_{i,l,m}^{born} - b_{i,l,m}^{ctn} \quad \forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, \quad (3.6)$$

$$\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} \left(b_{i,l,m}^{spl} + b_{i,l,m}^{tap} \right) \leq \sum_{j \in \Gamma^+(i)} X_{j,i} \quad \forall i \in V^*, \quad (3.7)$$

$$\sum_{m \in \mathcal{M}_l} b_{i,l,m}^{ctn} = \sum_{m \in \mathcal{M}_l} b_{i,l,m}^{tap} \quad \forall (i, l) \in V^* \times \mathcal{L}, \quad (3.8)$$

$$\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot k_{i,l,m}^{born} = \sum_{m \in \mathcal{M}_L} m \cdot w_{i,m} \quad \forall i \in V^*, \quad (3.9)$$

$$\sum_{m \in \mathcal{M}_L} w_{i,m} \leq 1 \quad \forall i \in V^*, \quad (3.10)$$

$$\sum_{j \in \Gamma^+(i)} X_{j,i} \leq 1 \quad \forall i \in V_N, \quad (3.11)$$

$$\sum_{j \in \Gamma^+(i)} X_{j,i} = 1 \quad \forall i \in V_D, \quad (3.12)$$

$$X_{i,j} \leq \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} k_{i,j,l,m} \quad \forall (i, j) \in E, \quad (3.13)$$

$$\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} k_{i,j,l,m} \leq X_{i,j} |V_D| \quad \forall (i, j) \in E, \quad (3.14)$$

$$X, w, b^{spl}, b^{tap}, b^{ctn}, u \in \{0, 1\}$$

$$k \in \{0, \dots, |V_D|\}, k^{born} \in \{0, \dots, M_L\}$$

Among the terms of the cost function, first comes the cost of the cables, then the cost of eventual welding boxes at a node, while the last term corresponds to the price of welds. Equations (3.1) are flow conservation constraints related to active modules. They ensure that active modules entering each node either satisfy the demand or leave the node.

Constraints (3.2) make sure that cables entering a no-demand node which are neither tapped nor spliced are extended in the same way. The first term stands for the incoming cables, minus the eventual tapped or spliced one, and the second term for the outgoing cables, minus the eventual created ones. Indeed, if only the k variables are non-null, every cable arriving at the node will have the exact replica at the outcome of the node. Observe that if there is either a splicing or a tapping of a cable of type l with m active modules (either $b_{i,l,m}^{spl} = 1$

or $b_{i,l,m}^{tap} = 1$), then the downstream cables born after this operation and the continued cable (if a tapping occurs) can never have the same number of active modules (i.e., $k_{i,l,m}^{born} = 0$ and $b_{i,l,m}^{ctn} = 0$).

The three following equations are the equivalent equations concerning demand nodes. Constraints (3.3) ensure any cable with a number of active modules different from D_i is continued at node $i \in V_D$. Indeed, since the demand can only be served by one cable, this cable will have D_i active modules. Equations (3.4) ensure that more cables not affected in node i with D_i active modules arrive at the node i than the ones leaving it. In fact, if the demand is served by one full cable ($u_i = 0$), then all but one of the inequalities of type (3.4) are satisfied with equality. This is also expressed by constraints (3.5) which make sure the demand point is only satisfied by one cable, as specified in Section 2.1.

Constraints (3.6) express the fact that the number of cables going out of a node without coming from a tapping or splicing operation is nonnegative. Equations (3.7) ensure that at most one of the two operations of splicing and tapping can be performed in each node as specified in Section 2.1. It uses the fact that

$\sum_{j \in \Gamma^+(i)} X_{j,i} \leq 1$ by (3.11) and (3.12), and if no cable arrive at a node, it cannot

be host to a splicing or tapping operation. (3.8) make sure that after tapping the size of the continued cable does not change while the number of active modules changes (decreases). Constraints (3.9) and (3.10) ensure the variable w counts the number of welds done in a given point. Notice that then number of active modules in born cables is equal to the number of welds. Finally, equations (3.11), (3.12), (3.13) and (3.14) stand for the fact that the set of used arcs is a tree of the civil engineering infrastructure. Compared to the formulation in Section 2.2, the variables k have the role of the flow variables which ensure the connectivity of the tree.

In order to give more insight into the validity of the model, we prove in the following that the equations from the model imply the conservation of active modules.

Proposition 3.1.1 *Equations (3.1) to (3.14) imply the following*

$$\forall i \in V_N, \quad \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot (b_{i,l,m}^{tap} + b_{i,l,m}^{spl}) = \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot (k_{i,l,m}^{born} + b_{i,l,m}^{ctn}) \quad (3.15)$$

$$\forall i \in V_D, \quad \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot (b_{i,l,m}^{tap} + b_{i,l,m}^{spl}) = D_i \cdot u_i + \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot (k_{i,l,m}^{born} + b_{i,l,m}^{ctn}) \quad (3.16)$$

Proof. Multiplying (3.2) by m and summing gives

$$\forall i \in V_N, \quad \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot \left(\sum_{j \in \Gamma^+(i)} k_{j,i,l,m} - b_{i,l,m}^{tap} - b_{i,l,m}^{spl} \right) =$$

$$\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot \left(\sum_{j \in \Gamma^+(i)} k_{i,j,l,m} - k_{i,l,m}^{born} - b_{i,l,m}^{ctn} \right)$$

Then, using (3.1) to simplify gives (3.15).

Multiplying (3.3) and (3.4) by m and summing gives

$$\begin{aligned} \forall i \in V_D, \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot \left(\sum_{j \in \Gamma^+(i)} k_{j,i,l,m} - b_{i,l,m}^{tap} - b_{i,l,m}^{spl} \right) = \\ D_i - D_i \cdot u_i + \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot \left(\sum_{j \in \Gamma^+(i)} k_{i,j,l,m} - k_{i,l,m}^{born} - b_{i,l,m}^{ctn} \right) \end{aligned}$$

Then, using (3.1) gives (3.16). \square

Equations (3.15) and (3.16) show that module conservation applies to splicing and tapping operations, as specified in Section 2.1. Since with (3.7), at most one cable is either spliced or tapped, active modules from a tapped or spliced cable are used for created cables and an eventual module-served node. Apart from module conservation, demand nodes need further explanation.

Let us focus on a node $i \in V_D$ where $u_i = 1$. Then, (3.5) combined with (3.4) become

$$\begin{aligned} \forall l \in \mathcal{L} \text{ such that } D_i \in M_l, \sum_{j \in \Gamma^+(i)} k_{j,i,l,D_i} - b_{i,l,D_i}^{spl} - b_{i,l,D_i}^{tap} = \\ \sum_{j \in \Gamma^+(i)} k_{i,j,l,D_i} - k_{i,l,D_i}^{born} - b_{i,l,D_i}^{ctn} \end{aligned}$$

This induces that all cables unaffected by splicing or tapping operations are unchanged in i (similarly to (3.2)).

Let us focus on a node $i \in V_D$ where $u_i = 0$. Then, (3.5) combined with (3.4) tell us that only one of the equations (3.5) has a strict inequality, with a difference between left and right terms of 1. In other words, $\exists l_0 \in \mathcal{L} | D_i \in \mathcal{M}_{l_0}$ such that

$$\begin{aligned} \forall l \in \mathcal{L} \setminus \{l_0\}, \sum_{j \in \Gamma^+(i)} k_{j,i,l,D_i} - b_{i,l,D_i}^{spl} - b_{i,l,D_i}^{tap} = \sum_{j \in \Gamma^+(i)} k_{i,j,l,D_i} - k_{i,l,D_i}^{born} - b_{i,l,D_i}^{ctn} \\ \sum_{j \in \Gamma^+(i)} k_{j,i,l_0,D_i} - b_{i,l_0,D_i}^{spl} - b_{i,l_0,D_i}^{tap} = 1 + \sum_{j \in \Gamma^+(i)} k_{i,j,l_0,D_i} - k_{i,l_0,D_i}^{born} - b_{i,l_0,D_i}^{ctn} \end{aligned}$$

This shows that the demand is served in a cable served way by a cable of size l_0 . All other cables unaffected by tapping or splicing operations are continued with the same size and same number of active modules.

Finally, constraints (3.6) make sure that if a cable is created in a node, there is indeed, a cable of the same size and same number of active modules going out of this node. Combined with (3.2) and (3.5), it also ensures that if a cable is separated in a node, then there is an incoming cable with the same size and number of active modules.

With the notation $N_{ml} = \sum_{l \in \mathcal{L}} M_l$, this model holds $(|V| - 1)(3N_{ml} + M_L) + |V_D| + |E|$ boolean variables, $(|V| - 1 + |E|)N_{ml}$ integer variables and $(|V| - 1)(5 + 2N_{ml} + L) + |VD|(N_{ml} + 1) + 2|E|$ constraints.

Remark 3.1.1 *In some contexts (see for instance [54]), the use of a duct may not be free. In this case, the model can be adapted by including to the cost function the term $+\sum_{(i,j) \in E} X_{i,j} P_{i,j}^{duct}$, where $P_{i,j}^{duct}$ stands for the cost of using the duct described by arc (i, j) .*

3.1.3 Preprocessing

Optimal solutions for the linear program described above enjoy particular properties. More specifically, under optimality, it is possible to fix the value of some variables to zero.

A born cable having a number m of active module always has the minimum size able to hold m modules. Said another way, if m belongs to \mathcal{M}_{l-1} , there is no reason to create a cable of size l with m active modules. This leads to the following equations:

$$\forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_{l-1}, k_{i,l,m}^{born} = 0. \quad (3.17)$$

For the same reason, the number of cables at the root node satisfy similar equations.

$$\forall (j, l) \in \Gamma^+(r) \times \mathcal{L}, \forall m \in \mathcal{M}_{l-1}, k_{r,j,l,m} = 0 \quad (3.18)$$

Moreover, if a demand node is of degree 1, then there is no need for neither tapping nor splicing in it. This is summarized as follows:

$$\begin{aligned} \forall i \in V_D, \text{ if } |\Gamma^+(i)| = 1, \text{ then } \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, \\ k_{i,l,m}^{born} = 0, b_{i,l,m}^{ctn} = 0, b_{i,l,m}^{tap} = 0, b_{i,l,m}^{spl} = 0, u_i = 0. \end{aligned} \quad (3.19)$$

Remark 3.1.2 *In the real-life instances we have at our disposal, many demand points are of incidence degree one.*

3.1.4 Valid inequalities

Steiner tree related inequalities

The variable X describes an oriented tree of the civil engineering architecture covering r and V_D . Some reinforcements related to this structure are proposed hereinafter.

First, let us consider nodes without demand (called "Steiner nodes" in the classical STP). If an incoming arc is selected, then at least one outgoing arc has to be selected. Which gives:

$$\forall i \in V_N, \quad \sum_{j \in \Gamma^+(i)} X_{j,i} \leq \sum_{j \in \Gamma^+(i)} X_{i,j} \quad (3.20)$$

Let $STmin$ denote the minimum number of arcs of an arborescence covering V_D and r . $STmin$ can be computed by solving a standard steiner tree problem. We can get the following

$$\sum_{(i,j) \in E} X_{i,j} \geq STmin \quad (3.21)$$

Remark 3.1.3 *While computing $STmin$ is supposed to be NP-hard, it is not computationally expensive for the sizes of our instances.*

Also note that a solution using arc (i, j) cannot contain arc (j, i) (in other words there are no tours of length 2).

$$\forall (i, j) \in E, \quad X_{i,j} + X_{j,i} \leq 1 \quad (3.22)$$

What follows states that if an arc (i, j) is part of a the solution, and if i is not the root, then one of the incoming arcs of i must be part of the solution. It follows from the connectivity of the solution. Hence

$$\forall (i, j) \in E \text{ such that } i \neq r, \quad \sum_{j' \in \Gamma^+(i) \setminus \{j\}} X_{j',i} \geq X_{i,j} \quad (3.23)$$

Many other inequalities of this type can be found in the literature (see [10], [15], [16] or [25] for instance). However, here, the most numerous variables are not the ones related to this Steiner Tree structure of the solution (the $X_{i,j}$). This is why in the following we will focus more on inequalities specific to the problem.

Chvátal-Gomory cuts

The inequalities below are Gomory-Chvatal rounding cuts.

Proposition 3.1.2 *The following inequalities are valid for FCND.*

$$\forall m \in \mathcal{M}_L, \quad \sum_{j \in \Gamma^+(r)} \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} \sum_{m' \in \{m, \dots, M_l\}} k_{r,j,l,m'} \leq \left\lfloor \frac{\sum_{i \in V_D} D_i}{m} \right\rfloor \quad (3.24)$$

Proof. We can get from (3.1):

$$\sum_{j \in \Gamma^+(r)} \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot k_{r,j,l,m} = \sum_{i \in V_D} D_i$$

Hence, for $m \in \mathcal{M}_L$,

$\sum_{j \in \Gamma^+(r)} \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} \sum_{m' \in \{m, \dots, M_l\}} m \cdot k_{r,j,l,m'} \leq \sum_{i \in V_D} D_i$. The wanted inequality is then obtained by simple rounding. \square

Problem related inequalities

To improve the quality of the linear relaxation, we should add more valid inequalities. Some of them are presented here.

After a tapping operation, the continued cable has less active modules. This gives the following:

Proposition 3.1.3 *Every optimal solution of the FCND problem verifies*

$$\forall (i, l) \in V^* \times \mathcal{L}, \forall m' \in \mathcal{M}_l \setminus \{1\}, \quad \sum_{m \in \{1, \dots, m'-1\}} b_{i,l,m}^{ctn} \geq \sum_{m \in \{1, \dots, m'\}} b_{i,l,m}^{tap}. \quad (3.25)$$

The proposition below exploits the fact that, tapping and splicing can never simultaneously occur in the same node. This is true in particular for a given cable size l .

Proposition 3.1.4 *Every optimal solution of the FCND problem verifies*

$$\forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, b_{i,l,m}^{tap} + b_{i,l,m}^{ctn} + \sum_{l' \in \mathcal{L}} \sum_{m' \in \mathcal{M}_{l'}} b_{i,l',m'}^{spl} \leq 1. \quad (3.26)$$

Proof. Let us consider, in an optimal solution S of the FCND problem, $i \in V^*$, $l \in \mathcal{L}$, and $m \in \mathcal{M}_l$.

If a splicing is done in i , $\sum_{l' \in \mathcal{L}} \sum_{m' \in \mathcal{M}_{l'}} b_{i,l',m'}^{spl} = 1$, then by (3.7) (which specifies that at most one operation can be done in a node), we have $b_{i,l,m}^{tap} = b_{i,l,m}^{ctn} = 0$.

Otherwise, let us assume $b_{i,l,m}^{tap} = 1$, which means there is a tapping, and $b_{i,l,m}^{ctn} = 1$, a continued cable with the same number of active modules goes out of i . Then, the solution identical to S but with $b_{i,l,m}^{tap} = b_{i,l,m}^{ctn} = 0$ is cheaper and feasible, which contradicts our hypothesis. Hence the result. \square

Assume that m modules are welded in a node i where $m \in \{1, \dots, m'\}$, then some downstream cables will be born in i and each one of these cables has at most m' active modules. This gives us the proposition below

Proposition 3.1.5 *The following inequalities are valid for the FCND problem*

$$\forall i \in V_N, \forall m' \in \mathcal{M}_L, \sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} k_{i,l,m}^{born} \geq \sum_{m \in \{1, \dots, m'\}} w_{i,m}. \quad (3.27)$$

Writing that a welding box is installed only if either a splicing or a tapping occurs, we get equality (3.28).

$$\forall i \in V^*, \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} b_{i,l,m}^{spl} + b_{i,l,m}^{tap} = \sum_{m \in \mathcal{M}_L} w_{i,m}. \quad (3.28)$$

The next inequality just says that when a cable having m active modules is spliced in a node $i \in V_N$ (no demand in i), then exactly m modules should be welded.

$$\forall i \in V_N, \forall m \in \mathcal{M}_L, \sum_{l \in \mathcal{L} | m \in \mathcal{M}_l} b_{i,l,m}^{spl} \leq w_{i,m}. \quad (3.29)$$

Proof. Let us consider $i \in V_N$ and $m \in \mathcal{M}_L$ such that $\sum_{l \in \mathcal{L} | m \in \mathcal{M}_l} b_{i,l,m}^{spl} = 1$. Then, by (3.7), we know that there is no tapping operation done in i , and that no other splicing is done. Which means (3.15) becomes $\sum_{l \in \mathcal{L}} \sum_{m' \in \mathcal{M}_l} m' \cdot k_{i,l,m'}^{born} = \sum_{l \in \mathcal{L} | m \in \mathcal{M}_l} m \cdot b_{i,l,m}^{spl}$. By (3.9), we have the result. \square

We also know that for such non demand nodes, tapping a cable with m' active modules requires welding at least 1 and at most $m' - 1$ modules. This can be shown using (3.8) and (3.15). As a consequence, the following inequality holds.

Proposition 3.1.6 *The following inequalities are valid for the FCND problem.*

$$\forall i \in V_N, \forall l \in \mathcal{L}, \forall m' \in \mathcal{M}_l, b_{i,l,m'}^{tap} \leq \sum_{m \in \{1, m' - 1\}} w_{i, m' - m}. \quad (3.30)$$

Note that in a no-demand point at least two cables must come out of a splicing or tapping operation for it to be worth the money paid. This is noted by $\forall i \in V_N, \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} k_{i,l,m}^{born} + b_{i,l,m}^{ctn} \geq 2 \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} b_{i,l,m}^{tap} + b_{i,l,m}^{spl}$. Besides, the downstream cables have always less active modules than the upstream cables, which gives:

Proposition 3.1.7 *Every optimal solution of the FCND problem verifies*

$$\begin{aligned} \forall i \in V_N, \forall m' \in \mathcal{M}_L, \sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} k_{i,l,m}^{born} + b_{i,l,m}^{ctn} \geq \\ 2 \sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} b_{i,l,m}^{tap} + b_{i,l,m}^{spl} \end{aligned} \quad (3.31)$$

Similarly, this is also true for a cable-served demand point, which is noted

Proposition 3.1.8 *Every optimal solution of the FCND problem verifies*

$$\begin{aligned} \forall i \in V_D, \forall m' \in \mathcal{M}_L, u_i + \sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} k_{i,l,m}^{born} + b_{i,l,m}^{ctn} \geq \\ 2 \cdot \sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} b_{i,l,m}^{tap} + b_{i,l,m}^{spl} \end{aligned} \quad (3.32)$$

For a node $i \in V_N$, if the number of cables arriving at this node is strictly smaller than the number of arcs covered downstream of this node ($\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} \sum_{j \in \Gamma^+(i)} k_{j,i,l,m} < \sum_{j \in \Gamma^+(i)} X_{i,j}$), then this node must be the host of a splicing or tapping operation (i.e. $\sum_{m \in \mathcal{M}_L} w_{i,m} = 1$). Since $\sum_{j \in \Gamma^+(i)} X_{i,j} \leq |\Gamma^+(i)| - 1$, we have:

Proposition 3.1.9 *The following inequalities are valid for the FCND problem*

$$\forall i \in V_N, \left(\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} \sum_{j \in \Gamma^+(i)} k_{j,i,l,m} \right) + (|\Gamma^+(i)| - 2) \sum_{m \in \mathcal{M}_L} w_{i,m} \geq \sum_{j \in \Gamma^+(i)} X_{i,j} \quad (3.33)$$

Proof. Let us consider in a FCND solution $i \in V_N$ such that $\sum_{j \in \Gamma^+(i)} X_{i,j} \geq 1$. If $\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} \sum_{j \in \Gamma^+(i)} k_{j,i,l,m} \geq \sum_{j \in \Gamma^+(i)} X_{i,j}$, then the inequality is verified.

If $\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} \sum_{j \in \Gamma^+(i)} k_{j,i,l,m} < \sum_{j \in \Gamma^+(i)} X_{i,j}$. Let us assume that $\sum_{m \in \mathcal{M}_L} w_{i,m} = 0$. Then, by (3.28), it means that there is no operation done in i , which means there are no born or continued cables (see (3.15)). Therefore, (3.2) becomes $\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} \sum_{j \in \Gamma^+(i)} k_{j,i,l,m} = \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} \sum_{j \in \Gamma^+(i)} k_{i,j,l,m} < \sum_{j \in \Gamma^+(i)} X_{i,j}$. However, summing (3.13) on all outgoing arcs of i gives $\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} \sum_{j \in \Gamma^+(i)} k_{i,j,l,m} \geq \sum_{j \in \Gamma^+(i)} X_{i,j}$.

So $\sum_{m \in \mathcal{M}_L} w_{i,m} = 1$, hence the result. \square

Remark 3.1.4 *There are $|V_N|(2M_L + N_{ml} + 2) + 2|E| + 2|M_L| + (|V| - 1)(2N_{ml} + M_L)$ additional constraints. The effect of these inequalities is assessed in Section 3.4.*

3.2 Complexity

We first show that the FCND problem is NP-hard by reducing it to a Steiner tree problem. More precisely, it is NP-hard in a restricted context: one cable size and demands of 1 module. This implies the NP-hardness in the general case.

Proposition 3.2.1 *The FCND problem is NP-hard even with one cable size and demands of 1 module.*

Proof. Let us consider an instance of the Steiner Tree decision problem, a graph $G = (V, E)$, $T \subseteq V$ a set of terminals, and an integer $n \in \mathbb{N}$. The question associated to this instance is "is there a tree G' made of edges and nodes of G spanning T with at most n edges?"

Let us reduce it to the following FCND instance.

- A civil engineering structure identical to $G = (V, E)$ where all the edges have length one.
- A root node $r \subseteq T$.
- A demand of 1 active module for every node in $T \subseteq r$.
- A set of $|T| - 1$ cables of sizes $\{1, \dots, |T| - 1\}$. All cables have a cost per length unit equal to 1.
- Zero cost for welds and boxes.

The question associated to this instance is "is there a solution of cost at most n ?"

If the Steiner Tree instance is feasible. Then, we have a tree G' with at most n edges spanning T . Let us build the FCND solution where:

- The civil engineering structure used is G' .
- Demand nodes which are leaf nodes of G' are cable-served, served by a cable of size M_1 .
- If a non-root node $v \in V$ has an outgoing degree larger than 2 in G' , a splicing is done in v . One cable of size M_1 is deployed on (u, v) where u is the successor of v , it is spliced in v , and $|\Gamma^+(v)| - 1$ cables are born, one for each outgoing arc. If v is a demand node, the demand is served in a module-served way.
- All non leaf nodes v in V_D are module-served by a splicing. The number of born cables in this splicing is $|\Gamma^+(v)| - 1$ (see previous point).

First, let us note that in this solution only one cable is deployed on each arc of G' . Indeed, it is true for the arcs having leaf nodes as a target. By construction, it is also true for higher arcs.

The cost of this solution is only the cost for deploying cables. Since there is one cable of cost per length unit 1 on ducts of length 1, the cost on one duct is 1, and the total cost of this solution is at most n .

Hence the FCND instance is feasible.

If the FCND instance has a solution of cost at most n . This solution uses an arborescence of the civil engineering infrastructure. Furthermore, this arborescence spans the root r and all the demand nodes, that is all nodes in T . On each arc of this arborescence, at least one cable is deployed, which contributes to the total cost for at least 1. Hence there cannot be more than n arcs of G used in this arborescence. So the Steiner Tree decision problem was feasible.

Since the two instances are equivalent, we have the result. \square

We show in the following that the FCND problem is not approximable within a ratio of 2.

Proposition 3.2.2 *The FCND problem cannot be approximated within a ratio $2 - \epsilon$ for any $\epsilon > 0$.*

Proof. Let us consider $\epsilon > 0$.

Let us consider an instance $G^H = (V^H, E^H)$ of the Hamiltonian chain problem.

Let us reduce it to the following FCND instance:

- A civil engineering structure $G = (V, E)$ where $V = V^H \cup \{r\}$, $E = E^H \cup \{(r, v) | v \in V^H\}$ where v is a node in V^H . The length of the arcs going out of r is 1, the length of the arcs in E is 0.
- All the nodes in V^H have a demand of 1 active module.

- Only one cable size containing $|V^H|$ active modules is available. Its cost per length unit is 1.
- The cost of welds is constant $\forall m \in \{1, \dots, |V^H|\}, PW_m = 1$, the cost of boxes is zero.

The question we ask is "is there a solution of the FCND instance with a cost at most 1" ?

Let us assume this FCND instance has a solution of cost 1. Then, it means there is only one cable deployed on edge (r, v) , since the solution requires at least one cable. Furthermore, since a weld costs at least 1, there cannot be any welds in the solution. Since all demand nodes are served by only one cable, this implies that the nodes in V^H are all served in a module-served way by a tapping (see Fig. 3.3), except one. Since all demand nodes are served by a single cable, the path on which this cable is deployed goes through all the nodes in V^H . In other words, this path is a hamiltonian chain, and the hamiltonian chain instance was feasible.

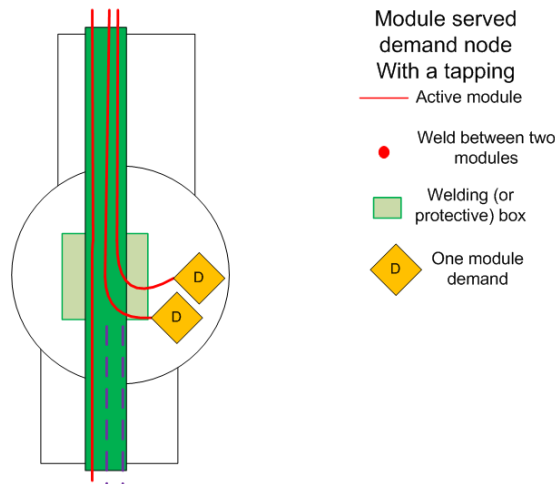


Figure 3.3: Module-served demand node with a tapping operation and without any born cables. A box is onstalled, there are no welds to be done.

Reciprocally, let us assume the Hamiltonian chain instance was feasible. Then, we can build the solution where

- Only one cable goes out of the root.
- This cable follows a hamiltonian chain in G^H . It serves all the nodes in the interior of the chain in a module-served way by a tapping operation.
- The last node of the hamiltonian chain is cable-served by this cable.

This solution has a cost of 1, since the only cost source is to deploy the cable on (r, v) .

Let us assume we have a polynomial time approximation scheme \mathcal{A} able to give a solution with a ratio $2 - \epsilon$. Here is a polynomial time algorithm able to decide the feasibility of any hamiltonian chain problem instance.

- Reduce the hamiltonian chain instance to the FCND instance described above.
- Run \mathcal{A} on this FCND instance.
- If the solution returned by \mathcal{A} has a cost strictly less than 2, then declare the hamiltonian chain instance feasible. Otherwise, declare it unfeasible.

If the cost c' of the solution given by \mathcal{A} verifies $c' < 2$, then the cost c^* of an optimal solution verifies $c^* < 2$. Since the cost of a solution is integer, $c^* \leq 1$. This implies hamiltonian chain instance is feasible. If the cost c' of the solution given by \mathcal{A} verifies $c' \geq 2$, the cost c^* of the optimal solution verifies $c^* \geq c' \frac{1}{2-\epsilon} > 1$. Since c^* is integer, it implies $c^* \geq 2$, and hamiltonian chain instance was not feasible. Therefore, this algorithm provides the answer of the hamiltonian chain problem.

Hence the result. □

3.3 Asymptotic behavior

In this section, we study the shape of the arborescences used in optimal solutions as a function of the costs. We do not know of any easy way to find an optimal solution arborescence without solving the problem. The problem of finding this arborescence is generally NP-hard (see Section 3.2). However, under some conditions, it relates to classical problems. In the following, we provide some insight about the arborescence selected in optimal solutions when separation costs are negligible and when cable separation costs are negligible.

We study two cases: first, the case where the cable deployment costs are negligible compared to cable separation costs. In this case, the arborescence selected is the shortest paths arborescence. Second, we study the case where the cable separation costs is negligible compared to cable deployment costs, and cables are large enough to contain all active modules. In this case, the arborescence selected is a solution of the Weighted Minimal Steiner Tree Problem.

Let us note, for $i \in V_D$, d_i^{pr} the length of the shortest path between the root node r and i in G . Let us note $C_i^{min} = \min\{C_l^{le} | l \in \mathcal{L}, M_l \geq D_i\}$ the smallest cable large enough to supply the demand in i .

Proposition 3.3.1 *Let us consider an FCND instance. If $\sum_{i \in V_D} d_i^{pr} \cdot C_i^{min} < PB_1$ (the cost of the smallest box is larger than the cost of the solution without*

separation), then in any optimal solution of the FCND instance, the arborescence selected is formed of the shortest paths from the root to each demand node.

Proof. Let us consider the FCND solution where one cable coming from the root serves each demand point by going through the shortest path. The size of each of these cables is known, it is the C_i^{min} for $i \in V_D$. The cost of this solution is $\sum_{i \in V_D} d_i^{pr} \cdot C_i^{min}$.

Furthermore,

- any solution with at least one cable separation will cost at least PB_1 ,
- any solution without cable separation but where a different arborescence is selected will be more expensive.

Hence the result. □

In the following, we note $D^{tot} = \sum_{i \in V_D} D_i$ the sum of all demands in the FCND instance. Let us note C^{ST} the length of the minimal weighted Steiner Tree spanning the demand nodes and the root. Since the number of possible trees spanning V_D is finite, there is a largest real number $\delta^{ST} > 0$ such that a tree spanning V_D and r has either a length of C^{ST} or a length larger than $C^{ST} + \delta^{ST}$. While in practice computing δ^{ST} can be hard, it is easily lower bounded. For instance, if the lengths of the edges are integer and dividable by $n \in \mathcal{N}$, then we have $\delta^{ST} \geq n$.

Proposition 3.3.2 *Let us consider a FCND instance. If*

- $M_1 \geq D^{tot}$ (the smallest cable contains more than D^{tot} modules)
- $\delta^{ST} \cdot C_1^{le} \leq (|V_D| - 1) \cdot (PB_1 + PW_{D^{tot}})$

Then any optimal solution of FCND uses as civil engineering structure an optimal solution of the Weighted Steiner Tree Problem.

Proof. Let us consider such a FCND instance. We can build a solution where the civil engineering structure used is the minimal weighted steiner tree spanning V_D and r . We call $G' = (V', A')$ this arborescence oriented from the root node. We build the following solution in which all arcs have only one cable of size M_1 (the smallest cable size).

- Leaf nodes are cable-served demand nodes, served by a cable of size M_1 .
- If a non-root node $v \in V'$ has an outgoing degree d larger than 2 in G' , a splicing is done in v . One cable of size M_1 is deployed on the only arc of A' whose target is v , it is spliced in v , and d cables are born, one for each outgoing arc. If v is a demand node, the demand is served in a module-served way.
- All non leaf nodes v in V_D are module-served by a splicing. The number of born cables in this splicing is the outgoing degree of v in G' (see previous point).

First, let us note that in this solution only one cable is deployed on each arc of G' . Indeed, it is true for the arcs having leaf nodes as a target. By construction, it is also true for higher arcs.

Then, let us note that there cannot be more than $|V_D| - 1$ splicing points. Indeed, let us consider a splicing in this solution. The spliced cable contains at least the active modules of one of the born cables and one additional demand point (either from a module-served demand point or from another born cable). Since at least one cable goes out of the root, there is at most $|V_D| - 1$ splicing operations.

This solution has a cost of $C^{ST} \cdot C_1^{le} + \epsilon$ where ϵ is the cost for boxes and wires. With the previous point, we have $\epsilon \leq (|V_D| - 1) \cdot (PB_1 + PW_{D^{tot}})$

Let us consider a solution of the FCND problem using an arborescence which is not a Minimal Weighted Steiner Tree of G . It has at least one cable on each arc it uses, so its cost is at least $(C^{ST} + \delta^{ST}) \cdot C_1^{le} > C^{ST} \cdot C_1^{le} + \epsilon$. So it cannot be an optimal solution of FCND.

Hence the result. \square

The next Section assesses the numerical complexity of the models as well as some properties of the solutions obtained.

3.4 Numerical results

The two models (the base model with variables filtered and the enhanced model with reinforcements) were tested on real-life instances. The solving algorithm was the default branch and bound from Cplex version 12.6.0.0. The solver configuration used was the default setting. The experiments were run on a computer composed of 4 processors of CPU 5110 and clocked at 1.6 GHz each. We consider here 6-fiber modules, and the cables have 1, 2, 4, 6, 8, 12, 18 or 24 modules. The instances are parts of the cities of Clermont-Ferrand and Arles (France), some of their key features are described in Table 3.1; "Cl" designs areas of Clermont-Ferrand, and "Ar" areas from Arles.

The costs used were real costs based on operator's internal studies that are currently used in network planning tools of Orange.

Table 3.1: Key features of the different instances

| instance | nodes | edges | demand nodes | total demand | overall length (m) |
|-----------|-------|-------|--------------|--------------|--------------------|
| zone 0 Cl | 83 | 82 | 26 | 38 | 1566.3 |
| zone 1 Cl | 82 | 85 | 25 | 40 | 1908.2 |
| zone 2 Cl | 77 | 79 | 24 | 40 | 1695.5 |
| zone 3 Cl | 70 | 73 | 20 | 28 | 2161.5 |
| zone 4 Cl | 81 | 87 | 24 | 34 | 1943.4 |
| zone 5 Cl | 74 | 75 | 22 | 58 | 2652.9 |
| zone 6 Cl | 57 | 59 | 14 | 20 | 1132.9 |
| zone 7 Cl | 64 | 64 | 13 | 59 | 1896.0 |
| zone 8 Cl | 84 | 86 | 21 | 35 | 3398.3 |
| zone 0 Ar | 127 | 127 | 45 | 61 | 5697.1 |
| zone 1 Ar | 190 | 220 | 38 | 55 | 35 289.2 |
| zone 2 Ar | 128 | 136 | 35 | 66 | 6941.6 |
| zone 3 Ar | 125 | 124 | 43 | 80 | 2917.1 |
| zone 4 Ar | 139 | 139 | 44 | 68 | 5039.1 |
| zone 5 Ar | 168 | 186 | 43 | 67 | 13 906.5 |
| zone 6 Ar | 229 | 249 | 35 | 68 | 14 525.0 |
| zone 7 Ar | 243 | 270 | 41 | 63 | 35 131.8 |
| zone 8 Ar | 353 | 398 | 68 | 78 | 56 776.8 |

The base model ((3.1) to (3.19)) is noted model (A) and the enhanced model ((3.1) to (3.32)) is noted model (B). The optimal solutions and continuous relaxation values are presented in Table 3.2. We also consider the greedy solution of FCND where neither tapping nor splicing is allowed. The obtained solution is obviously a feasible solution of the initial problem. The cost of this solution is also shown in Table 3.2. The character ”-” means the optimal solution could not be found with our computing power.

Table 3.2: Solution values

| instance | continuous relaxation model (A) | continuous relaxation model (B) | optimal solution | solution without splicing or tapping |
|-----------|---------------------------------|---------------------------------|------------------|--------------------------------------|
| zone 0 Cl | 1773.9 | 2062.5 | 2188.9 | 2596.7 |
| zone 1 Cl | 1845.6 | 2105.2 | 2241.0 | 2372.0 |
| zone 2 Cl | 2163.6 | 2520.1 | 2685.2 | 3013.1 |
| zone 3 Cl | 1224.4 | 1470.7 | 1551.1 | 1613.1 |
| zone 4 Cl | 1684.8 | 2007.8 | 2181.2 | 2627.5 |
| zone 5 Cl | 3572.5 | 3981.8 | 4155.8 | 5003.7 |
| zone 6 Cl | 992.2 | 1177.5 | 1289.1 | 1372.0 |
| zone 7 Cl | 1201.5 | 1281.8 | 1341.0 | 1348.1 |
| zone 8 Cl | 3114.2 | 3549.1 | 4093.9 | 6115.0 |
| zone 0 Ar | 4035.2 | 4927.4 | 5697.1 | 7395.2 |
| zone 1 Ar | 5544.1 | 7625.9 | 9350.1 | 11960.4 |
| zone 2 Ar | 4723.0 | 5325.7 | 5965.3 | 7647.7 |
| zone 3 Ar | 4746.8 | 5429.5 | 5945.3 | 7219.0 |
| zone 4 Ar | 3642.3 | 4545.9 | 4860.2 | 5623.4 |
| zone 5 Ar | 4293.3 | 5354.6 | 6347.6 | 8075.3 |
| zone 6 Ar | 6653.6 | 7503.8 | - | 12 766.1 |
| zone 7 Ar | 7533.8 | 10 401.8 | - | 15 421.3 |
| zone 8 Ar | 15 427.3 | 19 783.3 | - | 44 766.7 |

The comparison of the two models shows a significant improvement of the continuous relaxation value with the enhanced model, going in average from 77% to 91% of the solution value for instances where the solution is known. According to Table 3.2, when neither tapping nor splicing is allowed, the solution is in average 20% more expensive than the optimal solution, and more important on longer infrastructures.

We also report in Table 3.3 how the optimal solution's cost is decomposed: cable costs, cost of welds and welding box costs. In the optimal solutions, the overall cost of welds stands for in average 18% of the total cost while the cost of boxes is in average 3.5 % of the total cost. Note that for longer instances, the number of splicing and tapping points, as well as the costs of welds and welding boxes, is higher.

Table 3.3: Composition of the different cost sources in optimal solutions

| instance | cable cost | welds cost | protective box cost | number of tapping or splicing points |
|-----------|------------|------------|---------------------|--------------------------------------|
| zone 0 Cl | 1765.1 | 330 | 93.8 | 3 |
| zone 1 Cl | 1667.1 | 480 | 93.8 | 3 |
| zone 2 Cl | 1955.9 | 660 | 69.3 | 2 |
| zone 3 Cl | 1382.1 | 120 | 49.0 | 2 |
| zone 4 Cl | 1503.5 | 588 | 89.6 | 2 |
| zone 5 Cl | 3106.2 | 960 | 89.6 | 2 |
| zone 6 Cl | 949.8 | 270 | 69.3 | 2 |
| zone 7 Cl | 1256.4 | 60 | 24.5 | 1 |
| zone 8 Cl | 3254.7 | 660 | 179.2 | 4 |
| zone 0 Ar | 4321.4 | 1050 | 366.8 | 10 |
| zone 1 Ar | 8099.3 | 900 | 350.8 | 11 |
| zone 2 Ar | 4393.1 | 1344 | 228.2 | 6 |
| zone 3 Ar | 4373.0 | 1344 | 228.2 | 6 |
| zone 4 Ar | 3714.0 | 918 | 228.2 | 6 |
| zone 5 Ar | 5385.8 | 660 | 301.8 | 9 |

The computing times obtained with both models are displayed in Table 3.4. In order to avoid having too much constraints in model (B), those from (3.25) - (3.30) and (3.31) - (3.33) are removed from the branch and bound algorithm if they concern a node i such that in the initial relaxation $\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} (b_{i,l,m}^{tap} + b_{i,l,m}^{spl}) = 0$. In addition to the total computing time, we also provide the time before appearance of a first integer solution.

Table 3.4: Computation times

| instance | with model (A) | | with model (B) | |
|-----------|----------------|------------------------------|----------------|------------------------------|
| | time (s) | time until first integer (s) | time (s) | time until first integer (s) |
| zone 0 Cl | 12.1 | 10.2 | 7.2 | 4.1 |
| zone 1 Cl | 41.6 | 25.7 | 101.7 | 22.7 |
| zone 2 Cl | 87.8 | 34.0 | 62.8 | 52.9 |
| zone 3 Cl | 18.8 | 18.8 | 14.7 | 4.0 |
| zone 4 Cl | 39.5 | 21.7 | 146.2 | 81.5 |
| zone 5 Cl | 11.9 | 11.4 | 10.6 | 4.3 |
| zone 6 Cl | 45.7 | 22.2 | 33.4 | 4.6 |
| zone 7 Cl | 8.4 | 7.6 | 2.5 | 2.3 |
| zone 8 Cl | 4350.0 | 697.3 | 14667.4 | 417.8 |
| zone 0 Ar | 135.8 | 73.9 | 122.7 | 72.8 |
| zone 1 Ar | 567.5 | 208.9 | 561.9 | 463.5 |
| zone 2 Ar | 341.8 | 102.7 | 597.7 | 88.6 |
| zone 3 Ar | 66.0 | 19.2 | 98.5 | 23.6 |
| zone 4 Ar | 40.5 | 40.4 | 24.0 | 18.2 |
| zone 5 Ar | 4065.9 | 173.3 | - | 3151.9 |
| zone 6 Ar | - | - | - | 469.9 |
| zone 7 Ar | - | - | - | 12 001.0 |
| zone 8 Ar | - | - | - | - |

The computation time grows with the parameters described in table 3.1. The length of the underlying network has an influence as long distances between the splitter and the demand nodes triggers more splicing and tapping operations in the optimal solution, which means more variables from the ILP of section 3.1.2 will be non-null. Model (B) finds the optimal solution faster than the model (A) in 60% of the instances. It finds an integer solution faster than the model (A) in 75% of the cases.

In order to estimate the influence of the set of available cables \mathcal{L} , we solve the cable design problem with only 5 cable sizes: 1, 2, 6, 12 or 24 modules per cable. Table 3.5 displays the obtained results: the solution cost and the computing times when 5 cable sizes are considered. These results should be compared with those given in Table 3.2. The average use of cables in both cases (5 cable sizes and 8 cable sizes) are also reported. The costs used are the same than for previous computations.

Table 3.5: Influence of the set of cables available

| instance | with 5 cables sizes | | | with 8 cables sizes |
|-----------|---------------------|----------|------------------------------|------------------------------|
| | solution | time (s) | percentage of active modules | percentage of active modules |
| zone 0 Cl | 2206.8 | 4.7 | 95 | 95 |
| zone 1 Cl | 2248.5 | 14.7 | 95 | 95 |
| zone 2 Cl | 2685.2 | 27.3 | 95 | 95 |
| zone 3 Cl | 1590.0 | 16.9 | 89 | 93 |
| zone 4 Cl | 2195.4 | 40.6 | 93 | 95 |
| zone 5 Cl | 4345.2 | 14.5 | 84 | 89 |
| zone 6 Cl | 1293.1 | 13.0 | 94 | 94 |
| zone 7 Cl | 1489.3 | 1.7 | 87 | 98 |
| zone 8 Cl | 4162.5 | 1111.1 | 89 | 93 |
| zone 0 Ar | 5912.9 | 65.1 | 88 | 90 |
| zone 1 Ar | 9757.0 | 382.9 | 83 | 83 |
| zone 2 Ar | 5996.6 | 122.6 | 90 | 91 |
| zone 3 Ar | 5975.5 | 59.4 | 90 | 99 |
| zone 4 Ar | 4904.7 | 16.8 | 91 | 91 |
| zone 5 Ar | 6507.2 | 17 964.5 | 93 | 92 |
| zone 6 Ar | - | - | - | - |
| zone 7 Ar | - | - | - | - |
| zone 8 Ar | - | - | - | - |

Observe that the cost is in average 2% higher in the 5 cable solution than in the 8 cable solution. The chosen subset of available cables gives a pretty good approximation of the optimal solution. As for the computing time, solving the problem with only 5 cable sizes is generally much faster than solving it with 8 sizes. Consequently, one can use this technique to get a sub-optimal solution in a shorter time.

Regarding environmental aspects, it is desirable to use as much modules as possible in the used cables. The average $\frac{\text{active modules}}{\text{overall modules}}$ ratio is 93 % in the 8 cables solution, while it goes down to an average of 90 % in the 5 cables solution.

3.5 Conclusion

This chapter addresses the Fiber Cables Network Design problem. We proposed to tackle it thanks to integer programming and valid inequalities. By studying the solutions, we noticed that the introduction of splicing and tapping techniques enable significant cost savings. This justifies our focus on fiber cables.

On a theoretical point of view, the problem is NP-hard and hard to approx-

imate. Furthermore, it is linked to the Steiner Tree problem, since the design of a fiber cables network is done "inside" a Steiner Tree, in the sense that the used civil engineering structure must be an arborescence spanning all demand points. The relation with the Steiner Tree problem can be seen in the asymptotic behavior of FCND for negligible cable separation costs. The relation with the shortest paths arborescence is highlighted by the behavior of FCND for negligible cable deployment costs.

On an algorithmic point of view, the integer programming solution proposed was able to solve successfully the smallest real-life instances. The valid inequalities introduced significantly decrease the computation times. However, the largest real-life instances stay out of reach. This justifies to study the problem independently from elements such as splitter location, which would make the problem even harder to tract. This lead us to tackle the FCNDA problem in Chapter 4, which is a restriction of FCND in an arborescence, for which we can expect to have more instances within our reach.

Chapter 4

Fiber Cables Network Design in an Arborescence Problem

Chapter Abstract

This chapter tackles the FCNDA problem. It can be seen as a restriction of FCND where the civil engineering structure available is a tree. Recall that it can be stated as follows. Given an arborescence describing the civil engineering infrastructure, the decision problem consists in selecting a set of cables with enough active modules to serve the demand on each arc and associated separation points on the nodes. Separation points should conserve the number of active modules so that active modules are not interrupted from the fiber source to the demand. While minimizing the total cost (cable layout, welding boxes and welds), each demand node should be either cable served or module served by a single cable, and at most one separation point per node is allowed.

We study the complexity of the problem in Section 4.1: we prove that it is NP-hard, inapproximable, and that it cannot be easily modeled using a certain set of variables. This leads us to introduce the two integer linear programming models proposed in Sections 4.2.2 and 4.3.2. The first one is path based and uses a very different approach than what was done for the FCND, while the second one is arc-based and inspired by the model of Section 3.1. Variable fixings and valid inequalities are proposed for both models. Finally, computational experiments are presented in Section 4.4.

Hypothesis

An arborescence of the civil engineering structure is given, as well as demand points, and the number of fiber modules to be brought at each demand. This implies that the number of modules going through each duct is known (see example from Fig. 4.1).

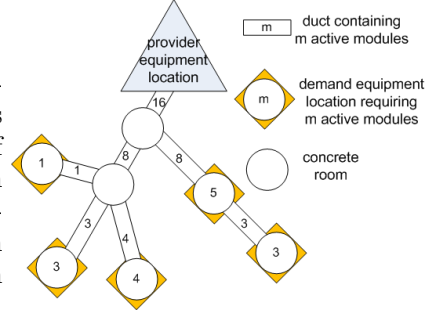


Figure 4.1: civil engineering and demand location example

In this problem, one has the possibility to:

- Use two different separation techniques: splicing and tapping. At most one of these operations can be done in a given node. The separation techniques conserve the number of active modules.
- Serve the demand in one of two different fashions: cable-served when one full cable is used, or module-served when modules from a cable are used.

Backfeed is not allowed, and the maintenance constraint does not apply.

4.1 Complexity

4.1.1 Proof of NP-completeness

Let us prove that the decision version of FCNDA is NP-complete even with 1 cable size and 2 potential splicing or tapping locations. We consider the Number Partitioning Problem (NPP), which is proven to be NP-complete in [34], and we prove that it can be reduced in polynomial time to a particular instance of FCNDA with an equivalent answer.

(NPP):

Instance: We have a set of N integers n_i ($\sum_{i \in \{1, \dots, N\}} n_i$ is supposed to be even).

Question: Is there a subset $S \subseteq \{1, \dots, N\}$ such that $\sum_{i \in S} n_i = \sum_{i \notin S} n_i$?

We consider an instance of (NPP) that we associate to the following FCNDA instance.

Let (V, A) be an arborescence describing the civil engineering structure ($V = \{r, 0, 1\} \cup \{v_i | i \in \{1, \dots, N\}\}$, $A = \{(r, 0); (0, 1); (1, v_i) | i \in \{1, \dots, N\}\}$); only one type of cable with a number of modules $M_1 = \frac{1}{2} \sum_{i \in \{1, \dots, N\}} n_i$ is available, its cost per length unit is $C_1^{le} = 1$. The length of all arcs of the arborescence are zero, except $(r, 0)$ which is of length 1. This means the cost of a cable created in r

is 1, and the cost of the other ones is 0. The number of active modules associated with each arc are: $m_{(r,0)}^{act} = m_{(0,1)}^{act} = \sum_{i=1}^N n_i; \forall i \in \{1, \dots, N\}, m_{(1,v_i)}^{act} = n_i$, which means that the demand points are the $v_i, i \in \{1, \dots, N\}$ and have respective demands n_i . This network is represented in Fig. 4.2a. We consider a zero cost for welding and welding boxes.

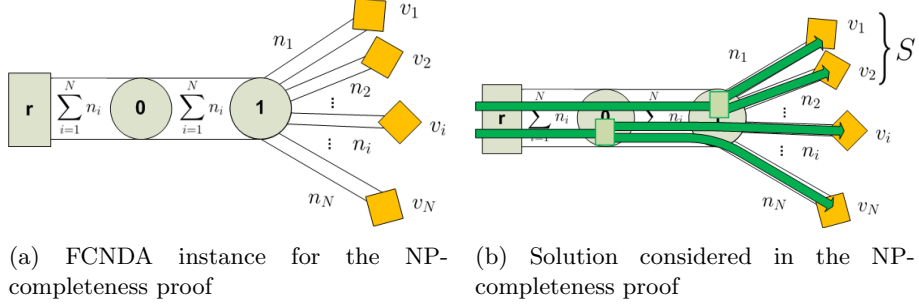


Figure 4.2: Illustration of the instance and solution considered for the NP-completeness proof

The question associated to this FCNDA instance is "Is there a cabling solution cheaper than 2?"

Let us first assume that (NPP) is feasible: $\exists S \subseteq \{1, \dots, N\}$ such that $\sum_{i \in S} n_i = \sum_{i \notin S} n_i$. We then build the following cabling solution:

- Two cables holding only active modules are installed on link $(r, 0)$.
- In node 0, one incoming cable is spliced into $N - |S|$ born cables. The born cables have a number of active modules $n_i, i \notin S$ and serve the demand nodes $(v_i)_{i \notin S}$.
- On link $(0, 1)$, one cable coming from r with only active modules, and $N - |S|$ cables serving demand nodes in $\{v_i | i \notin S\}$ are installed.
- In node 1, the incoming cable with only active modules is spliced into $|S|$ born cables. The born cables have n_i active modules and serve the demand nodes $(v_i)_{i \in S}$.
- One cable is installed on each link $(1, v_i)$.

Since the number of active modules is conserved in each splicing, the cabling solution described above is feasible (it is illustrated in Fig. 4.2b). Its cost is equal to 2, as the cables created in r have a cost of 1, and the other ones have a cost of 0.

Inversely, let us assume that (NPP) is not feasible: then, the solution described above is not possible anymore. One cable is not large enough to cover link $(r, 0)$, it cannot contain all the required active modules. Let us assume there is a solution with only two cables on $(r, 0)$. Since their combined number of modules is $\sum_{i \in \{1, \dots, N\}} n_i$, they both hold only active modules. If one

of them directly served the demand without enduring any operation, then the (NPP) instance was trivially feasible (one of the n_i is half the total sum). So both of them endure a splicing operation, one in node 0, the other in node 1. Let us consider the cables created in 1. They serve a subset S_1 of the demand nodes, and have a respective number of active modules of $n_i, i \in S_1$. Since the number of active modules in a splicing operation is conserved, we have $\sum_{i \in S_1} n_i = \frac{1}{2} \sum_{i \in \{1, \dots, N\}} n_i$ and the (NPP) instance was feasible.

Consequently, at least 3 cables need to be installed on arc $(r, 0)$, and such solution has a cost of at least 3.

4.1.2 Inapproximability

Proposition 4.1.1 *The FCNDA problem cannot be approximated within a ratio $\frac{3}{2} - \epsilon$ for any $\epsilon > 0$.*

Proof. Let us consider $\epsilon > 0$. Let us assume there is a polynomial time approximation algorithm \mathcal{A} which is able to give a solution with a ratio $\frac{3}{2} - \epsilon$. Let us consider an instance of (NPP).

Here is a polynomial time algorithm which can solve any instance of (NPP):

- Associate the FCNDA instance described in Section 4.1.1 to this (NPP) instance.
- Compute the cost of the solution given by the approximation algorithm \mathcal{A} .
- If this cost is higher or equal than 3, answer false. Otherwise, answer true.

If the cost c' of the solution given by \mathcal{A} verifies $c' < 3$, then the cost c^* of an optimal solution verifies $c^* < 3$. Since the cost of a solution is integer, $c^* \leq 2$. This implies (see Section 4.1.1) that the (NPP) instance is feasible. If the cost c' of the solution given by \mathcal{A} verifies $c' \geq 3$, the cost c^* of the optimal solution verifies $c^* \geq c' \frac{2}{3-2\epsilon} > 2$. Since c^* is integer, it implies $c^* \geq 3$, and the (NPP) instance is not feasible. Hence, this algorithm provides the answer of the (NPP) problem. \square

4.1.3 On the possible sets of decision variables

The models described later in Sections 4.2 and 4.3 use several types of variables. We can partly justify this choice by showing that a "minimalist" choice of variables leads to a formulation where one cannot assess the feasibility of a solution in a polynomial time with respect to the number of variables. Let us introduce $\mathcal{L} = \{1, \dots, L\}$ the set of available cables types of a FCNDA instance, V_D the set of demand nodes, and M_L the size of the largest cable. The number of modules weld in a node varies between 0 and M_L , while the number of cables going through an arc varies between 1 and $|V_D|$. Let us assume the FCNDA problem has a formulation using the set of variables $x = (x^{ca}, x^{box}, x^{we})$ described as follows:

- $\forall (i, j) \in A, \forall l \in \mathcal{L}, x_{(i,j),l}^{ca}$ denotes the number of cables of size l going through arc (i, j) . Its value varies between 0 and $|V_D|$
- $\forall i \in V^*, \forall l \in \mathcal{L}, x_{i,l}^{box}$ is the binary variable equal to 1 iff there is a welding box for a cable of size l installed in node i .
- $\forall i \in V^*, \forall m \in \{1, \dots, M_L\}, x_{i,m}^{we}$ is the binary variable equal to 1 iff there are m welds performed in node i .

These variables seem required to express the objective function. For a given instance I of the problem, $|I|$ denotes the instance size.

Proposition 4.1.2 *There is no algorithm \mathcal{F} and polynomial P such that for every instance I of the problem, for any value of x , $\mathcal{F}(I, x)$ answers the question "is x feasible?" (is there a feasible solution of variables x ?) in a time smaller than $P(|I|)$, unless $\mathcal{P} = \mathcal{NP}$.*

Proof. Let us assume we have such an algorithm. We then propose a polynomial time algorithm able to solve all instances of (NPP). Let us consider I^{NPP} an instance of this problem.

- First, associate to the NPP instance the instance of the FCNDA problem described in Section 4.1.1, which we note I^{FCNDA} , with the same notation. We have a polynomial Q such that $|I^{FCNDA}| \leq Q(|I^{NPP}|)$.
- Second, look for the feasibility of all solutions with two cables on the arc $(r, 0)$. This can be done using algorithm \mathcal{F} by testing the sets of variables with values $x_{(r,0),1}^{ca} = 2$; $x_{(1,v_i),1}^{ca} = 1$ for all $i \in \{1, \dots, N\}$; $x_{0,1}^{box} = x_{1,1}^{box} = 1$; for $m = \frac{1}{2} \sum_{i \in \{1, \dots, N\}} n_i, x_{0,m}^{we} = x_{1,m}^{we} = 1$; for all $m' \neq m, x_{0,m'}^{we} = x_{1,m'}^{we} = 0$. The only variable for which all possible values are tested is $x_{(0,1),1}^{ca}$. Feasibility of these solutions can be verified in $\mathcal{O}(|I^{FCNDA}|P(|I^{FCNDA}|)) = \mathcal{O}(Q(|I^{NPP}|)P(Q(|I^{NPP}|)))$.
- If one of these solutions is feasible, then declare I^{NPP} feasible.

This algorithm is valid since one of the tested solutions is feasible iff I^{NPP} is feasible. Indeed, let us assume the I^{NPP} instance is feasible. Then, there is a solution to I^{FCNDA} with two cables coming out of the root (see Section 4.1.1). This implies that one of the tested values for x is feasible. Inversely, let us assume at least one of the tested values of x are feasible. Then, I^{FCNDA} has at least one solution with two cables coming out of the root. This implies, by Section 4.1.1, that I^{NPP} is feasible. Hence the result. \square

The link between this proposition and the formulations one can expect is the following. In most formulations used to tackle NP-hard problems, one implicitly assumes that given a set of variable values, one can assess its feasibility in polynomial time. Indeed:

- Any ILP formulation of the kind $Ax \leq b, x \in \mathbb{Z}$ where A is a matrix and b is a vector of appropriate size, with constraints separable in polynomial time, gives a straightforward feasibility verification algorithm: one can check whether the constraints are violated or not.
- Let us consider harder formulations, such as quadratically constrained programming. Any constraint of the form $x^T Qx + a^T x \leq b$ can be checked in polynomial time for a given x . Which implies no formulation of this kind with a polynomial number of constraints can be expected. Similarly, other non-linearly constrained formulations can be excluded as well.

We do not intend to give an exhaustive list of the formulations used in the literature. However, these elements are enough to suggest that using the variables x as defined above is a very hard route to follow. We deduce from Proposition 4.1.2 that an ILP formulation using the variables above will contain constraints that are difficult to separate. This leads us to add more variables and propose the next two formulations of Sections 4.2 and 4.3.

4.2 Path-based integer programming formulation

4.2.1 Notation

An instance of the FCNDA problem is noted as follows. An arborescence $G = (V, A)$ describes the selected civil engineering structure where V denotes the set of concrete rooms (where cable separation can be performed) and A represents the set of ducts. Let r be the root of G and let $V^* = V \setminus \{r\}$. The set of paths in G is denoted by \mathcal{P} . Given any path $p \in \mathcal{P}$, $s(p)$ is the first node of p (source) and $t(p)$ is the last node (target). Each node $i \in V$ has a demand of D_i active modules ($D_i \geq 0$). $V_D \subset V^*$ denotes the set of nodes with a non-zero demand, while $V_N \subset V^*$ denotes the set of nodes with zero demand. Each arc $(i, j) \in A$ has a length $\Delta_{(i,j)} > 0$. For $i \in V$, we denote by $\Gamma^+(i)$ its set of successors $\{j \in V \mid (i, j) \in A\}$, and for $i \in V^*$, let $\gamma(i)$ be the unique predecessor of i . The number of active modules going through $(i, j) \in A$, denoted by $m_{(i,j)}^{act}$, can be calculated recursively starting with pendant vertices i (leaf nodes) by assigning D_i to $m_{(\gamma(i),i)}^{act}$ and writing that $m_{(i,j)}^{act} = \sum_{j' \in \Gamma^+(j)} m_{(j,j')}^{act} + D_j$.

The available cables are from a discrete set $\mathcal{L} = \{1, \dots, L\}$, each cable type being characterized by a number of modules M_l with $l \in \{1, \dots, L\}$. \mathcal{L} is ordered with respect to M , i.e. $\forall (l, l') \in \mathcal{L}^2, l < l' \iff M_l < M_{l'}$. The set of possible numbers of active modules inside a cable of type l is denoted by $\mathcal{M}_l = \{1, \dots, M_l\}$. Each cable has a cost per length unit C_l^{le} , while the cost of a welding box of type l is PB_l (recall that the size of the welding box depends of the size of the separated cable), and the cost for welding m modules is PW_m , all costs being strictly positive and increasing with respect to size.

The decision variables of the model are described below:

- $\forall p \in \mathcal{P}, \forall l \in \mathcal{L}, b_{p,l}^{born,sep} \in \{0, 1\}$: the binary variable equal to 1 iff there is a cable going through path p , born in $s(p)$ and separated in $t(p)$ (either spliced or tapped). Recall that since there is at most one cable that can be separated in a concrete room, $b_{p,l}^{born,sep}$ is binary. Also recall that $s(p)$ can be equal to the root r of G .
- $\forall p \in \mathcal{P}, \forall l \in \mathcal{L}, b_{p,l}^{ctn,sep} \in \{0, 1\}$: the binary variable equal to 1 iff there is a cable going through path p , which is the continuation of a cable tapped in $s(p)$ and separated in $t(p)$. Observe that $s(p) \neq r$.
- $\forall p \in \mathcal{P}$ such that $t(p) \in V_D, \forall l \in \mathcal{L}, b_{p,l}^{born,dem} \in \{0, 1\}$: the binary variable equal to 1 iff the demand of node $t(p)$ is served by a cable on path p in a cable-served way, and this cable is a cable born in $s(p)$. Its number of active modules is then $D_{t(p)}$.
- $\forall p \in \mathcal{P}$ such that $t(p) \in V_D, \forall l \in \mathcal{L}, b_{p,l}^{ctn,dem} \in \{0, 1\}$: the binary variable equal to 1 iff the demand of node $t(p)$ is served by a cable on path p in a cable-served way, and this cable is the continuation of a cable tapped in $s(p)$. Its number of active modules is then $D_{t(p)}$.
- $\forall p \in \mathcal{P}, m_p^{born,sep} \in \{0, \dots, M_L\}$: the number of active modules of the cable going through path p , born in $s(p)$ and either spliced or tapped in $t(p)$. Notice that such a cable cannot serve a demand in a cable-served way but might serve it in a module-served way.
- $\forall p \in \mathcal{P}, m_p^{ctn,sep} \in \{0, \dots, M_L - 1\}$: the number of active modules of the cable going through path p , continuation of a cable tapped in $s(p)$ and either spliced or tapped in $t(p)$. Observe again that such a cable cannot serve a demand in a cable-served way but might serve it in a module-served way.
- $\forall i \in V^*, \forall m \in \mathcal{M}_L, w_{i,m} \in \{0, 1\}$: the binary variable equal to 1 iff there are m modules to be weld in node i .

For the sake of clarity, a description of a small cabling solution by the model is given in Fig. 4.3.

4.2.2 Formulation

The FCNDA problem can be formulated as follows

$$\begin{aligned}
\min \quad & \sum_{p \in \mathcal{P}} \sum_{(i,j) \in \mathcal{P}} \sum_{l \in \mathcal{L}} C_l^{le} \cdot \Delta_{(i,j)} \cdot (b_{p,l}^{born,sep} + b_{p,l}^{born,dem} + b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem}) \\
& + \sum_{i \in V^*} \sum_{m \in \mathcal{M}_L} PW_m \cdot w_{i,m} + \sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} PB_l \cdot (b_{p,l}^{born,sep} + b_{p,l}^{ctn,sep}) \quad (4.1)
\end{aligned}$$

such that

$$\begin{aligned}
& \sum_{p \in \mathcal{P} | t(p)=i} (m_p^{born,sep} + m_p^{ctn,sep}) = \\
& \sum_{p \in \mathcal{P} | s(p)=i} (m_p^{born,sep} + m_p^{ctn,sep}) \\
& + \sum_{p \in \mathcal{P} | s(p)=i, t(p) \in V_D} \sum_{l \in \mathcal{L}} D_{t(p)} \cdot (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) \quad \forall i \in V_N, \quad (4.2)
\end{aligned}$$

$$\begin{aligned}
& \sum_{p \in \mathcal{P} | t(p)=i} (m_p^{born,sep} + m_p^{ctn,sep}) \\
& + \sum_{p \in \mathcal{P} | t(p)=i} \sum_{l \in \mathcal{L}} D_i \cdot (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) = D_i \\
& + \sum_{p \in \mathcal{P} | s(p)=i} (m_p^{born,sep} + m_p^{ctn,sep}) \\
& + \sum_{p \in \mathcal{P} | s(p)=i, t(p) \in V_D} \sum_{l \in \mathcal{L}} D_{t(p)} \cdot (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) \quad \forall i \in V_D, \quad (4.3)
\end{aligned}$$

$$\sum_{l \in \mathcal{L}} M_l \cdot b_{p,l}^{born,sep} \geq m_p^{born,sep} \quad \forall p \in \mathcal{P}, \quad (4.4)$$

$$\sum_{l \in \mathcal{L}} (M_l - 1) \cdot b_{p,l}^{ctn,sep} \geq m_p^{ctn,sep} \quad \forall p \in \mathcal{P}, \quad (4.5)$$

$$\begin{aligned}
& \sum_{p \in \mathcal{P} | s(p)=i} b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem} \leq \\
& \sum_{p \in \mathcal{P} | t(p)=i} b_{p,l}^{born,sep} + b_{p,l}^{ctn,sep} \quad \forall i \in V^*, l \in \mathcal{L}, \quad (4.6)
\end{aligned}$$

$$\sum_{p \in \mathcal{P} | t(p)=i} \sum_{l \in \mathcal{L}} b_{p,l}^{born,sep} + b_{p,l}^{ctn,sep} \leq 1 \quad \forall i \in V^*, \quad (4.7)$$

$$\sum_{l \in \mathcal{L}} \sum_{p \in \mathcal{P} | t(p)=i} b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem} \leq 1 \quad \forall i \in V_D, \quad (4.8)$$

$$\begin{aligned}
& \sum_{m \in \mathcal{M}_L} m \cdot w_{i,m} = \sum_{p \in \mathcal{P} | i=s(p)} m_p^{born,sep} \\
& + \sum_{p \in \mathcal{P} | s(p)=i, t(p) \in V_D} \sum_{l \in \mathcal{L}} D_{t(p)} \cdot b_{p,l}^{born,dem} \quad \forall i \in V^*, \quad (4.9)
\end{aligned}$$

$$\sum_{m \in \mathcal{M}_L} w_{i,m} \leq 1 \quad \forall i \in V^*, \quad (4.10)$$

$$\begin{aligned}
& b_{p,l}^{born,sep}, b_{p,l}^{ctn,sep}, b_{p,l}^{born,dem}, b_{p,l}^{ctn,dem} \in \{0, 1\} \quad \forall p \in \mathcal{P}, \forall l \in \mathcal{L}, \\
& w_{i,m} \in \{0, 1\} \quad \forall i \in V^*, \forall m \in \mathcal{M}_L, \\
& m_p^{born,sep} \in \{0, \dots, M_L\}, m_p^{ctn,sep} \in \{0, \dots, M_L - 1\} \quad \forall p \in \mathcal{P}.
\end{aligned}$$

The cost function (4.1) is a combination of cable costs (first term), weld costs (second term) and welding box costs (third term).

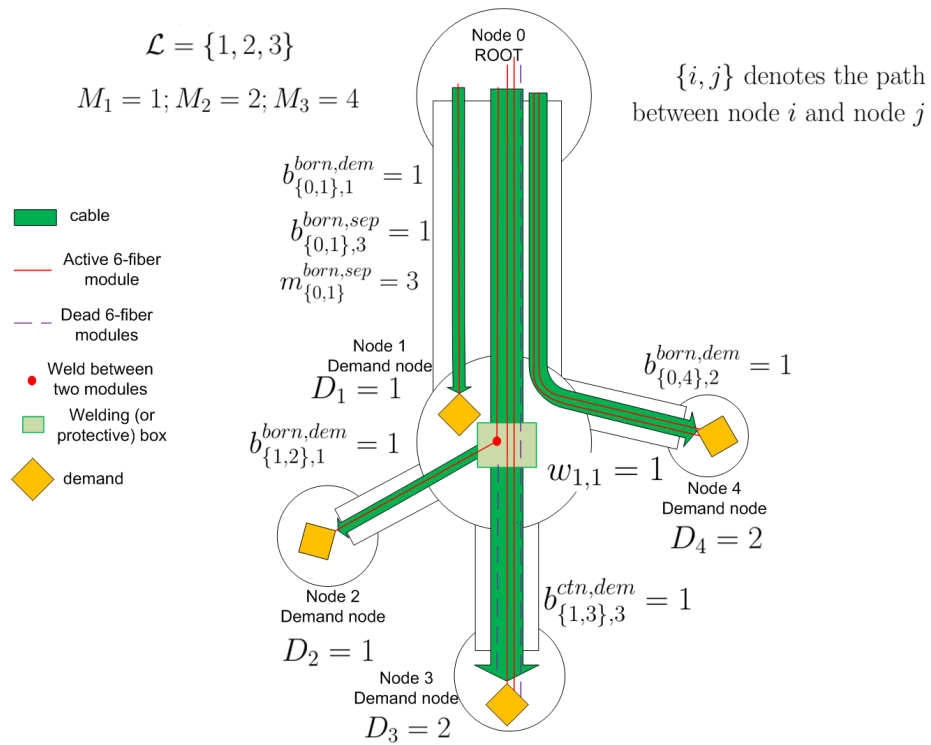


Figure 4.3: Path model illustration on a small example; a tapping is performed in node 1

Constraints (4.2) and (4.3) describe modules in separation points. Equations (4.2) ensure active modules conservation in no-demand separation points. Recall that the term $\sum_{p \in \mathcal{P} | s(p)=i, t(p) \in V_D} \sum_{l \in \mathcal{L}} D_{t(p)} \cdot b_{p,l}^{born,dem}$ is the number of active modules of an eventual cable born in i and serving $t(p)$, and $\sum_{p \in \mathcal{P} | s(p)=i, t(p) \in V_D} \sum_{l \in \mathcal{L}} D_{t(p)} \cdot b_{p,l}^{ctn,dem}$ is the number of active modules of an eventual cable continued from a tapping in i and serving $t(p)$. Equations (4.3) ensure active modules conservation in separation points with demand. In the case we have a cable-served node, $\sum_{p \in \mathcal{P} | t(p)=i} \sum_{l \in \mathcal{L}} D_i \cdot (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) = D_i$, and the equation regarding an eventual separation point is equivalent to (4.2). Otherwise, $\sum_{p \in \mathcal{P} | t(p)=i} \sum_{l \in \mathcal{L}} (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) = 0$, and we can see that the separated cable has the number of active modules required to both serve the demand and create the created cables. Note that the variables do not allow the number of modules in a cable to change unless a separation point is encountered, and that all cables serving the demand in a cable-served way have by definition the proper number of modules ($\sum_{p \in \mathcal{P} | s(p)=i, t(p) \in V_D} \sum_{l \in \mathcal{L}} D_{t(p)} \cdot b_{p,l}^{born,dem}$ or $\sum_{p \in \mathcal{P} | s(p)=i, t(p) \in V_D} \sum_{l \in \mathcal{L}} D_{t(p)} \cdot b_{p,l}^{ctn,dem}$). Hence, ensuring that modules are conserved in separation points with (4.2) and (4.3) is enough to ensure that modules are conserved in any node of the network. Equations (4.4) and (4.5) make sure that cables are large enough to hold their number of active modules. Recall that in the continuation of a tapped cable, there is at least one dead module. Constraints (4.6) define tapping operations properly. Whenever the continuation of a tapped cable of size l appears in i , there is a cable of the same size separated in i . Constraints (4.7) ensure that no more than one cable is either spliced or tapped in a node. It is valid in both V_N and V_D , since a cable of type $l \in \mathcal{L}$ on path $p \in \mathcal{P}$ serving a node in a cable-served way would be denoted either by $b_{p,l}^{born,dem}$ or by $b_{p,l}^{ctn,dem}$. Equations (4.8) ensure that demand nodes are served by at most one cable. Equations (4.9) and (4.10) make sure that the variable $w_{i,m}$ is equal to 1 iff there are m welds performed in node i .

Remark 4.2.1 *This ILP model holds $4|\mathcal{P}| \cdot L + M_L \cdot |V^*|$ binary variables, $2|\mathcal{P}|$ integer variables, and $(4 + L) \cdot |V^*| + |V_D| + 2|\mathcal{P}|$ constraints. The number of paths in an arborescence of n nodes varies between $n - 1$ (star graph, of the form (V, A) with $V = \{v_1, \dots, v_n\}$ and $A = \{(v_1, v_i) | i \in \{2, \dots, n\}\}$) and $\frac{n(n-1)}{2}$ (linear graph or path graph, of the form (V, A) with $V = \{v_1, \dots, v_n\}$ and $A = \{(v_{i-1}, v_i) | i \in \{2, \dots, n\}\}$).*

It is possible to set the value of some variables of this model by pre-processing the data.

4.2.3 Preprocessing

By definition, no continuation of a tapped cable comes from the root. Hence

$$\forall p \in \mathcal{P} \text{ such that } s(p) = r, \forall l \in \mathcal{L}, \quad b_{p,l}^{ctn,sep} = b_{p,l}^{ctn,dem} = 0$$

Demand nodes can only be served by one cable. This implies that if it is served in a cable-served way, the number of modules of the serving cable cannot be smaller than the required number of active modules. Besides, if the cable serving a node is the continuation of a tapped cable, it must have at least one dead module. Hence

$$\begin{aligned} \forall i \in V_D, \forall l \in \mathcal{L}, \text{ such that } D_i > M_l, \quad \sum_{p \in \mathcal{P} | t(p)=i} b_{p,l}^{born,dem} &= 0 \\ \forall i \in V_D, \forall l \in \mathcal{L}, \text{ such that } D_i + 1 > M_l, \quad \sum_{p \in \mathcal{P} | t(p)=i} b_{p,l}^{ctn,dem} &= 0 \end{aligned}$$

Similarly, if a demand node is served in a cable-served way by a born cable, then the size of this cable must be optimal with respect to the demand it serves. Indeed, if there exists a smaller cable size which is large enough to serve the demand, one could have chosen this smaller size instead, which gives a cheaper solution. This leads to

$$\forall p \in \mathcal{P} \text{ s.t. } t(p) \in V_D, \forall l \in \mathcal{L} \setminus \{1\}, \text{ if } M_{l-1} \geq D_{t(p)}, \text{ then } b_{p,l}^{born,dem} = 0$$

First, recall that cable-served demand nodes can only be served by one cable, holding exactly as many active modules as required. Furthermore, when some of these enter a separation point, they gather their active modules into one cable, its number of active modules being a sum of some of the demands downstream. Recursively, this means that all cables hold a number of active modules which is a sum of the demand in some demand nodes. Since the number of welds performed in a node is the sum of the number of active modules of cables born in this node, it is also a sum of some of the demands downstream. Hence, some values for this number are known to be impossible. Let us note $\forall i \in V, V_i^{ar}$ the arborescence rooted in i , i excluded, and $PM(i) := \{\sum_{j \in J} D_j | J \subseteq V_i^{ar}\}$ (PM stands for possible number of active modules).

$$\forall i \in V^*, \forall m \in \mathcal{M}_L \setminus PM(i), w_{i,m} = 0$$

The following propositions use some properties of the welding cost function PW .

Lemma 4.2.1 *From the concavity of PW and $PW_0 = 0$, we can derive:*

PW is subadditive,

$$\forall m_0 \in \mathcal{M}_L, \forall m_1 \leq m_0, \forall m_2 \leq m_1, PW_{m_0} - PW_{m_0-m_1} \geq PW_{M_L} - PW_{M_L-m_2}.$$

Proof. First, for m and m' in \mathcal{M}_L such that $m + m' \in \mathcal{M}_L$, by concavity, we have $PW_m \geq \frac{m}{m+m'} PW_{m+m'} + \frac{m'}{m+m'} PW_0$ and $PW_{m'} \geq \frac{m'}{m+m'} PW_{m+m'} + \frac{m}{m+m'} PW_0$. Summing gives the result.

Let us now consider $m_2 \leq m_1 \leq m_0 \in \mathcal{M}_L$. Since PW is growing, $PW_{M_L} - PW_{M_L-m_1} \geq PW_{M_L} - PW_{M_L-m_2}$. Besides, the concavity of PW gives $PW_{M_L-m_1} \geq \frac{m_1}{M_L+m_1-m_0} PW_{m_0-m_1} + \frac{M_L-m_0}{M_L+m_1-m_0} PW_{M_L}$ and $PW_{m_0} \geq \frac{M_L-m_0}{M_L+m_1-m_0} PW_{m_0-m_1} + \frac{m_1}{M_L+m_1-m_0} PW_{M_L}$. Summing gives the result. \square

The following proposition deals with cable-served nodes and the path on which the cable serving it lies. We can determine the cost of serving it by a continued cable on this path ($\sum_{(i,j) \in p} \Delta_{(i,j)} \cdot C_l^{le}$) and compare it to the cost of a born cable ($\sum_{(i,j) \in p} \Delta_{(i,j)} \cdot C_{l'}^{le} + PW_m$). If serving it by the continuation of a tapped cable is more expensive than serving with a born cable, the born cable will always be chosen, which makes sure no continuation of a cable tapped in $s(p)$ will serve it.

Proposition 4.2.2 *Let us consider $p \in \mathcal{P}$ such that $t(p) \in V_D$, $s(p) \neq r$, and let m be $D_{t(p)}$. Let consider $l \in \mathcal{L} \setminus \{1\}$ and assume $\exists l' < l$ such that $\sum_{(i,j) \in p} \Delta_{(i,j)} \cdot (C_l^{le} - C_{l'}^{le}) > PW_m$ and $M_{l'} \geq m$. Every optimal solution of the FCNDA problem verifies $b_{p,l}^{ctn,dem} = 0$.*

Proof. Let us consider an optimal solution of the FCNDA problem S , $p \in \mathcal{P}$ such that $t(p) \in V_D$, $s(p) \neq r$, $s(p) \neq r$, and $l \in \mathcal{L} \setminus \{1\}$. Let us note $m = D_{t(p)}$, and assume $\exists l' < l$ such that $\sum_{(i,j) \in p} \Delta_{(i,j)} \cdot (C_l^{le} - C_{l'}^{le}) > PW_m$, $M_{l'} \geq m$.

Let us assume $b_{p,l}^{ctn,dem} = 1$, $t(p)$ is served by the continuation of a cable tapped in $s(p)$. If $\forall m' \in \mathcal{M}_L$, $w_{s(p),m'} = 0$ (the case where no welds are performed in $s(p)$), let us consider the solution S' identical to S everywhere except for $w_{s(p),m} = 1$, $b_{p,l'}^{born,dem} = 1$ and $b_{p,l}^{ctn,dem} = 0$. S' is cheaper than S and feasible. Otherwise, $\exists m_0 \in \mathcal{M}_L$ such that $w_{s(p),m_0} = 1$. Besides, with (4.2) and (4.7), $m_0 \leq M_L - m$ (by counting the number of active modules, the number of welds performed in $s(p)$ cannot be larger than $M_L - m$). Let us consider the solution S' identical to S everywhere except for $b_{p,l}^{ctn,dem} = 0$, $b_{p,l'}^{born,dem} = 1$, $w'_{s(p),m_0+m} = 1$, $w_{s(p),m_0} = 0$. S' is the solution where a splicing is performed in $s(p)$ and a born cable of size l' is deployed on p . It is cheaper and feasible. Indeed, the difference δ between the cost of S and the cost of S' is $\delta = \sum_{(i,j) \in p} \Delta_{(i,j)} \cdot (C_l^{le} - C_{l'}^{le}) + PW_{m_0} - PW_{m_0+m}$, and by the subadditivity of PW , $\delta \geq \sum_{(i,j) \in p} \Delta_{(i,j)} \cdot (C_l^{le} - C_{l'}^{le}) - PW_m > 0$. Hence the result. \square

By considering a born cable not coming from the root, it is possible to determine the cost for creating this cable at the root. If it is cheaper to create it at the root than creating it downstream in a separation point, then in any optimal solution, this cable will be started at the root.

Proposition 4.2.3 *Let us consider $p \in \mathcal{P}$ such that $s(p) \neq r$ and let us note p' the only path such that $t(p') = t(p)$ and $s(p') = r$. For any $l \in \mathcal{L} \setminus \{1\}$, let m be $M_{l-1} + 1$. If $\sum_{(i,j) \in p' \setminus p} C_l^{le} \cdot \Delta_{(i,j)} < PW_{M_L} - PW_{M_L-m}$, then every optimal solution of the FCNDA problem verifies $b_{p,l}^{born,sep} = b_{p,l}^{born,dem} = 0$.*

Proof. Let S be an optimal solution. Let us consider $p \in \mathcal{P}$ such that $s(p) \neq r$ and let us note p' the only path such that $t(p') = t(p)$ and $s(p') = r$. Let us consider $l \in \mathcal{L} \setminus \{1\}$, and let us denote $m = M_{l-1} + 1$. Furthermore, let us assume $\sum_{(i,j) \in p' \setminus p} C_l^{le} \cdot \Delta_{(i,j)} < PW_{M_L} - PW_{M_L-m}$ and $b_{p,l}^{born,sep} = 1$ (there is a born cable on path p). We know that $m_1 = m_p^{born,sep} \geq M_{l-1} + 1$ (see also

(4.13)), which means, by (4.9) and (4.10), that $\exists! m_0 \geq m_p^{born,sep}, w_{s(p),m_0} = 1$ (m_0 denotes the number of welds performed in $s(p)$). It also means by (4.2) and (4.7) that $\exists! p_0 \in \mathcal{P}, m_{p_0}^{born,sep} + m_{p_0}^{ctn,sep} \geq m_0$ (p_0 is the path on which the cable separated in $s(p)$ lies).

If $m_{p_0}^{ctn,sep} > 0$ (the cable on p_0 is the continuation of a tapped cable). Then, $m_{p_0}^{born,sep} = 0$ and the solution S' identical to S everywhere except for $m_{p_0}^{ctn,sep} = m_{p_0}^{ctn,sep} - m_1, b_{p,l}^{born,sep} = 0, b_{p',l}^{born,sep} = 1, m_{p'}^{born,sep} = m_p^{born,sep}, w'_{s(p),m_0} = 0$ and $w'_{s(p),m_0-m_1} = 1$ is cheaper and feasible. S' is the solution where instead of installing a cable on p , one is installed on p' , which increases cable costs but reduces welds costs. Indeed, the difference D between the cost of S and the cost of S' is $D = PW_{m_0} - PW_{m_0-m_1} - \sum_{(i,j) \in p' \setminus p} C_l^{le} \cdot \Delta_{(i,j)}$ and by concavity of PW (see Lemma 4.2.1), $D \geq PW_{M_L} - PW_{M_L-m} - \sum_{(i,j) \in p' \setminus p} C_l^{le} \cdot \Delta_{(i,j)} > 0$. If $m_{p_0}^{born,sep} > 0$ (the cable on p_0 is a born cable). Then, $m_{p_0}^{ctn,sep} = 0$ and the solution identical to S everywhere except for $m_{p_0}^{born,sep} = m_{p_0}^{born,sep} - m_1, b_{p,l}^{born,sep} = 0, b_{p',l}^{born,sep} = 1, m_{p'}^{born,sep} = m_p^{born,sep}, w'_{s(p),m_0} = 0$ and $w'_{s(p),m_0-m_1} = 1$ is cheaper and feasible (the difference between the costs of S and S' is the same as in the previous case).

A similar reasoning can be conducted in the case $b_{p,l}^{born,dem} = 1$. Hence the result. \square

4.2.4 Valid inequalities

As will be shown in Table 4.3, the continuous relaxation of the ILP of Section 4.2.2 is not tight enough. Hereinafter we provide valid inequalities aiming to strengthen the continuous relaxation. The practical effectiveness of each family of valid inequalities is assessed in Table 4.5 through numerical experiments.

The following cuts express that any demand node i will be served by at least one cable, the size of which must be able to hold the demand.

Proposition 4.2.4 *Every optimal solution of the FCNDA problem verifies*

$$\forall i \in V_D, \sum_{p \in \mathcal{P} | t(p) = i} \left(\sum_{l \in \mathcal{L} | M_l \geq D_i} b_{p,l}^{born,dem} + \sum_{l \in \mathcal{L} | M_l \geq D_i + 1} (b_{p,l}^{born,sep} + b_{p,l}^{ctn,dem}) + \sum_{l \in \mathcal{L} | M_l \geq D_i + 2} b_{p,l}^{ctn,sep} \right) \geq 1 \quad (4.11)$$

Four different cases can occur on a given node i . First, if the node is served in a cable-served way by a born cable, this cable must contain at least D_i active modules. Second, if it is served in a cable-served way by a continued cable, then it has at least one dead module in addition to D_i active modules. Third, when the node is module-served by a born cable, then this cable has at least one more active module than what the demand requires, since new cables are created. Finally, if i is module-served by the continuation of a tapped cable, then this

cable contains at least one dead module, and at least one active module more than D_i .

The inequalities below state that at least two active modules must be present in any separated cable of an optimal solution. Indeed, any cable without active modules can be removed to give a cheaper feasible solution. Any cable with only one active module serves the demand.

Proposition 4.2.5 *Every optimal solution of the FCNDA problem satisfies*

$$\forall p \in \mathcal{P}, 2 \cdot \sum_{l \in \mathcal{L}} b_{p,l}^{ctn,sep} \leq m_p^{ctn,sep} \quad (4.12)$$

Proof. Let us assume that in an optimal solution S of the FCNDA problem, we have a path $p_1 \in \mathcal{P}$ such that $m_{p_1}^{ctn,sep} = 1$ and $\sum_{l \in \mathcal{L}} b_{p_1,l}^{ctn,sep} = 1$. This means there is a tapping done in the node $s(p_1)$, which is not the root, and let us note $m_1 \leq M_L - 1$ the number of welds done in $s(p_1)$. Our hypothesis also means that there is a cable with only one active module getting separated in $t(p_1)$ on path p_1 . Let us call p_2 the path on which continues the only cable created in $t(p_1)$. It is a born cable, since doing a tapping with only the continued cable coming out is impossible in an optimal solution (it would mean a protective box is installed for nothing). We then have $t(p_1) = s(p_2)$, and let us call $p_0 \in \mathcal{P}$ the only path such that $s(p_0) = s(p_1)$ and $t(p_0) = t(p_2)$. Let us consider the solution S' identical to S everywhere, but where we replace the two cables on p_1 and p_2 by a cable of size $l = 1$ on the path p_0 , born in $s(p_0)$. The variables values of S' different from those of S are:

- $w'_{s(p_0),m_1+1} = 1$, $w'_{s(p_0),m_1} = 0$ (we "shift" the weld that used to be done in $t(p_1)$ to $s(p_1)$)
- $\forall l \in \mathcal{L}, b_{p_1,l}^{ctn,sep'} = 0$; $m_{p_1}^{ctn,sep'} = 0$ (there is no more cable continued from a tapping done in $s(p_1)$ on the path p_1)
- $b_{p_0,1}^{born,dem'} = \sum_{l \in \mathcal{L}} b_{p_2,l}^{born,dem}$; $b_{p_0,1}^{born,sep'} = \sum_{l \in \mathcal{L}} b_{p_2,l}^{born,sep}$ (the cable on p_0 in S' has the same role than the born cable on p_2 in S , one of the sums is equal to 1).
- $m_{p_0}^{born,dem'} = m_{p_2}^{born,dem}$; $m_{p_0}^{born,sep'} = m_{p_2}^{born,sep}$ (the cable on p_0 in S' has the same role than the born cable on p_2 in S , one of the sums is equal to 1).

Since we use the smallest cable size, the overall cost of cables of S' is smaller than the one of S . Furthermore, by the concavity of PW , doing $m_1 + 1$ welds in $s(p_1)$ is cheaper than doing m_1 welds in $s(p_1)$ and one weld in $t(p_1)$. Which means S' is cheaper than S . Hence the result. \square

The inequalities below deal with the relation between the size of a born cable and its number of active modules.

Proposition 4.2.6 *Every optimal solution of the FCNDA problem verifies*

$$\forall p \in \mathcal{P}, m_p^{born,sep} \geq \sum_{l \in \mathcal{L}} b_{p,l}^{born,sep} + \sum_{l \in \mathcal{L} \setminus \{1\}} M_{l-1} \cdot b_{p,l}^{born,sep} \quad (4.13)$$

The intuition is that born cables have an optimal size regarding the number of active modules they hold: if a cable of size l is created, it must contain at least $M_{l-1} + 1$ active modules; otherwise, one can replace this cable with a smaller one of size $l - 1$, leading to a solution of lower cost.

Let us now focus on no-demand nodes. A cable is separated iff there are welds in this node. Note that this inequality cannot be generalized to V_D . Indeed, it is possible to have a tapping in a module-served demand node without creating any born cables, which means no welds are performed.

Proposition 4.2.7 *Every optimal solution of the FCNDA problem satisfies*

$$\forall i \in V_N, \sum_{m \in \mathcal{M}_L} w_{i,m} = \sum_{p \in \mathcal{P} | t(p)=i} \sum_{l \in \mathcal{L}} (b_{p,l}^{born,sep} + b_{p,l}^{ctn,sep}) \quad (4.14)$$

The following proposition asserts that if $m \leq m'$ welds are performed in a given node, then there will be born cables going out of this node with at most m active modules. Since these are born cables, their size cannot exceed the smallest size able to contain m' active modules (denoted by l_1).

Proposition 4.2.8 *Every optimal solution of the FCNDA problem verifies*

$$\forall i \in V^*, \forall m' \in \mathcal{M}_L, \text{ if } l_1 = \min\{l \in \mathcal{L} | M_l \geq m'\}, \text{ then} \\ \sum_{m \in \{1, \dots, m'\}} w_{i,m} \leq \sum_{p \in \mathcal{P} | s(p)=i} \sum_{l \in \mathcal{L} | l \leq l_1} (b_{p,l}^{born,dem} + b_{p,l}^{born,sep}) \quad (4.15)$$

Let us focus on degree 2 nodes. What follows implies that in a sequence of adjacent nodes of degree 2 without demand, a separation point will be preferably made either in the last node of the sequence, or upstream of the sequence.

Proposition 4.2.9 *Every optimal solution of the FCNDA problem satisfies*

$$\forall (i, j) \in A, \text{ such that } i \in V_N \text{ and } |\Gamma^+(i)| = 1, \sum_{m \in \mathcal{M}_L} w_{j,m} \geq \sum_{m \in \mathcal{M}_L} w_{i,m} \quad (4.16)$$

Proof. Let us assume there is a separation point in a degree 2 no-demand node i and no separation point in its successor j (Fig. 4.4 center).

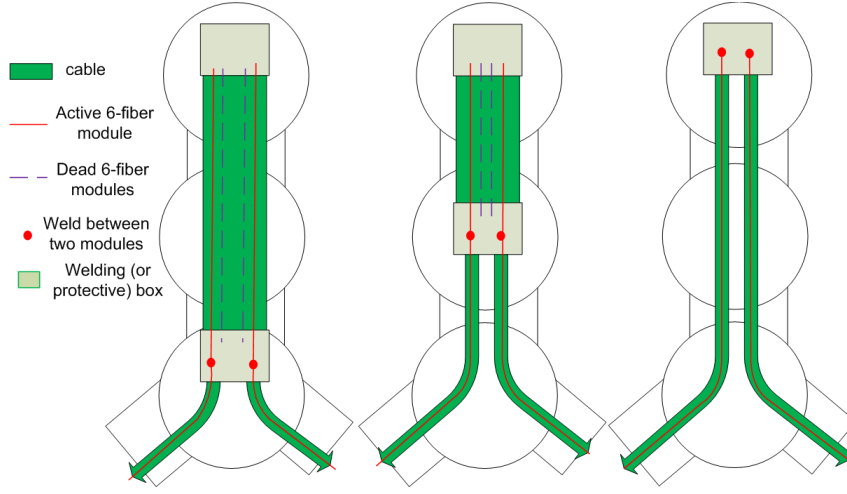


Figure 4.4: illustration of the proof of (4.16)

Let us distinguish two cases.

- If the sum of the cost per length unit of cables created in i , equal to $\sum_{p \in \mathcal{P} | s(p)=i} \sum_{l \in \mathcal{L}} C_l^{le} \cdot (b_{p,l}^{born,sep} + b_{p,l}^{ctn,sep} + b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem})$, is larger than the cost per length unit of the cable separated in i , then one can "shift" the separation point downstream to j and find a cheaper solution (Fig. 4.4 left).

- Let us consider the case where the sum of the costs per length unit of all created cables is smaller than the cost per length unit of the separated cable. First, note that this isn't possible if a tapping was performed in i . Indeed, in a tapping, the continued cable is as expensive as the tapped cable, and there are additional born cables. So we have a splicing operation in i . Let us consider the last separation point encountered by this spliced cable (upstream of i). Let us consider the solution where, instead of creating this cable (whether it was born or not), we create the downstream cables (Fig. 4.4 right). First, note that in this new solution, the overall cable cost is lower by hypothesis. Then, note that the welds that were in i are performed in another node instead. Since PW is subadditive, the overall price of welds is then lower. This gives a cheaper feasible solution. Hence the result. \square

The inequalities below state that if there are $m \geq m'$ welds performed in a given node, then there is a separated cable containing at least m' active modules. Furthermore, if this cable has been tapped, it contains at least an additional dead module.

Proposition 4.2.10 *The following constraints are valid inequalities for the FC-NDA problem.*

$$\forall i \in V^*, \forall m' \in \mathcal{M}_L, \sum_{m \geq m'} w_{i,m} \leq$$

$$\sum_{p \in \mathcal{P} | t(p)=i} \left(\sum_{l \in \mathcal{L} | M_l \geq m'} b_{p,l}^{born,sep} + \sum_{l \in \mathcal{L} | M_l \geq m'+1} b_{p,l}^{ctn,sep} \right) \quad (4.17)$$

Equations (4.18) assert that the required number of active modules in any arc is respected. It is redundant with the model.

Proposition 4.2.11 *Let us first show that constraints (4.2) to (4.10) induce the following equations:*

$$\forall (i, j) \in A, m_{(i,j)}^{act} = \sum_{p \in \mathcal{P}, (i,j) \in p} \left(m_p^{born,sep} + m_p^{ctn,sep} + \sum_{l \in \mathcal{L}} D_{t(p)} \cdot (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) \right) \quad (4.18)$$

Proof. First, (4.18) is clear for any arc (i, j) such that $\Gamma^+(j) = \emptyset$. Indeed, in that case, $j \in V_D$, otherwise the node j does not contribute to the problem. Then, (4.3) becomes $\sum_{p \in \mathcal{P}, t(p)=j} \sum_{l \in \mathcal{L}} D_{t(p)} \cdot (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) = \sum_{p \in \mathcal{P}, (i,j) \in p} \sum_{l \in \mathcal{L}} D_{t(p)} \cdot (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) = D_j = m_{(i,j)}^{act}$. Validity of (4.18) can then be deduced by induction. Let us consider $(\gamma(i), i) \in A$, such that $i \in V_D$ and $\forall j \in \Gamma^+(i)$, $\sum_{p \in \mathcal{P}, (i,j) \in p} (m_p^{born,sep} + m_p^{ctn,sep} + \sum_{l \in \mathcal{L}} D_{t(p)} \cdot (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem})) = m_{(i,j)}^{act}$. By summing these equations for all arcs (i, j) with $j \in \Gamma^+(i)$, as well as (4.3) for i , we have

$$\begin{aligned} & \sum_{p \in \mathcal{P} | t(p)=i} (m_p^{born,sep} + m_p^{ctn,sep}) + \sum_{p \in \mathcal{P} | t(p)=i} \sum_{l \in \mathcal{L}} D_i \cdot (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) + \\ & \sum_{j \in \Gamma^+(i)} \sum_{p \in \mathcal{P}, (i,j) \in p} \left(m_p^{born,sep} + m_p^{ctn,sep} + \sum_{l \in \mathcal{L}} D_{t(p)} \cdot (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) \right) = \\ & \sum_{j \in \Gamma^+(i)} m_{(i,j)}^{act} + D_i + \sum_{p \in \mathcal{P} | s(p)=i} (m_p^{born,sep} + m_p^{ctn,sep}) \\ & + \sum_{p \in \mathcal{P} | s(p)=i, t(p) \in V_D} \sum_{l \in \mathcal{L}} D_{t(p)} \cdot (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) \end{aligned}$$

Recalling that $\{p \in \mathcal{P} | \exists j \in \Gamma^+(i), (i, j) \in p\} \cup \{p \in \mathcal{P} | t(p) = i\} = (\{p \in \mathcal{P} | (\gamma(i), i) \in p\} \cup \{p \in \mathcal{P} | s(p) = i\})$, and $m_{(\gamma(i), i)}^{act} = D_i + \sum_{j \in \Gamma^+(i)} m_{(i,j)}^{act}$ we have the result. A similar calculation can show (4.18) for $i \in V_N$ by using (4.2). \square

Let us introduce MIR (mixed integer rounding) type inequalities. Let us consider $a \in A$. Let R be the function defined on $\mathbb{R}_{>0}^2$ by $R(\alpha, \beta) = \beta - \alpha \cdot \left\lfloor \frac{\beta}{\alpha} \right\rfloor$, the remainder of the division of β by α . Let μ_a be the function defined on $\mathbb{R}_{>0}^2$ by $\mu_a(\alpha, \beta) = R(\alpha, \beta) \cdot \left\lfloor \frac{\beta}{\alpha} \right\rfloor + \min(R(\alpha, \beta), R(\alpha, m_a^{act}))$.

Remark 4.2.2 *Most inequalities from the model are from one of two cases. By noting them $\sum_{i \in I} a_i x_i \leq b$, we either have $b = 0$ or $\forall i \in I, a_i = b$. In both cases, the application of MIR cuts does not provide any improvement.*

Proposition 4.2.12 *The following constraints are valid inequalities for the FC-NDA problem*

$$\begin{aligned} \forall a \in A, \forall y \in \mathbb{R}, \quad & \sum_{p \in \mathcal{P}|a \in p} \sum_{l \in \mathcal{L}} \left(\mu_a(y, \min(m_a^{act}, M_l)) \cdot b_{p,l}^{born,sep} + \right. \\ & \left. \mu_a(y, \min(m_a^{act}, M_l - 1)) \cdot b_{p,l}^{ctn,sep} \right) \\ & + \sum_{p \in \mathcal{P}|a \in p} \sum_{l \in \mathcal{L}} \mu_a(y, D_{t(p)}) \cdot (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) \geq \mu_a(y, m_a^{act}) \end{aligned} \quad (4.19)$$

Proof. Note that for any arc $a \in A$, the terms of (4.4) can be upper-bounded thanks to constraints (4.5) and (4.18). This gives $\sum_{p \in \mathcal{P}|a \in p} \sum_{l \in \mathcal{L}} M_l \cdot b_{p,l}^{born,sep} + (M_l - 1) \cdot b_{p,l}^{ctn,sep} + \sum_{p \in \mathcal{P}|a \in p} \sum_{l \in \mathcal{L}} D_{t(p)} \cdot (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) \geq m_a^{act}$. Besides, we can note that the coefficients of the left hand side larger than m_a^{act} can be lowered to m_a^{act} , the inequation defines the same set of integer solutions.

This gives $\sum_{p \in \mathcal{P}|a \in p} \sum_{l \in \mathcal{L}} \min(M_l, m_a^{act}) \cdot b_{p,l}^{born,sep} + \min(M_l - 1, m_a^{act}) \cdot b_{p,l}^{ctn,sep} + \sum_{p \in \mathcal{P}|a \in p} \sum_{l \in \mathcal{L}} D_{t(p)} \cdot (b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) \geq m_a^{act}$.

Dividing by some $y \in \mathbb{R}$ and applying mixed integer rounding such as in [9], gives the result. \square

Let us consider $i \in V^*$, and let us define, for $n \in \{1, \dots, |\Gamma^+(i)|\}$, $\eta_{i,n} = \min \{ \sum_{j \in J} m_{(i,j)}^{act} \mid J \subseteq \Gamma^+(i), |J| = n \}$. Intuitively, if a node $i \in V^*$ is of outgoing degree d , then either there are at least d cables upstream of i , or new cables are created. If new cables are created, then they must be able to hold the number of active modules of at least the outgoing arcs of i with the smallest number of active modules. This enables us to give a lower bound for the number of welds to be performed, expressed in the inequalities below.

Proposition 4.2.13 *The following constraints are valid inequalities for the FC-NDA problem*

$$\begin{aligned} \forall i \in V^*, \quad & \sum_{p \in \mathcal{P}|i \in p \setminus \{s(p), t(p)\}} \sum_{l \in \mathcal{L}} (b_{p,l}^{born,sep} + b_{p,l}^{born,dem} + b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem}) \\ & + \sum_{p \in \mathcal{P}|s(p)=i} \sum_{l \in \mathcal{L}} (b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem}) \\ & + \sum_{n \in \{1, \dots, |\Gamma^+(i)|\}} \sum_{m \in \mathcal{M}_L \mid m \geq \eta_{i,n}} w_{i,m} \geq |\Gamma^+(i)| \end{aligned} \quad (4.20)$$

Proof. Let us consider $i \in V^*$. In the case $\sum_{p \in \mathcal{P}|i \in p \setminus \{s(p), t(p)\}} \sum_{l \in \mathcal{L}} (b_{p,l}^{born,sep} + b_{p,l}^{born,dem} + b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem}) \geq |\Gamma^+(i)|$, the inequality is immediate.

Let us assume $\sum_{p \in \mathcal{P}|i \in p \setminus \{s(p), t(p)\}} \sum_{l \in \mathcal{L}} (b_{p,l}^{born,sep} + b_{p,l}^{born,dem} + b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem}) = n_0 < |\Gamma^+(i)|$ (there are not as many cables going through i unchanged than arcs

going out of i). This implies there are cables created in i .

- If the new cables are created out of a tapping. First of all, it means

$\sum_{p \in \mathcal{P}|s(p)=i} \sum_{l \in \mathcal{L}} (b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem}) = 1$. Besides, one cable created in i can be the continuation of a tapped cable. The other ones cover the remaining $|\Gamma^+(i)| - n_0 - 1$ arcs. The sum of their number of active modules must hence be at least $\eta_{i,|\Gamma^+(i)|-n_0-1}$. So since these are all born cables, by (4.9), this implies $\exists m_0 \geq \eta_{i,|\Gamma^+(i)|-n_0-1}$, $w_{i,m_0} = 1$. Let us consider the term $\sum_{n \in \{1, \dots, |\Gamma^+(i)|\}} \sum_{m \in \mathcal{M}_L | m \geq \eta_{i,n}} w_{i,m}$.

For $n \in \{1, \dots, |\Gamma^+(i)| - n_0 - 1\}$, $m_0 \geq \eta_{i,n}$, so $\sum_{m \in \mathcal{M}_L | m \geq \eta_{i,n}} w_{i,m} = 1$. Hence,

by summing all these terms, $\sum_{n \in \{1, \dots, |\Gamma^+(i)|\}} \sum_{m \in \mathcal{M}_L | m \geq \eta_{i,n}} w_{i,m} \geq |\Gamma^+(i)| - n_0 - 1$,

and the inequality holds.

- If the cables are created out of a splicing, then born cables cover the remaining $|\Gamma^+(i)| - n_0$ arcs. The sum of their number of active modules must hence be at least $\eta_{i,|\Gamma^+(i)|-n_0}$. So by (4.9), this implies $\exists m_0 \geq \eta_{i,n}$, $w_{i,m_0} = 1$. With a reasoning similar to the previous case, it means

$\sum_{n \in \{1, \dots, |\Gamma^+(i)|\}} \sum_{m \in \mathcal{M}_L | m \geq \eta_{i,n}} w_{i,m} \geq |\Gamma^+(i)| - n_0$, and the inequality holds. \square

The following proposition asserts that in any separation point of a no-demand node, at least two cables are created (born or continuing a tapped cable), and the created cables cannot be larger than the separated one.

Proposition 4.2.14 *Every optimal solution of the FCNDA problem verifies*

$$\forall i \in V_N, \forall l' \in \mathcal{L}, \sum_{p \in \mathcal{P}|s(p)=i} \sum_{l \leq l'} (b_{p,l}^{born,sep} + b_{p,l}^{ctn,sep} + b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) \geq 2 \cdot \sum_{p \in \mathcal{P}|t(p)=i} \sum_{l \leq l'} (b_{p,l}^{born,sep} + b_{p,l}^{ctn,sep}) \quad (4.21)$$

A part of the proof shows that any separation point in which only one cable is created can be replaced. The most complicated case is when a big cable is spliced into a smaller cable. In that case, the smaller cable could have been created upstream.

Proof. Let us consider, in an optimal solution S of the FCNDA problem, $i \in V_N$ and $l' \in \mathcal{L}$. Let us assume that $\sum_{p \in \mathcal{P}|s(p)=i} \sum_{l \leq l'} (b_{p,l}^{born,sep} + b_{p,l}^{ctn,sep} + b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) < 2 \sum_{p \in \mathcal{P}|t(p)=i} \sum_{l \leq l'} (b_{p,l}^{born,sep} + b_{p,l}^{ctn,sep})$. With (4.7), it means $\sum_{p \in \mathcal{P}|t(p)=i} \sum_{l \leq l'} (b_{p,l}^{born,sep} + b_{p,l}^{ctn,sep}) = 1$, which implies $\sum_{p \in \mathcal{P}|s(p)=i} \sum_{l \leq l'} (b_{p,l}^{born,sep} + b_{p,l}^{ctn,sep} + b_{p,l}^{born,dem} + b_{p,l}^{ctn,dem}) = \sum_{p \in \mathcal{P}|s(p)=i} \sum_{l \leq l'} (b_{p,l}^{born,sep} + b_{p,l}^{born,dem}) = 1$ (only one cable is created). Let us note $p_2 \in \mathcal{P}$ and $l_2 \in \mathcal{L}$ the only couple such that $s(p_2) = i$ and $b_{p_2,l_2}^{born,sep} + b_{p_2,l_2}^{born,dem} = 1$. Let us note $p_1 \in \mathcal{P}$ and $l_1 \in \mathcal{L}$ the only couple such that $t(p_1) = i$ and $b_{p_1,l_1}^{born,sep} + b_{p_1,l_1}^{ctn,sep} = 1$, and $p_0 \in \mathcal{P}$ the only path such that $s(p_0) = s(p_1)$

and $t(p_0) = t(p_2)$.

Let us assume $l_2 \geq l_1$ (the created cable is at least as large as the separated one). The number of active modules of the created cable is equal to the one of the spliced cable. This splicing is then useless. So $l_2 < l_1$ (the created cable is smaller than the separated one). The spliced cable is then the continuation of a tapped cable. Otherwise, its number of active modules would be larger than $M_{l_2} + 1$ (see (4.13)). So the cable on p_1 is the continuation of a tapped cable. Let us replace both cables by only one cable born in $s(p_0)$ lying on p_0 , of size l_2 , having the same number of active modules. First, the cost of this cable is less than the cost of the two previous ones. Moreover, the cost of the protective box in i is saved. Finally, the welds performed in i are done in $s(p_0)$ instead. Since PW is subadditive, the overall cost of welds is smaller. This leads to a cheaper solution. Hence the result. \square

Remark 4.2.3 *This model can be modified to take into account duct capacities. Indeed, for an arc $(i, j) \in A$ and a given cable type $l \in \mathcal{L}$, the number of cables of type l going through (i, j) is given by $\sum_{p \in \mathcal{P}|(i,j) \in p} (b_{p,l}^{born,sep} + b_{p,l}^{born,dem} + b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem})$. Hence, capacity constraints can be added to the model. For instance, the simple case where all cables have the same diameter and one can install at most $n_{(i,j)}$ cables along arc (i, j) can be implemented by adding the constraints*

$$\forall (i, j) \in A, \sum_{l \in \mathcal{L}} \sum_{p \in \mathcal{P}|(i,j) \in p} (b_{p,l}^{born,sep} + b_{p,l}^{born,dem} + b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem}) \leq n_{(i,j)}$$

The next section introduces an arc-node model for the FCNDA.

4.3 Arc based integer programming formulation

The following model is the adaptation of the model from Section 3.1 where the civil engineering graph is a tree. It is possible to exploit the properties of the civil engineering arborescence, especially when introducing valid inequalities.

4.3.1 Notation

We introduce the following decision variables:

- $\forall (i, j) \in A, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, k_{i,j,l,m} \in \{0, \dots, |V_D|\}$: the number of cables with M_l fiber modules from which m are active in (i, j) .
- $\forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, b_{i,l,m}^{spl} \in \{0, 1\}$: the binary variable that is equal to 1 iff we splice in node i a cable with M_l modules of fibers from which m are active modules.
- $\forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, k_{i,l,m}^{born} \in \{0, \dots, M_L\}$: the number of cables born in node i having M_l modules of fibers from which m are active modules.

- $\forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, b_{i,l,m}^{tap} \in \{0, 1\}$: the binary variable that is equal to 1 iff we tap a cable in node i having M_l modules of fibers from which m are active modules.
- $\forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, b_{i,l,m}^{ctn} \in \{0, 1\}$: the binary variable equal to 1 iff the continuation of a tapped cable in node i has M_l modules of fibers from which m are active modules.
- $\forall i \in V^*, \forall m \in \mathcal{M}_L, w_{i,m} \in \{0, 1\}$: the binary variable that is equal to 1 iff m modules are weld in either a tapping or a splicing on node i .
- $\forall i \in V_D, u_i \in \{0, 1\}$: the binary variable that is equal to 1 iff i is served in a module-served way.

For the sake of clarity, a description of a small cabling solution by the model is given in Fig. 4.5.

4.3.2 Formulation

The fiber cable network design problem can be cast into an integer linear program as follows:

$$\begin{aligned}
\min \quad & \sum_{(i,j) \in A} \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} C_l^{le} \cdot \Delta_{(i,j)} \cdot k_{i,j,l,m} \\
& + \sum_{i \in V^*} \sum_{m \in \mathcal{M}_L} PW_m \cdot w_{i,m} \\
& + \sum_{i \in V^*} \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} PB_l \cdot (b_{i,l,m}^{spl} + b_{i,l,m}^{tap}) \tag{4.22}
\end{aligned}$$

such that

$$\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot k_{i,j,l,m} = m_{(i,j)}^{act} \quad \forall (i,j) \in A, \tag{4.23}$$

$$\sum_{m \in \mathcal{M}_l} b_{i,l,m}^{tap} = \sum_{1 \leq m \leq M_l - 1} b_{i,l,m}^{ctn} \quad \forall i \in V^*, \forall l \in \mathcal{L}, \tag{4.24}$$

$$\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} b_{i,l,m}^{spl} + b_{i,l,m}^{tap} \leq 1 \quad \forall i \in V^*, \tag{4.25}$$

$$\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot k_{i,l,m}^{born} = \sum_{m \in \mathcal{M}_L} m \cdot w_{i,m} \quad \forall i \in V^*, \tag{4.26}$$

$$\sum_{m \in \mathcal{M}_L} w_{i,m} \leq \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} (b_{i,l,m}^{spl} + b_{i,l,m}^{tap}) \quad \forall i \in V^*, \tag{4.27}$$

$$k_{\gamma(i),i,l,m} - b_{i,l,m}^{spl} - b_{i,l,m}^{tap} = \sum_{j \in \Gamma^+(i)} k_{i,j,l,m} - k_{i,l,m}^{born} - b_{i,l,m}^{ctn} \tag{4.28}$$

$\forall i \in V_N, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l,$

$$\begin{aligned}
& k_{\gamma(i),i,l,m} - b_{i,l,m}^{spl} - b_{i,l,m}^{tap} = \\
& \sum_{j \in \Gamma^+(i)} k_{i,j,l,m} - k_{i,l,m}^{born} - b_{i,l,m}^{ctn} \\
& \forall i \in V_D, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l \text{ such that } m \neq D_i, \tag{4.29}
\end{aligned}$$

$$\begin{aligned}
& k_{\gamma(i),i,l,D_i} - b_{i,l,D_i}^{spl} - b_{i,l,D_i}^{tap} \geq \\
& \sum_{j \in \Gamma^+(i)} k_{i,j,l,D_i} - k_{i,l,D_i}^{born} - b_{i,l,D_i}^{ctn} \\
& \forall i \in V_D, \forall l \in \mathcal{L} \text{ such that } D_i \in \mathcal{M}_l, \tag{4.30}
\end{aligned}$$

$$\begin{aligned}
& \sum_{l \in \mathcal{L} | D_i \in \mathcal{M}_l} (k_{\gamma(i),i,l,D_i} - b_{i,l,D_i}^{spl} - b_{i,l,D_i}^{tap}) = \\
& 1 - u_i + \sum_{l \in \mathcal{L} | D_i \in \mathcal{M}_l} \left(\sum_{j \in \Gamma^+(i)} k_{i,j,l,D_i} - k_{i,l,D_i}^{born} - b_{i,l,D_i}^{ctn} \right) \quad \forall i \in V_D, \tag{4.31}
\end{aligned}$$

$$\begin{aligned}
& 0 \leq \sum_{(i,j) \in \Gamma^+(i)} k_{i,j,l,m} - k_{i,l,m}^{born} - b_{i,l,m}^{ctn} \\
& \forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, \tag{4.32}
\end{aligned}$$

$$\begin{aligned}
& k_{i,j,l,m} \in \{0, \dots, |V_D|\}; k_{i,l,m}^{born} \in \{0, \dots, M_L\} \\
& \forall (i,j) \in A, \forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, \\
& b_{i,l,m}^{tap} \in \{0, 1\}; b_{i,l,m}^{spl} \in \{0, 1\}; b_{i,l,m}^{ctn} \in \{0, 1\}; w_{i,m} \in \{0, 1\} \\
& \forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l.
\end{aligned}$$

The cost function (4.22) is a combination of cable costs (first term), weld costs (second term) and welding boxes costs (third term).

Equations (4.23) ensure, with the definition of m^{act} , conservation of the number of active modules at each node (they either leave the node or serve the demand). Constraints (4.24) control the dimensioning of continued cables with respect to the original tapped ones (it has the same size). Constraints (4.25) ensure that at most one of the two operations of splicing and tapping can be performed in each node. Constraints (4.26) and (4.27) ensure a proper counting of the number of welds in each node, based on the number of active modules in born cables.

Constraints (4.28) make sure that incoming cables of a no-demand node which are neither tapped nor spliced are continued with the same size and same number of active modules. The first term stands for the incoming cables, minus the eventual separated cable, and the second term for the outgoing cables, minus the eventual created ones. Together, (4.23) and (4.28) ensure module conservation in splicing and tapping operations of no-demand nodes (see (4.33)).

Constraints (4.29), (4.30) and (4.31) are related to demand nodes. If the number of active modules of a cable is not equal to the demand, the behaviour of the cables is similar to the one of no-demand nodes: it is either spliced, tapped or leaves the node. (4.30) expresses the fact that if the number of active modules of a cable is equal to the demand, then, among the cables unaffected

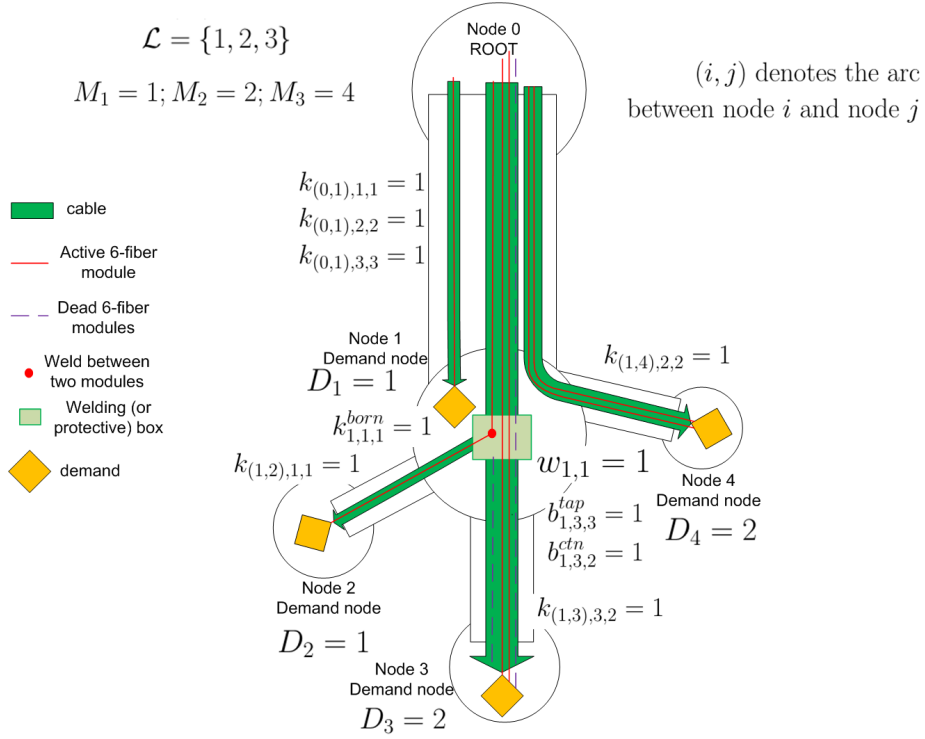


Figure 4.5: Arc model illustration on a small example; a tapping is performed in node 1

by tapping or splicing, there are always more cables incoming in the node than outgoing. The ones unaffected by the demand or a tapping and splicing are continued identically. Equations (4.31) ensure demand nodes are either served by only one cable or by modules from a tapping/splicing operation, and that this cable has the right number of active modules. Together, constraints (4.23), (4.29) and (4.31) ensure module conservation in splicing and tapping operations (see (4.34)).

In order to give more insight into on the validity of the models, we prove in the following that the equations from the model imply the conservation of active modules.

Proposition 4.3.1 *Equations (4.23) to (4.32) imply the following*

$$\forall i \in V_N, \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot (b_{i,l,m}^{tap} + b_{i,l,m}^{spl}) = \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot (k_{i,l,m}^{born} + b_{i,l,m}^{ctn}) \quad (4.33)$$

$$\forall i \in V_D, \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot (b_{i,l,m}^{tap} + b_{i,l,m}^{spl}) = D_i \cdot u_i + \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot (k_{i,l,m}^{born} + b_{i,l,m}^{ctn}) \quad (4.34)$$

Proof. Multiplying (4.28) by m and summing gives

$$\begin{aligned} \forall i \in V_N, \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot (k_{\gamma(i),i,l,m} - b_{i,l,m}^{tap} - b_{i,l,m}^{spl}) = \\ \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot \left(\sum_{j \in \Gamma^+(i)} k_{i,j,l,m} - k_{i,l,m}^{born} - b_{i,l,m}^{ctn} \right) \end{aligned}$$

Then, using (4.23) and the fact that $m_{\gamma(i),i}^{act} = \sum_{j \in \Gamma^+(i)} m_{(i,j)}^{act}$ gives (4.33).

Multiplying (4.29) and (4.31) by m and summing gives

$$\begin{aligned} \forall i \in V_D, \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot (k_{\gamma(i),i,l,m} - b_{i,l,m}^{tap} - b_{i,l,m}^{spl}) = \\ D_i - D_i \cdot u_i + \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} m \cdot \left(\sum_{j \in \Gamma^+(i)} k_{i,j,l,m} - k_{i,l,m}^{born} - b_{i,l,m}^{ctn} \right) \end{aligned}$$

Then, using (4.23) and the fact that $m_{\gamma(i),i}^{act} = D_i + \sum_{j \in \Gamma^+(i)} m_{(i,j)}^{act}$ gives (4.34).

□

Equations (4.33) and (4.34) show that module conservation applies to splicing and tapping operations, as specified in Section 2.1. Since with (4.25), at most one cable is either spliced or tapped, active modules from a tapped or spliced cable are used for created cables and an eventual module-served node. Apart from module conservation, demand nodes need further explanation.

Let us focus on a node $i \in V_D$ where $u_i = 1$. Then, (4.30) combined with (4.31) become

$$\begin{aligned} \forall l \in \mathcal{L} \text{ such that } D_i \in M_l, k_{\gamma(i),i,l,D_i} - b_{i,l,D_i}^{spl} - b_{i,l,D_i}^{tap} = \\ \sum_{j \in \Gamma^+(i)} k_{i,j,l,D_i} - k_{i,l,D_i}^{born} - b_{i,l,D_i}^{ctn} \end{aligned}$$

This induces that all cables unaffected by splicing or tapping operations are unchanged in i (similarly to (4.28)).

Let us focus on a node $i \in V_D$ where $u_i = 0$. Then, (4.30) combined with (4.31) tell us that only one of the equations (4.30) has a strict inequality, with a difference between left and right terms of 1. In other words, $\exists l_0 \in \mathcal{L} | D_i \in \mathcal{M}_{l_0}$ such that

$$\begin{aligned} \forall l \in \mathcal{L} \setminus \{l_0\}, k_{\gamma(i),i,l,D_i} - b_{i,l,D_i}^{spl} - b_{i,l,D_i}^{tap} = \sum_{j \in \Gamma^+(i)} k_{i,j,l,D_i} - k_{i,l,D_i}^{born} - b_{i,l,D_i}^{ctn} \\ k_{\gamma(i),i,l_0,D_i} - b_{i,l_0,D_i}^{spl} - b_{i,l_0,D_i}^{tap} = 1 + \sum_{j \in \Gamma^+(i)} k_{i,j,l_0,D_i} - k_{i,l_0,D_i}^{born} - b_{i,l_0,D_i}^{ctn} \end{aligned}$$

This shows that the demand is served in a cable served way by a cable of size l_0 . All other cables unaffected by tapping or splicing operations are continued with the same size and same number of active modules.

Finally, constraints (4.32) make sure that if a cable is created in a node, there is

indeed, a cable of the same size and same number of active modules going out of this node. Combined with (4.28) and (4.30), it also ensures that if a cable is separated in a node, then there is an incoming cable with the same size and number of active modules.

Remark 4.3.1 *Introducing notation $N_{ml} = \sum_{l \in \mathcal{L}} M_l$, this model holds $(|V| - 1)(3N_{ml} + M_L) + |V_D|$ boolean variables, $(|V| - 1 + |E|)N_{ml}$ integer variables and $(|V| - 1)(5 + 2N_{ml} + L) + |V_D|(N_{ml} + 1)$ constraints. Compared to the model of Section 4.2.2 ((4.1) - (4.10)), it can hold either more or less variables, depending on the values $L, (M_l)_{l \in \mathcal{L}}$ and on G (see Table 4.2 and Remark 4.2.1).*

4.3.3 Preprocessing

We can fix the value of some variables.

First of all, no cable holding only one active module can endure a separation. Tapping of a cable holding one active module is useless. In the case of a splicing operation, only one cable can be born, and must be strictly smaller. Such an operation could be avoided by creating a smaller cable instead (see (4.21) and its justification) in the upstream separation point. In other words, $\forall i \in V^*, \forall l \in \mathcal{L}, b_{i,l,1}^{tap} = b_{i,l,1}^{spl} = 0$.

Notice that the feasibility of (4.23) implies that some variables are equal to zero. This illustrates the fact that no cable can hold more active modules than the arc it goes through.

$$\begin{aligned} \forall (i, j) \in A, \forall m \in \mathcal{M}_L \text{ such that } m > m_{(i,j)}^{act}, \quad \sum_{l \in \mathcal{L} | m \in \mathcal{M}_l} k_{i,j,l,m} &= 0 \\ \forall i \in V^*, \forall m \in \mathcal{M}_L \text{ such that } m > m_{(\gamma(i),i)}^{act}, \\ w_{i,m} = \sum_{l \in \mathcal{L} | m \in \mathcal{M}_l} b_{i,l,m}^{spl} = \sum_{l \in \mathcal{L} | m \in \mathcal{M}_l} b_{i,l,m}^{tap} = \sum_{l \in \mathcal{L} | m \in \mathcal{M}_l} b_{i,l,m}^{ctn} = \sum_{l \in \mathcal{L} | m \in \mathcal{M}_l} k_{i,l,m}^{born} &= 0 \end{aligned}$$

In any optimal solution, no cable begins with a size larger than what it needs to hold its active modules. This can be written

$$\forall l \in \mathcal{L}, \forall m \in \mathcal{M}_{l-1}, \quad \sum_{j \in \Gamma^+(r)} k_{r,j,l,m} = 0$$

(recall that r stands for the root node of the arborescence)

$$\forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_{l-1}, k_{i,l,m}^{born} = 0$$

Similarly to the model of Section 4.2.2, a cable will hold a number of active modules equal to the sum of a subset of the demands in the underlying arborescence. Recall from Section 4.2.3 that $\forall i \in V, V_i^{ar}$ denotes the arborescence rooted in i , i excluded, and $PM(i) = \{\sum_{j \in J} D_j | J \subseteq V_i^{ar}\}$. Let us note $PM'(i) = \{\sum_{j \in J} D_j | J \subseteq (V_i^{ar} \cup \{i\})\}$. A cable going through any arc (i, j) must hold a number of active modules which is a sum of some demands from

nodes located in the arborescence rooted in j , j included. This leads to the following simplifications:

$$\begin{aligned}
& \forall i \in V^*, \forall m \in \mathcal{M}_L \setminus PM(i), w_{i,m} = 0 \\
& \forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l \setminus (\cup_{j \in \Gamma^+(i)} PM'(j)), b_{i,l,m}^{ctn} = k_{i,l,m}^{born} = 0 \\
& \forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l \setminus PM'(i), b_{i,l,m}^{tap} = b_{i,l,m}^{spl} = 0 \\
& \forall (i,j) \in A, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l \setminus PM'(j), k_{i,j,l,m} = 0
\end{aligned}$$

4.3.4 Valid inequalities

Valid inequalities are added to tighten the relaxation of the model of Section 4.3.2.

Let us consider tapping operations in an optimal solutions. The continuation of a tapped cable always has strictly less active modules than the tapped cable. This gives us the following proposition.

Proposition 4.3.2 *Every optimal solution of the FCNDA problem verifies*

$$\forall i \in V^*, \forall l \in \mathcal{L}, \forall m' \in \mathcal{M}_l \setminus \{1\}, \sum_{m \in \{1, \dots, m'-1\}} b_{i,l,m}^{ctn} \geq \sum_{m \in \{1, \dots, m'\}} b_{i,l,m}^{tap} \quad (4.35)$$

The proposition below asserts the simultaneous appearance of tapping and splicing operations with welds in no-demand nodes. Its meaning is similar to (4.14).

Proposition 4.3.3 *The following equalities are valid for the FCNDA problem*

$$\forall i \in V_N, \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} b_{i,l,m}^{spl} + b_{i,l,m}^{tap} = \sum_{m \in \mathcal{M}_L} w_{i,m} \quad (4.36)$$

If a cable holding m' active modules is tapped, there is at least 1 and at most $m' - 1$ welds. If a cable holding $m \in \{1, \dots, m' - 1\}$ modules is spliced in a no-demand node, then all its active modules are used to create born cables, and m welds are performed, leading to the following proposition.

Proposition 4.3.4 *The following constraints are valid inequalities for the FCNDA problem*

$$\begin{aligned}
& \forall i \in V_N, \forall m' \in \mathcal{M}_L, \sum_{l \in \mathcal{L} | m' \in \mathcal{M}_l} b_{i,l,m'}^{tap} + \\
& \sum_{m \in \{1, \dots, m'-1\}} \sum_{l \in \mathcal{L} | m \in \mathcal{M}_l} b_{i,l,m}^{spl} \leq \sum_{m \in \{1, \dots, m'-1\}} w_{i,m} \quad (4.37)
\end{aligned}$$

For any $i \in V^*$ and $n \in \{1, \dots, |\Gamma^+(i)|\}$, let $\eta_{i,n}$ be $\min\{\sum_{j \in J} m_{(i,j)}^{act} | J \subseteq \Gamma^+(i), |J| = n\}$ (see (4.20)). The following constraints express the fact that if there are not enough cables coming to a node to have one for each outgoing arc, then there must be a separation operation in this node. Its meaning is similar to (4.20).

Proposition 4.3.5 *The following constraints are valid inequalities for the FC-NDA problem*

$$\begin{aligned} \forall i \in V_N, \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} k_{\gamma(i),i,l,m} - \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} b_{i,l,m}^{spl} \\ + \sum_{n \in \{1, \dots, |\Gamma^+(i)|\}} \sum_{m \in \mathcal{M}_L | m \geq \eta_{i,n}} w_{i,m} \geq |\Gamma^+(i)| \end{aligned} \quad (4.38)$$

$$\begin{aligned} \forall i \in V_D, \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} k_{\gamma(i),i,l,m} - \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} b_{i,l,m}^{spl} \\ + \sum_{n \in \{1, \dots, |\Gamma^+(i)|\}} \sum_{m \in \mathcal{M}_L | m \geq \eta_{i,n}} w_{i,m} \geq |\Gamma^+(i)| + 1 - u_i \end{aligned} \quad (4.39)$$

Proof. Let us consider $i \in V_N$.

Let us start with the case where $\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} b_{i,l,m}^{spl} + b_{i,l,m}^{tap} = 0$ (there is no separation point in i). Then, the cables going through $\alpha(i)$ are exactly the cables going through i without enduring any operation in i . So using the variables of Section 4.2.1 we have $\sum_{l \in \mathcal{L}} \sum_{p \in \mathcal{P} | i \in p \setminus \{s(p), t(p)\}} (b_{p,l}^{born,sep} + b_{p,l}^{born,dem} + b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem}) = \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} k_{\gamma(i),i,l,m}$ and $\sum_{p \in \mathcal{P} | s(p)=i} \sum_{l \in \mathcal{L}} (b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem}) = 0$. With (4.20), we have the result.

Let us consider the case where there is a splicing in i , which gives $\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} b_{i,l,m}^{spl} = 1$. The cables going through $\alpha(i)$ are the cables going through i without enduring any operation in i and the cable spliced in i . So using the variables of Section 4.2.1 gives $\sum_{l \in \mathcal{L}} \sum_{p \in \mathcal{P} | i \in p \setminus \{s(p), t(p)\}} (b_{p,l}^{born,sep} + b_{p,l}^{born,dem} + b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem}) = \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} k_{\gamma(i),i,l,m} - \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} b_{i,l,m}^{spl}$ and $\sum_{p \in \mathcal{P} | s(p)=i} \sum_{l \in \mathcal{L}} (b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem}) = 0$. By (4.20), we have the result.

Let us consider the case where there is a tapping in i , which gives $\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} b_{i,l,m}^{tap} = 1$. The cables going through $\alpha(i)$ are the cables going through i without enduring any operation in i and the cable tapped in i . Besides, the continuation of this tapped cable leaves the node i , which gives, with the variables of Section 4.2.1, $\sum_{p \in \mathcal{P} | s(p)=i} \sum_{l \in \mathcal{L}} (b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem}) = 1$.

So $\sum_{l \in \mathcal{L}} \sum_{p \in \mathcal{P} | i \in p \setminus \{s(p), t(p)\}} (b_{p,l}^{born,sep} + b_{p,l}^{born,dem} + b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem}) + \sum_{p \in \mathcal{P} | s(p)=i} \sum_{l \in \mathcal{L}} (b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem}) = \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} k_{\gamma(i),i,l,m}$. With (4.20), we have the result.

The case where $i \in V_D$ can be proved in a similar way. If $u_i = 1$, then the node is module-served and the equations related to it are similar. Otherwise, one of the cables going through $\alpha(i)$ is used to serve the demand. \square

What follows states that in any optimal solution, at least two cables are created in a separation point of a no-demand node. Furthermore, these cables always have less active modules. It is somehow similar to (4.21).

Proposition 4.3.6 *Every optimal solution of the FCNDA problem verifies*

$$\begin{aligned} \forall i \in V_N, \forall m' \in \mathcal{M}_L, \sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} k_{i,l,m}^{born} + b_{i,l,m}^{ctn} \geq \\ 2 \sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} b_{i,l,m}^{tap} + b_{i,l,m}^{spl} \end{aligned} \quad (4.40)$$

Proof. Let us consider $i \in V_N$ and an optimal solution of the FCNDA problem. In the case $\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} b_{i,l,m}^{tap} + b_{i,l,m}^{spl} = 0$ (no cable is separated in i), the inequalities are immediate. Otherwise, by (4.25), we have a unique m_0 for which $\sum_{l \in \mathcal{L} | M_l \geq m_0} b_{i,l,m_0}^{tap} + b_{i,l,m_0}^{spl} = 1$ (there is a separated cable of size m_0). Furthermore, at least two cables must be created in a separation point in i (see (4.21) with $l' = L$). So we have $\sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} k_{i,l,m}^{born} + b_{i,l,m}^{ctn} \geq 2 \sum_{l \in \mathcal{L}} \sum_{m \in \mathcal{M}_l} b_{i,l,m}^{tap} + b_{i,l,m}^{spl} = 1$. Besides, with active module conservation, the created cables have less modules than the separated cable. So for any $m' < m_0$, $\sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} (b_{i,l,m}^{tap} + b_{i,l,m}^{spl}) = 0$ and we have (4.40). And for $m' \geq m_0$, we have $\sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} k_{i,l,m}^{born} + b_{i,l,m}^{ctn} \geq 2$, which gives (4.40). Hence the result. \square

The following proposition asserts that in any optimal solution, at least two cables are created in a separation point of a cable-served demand node.

Proposition 4.3.7 *Every optimal solution of the FCNDA problem verifies*

$$\begin{aligned} \forall i \in V_D, \forall m' \in \mathcal{M}_L, u_i + \sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} k_{i,l,m}^{born} + b_{i,l,m}^{ctn} \geq \\ 2 \sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} b_{i,l,m}^{tap} + b_{i,l,m}^{spl} \end{aligned} \quad (4.41)$$

Proof. Let us consider in an optimal solution S a node $i \in V_D$, and $m' \in \mathcal{M}_L$. If $u_i = 0$, the equations describing this node are similar to those describing no-demand nodes, and we get the result by (4.40).

If $u_i = 1$, let us assume that $u_i + \sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} k_{i,l,m}^{born} + b_{i,l,m}^{ctn} < 2 \sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} b_{i,l,m}^{tap} + b_{i,l,m}^{spl}$.

This means $\sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} k_{i,l,m}^{born} + b_{i,l,m}^{ctn} = 0$ and

$\sum_{l \in \mathcal{L}} \sum_{m \in \{1, \dots, \min(m', M_l)\}} b_{i,l,m}^{tap} + b_{i,l,m}^{spl} = 1$ (there is a separation point without any created cable, only used for serving the demand in i in a module-served way). Then, the solution S' identical to S everywhere except for $u'_i = 0$ and $\forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, b'_{i,l,m}{}^{spl} = b'_{i,l,m}{}^{tap} = 0$ is cheaper and feasible, which contradicts our hypothesis. Hence the result. \square

Let us introduce mixed integer rounding cuts. Let us consider $(i, j) \in A$. Similarly to (4.19), let R be the function defined on $\mathbb{R}_{>0}^2$ by $R(\alpha, \beta) = \beta - \alpha \cdot \left\lfloor \frac{\beta}{\alpha} \right\rfloor$, the remainder of the division of β by α . Let $\mu_{(i,j)}$ be the function defined on $\mathbb{R}_{>0}^2$ by $\mu_{(i,j)}(\alpha, \beta) = R(\alpha, \beta) \cdot \left\lfloor \frac{\beta}{\alpha} \right\rfloor + \min(R(\alpha, \beta), R(\alpha, m_{(i,j)}^{act}))$.

Proposition 4.3.8 *The following constraints are valid inequalities for the FC-NDA problem*

$$\forall (i, j) \in A, \sum_{l \in \mathcal{L}} \mu_{i,j}(y, m) \cdot k_{i,j,l,m} \geq \mu_{i,j}(y, m_{(i,j)}^{act}) \quad (4.42)$$

Proof. Dividing (4.23) by some $y \in \mathbb{R}$ and applying mixed integer rounding such as in [9] gives the result. \square

Remark 4.3.2 *This model can be adapted to integrate duct capacities. This can be done similarly to the ILP of Section 4.2.2: the number of cables of a given size $l \in \mathcal{L}$ going through a given duct $(i, j) \in A$ is equal to $\sum_{m \in \mathcal{M}_l} k_{i,j,l,m}$.*

4.4 Computational results

4.4.1 Test description and setting

We assessed the models on real-life instances from the French cities of Arles (denoted “Ar”), Carcassonne (“Ca”), Clermont-Ferrand (“Cl”) and Etrepilly (“Et”). For each city, distribution zones are denoted with a “D” followed by its index, and the transport zone is denoted with a “T” (transport and distribution zones are illustrated in Fig. 2.1). Table 4.1 provides key features of the different instances. The first column reports the number of paths of the civil engineering arborescence, the second one the number of arcs, the third column (labeled “deg 3+”) stands for the number of nodes of degree 3 or more. The fourth column displays the number of demand nodes, the fifth reports the total demand while the last column provides the average distance from the root to demand points.

Distribution zones are to be covered with cables containing 6-fiber modules, using cables of 1, 2, 4, 6, 8, 12, 18 or 24 modules, while transport zones are to be covered with cables containing 12-fiber modules, using cables of 1, 2, 4, 8, 12, 16 or 20 modules. The costs used in this Section are real costs currently charged to Orange by the subcontractors in charge of deploying the network.

The algorithm used to solve the ILP is the branch and bound from Cplex version 12.6.0.0. in its default configuration, unless specified otherwise. The experiments were run on a computer composed of 4 processors Intel Xeon of CPU 5110 and clocked at 1.6 GHz each.

In the following, the expression “base path model” refers to the model from Section 4.2.2 only including preprocessing of Section 4.2.3, while “enhanced path model” refers to the one from Section 4.2.2 with the valid inequalities from Section 4.2.4. Similarly, the expression “base arc-node model” refers to the model from Section 4.3.2 including preprocessing from Section 4.3.3 while “enhanced arc-node model” refers to the one from Section 4.3.2 with all valid inequalities from Section 4.3.4.

This Section is organized as follows. A benchmark for performances is introduced in Section 4.4.2. Base models are compared in Section 4.4.2, while

Table 4.1: key features of the instances

| instance | features | | | | | |
|----------|-----------------|----------------|--------|------------------------|--------------|----------------------|
| | number of paths | number of arcs | deg 3+ | number of demand nodes | total demand | average distance (m) |
| Et D 1 | 1776 | 153 | 30 | 48 | 67 | 1366.7 |
| Et D 2 | 1762 | 124 | 28 | 39 | 56 | 2961.9 |
| Et D 3 | 4590 | 196 | 36 | 50 | 63 | 3259.0 |
| Or D 1 | 1284 | 181 | 39 | 57 | 75 | 275.2 |
| Or D 2 | 885 | 123 | 22 | 41 | 62 | 366.4 |
| Or D 3 | 1456 | 152 | 38 | 52 | 66 | 447.7 |
| Or D 4 | 856 | 94 | 20 | 29 | 56 | 296.5 |
| Or D 5 | 820 | 109 | 24 | 38 | 64 | 248.5 |
| Or D 6 | 1409 | 172 | 37 | 57 | 76 | 432.1 |
| Or D 7 | 1389 | 146 | 33 | 43 | 65 | 277.6 |
| Ar D 1 | 1180 | 113 | 35 | 45 | 61 | 325.0 |
| Ar D 2 | 979 | 103 | 29 | 38 | 55 | 602.1 |
| Ar D 3 | 1299 | 103 | 21 | 35 | 66 | 368.8 |
| Ar D 4 | 1256 | 123 | 26 | 43 | 80 | 276.3 |
| Ar D 5 | 1037 | 129 | 25 | 44 | 68 | 242.8 |
| Ar D 6 | 1453 | 137 | 28 | 43 | 67 | 358.6 |
| Ar D 7 | 3017 | 139 | 26 | 35 | 68 | 587.9 |
| Ar D 8 | 1732 | 163 | 30 | 41 | 63 | 691.2 |
| Ar D 9 | 3273 | 219 | 53 | 68 | 78 | 1384.7 |
| Ar T | 2673 | 204 | 54 | 65 | 65 | 1505.7 |
| Or T | 4079 | 289 | 81 | 95 | 95 | 1264.7 |
| Et T | 1605 | 110 | 30 | 35 | 35 | 3137.3 |
| Ca T | 4091 | 283 | 75 | 88 | 88 | 1999.7 |

enhanced models are assessed in Section 4.4.2. The optimal solution characteristics are discussed in Section 4.4.3 from an operational point of view. Possible restrictions of the problem are considered in Section 4.4.4.

4.4.2 Model comparison

Table 4.2 reports the number of variables, constraints and valid inequalities of each model. The numbers of valid inequalities mentioned includes all families from Section 4.2.4 (in column “path model”) and from Section 4.3.4 (in column “arc model”).

The number of variables for both models is quite similar, the path model having more variables in most of the cases (17 instances over 23). Note that instances with significantly more paths than arcs have significantly more variables for the path model. The path model has less constraints than the arc-node model in all instances.

Table 4.2: key features of the models

| instance | variables | | constraints | | valid inequalities | |
|----------|-----------|------------|-------------|------------|--------------------|------------|
| | arc model | path model | arc model | path model | arc model | path model |
| Et D 1 | 21135 | 32333 | 29565 | 5508 | 39777 | 8610 |
| Et D 2 | 21238 | 32623 | 23665 | 5087 | 31769 | 7834 |
| Et D 3 | 31893 | 82749 | 36400 | 11630 | 50060 | 16019 |
| Or D 1 | 18385 | 23122 | 34487 | 4845 | 46278 | 8034 |
| Or D 2 | 13937 | 16212 | 23328 | 3299 | 31019 | 5583 |
| Or D 3 | 19105 | 27096 | 29054 | 4812 | 38510 | 7692 |
| Or D 4 | 13484 | 15498 | 18341 | 2929 | 24766 | 4794 |
| Or D 5 | 12705 | 14965 | 21307 | 3034 | 28259 | 5060 |
| Or D 6 | 16783 | 25566 | 33020 | 4987 | 44021 | 8120 |
| Or D 7 | 16829 | 24914 | 27229 | 4585 | 36777 | 7336 |
| Ar D 1 | 19596 | 22323 | 23958 | 3917 | 31489 | 6360 |
| Ar D 2 | 45052 | 18325 | 33695 | 4264 | 47248 | 7517 |
| Ar D 3 | 23920 | 23649 | 23361 | 4157 | 31762 | 6893 |
| Ar D 4 | 16206 | 22947 | 23480 | 4043 | 31011 | 6729 |
| Ar D 5 | 17936 | 18994 | 25838 | 3774 | 34516 | 6348 |
| Ar D 6 | 27727 | 26295 | 30489 | 4953 | 41773 | 8055 |
| Ar D 7 | 57101 | 53969 | 39824 | 8805 | 57043 | 13758 |
| Ar D 8 | 49316 | 30683 | 42562 | 6409 | 60552 | 10658 |
| Ar D 9 | 77701 | 60753 | 62544 | 10838 | 88031 | 17123 |
| Ar T | 19124 | 43193 | 32312 | 7655 | 43064 | 11068 |
| Or T | 27764 | 66253 | 45962 | 11432 | 60997 | 16391 |
| Et T | 11416 | 25981 | 17420 | 4455 | 23220 | 6347 |
| Ca T | 25893 | 65938 | 44686 | 11383 | 59745 | 16218 |

Greedy algorithm

As a benchmark, a greedy algorithm providing a sub-optimal solution is proposed, in which all the demand nodes are cable-served (cf Fig. 2.5), and no separation points are considered (cf Fig. 2.3). The size of the cable serving each demand node is then the smallest one that is greater than the demand. The only computation effort consists in computing, for each $i \in V_D$, $\min\{l \in \mathcal{L} | M_l \geq D_i\}$.

Base models

For computational tests on both base models, a limit on the computation time was set at 600 seconds. Table 4.3 reports the results obtained with the path and arc-node base models (the column “solution” reports the value of the best solution found), as well as the greedy solution without separation described above (in the column labeled “greedy solution”). As performance indicators, we report the ratio between the initial relaxation and the optimal solution (in

the column labeled “relaxation”; the optimal solution is displayed in Table 4.4), the computation time (“time”), the number of branches explored during the branch-and-bound process (“branch”), the final gap (“final gap”), and the time spent until the appearance of the first integer solution (“time to first integer”). Since the greedy solution is computed in a very short time (less than a second for all instances), the associated computation time is not reported.

The best solutions found with the base models are on average roughly 25 % cheaper than the greedy solutions. This shows that allowing tapping and splicing operations enables much savings. Besides, note that the cost of a greedy solution grows with respect to the average root to demand points distance. This is due to the fact that their cost is composed only of cable deployment cost, which grows linearly with the length of the cables.

Comparing base model performances, the main observation is that the arc-node model was overall more efficient. It was able to solve 22 instances out of 23 optimally within the time limit, while the base path model was able to solve only 3 of them. In addition, the best solution found by the arc-node model was always better than the one given by the path model.

As for the resolution process, we provide insight on the better performances of the base arc-node model. Its initial relaxation was always better than the one of the base path model. There is a significantly larger number of nodes in the branch and bound procedure for the path model compared to the arc-node model (on average about 7 times less for the arc-node model), which explains the larger computation times. This can be seen as a consequence of the smaller value of the relaxation. On a side note, the first integer solution appeared on average 1 second into the resolution process in the path model, while it took on average 10 seconds with the arc-node model.

Enhanced models

For computational tests with the enhanced model, no time limit was set. We tried different ways of including the valid inequalities into the solving process by cutting planes. With the exception of mixed integer cuts, including all families completely at the beginning of the branch-and-bound turned out to be the most efficient method. One possible explanation is that the relaxation of the base models was too weak, especially for the path model. Since the number of valid inequalities is not huge, adding them at the root turns out to be a reasonable strategy. Regarding Mixed Integer Rounding cuts (4.19) and (4.42), for an arc $a \in A$, we picked cuts with values of y in $\{\frac{m_a^{act}}{n+0.001} | n \in \{1, \dots, m_a^{act}\}\}$. These values are chosen to be the smallest values of y for a given constant right hand side. The ones which increase the initial relaxation were added at the root of the Branch-and-Bound algorithm (after the initial relaxation is computed, the violated cuts are added). The left part of Table 4.4 displays the results obtained with the valid inequalities. The indicators are denoted in the same way than in Table 4.3. The final gap being always equal to zero, it is not reported. The optimal solution is displayed separately and the times of appearance of the first integer solutions are not reported.

Table 4.3: Results of base models

| instance | arc-node model | | | | | | path model | | | | | | greedy solution | optimal solution |
|----------|----------------|----------|--------|---------------|----------|-----------------------|----------------|----------|--------|---------------|----------|-----------------------|-----------------|------------------|
| | relaxation (%) | time (s) | branch | final gap (%) | solution | time to first integer | relaxation (%) | time (s) | branch | final gap (%) | solution | time to first integer | | |
| Et D 1 | 95.1 | 83 | 5130 | 0.0 | 15958.1 | 12 | 70.3 | 600 | 44176 | 7.2 | 16419.4 | 1 | 33 640.1 | 15 958.1 |
| Et D 2 | 96.0 | 18 | 1665 | 0.0 | 27222.4 | 8 | 78.9 | 600 | 46183 | 6.4 | 27232.4 | 1 | 60 866.4 | 27 222.4 |
| Et D 3 | 93.8 | 129 | 7743 | 0.0 | 32 435.9 | 13 | 64.8 | 600 | 8725 | 23.4 | 37615.0 | 5 | 78 298.2 | 32 435.9 |
| Or D 1 | 96.6 | 5 | 0 | 0.0 | 6803.2 | 3 | 38.3 | 600 | 46827 | 5.2 | 6859.4 | 0 | 7732.2 | 6803.2 |
| Or D 2 | 94.9 | 8 | 426 | 0.0 | 6363.8 | 5 | 48.6 | 159 | 35844 | 0.0 | 6363.8 | 0 | 8348.0 | 6363.8 |
| Or D 3 | 92.1 | 102 | 5293 | 0.0 | 7932.6 | 11 | 41.8 | 600 | 39633 | 19.0 | 8969.4 | 0 | 11 408.0 | 7932.6 |
| Or D 4 | 95.1 | 9 | 179 | 0.0 | 3941.9 | 2 | 44.1 | 600 | 65869 | 5.5 | 3983.9 | 0 | 4789.1 | 3941.9 |
| Or D 5 | 98.0 | 3 | 0 | 0.0 | 4230.3 | 1 | 42.7 | 337 | 72734 | 0.0 | 4230.3 | 0 | 4935.8 | 4230.3 |
| Or D 6 | 94.5 | 27 | 6147 | 0.0 | 8941.4 | 3 | 42.6 | 600 | 45171 | 10.0 | 9196.3 | 0 | 11 954.2 | 8941.4 |
| Or D 7 | 95.0 | 8 | 264 | 0.0 | 5207.0 | 2 | 43.4 | 600 | 48599 | 6.8 | 5284.2 | 0 | 6276.9 | 5206.7 |
| Ar D 1 | 93.8 | 19 | 1982 | 0.0 | 5738.2 | 8 | 44.5 | 600 | 45608 | 7.3 | 5846.3 | 0 | 7395.2 | 5738.2 |
| Ar D 2 | 94.4 | 17 | 2208 | 0.0 | 9350.1 | 7 | 46.0 | 600 | 58225 | 3.0 | 9365.0 | 0 | 11 960.4 | 9350.1 |
| Ar D 3 | 92.8 | 554 | 20763 | 0.0 | 5965.3 | 23 | 52.6 | 600 | 38606 | 8.1 | 6209.7 | 0 | 7647.7 | 5965.3 |
| Ar D 4 | 93.0 | 61 | 5861 | 0.0 | 5945.3 | 18 | 51.4 | 600 | 47777 | 4.0 | 6047.2 | 0 | 7219.0 | 5945.3 |
| Ar D 5 | 97.5 | 12 | 227 | 0.0 | 4860.2 | 6 | 43.2 | 56 | 9116 | 0.0 | 4860.2 | 0 | 5623.4 | 4860.2 |
| Ar D 6 | 94.3 | 26 | 6759 | 0.0 | 6347.6 | 4 | 43.0 | 600 | 37306 | 13.5 | 6770.4 | 0 | 8075.3 | 6347.6 |
| Ar D 7 | 91.9 | 600 | 16898 | 2.6 | 8355.4 | 19 | 47.0 | 600 | 5536 | 31.7 | 10945.3 | 2 | 12 766.1 | 8295.9 |
| Ar D 8 | 95.8 | 14 | 714 | 0.0 | 11489.7 | 11 | 50.6 | 600 | 27869 | 10.9 | 12052.8 | 1 | 15 421.3 | 11489.7 |
| Ar D 9 | 91.1 | 185 | 6000 | 0.0 | 23086.8 | 52 | 40.2 | 600 | 5002 | 39.9 | 33545.2 | 2 | 44 766.7 | 23086.8 |
| Ar T | 97.3 | 13 | 894 | 0.0 | 78115.9 | 0 | 69.7 | 600 | 36720 | 2.3 | 78674.9 | 1 | 86 124.5 | 78115.9 |
| Or T | 97.7 | 62 | 4007 | 0.0 | 98316.1 | 1 | 66.6 | 600 | 13579 | 5.6 | 101412.2 | 2 | 105 726.0 | 98316.1 |
| Et T | 95.4 | 12 | 1004 | 0.0 | 79829.4 | 0 | 73.6 | 600 | 38349 | 8.0 | 82834.1 | 1 | 96 628.0 | 79829.4 |
| Ca T | 97.8 | 38 | 1861 | 0.0 | 138320.4 | 1 | 68.0 | 600 | 8175 | 6.5 | 144233.2 | 2 | 154 858.4 | 138320.4 |

Both enhanced models solved all instances to optimality. The enhanced arc-node model proved faster on 18 of our 23 instances, solving instances in 26 seconds on average, while the enhanced path model needed on average 54 seconds. This was partly due to the number of explored branches during the branch-and-bound process, with an average around 1400 for the arc-node model and around 4500 for the path model. Compared to the base models, the valid inequalities were more beneficial to the path model, since the initial relaxation to optimal solution ratio jumped from 55.7 % to 98.0 %, while the average improvement for the arc-node was about 2.5 %, with an initial relaxation to optimal solution ratio of 97.3 %. One can explain this by the fact that the base arc-node model relaxation was already tight.

Table 4.4: Results of enhanced models

| instance | arc-node model | | | path model | | |
|----------|----------------|----------|--------|----------------|----------|--------|
| | relaxation (%) | time (s) | branch | relaxation (%) | time (s) | branch |
| Et D 1 | 97.2 | 10 | 120 | 98.1 | 16 | 3851 |
| Et D 2 | 98.4 | 5 | 0 | 99.8 | 8 | 922 |
| Et D 3 | 96.2 | 13 | 0 | 98.6 | 56 | 4677 |
| Or D 1 | 97.0 | 4 | 0 | 97.8 | 4 | 0 |
| Or D 2 | 97.8 | 3 | 0 | 99.3 | 1 | 0 |
| Or D 3 | 95.2 | 21 | 1244 | 95.4 | 105 | 15 179 |
| Or D 4 | 97.0 | 4 | 0 | 97.4 | 6 | 330 |
| Or D 5 | 98.4 | 3 | 0 | 98.8 | 4 | 236 |
| Or D 6 | 96.3 | 23 | 3924 | 96.2 | 47 | 8537 |
| Or D 7 | 96.9 | 9 | 79 | 98.9 | 9 | 535 |
| Ar D 1 | 96.9 | 15 | 816 | 96.6 | 22 | 2873 |
| Ar D 2 | 97.2 | 16 | 773 | 99.9 | 16 | 4556 |
| Ar D 3 | 95.2 | 102 | 4986 | 94.8 | 147 | 16 157 |
| Ar D 4 | 95.0 | 63 | 8797 | 95.7 | 25 | 5250 |
| Ar D 5 | 98.2 | 10 | 83 | 98.4 | 2 | 0 |
| Ar D 6 | 96.8 | 20 | 2357 | 98.0 | 11 | 1313 |
| Ar D 7 | 95.5 | 198 | 9506 | 96.4 | 414 | 30 041 |
| Ar D 8 | 98.3 | 5 | 0 | 99.5 | 6 | 0 |
| Ar D 9 | 95.9 | 18 | 0 | 98.3 | 336 | 10 628 |
| Ar T | 99.7 | 1 | 0 | 99.8 | 1 | 0 |
| Or T | 99.4 | 3 | 0 | 99.6 | 2 | 0 |
| Et T | 99.9 | 1 | 0 | 99.9 | 1 | 0 |
| Ca T | 99.3 | 2 | 0 | 99.3 | 6 | 0 |

The average influence of each family of valid inequalities considered individually is displayed in Table 4.5. For any family of valid inequalities, let us note "B1R" the value of the initial relaxation of the base model to which we add only this family, "BMR" the value of the initial relaxation of the base model, "EMR" the value of the initial relaxation of the enhanced model, "E1R" the value of

the initial relaxation of the enhanced model from which we remove this family, and "OS" the optimal solution. For each reinforcement, the value displayed in the line labeled "gap closure base model" is $\frac{B1R-BMR}{OS-BMR}$: it displays the influence of this reinforcement alone compared to the base model. The value displayed in the line labelled "gap closure enhanced model" is $\frac{EMR-E1R}{OS-EMR}$: it displays the influence of this reinforcement alone compared to the enhanced model. Valid inequalities (4.38) and (4.39), as well as (4.40) and (4.41), are displayed together due to their similar meaning.

Table 4.5: Influence of the different families of valid inequalities

| | path model | | | | | |
|--------------------------------|----------------|--------|--------|----------------|----------------|--------|
| valid inequality | (4.11) | (4.12) | (4.13) | (4.14) | (4.15) | (4.16) |
| gap closure base model (%) | 37.1 | 63.0 | 42.2 | 35.8 | 35.6 | 35.6 |
| gap closure enhanced model (%) | 32.4 | 33.3 | 32.0 | 32.0 | 31.5 | 51.0 |
| valid inequality | (4.17) | (4.19) | (4.20) | (4.21) | | |
| gap closure base model (%) | 36.0 | 45.6 | 38.7 | 35.6 | | |
| gap closure enhanced model (%) | 33.2 | 72.2 | 40.6 | 31.9 | | |
| | arc-node model | | | | | |
| valid inequality | (4.35) | (4.36) | (4.37) | (4.38), (4.39) | (4.40), (4.41) | (4.42) |
| gap closure base model (%) | 25.8 | 2.6 | 13.7 | 15.2 | 5.8 | 50.5 |
| gap closure enhanced model (%) | 28.5 | 5.2 | 0.6 | 5.1 | 3.0 | 50.3 |

All families of valid inequalities have some impact on the initial relaxation. Besides, none is completely dominated by the others combined, as shown by the gap closure to the enhanced model. Note that mixed integer rounding inequalities (families (4.19) and (4.42)) have the biggest influence on the relaxation for both models. Valid inequalities related to the path model have much more effect than the ones related to the arc-node model. This was expected, since the initial relaxation of the base arc-node model was already tighter than the one of the base path model.

Sensitivity to the number of cables

The number of variables of each model depends directly on the cable sizes considered. To highlight the impact of those parameters on the enhanced models, we introduced new cable sizes of 30, 36, 42 and 48 modules for the distribution zones, and new cable sizes of 24, 28, 32 and 40 modules for the transport zones. The results are displayed in Table 4.6.

The initial relaxation for both models stayed in the same range than for the 8 cables data, with an average of 98.1 % for the path model, and of 97.4 % for the arc-node model. This ratio was larger for the path model for 21 of our instances. Furthermore, the path model was faster for 20 instances out of 23, with an average computation time of 57 seconds, while the average computation time of the arc-node model was 88 seconds.

One can assume this tendency is due to the variation of the number of variables with respect to cable sizes. For the path model, this number grows affinely with the number of cable sizes (L , see Section 4.2.1), while for the arc-node model, this number grows linearly with the sum of cable sizes ($\sum_{l \in \mathcal{L}} M_l$, see Section 4.3.2). This suggests that with a lot of different available cables, one should favor the path approach.

Table 4.6: Performance of enhanced models with 4 additional cable types

| instance | path model | | | arc node model | | |
|----------|------------|------|----------|----------------|------|----------|
| | relaxation | time | branches | relaxation | time | branches |
| Et D 1 | 97.9 | 24 | 4688 | 97.1 | 24 | 144 |
| Et D 2 | 99.8 | 7 | 447 | 98.4 | 15 | 0 |
| Et D 3 | 98.9 | 121 | 8618 | 98.1 | 44 | 1795 |
| Or D 1 | 97.8 | 5 | 0 | 97.0 | 13 | 0 |
| Or D 2 | 99.3 | 1 | 0 | 97.8 | 9 | 0 |
| Or D 3 | 95.8 | 81 | 7472 | 95.2 | 66 | 3410 |
| Or D 4 | 97.5 | 6 | 315 | 97.0 | 9 | 0 |
| Or D 5 | 99.1 | 5 | 135 | 98.4 | 12 | 0 |
| Or D 6 | 95.9 | 21 | 5269 | 96.5 | 103 | 9279 |
| Or D 7 | 98.8 | 10 | 649 | 96.9 | 19 | 0 |
| Ar D 1 | 96.1 | 36 | 5121 | 96.7 | 61 | 3371 |
| Ar D 2 | 97.1 | 48 | 6918 | 96.8 | 60 | 3593 |
| Ar D 3 | 94.0 | 59 | 3168 | 93.6 | 416 | 5744 |
| Ar D 4 | 97.1 | 9 | 0 | 97.5 | 259 | 3593 |
| Ar D 5 | 98.8 | 2 | 0 | 98.2 | 11 | 0 |
| Ar D 6 | 97.7 | 14 | 1267 | 96.5 | 45 | 3917 |
| Ar D 7 | 97.5 | 608 | 21 703 | 95.4 | 758 | 9819 |
| Ar D 8 | 99.4 | 8 | 0 | 98.3 | 17 | 0 |
| Ar D 9 | 99.3 | 234 | 8173 | 98.2 | 46 | 0 |
| Ar T | 99.8 | 2 | 0 | 99.7 | 6 | 0 |
| Or T | 99.5 | 5 | 0 | 99.4 | 10 | 0 |
| Et T | 99.9 | 1 | 0 | 99.9 | 3 | 0 |
| Ca T | 99.3 | 7 | 0 | 99.3 | 10 | 0 |

Sensitivity to the demand

To assess the influence of the number of demand nodes, we built fictive instances based on the civil engineering structure of Orvault distribution instances (labelled "Or D"), on which a demand of 1 active module has been added at some nodes previously without demand. The number of demand nodes added is a percentage of the existing number of demand nodes. The results related to the average computation time are displayed in Figure 4.6.

The impact on the computation time is significant, the computation times for of the arc-node and path models being multiplied respectively by more than 30 and 6. For the arc-node model, the computation time grows from 10 seconds (reached with no fictive demand nodes) to 357 seconds (with 90% additional demand nodes). Regarding the path model, the computation time grows from 54 seconds (reached with no fictive demand nodes) to 249 seconds (with 90% additional demand nodes).

This significant growth can be explained for the arc-node model by the additional variable u which is defined on demand nodes, as well as the relative sim-

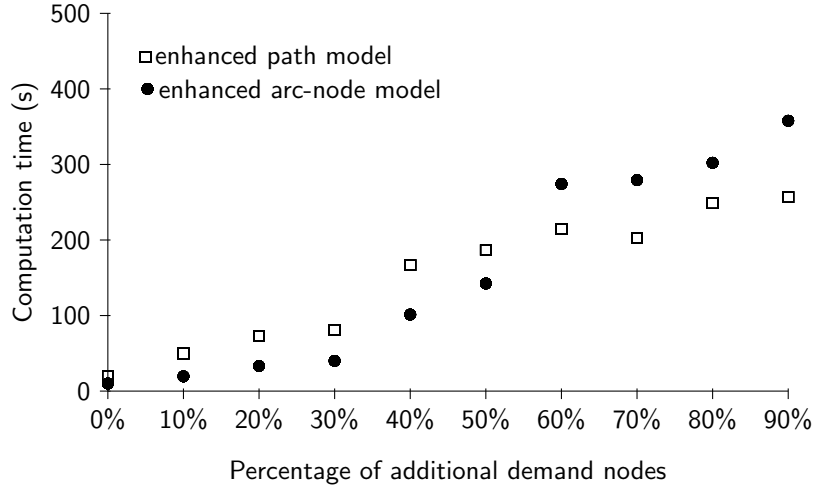


Figure 4.6: Evolution of the average computation time with respect to the number of demand nodes on fictive instances

plicity of equations (4.28) for no-demand nodes compared to equations (4.29), (4.30) and (4.31) for nodes with demand. As for the path model, it can be explained by a larger number of variables on nodes with demand. Indeed, for a demand node i and a path p finishing in i ($t(p) = i$), for $l \in \mathcal{L}$, we define variables $b_{p,l}^{born,dem}$ and $b_{p,l}^{ctn,dem}$ describing a cable serving the demand. These are not defined on a no-demand node.

4.4.3 Optimal design characteristics and operational recommendations

Table 4.7 reports some characteristics of the optimal designs. To gain insight in the structure of optimal deployments, we choose to display the number of splicing and tapping points, the number of separation points located on nodes of degree 3 or more (column “separation in deg 3+”), the number of cables going out of the root, and the percentages of cable deployment and separation costs (cable separation cost being the addition of boxes and welds costs).

Table 4.7: Solutions features

| instance | splicing points | tapping points | separation in deg 3+ | cables at the root | cable deployment cost | cable separation cost |
|----------|-----------------|----------------|----------------------|--------------------|-----------------------|-----------------------|
| Et D 1 | 0 | 13 | 13 | 14 | 87.5 | 12.5 |
| Et D 2 | 1 | 7 | 8 | 3 | 91.7 | 8.3 |
| Et D 3 | 1 | 16 | 17 | 11 | 92.7 | 7.3 |
| Or D 1 | 0 | 9 | 9 | 30 | 83.2 | 16.8 |
| Or D 2 | 0 | 9 | 9 | 17 | 80.0 | 20.0 |
| Or D 3 | 0 | 14 | 14 | 18 | 78.0 | 22.0 |
| Or D 4 | 0 | 7 | 7 | 9 | 80.4 | 19.6 |
| Or D 5 | 0 | 5 | 5 | 19 | 81.7 | 18.3 |
| Or D 6 | 1 | 10 | 10 | 13 | 76.5 | 23.5 |
| Or D 7 | 0 | 9 | 9 | 23 | 82.5 | 17.5 |
| Ar D 1 | 0 | 8 | 8 | 10 | 74.9 | 25.1 |
| Ar D 2 | 0 | 11 | 10 | 12 | 86.3 | 13.4 |
| Ar D 3 | 0 | 4 | 4 | 7 | 73.6 | 16.4 |
| Ar D 4 | 0 | 6 | 6 | 11 | 73.6 | 16.4 |
| Ar D 5 | 0 | 5 | 5 | 21 | 77.1 | 22.9 |
| Ar D 6 | 0 | 7 | 6 | 25 | 84.4 | 15.6 |
| Ar D 7 | 0 | 9 | 9 | 4 | 75.4 | 24.6 |
| Ar D 8 | 0 | 10 | 10 | 12 | 86.3 | 13.7 |
| Ar D 9 | 1 | 24 | 25 | 5 | 82.9 | 17.1 |
| Ar T | 1 | 11 | 12 | 31 | 90.6 | 9.4 |
| Or T | 1 | 14 | 15 | 58 | 91.8 | 8.2 |
| Et T | 2 | 6 | 8 | 9 | 92.4 | 7.6 |
| Ca T | 4 | 17 | 21 | 40 | 92.0 | 8.0 |

First, note that despite inequality (4.16), which states that separation points are performed rather in the last node of a series of adjacent degree 2 nodes without demand, 1.2% of separation points are located in nodes of degree 2 without demand. This illustrates the necessity of keeping such nodes. Theoretically it is imposed by the engineering rule allowing at most one cable separation in a node. However, this could be used for designing a presumably good quality heuristic solution where no separation points would be allowed in nodes of degree 2. It can be implemented by merging arcs of such series, thus reducing the problem size.

On a different note, observe that 95% of the separation points of optimal designs are tapping points. By comparing the instances for which splicing points are chosen to the average root-demand distance (see Table 4.1), we can note that the splicing is usually chosen for the instances where the demands are located far away from the root. This illustrates the asymptotic behavior of the problem regarding distances. For short distances, cables cost becomes negligible, and no separation will be done. For long distances, welds costs become negligible, and splicing operations will be chosen. This could be used to simplify the problem

with small losses for short instances. Excluding the splicing option can ease decision making while providing a solution close to optimal.

To get a deeper insight regarding solution shapes, the number of cables coming out of the root node is worth being discussed for two main reasons. First, it shows that despite the economy of scale of cable capacity, the optimal solution is often far from using only the biggest cables available, even at the root. Furthermore, it illustrates some diversity of the optimal design depending on the instance.

As for cost sources, the most expensive part of the solutions is always the cable deployment cost, which stands on average for 83 % of the total cost, with some variation. Despite the fact that cable separation costs are relatively low, the impact of separation techniques is quite high as already shown by the greedy solution (cheapest solution obtained without separation). These observations should be useful to decision makers in charge of strategic choices prior to the network design itself (choice of the subcontractor, cable and weld cost, cable manufacturer, ...).

4.4.4 Restricted problem analysis

Section 4.4.3 suggests that a solution where only tapping is allowed would be of good quality. We studied two restricted cases: when only tapping is allowed, and when only splicing is allowed. The cheapest solution with only splicing can be obtained by giving zero value to some variables:

- For the path model, $\forall p \in \mathcal{P}, \forall l \in \mathcal{L}, b_{p,l}^{ctn,sep} = b_{p,l}^{ctn,dem} = 0;$
 $\forall p \in \mathcal{P}, m_p^{ctn,sep} = 0$
- For the arc-node model, $\forall i \in V^*, \forall l \in \mathcal{L}, \forall \mathcal{M}_l, b_{i,l,m}^{tap} = b_{i,l,m}^{ctn} = 0$

The cheapest solution with only tapping allowed can be obtained as follows:

- For the path model, inequality (4.6) becomes
 $\forall i \in V^*, \forall l \in \mathcal{L}, \sum_{p \in \mathcal{P} | s(p)=i} b_{p,l}^{ctn,sep} + b_{p,l}^{ctn,dem} = \sum_{p \in \mathcal{P} | t(p)=i} b_{p,l}^{born,sep} + b_{p,l}^{ctn,sep}$. It specifies that in every separation point there should be a continued cable, implying that this separation point is a tapping.
- For the arc-node model, one can give value 0 to the following variables
 $\forall i \in V^*, \forall l \in \mathcal{L}, \forall m \in \mathcal{M}_l, b_{i,l,m}^{spl} = 0$

All inequalities developed in Sections 4.2.4 and 4.3.4 are still valid for the restricted problems. The results are displayed in Table 4.8.

Table 4.8: Restricted problem solutions

| instance | only splicing allowed | | | only tapping allowed | | |
|----------|---------------------------|-------------------------------|------------------|---------------------------|-------------------------------|------------------|
| | path model time (s) | arc-node model time (s) | solution cost | path model time (s) | arc-node model time (s) | solution cost |
| Et D 1 | 2 | 3 | 16 460.7 | 14 | 12 | 15 958.1 |
| Et D 2 | 1 | 1 | 27 710.3 | 7 | 2 | 27 245.7 |
| Et D 3 | 9 | 5 | 33 129.2 | 103 | 11 | 32 696.4 |
| Or D 1 | 1 | 1 | 7003.8 | 5 | 3 | 6803.2 |
| Or D 2 | 0 | 2 | 6748.6 | 1 | 2 | 6363.8 |
| Or D 3 | 6 | 12 | 8541.3 | 72 | 15 | 7932.6 |
| Or D 4 | 1 | 4 | 4138.8 | 7 | 3 | 3941.9 |
| Or D 5 | 0 | 2 | 4459.9 | 4 | 3 | 4233.8 |
| Or D 6 | 3 | 6 | 9318.1 | 17 | 16 | 9017.2 |
| Or D 7 | 1 | 2 | 5577.0 | 9 | 5 | 5206.7 |
| Ar D 1 | 2 | 7 | 6087.3 | 20 | 13 | 5738.2 |
| Ar D 2 | 4 | 8 | 9869.0 | 14 | 10 | 9350.1 |
| Ar D 3 | 3 | 4 | 6125.8 | 119 | 94 | 5965.3 |
| Ar D 4 | 4 | 37 | 6461.9 | 20 | 50 | 5945.3 |
| Ar D 5 | 0 | 2 | 5081.8 | 1 | 3 | 4860.2 |
| Ar D 6 | 2 | 5 | 6544.2 | 11 | 14 | 6347.6 |
| Ar D 7 | 8 | 7 | 8638.6 | 398 | 181 | 8295.9 |
| Ar D 8 | 2 | 3 | 12 248.4 | 5 | 4 | 11 519.7 |
| Ar D 9 | 42 | 21 | 24 422.8 | 249 | 12 | 23 112.5 |
| Ar T | 1 | 3 | 79 288.2 | 2 | 10 | 78 116.7 |
| Or T | 2 | 7 | 99 450.6 | 3 | 12 | 98 373.3 |
| Et T | 1 | 3 | 80 602.2 | 1 | 5 | 80 239.3 |
| Ca T | 3 | 7 | 139 572.6 | 7 | 32 | 138 799.2 |

The cost of the solutions with only tapping was on average 0.2 % more expensive than the optimal solution, while the cost of the solution with only splicing was on average 4.2% more expensive than the optimal solution. It was expected that the solution with only tapping would be of better quality than the solution with only splicing (see column 2 and 3 of Table 4.7).

Let us focus on the computation time of the problem with only splicing. The average time was of 7 seconds with the arc-node model, and 5 seconds for the path model. These are significantly lower than the computation times of the nominal (unconstrained, tapping and splicing allowed) problem, which were on average of 26 and 54 seconds.

The computation times for the problem with only tapping was higher than the computation time with only splicing: 24 seconds on average for the arc-node model and 57 seconds on average for the path model. The difference between the two models (path and arc-node) can be explained here by the difference in the formulation simplifications: for the arc-node model, the restriction was expressed by setting to zero some variables, which makes the computations

easier, while for the path model, the restriction was expressed by adding some constraints. Moreover, compared to the nominal problem, the computation times are comparable.

4.5 Conclusion

This chapter addresses the problem of fiber cables network design in tree shaped infrastructures via integer linear programming methods. Two main models are introduced, one of them is based on paths, while the other is based on arcs. The problem is also proven to be NP-hard, and hard to approximate within a ratio smaller than $\frac{3}{2}$.

The integer programming approaches proposed here are efficient enough to solve real-life instances. Each model is enhanced thanks to valid inequalities. The enhancements introduced sensitively improve the computational performances.

Regarding cost savings, by introducing the possibility of doing some splicing and tapping operations, one can save around 25% of the total cost. This is comparable to the savings obtained with the FCND problem, for which more possibilities were available.

The work done in this chapter will be the basis for the implementation of a decision aid tool at Orange. Indeed, it allows significant cost savings, while not being too time-expensive on real-life instances. Furthermore, additional engineering rules can be added for both models if needed, such as duct capacities for instance. This is especially true for the path-based formulation. However, some operations effectively used by technicians are out of reach with this model. The following chapter tackles backfeed, which is a possible operation not covered by these formulations.

Chapter 5

Backfeed in Fiber Cables Network Design

Chapter Abstract

This chapter tackles the problem of Backfeed Fiber Cables Network Design (BFCND). Recall that it can be stated as follows: given a civil engineering arborescence, a root node, demand nodes, a set of available cables and the associated costs, find the fiber cables network design with minimal cost, while respecting the engineering rules listed in Section 2.1. At most one cable can be spliced in a cabinet (no tapping is allowed). One full cable must serve each demand (cable-served way only). Cables can be deployed on a path going partly "upstream" the civil engineering structure.

An integer programming based solution is proposed, and some associated valid inequalities are introduced in Section 5.1. The problem is proven to be NP-hard in Section 5.2. The formulation is assessed on real-life instances in Section 5.3.

Hypothesis

An arborescence of the civil engineering structure is given, as well as demand points, and the number of fiber modules to be brought at each demand. This implies that the number of modules going through each duct is known. In this problem, the only cable separation technique allowed is splicing. Furthermore, all demand nodes are assumed to be cable-served leaf nodes.

We allow for the cables to be deployed upstream the civil engineering structure, which is referred to as "backfeed" (see Fig. 5.1).

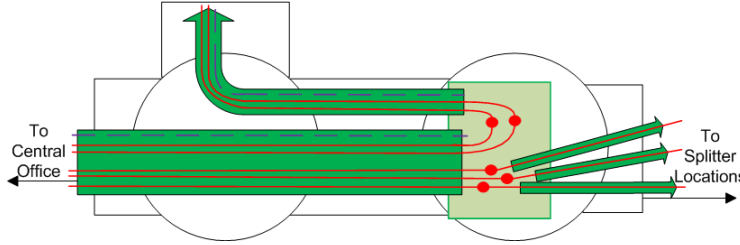


Figure 5.1: Illustration of backfeed

The maintenance constraint does not apply.

5.1 Integer Programming Formulation

5.1.1 Formulation

A non-oriented tree $G = (V, E)$ describes the selected civil engineering structure where V denotes the set of concrete rooms and E represents the set of ducts. Let r be a pendant node representing the CO and let $V^* = V \setminus r$. The set of oriented cable paths in G is denoted by \mathcal{P} . Given any $p \in \mathcal{P}$, $s(p) \in V$ is the origin node of p and $t(p) \in V^*$ is its target node. $V_D \subset V^*$ denotes the set of pendant vertices of G excluding r and $V_N = V^* \setminus V_D$. Each pendant node $i \in V_D$ has a non-null demand D_i .

The cables available are from a discrete set $\mathcal{L} = \{1, \dots, L\}$, each cable type being characterized by a number of modules M_l with $l \in \{1, \dots, L\}$. \mathcal{L} is ordered with respect to M , i.e. $\forall (l, l') \in \mathcal{L}^2, l \leq l' \iff M_l \leq M_{l'}$. The set of possible numbers of active modules inside a cable of type l is denoted by $\mathcal{M}_l = \{1, \dots, M_l\}$. The cost for laying out a cable of size $l \in \mathcal{L}$ on path $p \in \mathcal{P}$ is $C_{p,l}$, while the cost of a welding box of type l is PB_l , and the cost for welding m modules is PW_m , all costs being strictly positive.

The decision variables of the model are described below:

- $\forall p \in \mathcal{P}, \forall l \in \mathcal{L}, b_{p,l}$ is the binary variable equal to 1 iff there is a cable of size l set in path p , born in $s(p)$ and either spliced or serving a demand in

$t(p)$ (it is spliced if $t(p) \in V_N$ and serving a demand if $t(p) \in V_D$). Recall that $s(p)$ can be equal to the root r .

- $\forall p \in \mathcal{P}, m_p$ denotes the number of active modules of the cable laid out in path p , born in $s(p)$ and either spliced or serving a demand in $t(p)$.
- $\forall i \in V_N, \forall m \in \mathcal{M}_L, w_{i,m}$ is the binary variable equal to 1 iff there are m modules to be weld in node i .

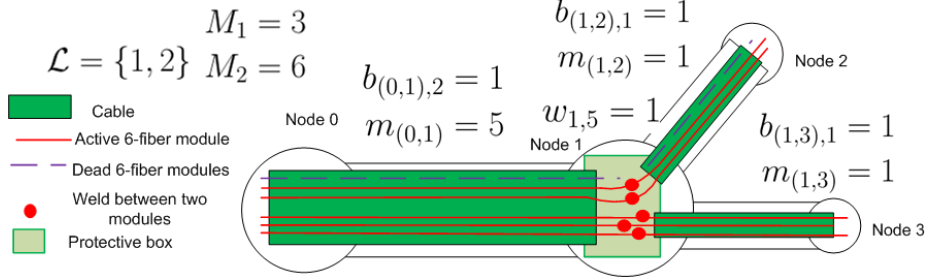


Figure 5.2: Illustration of the model on a small example. A cable is spliced in node 1.

For the sake of clarity, the model is illustrated in figure 5.2. The BFCND problem can be formulated as follows:

$$\min \sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} C_{p,l} \cdot b_{p,l} + \sum_{i \in V_N} \sum_{m \in \mathcal{M}_L} PW_m \cdot w_{i,m} + \sum_{p \in \mathcal{P} | t(p) \in V_N} \sum_{l \in \mathcal{L}} PB_l \cdot b_{p,l}$$

such that

$$\forall i \in V_N, \sum_{p \in \mathcal{P} | t(p)=i} \sum_{l \in \mathcal{L}} b_{p,l} \leq 1 \quad (5.1)$$

$$\forall i \in V_D, l_1 = \min\{l \in \mathcal{L} | M_l \geq D_i\}, \sum_{p \in \mathcal{P} | t(p)=i} b_{p,l_1} = 1$$

$$\sum_{p \in \mathcal{P} | t(p)=i} \sum_{l \in \mathcal{L} \setminus l_1} b_{p,l} = 0 \quad (5.2)$$

$$\forall p \in \mathcal{P} | t(p) \in V_D, m_p = b_{p,l_1} \cdot D_i \quad (5.3)$$

$$\forall p \in \mathcal{P}, \sum_{l \in \mathcal{L}} M_l \cdot b_{p,l} \geq m_p \quad (5.4)$$

$$\forall i \in V_N, \sum_{p \in \mathcal{P} | t(p)=i} m_p = \sum_{p \in \mathcal{P} | s(p)=i} m_p \quad (5.5)$$

$$\forall i \in V_N, \sum_{m \in \mathcal{M}_L} m \cdot w_{i,m} = \sum_{p \in \mathcal{P} | i=s(p)} m_p \quad (5.6)$$

$$\forall i \in V_N, \sum_{m \in \mathcal{M}_L} w_{i,m} \leq 1 \quad (5.7)$$

The cost function is the sum of the cost of cable deployment, the cost of welds, and the cost of protective boxes.

Equations (5.1) ensure that at most one cable is spliced in a no-demand node. Constraints (5.2) ensure the demand is served by exactly one cable. The size of this cable is known, it is the smallest size which can contain the demand. Constraints (5.3) ensure that the cable serving the demand holds enough active modules. (5.4) make sure that cables are large enough to hold their number of active modules. Equations (5.5) ensure active modules conservation in a splicing. Note that as long as (5.1) and (5.4) stand, then at most one of the terms of the left side of (5.5) is non-null. Finally, (5.6) and (5.7) ensure the number of welds to be done is properly counted.

5.1.2 Valid Inequalities

The equations below state that for a given number of active modules, the size of a cable must be the smallest one able to hold as many active modules.

Proposition 5.1.1 *Any optimal solution of the BFCND problem verifies*

$$\forall p \in \mathcal{P} | t(p) \in V_N, m_p \geq \sum_{l \in \mathcal{L}} b_{p,l} + \sum_{l \in \mathcal{L} \setminus 1} M_{l-1} \cdot b_{p,l} \quad (5.8)$$

Proof. Let us assume that in an optimal solution S of the problem, $\exists p \in \mathcal{P}$ such that $t(p) \in V_N, \exists l \in \mathcal{L}, m_p < \sum_{l \in \mathcal{L}} b_{p,l} + \sum_{l \in \mathcal{L} \setminus 1} M_{l-1} \cdot b_{p,l}$. This implies $\sum_{l \in \mathcal{L}} b_{p,l} = 1$, and we have $m_p \leq \sum_{l \in \mathcal{L} \setminus 1} M_{l-1} \cdot b_{p,l}$. Let us consider the solution S' identical to S everywhere but in $b'_{p,l} = 0$ and $b'_{p,l-1} = 1$. Then, the only constraint where terms change from S to S' is (5.4) in p , and we can see that S' is feasible. Since it is also cheaper, we have the result. \square

These equations denote the simultaneous appearance of welds and splicing points. Besides, they also take into account that in the splicing of a cable of size l , there are at least $M_{l-1} + 1$ and at most M_l welds.

Proposition 5.1.2 *The following inequalities are valid for the BFCND problem*

$$\forall i \in V_N, \forall l \in \mathcal{L}, \sum_{m \in \mathcal{M}_l \setminus \mathcal{M}_{l-1}} w_{i,m} = \sum_{p \in \mathcal{P} | t(p)=i} b_{p,l} \quad (5.9)$$

Proof. Let us consider an optimal solution S of the problem. Let us consider $i \in V_N$ and $l \in \mathcal{L}$. If $\sum_{p \in \mathcal{P} | t(p)=i} b_{p,l} = 1$, then (5.4) and (5.8) give $M_{l-1} + 1 \leq m_p \leq M_l$. Using (5.5) and (5.6) gives $\sum_{m \in \mathcal{M}_l \setminus \mathcal{M}_{l-1}} w_{i,m} = 1$. Reciprocally, if $\sum_{m \in \mathcal{M}_l \setminus \mathcal{M}_{l-1}} w_{i,m} = 1$, then (5.5) and (5.6) give $M_{l-1} + 1 \leq \sum_{p \in \mathcal{P} | t(p)=i} m_p \leq M_l$ which implies $\sum_{p \in \mathcal{P} | t(p)=i} b_{p,l} = 1$. \square

Let us note $G' = (V, A)$ the arborescence obtained by orienting G with r as a root. For $i \in V$, we note $\Gamma^+(i) = \{j \in V \mid (i, j) \in A\}$ and $\forall i \in V^*$, $\gamma(i)$ denotes the unique predecessor of i . A lower bound of the number of active modules going through $(i, j) \in A$, denoted by $m_{i,j}^{act}$, can be calculated recursively starting with pendant vertices i by assigning D_i to $m_{\gamma(i)i}^{act}$ and writing that $m_{i,j}^{act} = \sum_{j' \in \Gamma^+(j)} m_{j,j'}^{act}$.

Proposition 5.1.3 *The following equations are valid inequalities for the BFCND problem*

$$\forall a \in A, \forall x \in \mathbb{R} \mid x > 0, \quad \sum_{p \in \mathcal{P} \mid a \in p} \sum_{l \in \mathcal{L}} \left\lceil \frac{M_l}{x} \right\rceil b_{p,l} \geq \left\lceil \frac{m_a^{act}}{x} \right\rceil \quad (5.10)$$

Proof. By starting from the number of active modules going through each edge, we have $\sum_{p \in \mathcal{P} \mid (i,j) \in p} m_p \geq m_{i,j}^{act}$. Furthermore, we can upper bound the left hand term using (5.4), which gives

$$\sum_{p \in \mathcal{P} \mid (i,j) \in p} \sum_{l \in \mathcal{L}} b_{p,l} \geq m_{i,j}^{act}$$

Dividing by some $x \in \mathbb{R}$ and rounding gives the result. \square

Remark 5.1.1 *If $\mathcal{P} = \mathcal{P}^A$ (no backfeed is allowed), some variables can be eliminated from the model. In this case, it is not possible to have a number of welds in i larger than $m_{\gamma(i),i}^{act}$. In other words, $\forall m \in \mathcal{M}_L$, if $m > m_{\gamma(i),i}^{act}$, then $w_{i,m} = 0$.*

5.2 Complexity

Let us show that BFCND is NP-hard. The next proof uses the same reduction than the one used in Section 4.1.

Proposition 5.2.1 *BFCND is NP-complete even with 1 cable size and 2 potential splicing locations.*

Proof. We consider the Number Partitioning Problem (NPP), which is proven to be NP-complete in [34], and we prove here that it can be reduced in polynomial time to a particular instance of BFCND with an equivalent answer.

(NPP):

Instance: We have a set of N integers n_i ($\sum_{i \in \{1, \dots, N\}} n_i$ is supposed to be even).

Question: Is there a subset $S \subseteq \{1, \dots, N\}$ such that $\sum_{i \in S} n_i = \sum_{i \notin S} n_i$?

We consider an instance of (NPP) that we associate to the following BFCND instance.

Let (V, A) be an arborescence describing the civil engineering structure ($V = \{r, 0, 1\} \cup \{v_i \mid i \in \{1, \dots, N\}\}$, $A = \{(r, 0); (0, 1); (1, v_i) \mid i \in \{1, \dots, N\}\}$); only one type of cable with a number of modules $M_1 = \frac{1}{2} \sum_{i \in \{1, \dots, N\}} n_i$ is available, its cost per length unit is $C_1 = 1$. The length of all arcs of the arborescence are zero, except $(r, 0)$ which is of length 1. This means the cost of a cable created in

r is 1, and the cost of the other ones is 0. The number of active modules associated with each arc are: $m_{(r,0)}^{act} = m_{(0,1)}^{act} = \sum_{i=1}^N n_i; \forall i \in \{1, \dots, N\}, m_{(1,v_i)}^{act} = n_i$, which means that the demand points are the $v_i, i \in \{1, \dots, N\}$ and have respective demands n_i . This network is represented in Fig. 5.3a. We consider a zero cost for welding and welding boxes.

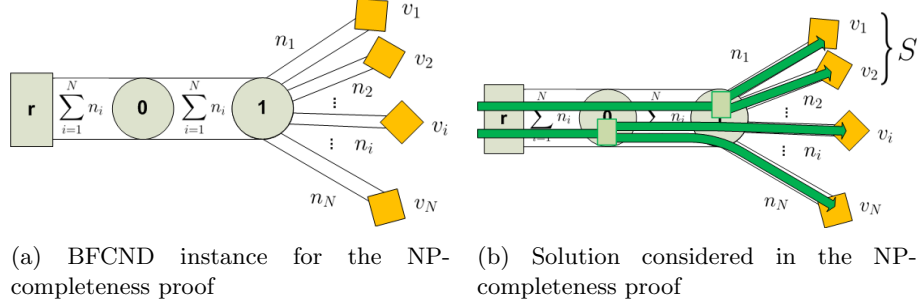


Figure 5.3: Illustration of the instance and solution considered for the NP-completeness proof

The question associated to this BFCND instance is "Is there a cabling solution cheaper than 2?"

Let us first assume that (NPP) is feasible: $\exists S \subseteq \{1, \dots, N\}$ such that $\sum_{i \in S} n_i = \sum_{i \notin S} n_i$. We then build the following cabling solution:

- Two cables holding only active modules are installed on link $(r, 0)$.
- In node 0, one incoming cable is spliced into $N - |S|$ born cables. The born cables have a number of active modules $n_i, i \notin S$ and serve the demand nodes $(v_i)_{i \notin S}$.
- On link $(0, 1)$, one cable coming from r with only active modules, and $N - |S|$ cables serving demand nodes in $\{v_i | i \notin S\}$ are installed.
- In node 1, the incoming cable with only active modules is spliced into $|S|$ born cables. The born cables have n_i active modules and serve the demand nodes $(v_i)_{i \in S}$.
- One cable is installed on each link $(1, v_i)$.

Since the number of active modules is conserved in each splicing, the cabling solution described above is feasible (it is illustrated in Fig. 5.3b). Its cost is equal to 2, as the cables created in r have a cost of 1, and the other ones have a cost of 0.

Inversely, let us assume that (NPP) is not feasible: then, the solution described above is not possible anymore. One cable is not large enough to cover link $(r, 0)$, it cannot contain all the required active modules. Let us assume there is a solution with only two cables on $(r, 0)$. Since their combined number of modules is $\sum_{i \in \{1, \dots, N\}} n_i$, they both hold only active modules. If one

of them directly served the demand without enduring any operation, then the (NPP) instance was trivially feasible (one of the n_i is half the total sum). So both of them endure a splicing operation, one in node 0, the other in node 1. Let us consider the cables created in 1. They serve a subset S_1 of the demand nodes, and have a respective number of active modules of $n_i, i \in S_1$. Since the number of active modules in a splicing operation is conserved, we have $\sum_{i \in S_1} n_i = \frac{1}{2} \sum_{i \in \{1, \dots, N\}} n_i$ and the (NPP) instance was feasible.

Consequently, at least 3 cables need to be installed on arc $(r, 0)$, and such solution has a cost of at least 3.

Hence the two instances are equivalent, and BFCND is NP-hard. \square

The results proven for FCNDA in Section 4.1 can also be proven for BFCND using this reduction.

In the following, "base model" denotes the model with equations (5.1) to (5.7) while "enhanced model" denotes the model with equations (5.1) to (5.9). Let us denote $G' = (V, A)$ the arborescence obtained by orienting G with r as a root. In the following, \mathcal{P}^A denotes all paths in the arborescence G' while \mathcal{P}^E denotes all oriented paths in the non-oriented tree G .

5.3 Numerical Results

The models were assessed on real-life instances extracted from french cities with various population densities. The costs used were those currently paid by the operator Orange to sub-contractors in charge of the network deployment. We use 7 different types of cables, which hold 1,2,4,8,12,16 or 20 modules. The instances were solved with the Branch-and-Bound algorithm of Cplex 12.6.0 in its default configuration. Key features of each instance are presented in table 5.1.

Table 5.1: Instances key features

| instance | features | | |
|-------------|----------|---------|-------------------|
| | $ V_N $ | $ V_D $ | $ \mathcal{P}^A $ |
| Arles | 139 | 65 | 2673 |
| Clermont | 194 | 95 | 4079 |
| Lannion | 75 | 35 | 1605 |
| Carcassonne | 195 | 88 | 4091 |
| Brest | 153 | 74 | 2744 |
| Lorient | 129 | 60 | 2405 |
| Merignac | 86 | 41 | 1356 |

Table 5.2 reports algorithmic results when no backfeed is allowed. The column labeled "LR" displays the ratio $\frac{\text{linear relaxation}}{\text{optimal solution}}$, the column labeled "time" displays the computation time for each optimal solution, the column labeled

"branches" displays the number of nodes explored by the Branch-and-Bound algorithm, and the column labeled "gap closure" reports the ratio $\frac{\text{enhanced model LR} - \text{base model LR}}{\text{optimal solution} - \text{base model LR}}$.

Table 5.2: Algorithmic results

| instance | base model | | | enhanced model | | |
|-------------|------------|----------|----------|-----------------|----------|----------|
| | LR (%) | time (s) | branches | gap closure (%) | time (s) | branches |
| Arles | 85.4 | 181 | 74122 | 15.6 | 10 | 0 |
| Clermont | 84.8 | 544 | 129 606 | 16.0 | 23 | 0 |
| Lannion | 87.8 | 456 | 204 158 | 17.8 | 9 | 0 |
| Carcassonne | 84.0 | 159 | 40 616 | 21.4 | 22 | 0 |
| Brest | 85.9 | 141 | 47 848 | 21.7 | 9 | 0 |
| Lorient | 87.0 | 457 | 250 109 | 19.8 | 13 | 70 |
| Merignac | 83.9 | 6 | 7669 | 27.5 | 2 | 0 |

The model was able to solve all the instances proposed in less than 600 seconds, and the valid inequalities made the resolution easier. The average computation time was reduced from 277 seconds with the base model to 13 seconds with the enhanced model. As for the computational side, the proposed valid inequalities tighten the linear relaxation. The gap between the linear relaxation and the optimal solution is reduced in average by 20 % in the enhanced model.

As for operational aspects, we define three different backfeed policies:

I No backfeed allowed. In this case, $\mathcal{P} = \mathcal{P}^A$.

II Backfeed allowed for cables serving the demand only. In this case, $\mathcal{P} = \mathcal{P}^A \cup \{p \in \mathcal{P}^E, t(p) \in V_D\}$. This is the choice currently favored by Orange.

III Backfeed allowed for every cable. This means $\mathcal{P} = \mathcal{P}^E$ ($\mathcal{P}^A \subset \mathcal{P}^E$).

The cost and computation time of the optimal solutions obtained with each strategy are reported in table 5.3.

The solution without backfeed was in average 25 % more expensive than the one with all backfeed allowed. This shows how allowing this technique can be useful. On a different note, the cost of the solution with only last cable backfeed allowed is almost the same as the one with all backfeed allowed, changing in only 2 instances and being in average only 0.3% more expensive. This shows that most backfeed is done closer to the splitters than to the central office. One may interpret this as a consequence of the concave cost of cables with respect to the diameter: the relatively small linear cost of large cables coming out of the root makes them go close to demand points without backfeed, and these cables are then spliced into several smaller ones serving the demand with backfeed.

The computation time in each case increases with the number of paths. It is in average 12 seconds with policy I, 37 seconds with policy II and 387 seconds with policy III. This was expected since the number of variables increases

Table 5.3: Policies assessment

| instance | Policy I | | Policy II | | Policy III | |
|-------------|----------|-----------|-----------|-----------|------------|-----------|
| | time | solution | time | solution | time | solution |
| Arles | 10 | 88 880.4 | 21 | 68 064.8 | 112 | 68 064.8 |
| Clermont | 23 | 113 100.5 | 94 | 84 717.9 | 828 | 84 331.0 |
| Lannion | 9 | 88 835.6 | 11 | 75 791.1 | 46 | 75 791.1 |
| Carcassonne | 22 | 157 347.2 | 71 | 130 922.2 | 988 | 130 922.2 |
| Brest | 141 | 101 140.1 | 39 | 82 447.5 | 576 | 82 447.5 |
| Lorient | 13 | 97 899.2 | 20 | 76 720.6 | 134 | 76 720.6 |
| Merignac | 2 | 55 486.5 | 6 | 46 008.3 | 27 | 45 132.5 |

significantly with looser policies. In case of instances are too large to be solved with policy III, one can assume that choosing policy II instead would give a very good sub-optimal solution quickly.

5.4 Conclusion

The introduction of backfeed into the cable network optimization process enables significant savings at the expense of harder computations. The effect of the valid inequalities provided for our integer programming formulation suggests that it can still be further tightened in order to reduce computation times.

On an operational level, backfeed makes network operations, administration and maintenance activities more complex. This leads us to the following chapter, in which we introduce additional constraints meant to facilitate these future activities.

Chapter 6

Easy-Maintenance Fiber Cables Network Design

Chapter Abstract

This chapter deals with the maintenance constraint introduced in Section 2.1, and illustrated in Figure 6.1. It tackles two problems which differ by the introduction of this constraint, the unconstrained problem referred to as SFCND and the constrained problem referred to as ESFCND. SFCND can be defined as a restriction of FCNDA where no tapping is allowed. ESFCND can be defined as a restriction of SFCND where the maintenance constraint has to be respected. Recall that the maintenance constraint can be formulated as follows: cables going through a given duct must be born in the same cabinet (see Fig. 2.8).

For each problem, a specific integer programming solution is proposed in Section 6.1. Valid inequalities are also introduced for both problems. Then, Section 6.2 presents two algorithms for the resolution of ESFCND: an exact dynamic programming algorithm in Section 6.2.1 as well as a FPTAS in Section 6.2.2. The theoretical complexities of both problems are proven and compared in different contexts in Section 6.3. Finally, the efficiency of the models as well as the properties of the solutions are assessed in Section 6.4.

Hypothesis

In the two problems considered in this chapter, an arborescence of the civil engineering structure is given, as well as demand points, and the number of fiber modules to be brought at each demand. This implies that the number of modules going through each duct is known.

In this problem, one has the possibility to:

- Use the splicing technique.
- Serve the demand in one of two different fashions: cable-served when one full cable is used, or module-served when modules from a cable are used.

Backfeed is not allowed.

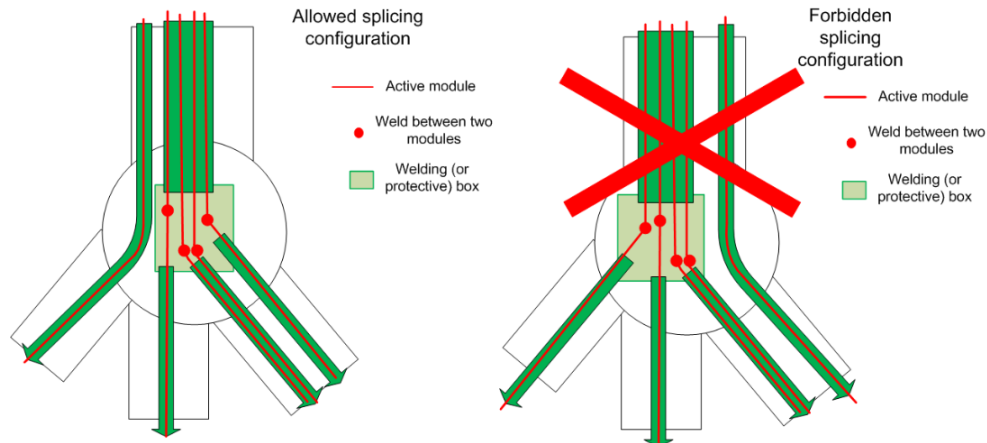


Figure 6.1: Left: Splicing configuration respecting the maintenance rule. On all ducts, cables are born in the same cabinet; Right: Possible splicing configuration in conflict with the maintenance rule. On the bottom-right duct, two different cables are born in different cabinets.

The maintenance rule (see Fig. 6.1) applies only to ESFCND. Recall that it is the following: "cables going through the same duct must be born in the same cabinet".

6.1 Integer Programming

6.1.1 SFCND

Notation and formulation

The following notations will also be used in section 6.2.

An arborescence $G = (V, A)$ describes the civil engineering infrastructure, V the cabinets and A the ducts, and its root $r \in V$ denotes the fiber source (CO or splitter location). For any $i \in V$, $D_i \in \mathbb{N}$ denotes the demand (number of active modules required) in node i . We define $V^* = V \setminus r$, the set of demand nodes is noted $V_D = \{v \in V, D_v > 0\}$, the set of nodes without demand $V_N = V^* \setminus V_D$. Each arc $(i, j) \in A$ has a length $\Delta_{(i,j)} > 0$ and must contain $m_{i,j}^{act}$ active modules ($m_{i,j}^{act}$ being known, since we are in an arborescence). For $i \in V$, we denote $\Gamma^+(i)$ the set successors of i and $\gamma(i)$ its predecessor.

We have L different cable types at our disposal, we note $\mathcal{L} = \{1, \dots, L\}$ the set of cables. Cables of type $l \in \mathcal{L}$ have a size of $M_l \in \mathbb{N}$ modules, and for $l \in \mathcal{L}$, we note $\mathcal{M}_l = \{1, \dots, M_l\}$ (the range of possible number of active modules in a cable of type l).

For $l \in \mathcal{L}$, let us define C_l^{le} the cost per length unit of a cable of size l , and PB_l the cost of a box of size l . For $m \in \mathcal{M}_L$, let us define the cost of the smallest cable able to contain m active modules $C_m^{min} = C_{l_1}^{le}$ where $l_1 = \min\{l \in \mathcal{L}, m \leq M_l\}$, and PW_m the cost for welding m modules.

We introduce \mathcal{P} the set of directed paths of G , and for $p \in \mathcal{P}$, we note by $s(p)$ its source node, $t(p)$ its target node, and Δ_p its length (which extends Δ from A to \mathcal{P}).

We define the following variables:

- $\forall l \in \mathcal{L}, \forall p \in \mathcal{P}, k_{p,l}^{spl} \in \{0, 1\}$ the binary variable equal to 1 iff there is a cable of size l on path p spliced in $t(p)$.
- $\forall p \in \mathcal{P}, k_p^{dem} \in \{0, 1\}$ the binary variable equal to 1 iff there is a cable on path p serving the demand in $t(p)$ in a cable-served way. Its size is known, it is $\min\{l \in \mathcal{L} | M_l \geq D_{t(p)}\}$.
- $\forall p \in \mathcal{P}, m_p^{spl} \in \{0, \dots, M_L\}$ the number of active modules of the cable on path p spliced in $t(p)$.
- $\forall i \in V^*, \forall m \in \mathcal{M}_L, w_{i,m}$ the binary variable equal to 1 iff m welds are done in node i .

The problem can be formulated as follows:

$$\begin{aligned} \min \sum_{p \in \mathcal{P}} \Delta_p \cdot \left(C_{D_{t(p)}}^{min} \cdot k_p^{dem} + \sum_{l \in \mathcal{L}} C_l^{le} \cdot k_{p,l}^{spl} \right) \\ + \sum_{i \in V_N} \sum_{m \in \mathcal{M}_L} PW_m \cdot w_{i,m} + \sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}} PB_l \cdot k_{p,l}^{spl} \end{aligned}$$

such that

$$\sum_{p \in \mathcal{P} | t(p)=i} \sum_{l \in \mathcal{L}} k_{p,l}^{spl} \leq 1 \quad \forall i \in V^*, \quad (6.1)$$

$$\sum_{p \in \mathcal{P} | t(p)=i} k_p^{dem} \leq 1 \quad \forall i \in V_D, \quad (6.2)$$

$$\sum_{l \in \mathcal{L}} M_l \cdot k_{p,l}^{spl} \geq m_p^{spl} \quad \forall p \in \mathcal{P}, \quad (6.3)$$

$$\begin{aligned} \sum_{p \in \mathcal{P} | t(p)=i} m_p^{spl} &= D_i \cdot (1 - \sum_{p \in \mathcal{P} | t(p)=i} k_p^{dem}) \\ + \sum_{p \in \mathcal{P} | s(p)=i} (m_p^{spl} + D_{t(p)} k_p^{dem}) & \quad \forall i \in V^*, \end{aligned} \quad (6.4)$$

$$\sum_{m \in \mathcal{M}_L} m \cdot w_{i,m} = \sum_{p \in \mathcal{P} | i=s(p)} (m_p^{spl} + D_{t(p)} \cdot k_p^{dem}) \quad \forall i \in V^*, \quad (6.5)$$

$$\sum_{m \in \mathcal{M}_L} w_{i,m} \leq 1 \quad \forall i \in V_N, \quad (6.6)$$

$$k^{dem}, k^{spl}, w \in \{0, 1\}; m^{spl} \in \{0, \dots, M_L\}$$

In the cost function, the first term stands for the cost of cables, the second term for the cost of welds, and the last term for the cost of boxes. Equations (6.1) ensure at most one cable is spliced in a node. Constraints (6.2) a most one cable serves the demand in a cable-served way. Equations (6.3) make sure that spliced cables are large enough to contain their number of active modules. Constraints (6.4) are active module conservation equations. The left hand side term stands for the number of modules of the spliced cable. The first right side hand term is the number of modules necessary to serve the demand, in case it is not cable-served. The last term is the number of active modules of born cables. Finally, (6.5) and (6.6) ensure that w counts the number of welds to be done in each node.

Remark 6.1.1 *It is possible to fix the value of some variables. First, notice that leaf nodes are demand nodes. These nodes will be served in a cable-served way, and no operation will be done inside them. This gives, for all nodes $i \in V_D$ such that $|\Gamma^+(i)| = 0$:*

$$\begin{aligned} \forall m \in \mathcal{M}_L, w_{i,m} &= 0 \\ \forall p \in \mathcal{P} | t(p) = i, \forall l \in \mathcal{L}, k_{p,l}^{spl} &= 0 \end{aligned}$$

Furthermore, the number of welds done in a node cannot exceed the number of active modules going out of this node. This gives:

$$\forall i \in V^*, \forall m \in \mathcal{M}_L, \text{ if } m > \sum_{j \in \Gamma^+(i)} m_{i,j}^{act}, w_{i,m} = 0$$

Valid inequalities

We propose here several valid inequalities to tighten the formulation.

Let us define, for all $m \in \mathbb{N}$, the minimum cost per length unit of a set of cables able to contain m active modules denoted by $LB(m)$. For a given m , $LB(m) = \{\min \sum_{l \in \mathcal{L}} C_l^{le} \cdot n_l \mid \sum_{l \in \mathcal{L}} M_l \cdot n_l \geq m, n \in \mathbb{N}^L\}$.

Proposition 6.1.1 *The following inequalities are valid for the ESFCND problem:*

$$\forall (i, j) \in A, \quad \sum_{p \in \mathcal{P} | (i, j) \in p} \left(\sum_{l \in \mathcal{L}} (C_l^{le} \cdot k_{p,l}^{spl}) + C^{min}(D_{t(p)}) \cdot k_p^{dem} \right) \geq LB(m_{i,j}^{act}) \quad (6.7)$$

The left hand side is the cost per length unit of the cables going through (i, j) .

Let us consider a path $p \in \mathcal{P}$ such that $t(p) \in V_D$ and $s(p) \neq r$. If there is a cable deployed on p , born in $s(p)$ and serving the demand in $t(p)$, then we know there is a splicing operation done in $s(p)$. Furthermore, there is at least $D_{t(p)}$ welds in this operation, since the cable serving $t(p)$ contains $D_{t(p)}$ active modules.

Proposition 6.1.2 *The following valid inequalities are valid for the ESFCND problem:*

$$\forall p \in \mathcal{P} | t(p) \in V_D \text{ and } s(p) \neq r, k_p^{dem} \leq \sum_{m \geq D_{t(p)}} w_{s(p),m} \quad (6.8)$$

Proof. Let consider a path $p \in \mathcal{P}$ such that $t(p) \in V_D$, $s(p) \neq r$, and $k_p^{dem} = 1$ ($t(p)$ is cable served by a cable on p). By (6.5), it gives $\sum_{m \in \mathcal{M}_L} m \cdot w_{s(p),m} \geq D_{t(p)}$ (there are at least $D_{t(p)}$ welds done in $s(p)$). Which means, with (6.6), $\exists! m_0 \geq D_{t(p)}, w_{s(p),m_0} = 1$. Hence the result. \square

6.1.2 ESFCND

ESFCND can be solved by using the same variables as in Section 6.1.1. The cost function is the same, the set of feasible solutions is described by constraints (6.1) to (6.6) to which we add the maintenance constraints described below:

$$\forall (p, p') \in \mathcal{P}^2 \text{ such that } s(p) \neq s(p') \text{ and } \exists a \in A, a \in p \text{ and } a \in p', \quad k_p^{dem} + k_{p'}^{dem} \leq 1 \quad (6.9)$$

$$\sum_{l \in \mathcal{L}} k_{p,l}^{spl} + \sum_{l \in \mathcal{L}} k_{p',l}^{spl} \leq 1 \quad (6.10)$$

$$\sum_{l \in \mathcal{L}} k_{p,l}^{spl} + k_{p'}^{dem} \leq 1 \quad (6.11)$$

These constraints ensure that on two paths which have different origins but an arc in common, there can be only one cable. Constraints (6.9) ensure it in

the case the two cables are serving the demand. Constraints (6.10) in the case both cables are spliced (at most one term in the sum $\sum_{l \in \mathcal{L}} k_{p,l}^{sp}$ is equal to 1, since there can be at most one splicing operation in $t(p)$, the same goes for p'). Finally, constraints (6.11) in the case one of them is spliced and the other one serves the demand.

The next section introduces an alternative mixed integer programming approach for ESFCND, based on arcs rather than paths. It uses the properties of the problem, and has less variables and less constraints.

Notations and formulation

We keep the same notations for the problem instance. In addition, let us define for $(i, j) \in A$, $U_{i,j}$ an upper bound of the cost per length unit of the cables going through duct (i, j) .

We define the following variables:

- $\forall (i, j) \in A, x_{i,j} \in \{0, 1\}$ the binary variable equal to 1 iff the cables on arc (i, j) are born in i .
- $\forall (i, j) \in A, c_{i,j} \in \mathbb{R}$ the continuous variable equal to the cost per length unit of the cables on arc (i, j) .
- $\forall (i, j) \in A, z_{i,j} \in \mathbb{R}$ the continuous variable equal to $x_{i,j} \cdot c_{i,j}$.
- $\forall i \in V_D, u_i \in \{0, 1\}$ the binary variable equal to 1 iff the node i is module-served.
- $\forall i \in V^*, \forall m \in \mathcal{M}_L, w_{i,m}$ the binary variable equal to 1 iff m welds are done in node i (since its meaning is identical to Section 6.1.1, we keep the same name).
- $\forall i \in V^*, \forall l \in \mathcal{L}, y_{i,l}$ the binary variable equal to 1 iff a cable of size l is spliced in i .

The problem can be formulated as follows:

$$\begin{aligned} \min \quad & \sum_{i \in V^*} \sum_{m \in \mathcal{M}_L} PW_m \cdot w_{i,m} \\ & + \sum_{(i,j) \in A} \Delta_{(i,j)} \cdot c_{i,j} + \sum_{i \in V^*} \sum_{l \in \mathcal{L}} PB_l \cdot y_{i,l} \end{aligned} \tag{6.12}$$

such that

$$c_{\gamma(i),i} = \sum_{l \in \mathcal{L}} C_l^{le} y_{i,l} + \sum_{j \in \Gamma^+(i)} c_{i,j}$$

$$- \sum_{j \in \Gamma^+(i)} z_{i,j} + (1 - u_i) \cdot C_{D_i}^{min} \quad \forall i \in V_D, \quad (6.13)$$

$$c_{\gamma(i),i} = \sum_{l \in \mathcal{L}} C_l^{le} y_{i,l} + \sum_{j \in \Gamma^+(i)} c_{i,j} - \sum_{j \in \Gamma^+(i)} z_{i,j} \quad \forall i \in V_N, \quad (6.14)$$

$$\sum_{l \in \mathcal{L}} M_l \cdot y_{i,l} \geq D_i \cdot u_i + \sum_{j \in \Gamma^+(i)} m_{i,j}^{act} \cdot x_{i,j} \quad \forall i \in V_D, \quad (6.15)$$

$$\sum_{l \in \mathcal{L}} M_l \cdot y_{i,l} \geq \sum_{j \in \Gamma^+(i)} m_{i,j}^{act} \cdot x_{i,j} \quad \forall i \in V_N, \quad (6.16)$$

$$\sum_{l \in \mathcal{L}} y_{i,l} \leq 1 \quad \forall i \in V^*, \quad (6.17)$$

$$\sum_{m \in \mathcal{M}_L} m \cdot w_{i,m} = \sum_{j \in \Gamma^+(i)} m_{i,j}^{act} \cdot x_{i,j} \quad \forall i \in V^*, \quad (6.18)$$

$$\sum_{m \in \mathcal{M}_L} w_{i,m} \leq 1 \quad \forall i \in V^*, \quad (6.19)$$

$$z_{i,j} \geq c_{i,j} - U_{i,j} \cdot (1 - x_{i,j}) \quad \forall (i,j) \in A, \quad (6.20)$$

$$z_{i,j} \leq U_{i,j} \cdot x_{i,j} \quad \forall (i,j) \in A, \quad (6.21)$$

$$z_{i,j} \leq c_{i,j} \quad \forall (i,j) \in A, \quad (6.22)$$

$$u, w, x, y \in \{0, 1\}; c, z \in \mathbb{R}$$

The first term of the cost function denotes the cost of welds, the second term stands for the cost of cables, and the last term stands for the cost of boxes. Equations (6.13) ensure the cost per length unit of any arc is properly counted. The term $\sum_{l \in \mathcal{L}} C_l^{le} y_{i,l}$ stands for the cost of the cable spliced in i , if any. If for some arc $(i, j) \in A$ such that $j \in \Gamma^+(i)$ we have $x_{i,j} = 0$, then the cables on (i, j) come from $(\gamma(i), i)$ unchanged. Otherwise, they come from the splicing operation done in i . The last term stands for the cost of the cable serving the demand in i . Equations (6.14) are the equivalent concerning nodes without demand. Equations (6.15), (6.16) and (6.17) ensure the cable spliced in i is large enough to contain its active modules. The first term of the right hand side of (6.15) stands for modules serving the demand, the second term for modules of born cables. Constraints (6.18) and (6.19) ensure the variable $w_{i,m}$ is equal to 1 iff there are m welds done in node i . Finally, constraints (6.20), (6.21) and (6.22) ensure $\forall (i, j) \in A, z_{i,j} = x_{i,j} \cdot c_{i,j}$ (these are linearisation equations).

Remark 6.1.2 *It is possible to fix the value of some variables. Assuming there exists $i \in V^*$ and $m_1 \in \mathcal{M}_L$ such that $w_{i,m_1} = 1$, then by (6.18), we know there exists $S \subseteq \Gamma^+(i)$ such that $m_1 = \sum_{j \in S} m_{i,j}^{act}$. This gives by contraposition $\forall i \in V^*, \forall m \in \mathcal{M}_L$ if $m \notin \{\sum_{j \in S} m_{i,j}^{act} | S \subseteq \Gamma^+(i)\}$ then $w_{i,m} = 0$. It can be computed in $\mathcal{O}(|\Gamma^+(i)| \times M_L)$ (which is not a polynomial with respect to the instance size, provided M_L is not coded in an unary system).*

Valid Inequalities

The continuous relaxation of the formulation introduced above shows is weak, mostly due to the linearisation of z . We propose here several valid inequalities to tighten it.

In nodes without demand, if a cable of size l is spliced, then it has a number of active modules between M_l and $M_{l-1}+1$; otherwise one could install a smaller cable and obtain a cheaper solution. With the convention $M_0 = 0$ and $\mathcal{M}_0 = \emptyset$, this gives:

Proposition 6.1.3 *Every optimal solution of the ESFCND problem verifies*

$$\forall i \in V_N, \forall l \in \mathcal{L}, y_{i,l} = \sum_{m \in \mathcal{M}_l \setminus \mathcal{M}_{l-1}} w_{i,m} \quad (6.23)$$

Proof. Let us consider an optimal solution S of the ESFCND problem. Let us consider $i \in V^*$ and $l \in \mathcal{L}$ such that $y_{i,l} = 1$ (a box of size l is installed in i). This gives us $1 \leq \sum_{j \in \Gamma^+(i)} m_{i,j}^{act} \cdot x_{i,j}$ (there are cables born in i); otherwise we could obtain a cheaper solution by setting $y_{i,l}$ to 0.

Either (6.16) or (6.15) give us $M_l \geq \sum_{j \in \Gamma^+(i)} m_{i,j}^{act} \cdot x_{i,j}$. Furthermore, with (6.18) and (6.19), we can obtain $\exists m_0 \in \{1, \dots, M_l\}, w_{i,m_0} = 1$ (in other words, $m_0 \leq M_l$ welds are done in i).

If $l = 1$, we have the result.

Otherwise, let us assume $m_0 \leq M_{l-1}$. Then, the solution S' identical to S everywhere but in $y'_{i,l-1} = 1$ and $y'_{i,l} = 0$ is a feasible cheaper solution (it is the solution obtained by replacing the cable spliced in i by a smaller cable, leading to a smaller cost for boxes and cables). Which contradicts our hypothesis.

Hence the result. \square

With a reasonment similar to the one from Proposition 6.1.1 (see definition of LB), we can get a lower bound of the cost per length unit of the cables on each arc.

Proposition 6.1.4 *The following inequalities are valid for the ESFCND problem:*

$$\forall (i, j) \in A, c_{i,j} \geq LB(m_{i,j}^{act}) \quad (6.24)$$

If the cables on some arc $(i, j) \in A$ are born in i , then at least $m_{i,j}^{act}$ welds are done in node i . This implies what follows.

Proposition 6.1.5 *The following inequalities are valid for the ESFCND problem*

$$\forall (i, j) \in A, x_{i,j} \leq \sum_{m \in \mathcal{M}_L | m \geq m_{i,j}^{act}} w_{i,m} \quad (6.25)$$

Proof. Let us consider a solution of the ESFCND problem. Let us consider $(i, j) \in A$ such that $x_{i,j} = 1$. This implies, by (6.18) that $\sum_{m \in \mathcal{M}_L} m \cdot w_{i,m} \geq m_{i,j}^{act}$. Then, with (6.19), it follows that $\exists m_0 \geq m_{i,j}^{act}, w_{i,m_0} = 1$ (only one of the variables $w_{i,m}$ can be equal to 1). Hence the result. \square

6.2 Dynamic Programming for ESFCND

For any node $i \in V^*$, we introduce the additional notation $V^{pr}(i)$, which refers to the set of nodes on the path from the root to i , excluding i and including r .

6.2.1 Exact Algorithm

The ESFCND problem can be solved by Algorithm 1. To each node $i \in V^*$, and for each node $j \in V^{pr}(i)$, we associate to i a label $\langle j, C(i, j) \rangle \in V^{pr}(i) \times \mathbb{R}$ where $C(i, j)$ is the minimum cost of the network rooted in i plus the cost of the cables on the path from j to i , assuming these are born in node j .

Algorithm 1 Exact Resolution Algorithm for ESFCND

```

1: procedure INITIALISATION()
2:   for  $i \in V_D | \Gamma^+(i) = \emptyset$  do
3:     for  $j \in V^{pr}(i)$  do
4:       Add to  $i$  the label  $\langle j, C_{D_i}^{min} \cdot \Delta_p \rangle$  where  $p \in \mathcal{P}$  is the only path
       s.t.  $s(p) = j$  and  $t(p) = i$ .
5:     end for
6:     Declare  $i$  labeled.
7:   end for
8: end procedure
9: procedure RECURSION()
10:  while  $\exists r' \in \Gamma^+(r)$  such that  $r'$  has not been labeled do
11:    for every node  $i \in V^*$  such that all nodes in  $\Gamma^+(i)$  have been labeled
12:    do
13:      for  $j \in V^{pr}(i)$  do
14:         $\triangleright$  We select the operation in  $i$  minimizing the network cost.
        Add the label  $\langle j, C(i, j) \rangle$  to node  $i$  where
        
$$C(i, j) = \min_{S \subseteq \Gamma^+(i), u \in \{0,1\}} \sum_{k \in S} C(k, i) + \sum_{k \in \Gamma^+(i) \setminus S} C(k, j) + PW_m + \Delta_p \cdot C_{l_1}^{le} + \Delta_p \cdot C_{D_i}^{min} \cdot (1 - u) \quad (6.26)$$

        with 
$$\begin{cases} m = \sum_{k \in S} m_{i,k}^{act}; l_1 = \min\{l \in \mathcal{L} | M_l \geq u \cdot D_i + \sum_{k \in S} m_{i,k}^{act}\} \\ p \in \mathcal{P} \text{ is the only path such that } s(p) = j, t(p) = i \end{cases}$$

15:      end for
16:      Declare  $i$  labeled.
17:    end for
18:  end while
19: end procedure
20: procedure TERMINATION()
21:  return  $\sum_{r' \in \Gamma^+(r)} C(r', r)$ 
22: end procedure

```

The algorithm is initialized at leaf nodes (line 4), which are cable-served demand nodes, and where the size of the cable serving the demand is known.

For a node i such that all nodes in $\Gamma^+(i)$ have been labeled, and for $j \in V^{pr}(i)$, (6.26) computes the minimum cost network if the next operation is done in j . For $i \in V^*$ and $k \in \Gamma^+(i)$, $k \in S$ iff the cables going through arc (i, k) are born in node i . Similarly, the boolean u is equal to 1 iff the node i is module-served (its meaning is similar than the variable u_i in Section 6.1.2).

We propose to compute it with a brute-search algorithm on the set S and on u . For given nodes $i \in V^*$, $j \in V^{pr}(i)$, it can be done in $\mathcal{O}(|\Gamma^+(i)| \times 2^{|\Gamma^+(i)|+1})$.

Lemma 6.2.1 *Algorithm 1 runs in time $\mathcal{O}(2^{1+\max\Gamma} \times |V|^2)$ where $\max\Gamma$ denotes the maximal degree (number of successors) of a node in the graph.*

This can be shown by summing the operations done for each loop.

Remark 6.2.1 *This implies that if the maximal degree of nodes in the graph is bounded by a constant, then algorithm 1 runs in polynomial time.*

For a non-leaf node $i \in V^*$ and $j \in V^{pr}(i)$, when we compute (6.26), we do not consider the cost of the welds done in j . This comes later, while j is being labeled. It does not influence the network below, since all cables going through $(\gamma(i), i)$ are born in i . C^* is the sum of the following elements:

- the cost of the network in the arborescence rooted in i , including the cost of the welds and boxes in i (if any)
- the cost of cables deployed from i to j

This leads us to show the next proposition to show the validity of the algorithm.

Proposition 6.2.2 *Let us consider $i \in V^*$. When i is declared labeled in algorithm 1, there exists a node $j \in V^{pr}(i)$ such that in the label $\langle j, C(i, j) \rangle$, $C(i, j)$ describes the cost of the minimum ESFCND solution in the arborescence rooted in node i plus the cost of the cables on the path from j to i .*

We will start to prove it for leaf nodes, then recursively on higher nodes.

Proof. ★ Let us consider a leaf node i . In the minimum cost network, it is served in a cable-served way with a cable of type $l_1 = \min\{l \in \mathcal{L} | M_l \geq D_i\}$. This cable is born in some node $j \in V^{pr}(i)$, eventually the root. Let us call $p \in \mathcal{P}$ the only path such that $s(p) = j$ and $t(p) = i$. The label $\langle j, C(i, j) \rangle$ has a cost of $C_{D_i}^{min} \cdot \Delta_p$.

★ Let us consider a non-leaf node $i \in V^*$ such that all nodes in $\Gamma^+(i)$ have been labeled. In the minimal cost network, the cables going through arc $(\gamma(i), i)$ are all born in a node $j \in V^{pr}(i)$. Thanks to the maintenance constraint, we know that they are all born in the same node. Since all nodes $k \in \Gamma^+(i)$ have been labeled, for each of these nodes, there is a node $j_k \in V^{pr}(k)$ such that in the

label $\langle j_k, C(k, j_k) \rangle$, $C(k, j_k)$ describes the cost of the minimum cost network in the arborescence rooted in k plus the cost of the cables on the path from j_k to k . Furthermore, since the cables going through arc $(\gamma(i), i)$ are all born in j , we have either $j_k = j$ or $j_k = i$. Let us consider the label $\langle j, C(i, j) \rangle$ of node i . If in the minimal network i is module-served, then we will have $u = 0$ in the computation of (6.26). Furthermore, let us consider $k \in \Gamma^+(i)$. If $j_k = i$, we will have $k \in S$ in the computation of (6.26), and $k \in \Gamma^+(i) \setminus S$ otherwise. Hence the result. \square

The termination of the algorithm derives from Proposition 6.2.2. For each node $r' \in \Gamma^+(r)$, we have $V^{pr}(r') = \{r\}$. This implies, using this proposition, that in the label $\langle r, C(r', r) \rangle$, $C(r', r)$ is the cost of the minimum network cost in the arborescence rooted in r' plus the cost of the cables on (r, r') . Summing these values gives the minimum network cost.

The computation of (6.26) at each step is not done in polynomial time. There are many algorithms able to tackle it (dynamic programming, brute search, ...). We propose a way to tackle it in the next section which allows us to give an approximation in polynomial time, thus providing a polynomial time approximation algorithm.

6.2.2 Approximation algorithm

In this Section, we propose here a Fully Polynomial Time Approximation Scheme (FPTAS) for ESFCND, in the case where:

- The height of the arborescence describing the civil engineering is upper bounded by $H \in \mathbb{N}$.
- The number of intervals on which the cost of the welds PW is a linear function with respect to m is upper bounded by $F \in \mathbb{N}$ (recall that PW is defined to be piecewise linear).

We introduce the following additional notation. PW is decomposed into its linear components. For $f \in \{1, \dots, F\}$, we have successive integers B_f such that $\forall m \in \{B_f, \dots, B_{f+1}\}$, $PW_m = PW^{a,f} \times m + PW^{b,f}$.

A FPTAS for the knapsack problem is available in [41]. This algorithm \mathcal{A} gives, for an instance of the knapsack problem, and a number $\alpha > 1$, a solution S to the knapsack problem of cost C^{approx} where $C^{approx} \leq \alpha \times OPT$ and OPT is the optimal solution cost (here, we consider the minimization version of the knapsack problem, or "covering problem").

In algorithm 1, the computation of (6.26) is the only step which is not done in polynomial time. We propose to solve it with algorithm 2, which reformulates it as a series of knapsack problems. Then, each of the knapsack problems can be approximated thanks to the knapsack FPTAS.

Algorithm 2 Computation of (6.26)

```

1: procedure  $C(i, j)$  CALCULATION()
2:   Define  $C_{\min} := +\infty$ 
3:   for  $l \in \mathcal{L}$  do
4:     for  $f \in \{1, \dots, F\}$  do
5:       if  $(\{M_{l-1} + 1, \dots, M_l\}) \cap \{B_f, \dots, B_{f+1}\} \neq \emptyset$  then
6:          $m_1 := \max(M_{l-1} + 1, B_f)$ 
7:         Solve the following knapsack problems


$$C_1 = \min \sum_{k \in \Gamma^+(i)} (x_k \cdot C(k, i) + (1 - x_k) \cdot C(k, j))$$


$$+ PW^{a,f} \times \sum_{k \in \Gamma^+(i)} x_k \cdot m_{i,k}^{act} + PW_{b,f} + \Delta_{(i,j)} \cdot C_l^{le}$$


$$+ \Delta_{(i,j)} \cdot C^{min}(D_i) \tag{6.27}$$

such that 
$$\sum_{k \in \Gamma^+(i)} x_k \cdot m_{i,k}^{act} \geq m_1$$


8:       end if
9:       if  $(\{M_{l-1} + 1, \dots, M_l\}) \cap \{B_f - D_i, \dots, B_{f+1} - D_i\} \neq \emptyset$  then
10:         $m_2 := \max(M_{l-1} + 1, B_f) - D_i$ 


$$C_2 = \min \sum_{k \in \Gamma^+(i)} (x_k \cdot C(k, i) + (1 - x_k) \cdot C(k, j))$$


$$+ PW^{a,f} \times \sum_{k \in \Gamma^+(i)} x_k \cdot m_{i,k}^{act} + PW^{b,f} + \Delta_{(i,j)} \cdot C_l^{le} \tag{6.28}$$

such that 
$$\sum_{k \in \Gamma^+(i)} x_k \cdot m_{i,k}^{act} \geq m_2$$


$$x \in \{0, 1\}^{|\Gamma^+(i)|}$$


11:      end if
12:       $C^{min} := \min(C^{min}, C_1, C_2)$ 
13:    end for
14:  end for
15:   $C^{min} := \min(C^{min}, \sum_{k \in \Gamma^+(i)} C(k, j))$ 
16:  return  $C^{min}$ 
17: end procedure

```

The algorithm spans all possible cable sizes. For each cable size l , it computes the minimum cost splicing operation in which a cable of size l is spliced in i . (6.27) computes the minimal cost splicing in the case $u = 0$, and (6.28) computes the minimal cost splicing in the case $u = 1$. Finally, in line 15, it compares the best splicing obtained with the cost of continuing all cables.

The following lemma stems from the concavity of PW .

Lemma 6.2.3 $\forall (f, f') \in \{1, \dots, F\}^2$, if $f \leq f'$, then $\forall m \geq B_{f'}$, $PW^{a,f'} \times m + PW^{b,f'} \leq PW^{a,f} \times m + PW^{b,f}$

Proof. Let us assume $\exists (f_1, f_2) \in \{1, \dots, F\}^2$, with $f_1 \leq f_2$ and $\exists m \geq B_{f_2}$ such that $PW^{a,f_2} \times m + PW^{b,f_2} > PW^{a,f_1} \times m + PW^{b,f_1}$.

Since $PW^{b,f}$ is decreasing with respect to f , this means $PW^{a,f_2} > PW^{a,f_1}$, which contradicts the concavity of PW .

Hence the result. \square

From this lemma, we can get that if, for some $l \in \mathcal{L}$ and $f \in \mathcal{F}$, C_1 is reached for values of x_k such that $\sum_{k \in \Gamma^+(i)} m_{i,k}^{act} x_k > B_{f+1}$ (the values returned by the knapsack problem are higher than the range of welds we consider), then a lower value of C_1 can be reached for l and $f + 1$. A similar reasoning can be done for C_2 .

Let us consider $H \in \mathbb{N}$. Let us consider an instance of ESFCND where the civil engineering arborescence height is upper bounded by a constant H . We propose the following FPTAS for ESFCND.

Let us consider $\alpha > 1$. There is a polynomial time algorithm \mathcal{A} which approximates the knapsack within a ratio $\alpha^{\frac{1}{H}}$. Run algorithm \mathcal{A}' which is a variant of algorithm 1 where:

- Each computation of (6.26) is done with algorithm 2.
- In algorithm 2, each computation of (6.27) and (6.28) is approximated with algorithm \mathcal{A} .

This algorithm runs in polynomial time. Indeed, in algorithm 1, the only step which is not done in polynomial time is replaced by a polynomial time algorithm.

Proposition 6.2.4 *Algorithm \mathcal{A}' returns a cost v of the ESFCND problem such that $v \leq \alpha v^*$ where v^* is the cost of optimal solution of ESFCND.*

Proof. \star Let us consider a leaf node $i \in V_D$. The labels $C(i, j)$ for $j \in V^{pr}(i)$ have the same value in algorithm 1 and algorithm \mathcal{A}' .

\star Let us consider a non-leaf node $i \in V^*$ and $j \in V^{pr}(i)$. In the computation of (6.26) by algorithm 2, C^* is approximated with a ratio of $\alpha^{\frac{1}{H}}$. Its value is the sum of welds and boxes costs and of a linear combination of the values of $C(k, i)$ and $C(k, j)$ for $k \in \Gamma^+(i)$. So it multiplies the approximation ratios of the values of $C(k, i)$ and $C(k, j)$. Hence, each time a node is labeled, the approximation ratio of its labels are $\alpha^{\frac{1}{H}}$ time the approximation ratio of its children node.

Hence the global multiplicative ratio of this algorithm is α .

The next section assesses the complexity of SFCND and ESFCND.

6.3 Complexity

We show in Section 6.3.1 that SFCND is NP-hard even with 1 cable size and an upper bound on the node degree of 2, and in Section 6.3.2 that ESFCND is NP-hard.

6.3.1 SFCND

Let us consider the Number Partitioning Problem (NPP), which is shown to be NP-complete in [34].

Instance: A set of N strictly positive integers $\{n_i \in \mathbb{N} | i \in \{1, \dots, N\}\}$.

Question: Is there a partition of the integers $S \subseteq \{1, \dots, N\}$ such that $\sum_{i \in S} n_i = \sum_{i \notin S} n_i$?

We consider an instance of the NPP and associate it to the following SFCND instance: Let $G = (V, A)$ be an arborescence describing the civil engineering structure, $(V = \{r, 0, 1\} \cup \{v_i | i \in \{1, \dots, N\}\}, A = \{(r, 0); (0, 1); (1, v_1); (v_{i-1}, v_i) | i \in \{2, \dots, N\}\})$ (G is a chain graph), r is the fiber source. The demand nodes are $\{v_i, i \in \{1, \dots, N\}\}$ and have respective demands $n_i, i \in \{1, \dots, N\}$ modules. Only one type of cable is available, with size $M_1 = \frac{1}{2} \sum_{i \in \{1, \dots, N\}} n_i$. Its cost per length unit is $C_1 = 1$. The lengths of all arcs of the arborescence are null, except $(r, 0)$ which is of length 1. This means the cost of a cable born in r is 1, and the cost of the other ones is 0. The cost of welds and boxes is null.

The question associated to this SFCND instance is "Is there a cabling solution cheaper than 2 ?".

★ If (NPP) is feasible: $\exists S \subseteq \{1, \dots, N\}$ such that $\sum_{i \in S} n_i = \sum_{i \notin S} n_i$. We then build the following cabling solution:

- Two cables holding only active modules are installed on link $(r, 0)$.
- In node 0, one incoming cable is spliced into $N - |S|$ born cables. The born cables have a number of active modules $n_i, i \notin S$ and serve respectively the demand nodes $(v_i)_{i \notin S}$.
- In node 1, the cable coming from the root with only active modules is spliced into $|S|$ born cables. The born cables have n_i active modules and serve the demand nodes $(v_i)_{i \in S}$.

Since the number of active modules is conserved in each splicing, the cabling solution described above is feasible (it is illustrated in Fig. 6.2, as well as the instance). Its cost is equal to 2.

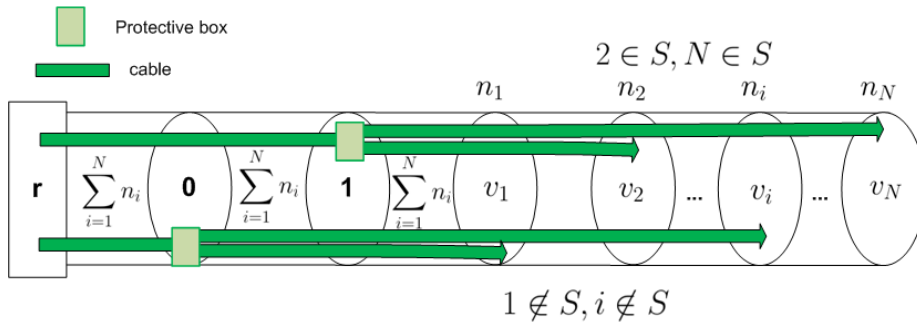


Figure 6.2: Instance and solution used in the complexity proof

★ If (NPP) is not feasible. Then, the solution described above is not possible anymore. One cable is not large enough to cover link $(r, 0)$. Two cables cannot cover $(r, 0)$ either, since they would both have only active modules, which would mean that the (NPP) problem was feasible. Consequently, at least 3 cables need to be installed on arc $(r, 0)$, and such a solution has a cost of at least 3.

Remark 6.3.1 *The solution illustrated in Fig. 6.2 is not valid for ESFCND, the maintenance rule is not respected in nodes 0 and 1.*

We showed in Section 4.1.1 that FCNDA is NP-hard. The same proof can be conducted for SFCND, and show that it is NP-hard in the following context

- One cable size available.
- Civil engineering arborescence height of 3.
- Null welding cost.

6.3.2 ESFCND

ESFCND can be shown to be NP-complete by reduction from the (NPP). With the same notations, let us consider an instance of the NPP and associate it to the following ESFCND instance. The civil engineering structure is described by the set of nodes is $V = \{r, 0\} \cup \{v_i | i \in \{1, \dots, N\}\}$; the set of arcs $A = \{(0, v_i) | i \in \{1, \dots, N\}\} \cup \{(r, 0)\}$; r is the fiber source, the nodes $\{v_i | i \in \{1, \dots, N\}\}$ have a demand of n_i modules. The length of all arcs except $(r, 0)$ is null. We have $N + 1$ cables available:

- N cables of sizes n_i modules and cost per length unit n_i
- A cable of size $\frac{1}{2} \sum_{i=1}^N n_i$ and cost per length unit $\frac{1}{2} \sum_{i=1}^N n_i - 1$

The cost of welds and boxes is null.

The question we ask is "is there a solution of cost at most $\sum_{i=1}^N n_i - 1$ "?

★ If (NPP) is feasible. Then, we have $S \subseteq \{1, \dots, N\}$ such that $\sum_{i \in S} n_i = \sum_{i \notin S} n_i$. We consider the solution of ESFCND where

- For $i \in \{1, \dots, N\}$, on each arc $(0, v_i)$, we lay down a cable of size n_i
- In the node 0, a cable of size $\frac{1}{2} \sum_{i \in \{1, \dots, N\}} n_i$ is spliced. Cables of size $n_i, i \in S$ are born, and serve the demand of nodes $v_i, i \in S$.
- On the arc $(r, 0)$, a cable of size $\frac{1}{2} \sum_{i \in \{1, \dots, N\}} n_i$ holding only active modules is deployed (the one spliced in 0); as well as $N - |S|$ cables of sizes $n_i, i \notin S$ which serve the demand in nodes $v_i, i \notin S$.

The cost of this solution is the cost of cables on arc $(r, 0)$ which is $\sum_{i \in \{1, \dots, N\}} n_i - 1$. It is illustrated in Fig. 6.3.

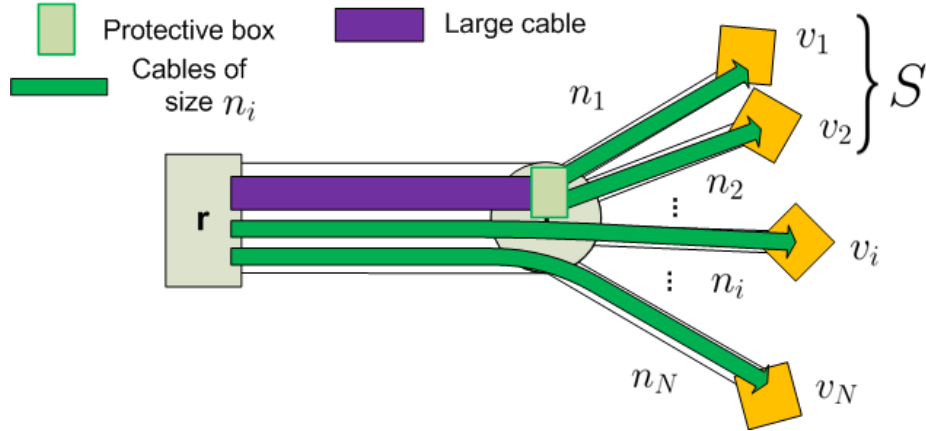


Figure 6.3: Instance and solution used in the complexity proof for ESFCND

★ If (NPP) is not feasible. In a minimal cost solution, the size of cables serving the demand is known. For a given $i \in \{1, \dots, N\}$, v_i is served by a cable of size n_i . Which leaves three types of solutions to consider.

The solution without splicing has a cost $\sum_{i \in \{1, \dots, N\}} n_i$. Each demand node is served by a cable coming directly from the root r .

Any solution where a cable of size $\frac{1}{2} \sum_{i \in \{1, \dots, N\}} n_i$ is spliced in 0 has a cost at least equal to $\sum_{i \in \{1, \dots, N\}} n_i$. Indeed, let us note $E \subseteq \{1, \dots, N\}$ the set such that cables of sizes $n_i, i \in E$ are born in 0. Since the NPP instance is not feasible, we have $\sum_{i \in E} n_i < \frac{1}{2} \sum_{i \in \{1, \dots, N\}} n_i$, so the cost of cables which are continued in 0 is $\sum_{i \notin E} n_i > \frac{1}{2} \sum_{i \in \{1, \dots, N\}} n_i$, and the total cost of the network is $\sum_{i \notin E} n_i + \frac{1}{2} \sum_{i \in \{1, \dots, N\}} n_i - 1 \geq \sum_{i \in \{1, \dots, N\}} n_i$.

Any solution where a smaller cable is spliced in 0 has a cost at least equal to $\sum_{i \in \{1, \dots, N\}} n_i$. Indeed, in any splicing of a cable of size n_i for a given $i \in \{1, \dots, N\}$, the spliced cable is at least as expensive than the born cables.

6.3.3 Synthesis

To the results proven here, we can add those deducible from Section 6.2. The restriction of ESFCND where there is an upper bound on the node degree can be solved in polynomial time, since in that case the computation of (6.26) can be done in polynomial time. This implies that it is also polynomial when more parameters are fixed. Furthermore, we showed in Section 6.2.2 that the problem admits a FPTAS under some conditions. As for SFCND, its NP-hardness in a restricted setting implies its NP-hardness in the more general cases. These results are summed up in table 6.1.

Table 6.1 shows a theoretical difference in the complexities of the two problems ESFCND and SFCND. We assess the numerical aspect of this difference in the next section.

Table 6.1: Complexity of the two problems in different contexts

| Problem | Complexity | | | |
|---------|------------|----------------|-----------------------|-------------------|
| | none | maximum degree | maximum degree L | F, H |
| SFCND | NP-hard | NP-hard | NP-hard | NP-hard |
| ESFCND | NP-hard | P | P | NP-hard, FPTAS |

6.4 Results

We assessed the solution methods on real-life instances taken from the city of Arles (France).

The cables available have a size of 1, 2, 4, 6, 8, 12, 18 or 24 modules. The resolution algorithm for the MIPs was the Cplex 12.6 default branch-and-bound algorithm.

Table 6.2: key features of the real-life instances

| instance | features | | | |
|----------|------------|------|--------------|--------------|
| | max degree | arcs | demand nodes | total demand |
| Ar 1 | 4 | 113 | 45 | 61 |
| Ar 2 | 6 | 103 | 38 | 55 |
| Ar 3 | 5 | 103 | 35 | 66 |
| Ar 4 | 6 | 123 | 43 | 80 |
| Ar 5 | 7 | 129 | 44 | 68 |
| Ar 6 | 6 | 137 | 43 | 67 |
| Ar 7 | 4 | 139 | 35 | 68 |
| Ar 8 | 5 | 163 | 41 | 63 |
| Ar 9 | 4 | 219 | 68 | 78 |

6.4.1 Models comparison

The results of the numerical experiments regarding the SFCND and ESFCND problem are displayed respectively in tables 6.3 and 6.4, "base model" always refers to the MIP without valid inequalities, and "enhanced model" to the MIP with valid inequalities. The columns of both tables are labeled as follows: "time" stands for the computation time; "CR" stands for the continuous relaxation as a ratio of the optimal solution; "Br" stands for the number of explored branches of the Branch and Bound algorithm.

Regarding SFCND, the valid inequalities have had a positive effect on the average computation time, which went down from 546 to 62 seconds. However, on most instances (8 out of 9), the MIP is solved faster without the valid in-

Table 6.3: Results for SFCND

| instance | base formulation | | enhanced formulation | |
|----------|------------------|--------|----------------------|--------|
| | time (s) | CR (%) | time (s) | CR (%) |
| Ar 1 | 8 | 90.3 | 16 | 91.0 |
| Ar 2 | 9 | 83.7 | 24 | 92.4 |
| Ar 3 | 17 | 92.2 | 22 | 93.3 |
| Ar 4 | 19 | 89.2 | 46 | 90.0 |
| Ar 5 | 1 | 94.9 | 2 | 95.2 |
| Ar 6 | 2 | 92.5 | 3 | 94.7 |
| Ar 7 | 13 | 92.4 | 29 | 93.7 |
| Ar 8 | 8 | 89.6 | 12 | 91.7 |
| Ar 9 | 4837 | 89.4 | 408 | 91.6 |

equalities. This suggest that they are more useful for instances that are hard to solve. Regarding the algorithm, the continuous relaxation goes from an average of 90.5 % to 92.6 %. The high relaxation of the base model can explain the mitigated impact of the inequalities on the performances.

Regarding ESFCND, all instances were easier to solve (computation times are displayed in milliseconds). The valid inequalities have had a beneficial effect on the computation time, all instances are solved faster with the enhanced formulation. The average computation time goes from 1730 to 329 ms. On an algorithmic level, the initial relaxation goes from an average of 13.2 % of the optimal solution cost to 87.3 % of the optimal solution cost. This has a significant impact on the number of nodes of the branch-and-bound algorithm, which goes from an average of 1100 branches to an average of 4 branches; 7 instances out of 9 were solved without branching. The dynamic programming approach was more efficient than the enhanced integer programming formulation, it solved 7 out of 9 instances faster.

6.4.2 Sensitivity analysis

Section 6.3 points to the maximal node degree as a key element of the problems complexity. Since the highest node degree of all real-life instances is between 4 and 7, we used fictive instances to assess the performances of each resolution technique when some of the nodes have a high degree. Their features are displayed in table 6.5.

Table 6.4: Results for ESFCND

| instance | base formulation | | | enhanced formulation | | | dynamic programming |
|----------|------------------|--------|------|----------------------|--------|----|---------------------|
| | time (ms) | CR (%) | Br | time (ms) | CR (%) | Br | time (ms) |
| Ar 1 | 1457 | 14.0 | 1191 | 305 | 89.2 | 0 | 324 |
| Ar 2 | 1174 | 17.8 | 462 | 239 | 86.6 | 0 | 239 |
| Ar 3 | 1317 | 13.6 | 153 | 318 | 81.7 | 0 | 66 |
| Ar 4 | 742 | 15.7 | 72 | 268 | 86.8 | 0 | 87 |
| Ar 5 | 746 | 18.2 | 0 | 477 | 89.2 | 0 | 88 |
| Ar 6 | 1477 | 15.5 | 66 | 238 | 91.8 | 0 | 110 |
| Ar 7 | 1667 | 9.7 | 1045 | 190 | 80.1 | 0 | 121 |
| Ar 8 | 1786 | 9.4 | 414 | 344 | 89.8 | 21 | 103 |
| Ar 9 | 5204 | 5.3 | 6302 | 507 | 90.8 | 9 | 306 |

Table 6.5: key features of the fictive instances

| instance | features | | | |
|----------|------------|------|--------------|--------------|
| | max degree | arcs | demand nodes | total demand |
| Fi 10 | 11 | 20 | 15 | 71 |
| Fi 11 | 12 | 22 | 16 | 84 |
| Fi 12 | 13 | 24 | 18 | 97 |
| Fi 13 | 14 | 26 | 19 | 112 |
| Fi 14 | 15 | 28 | 21 | 112 |
| Fi 15 | 16 | 30 | 22 | 127 |
| Fi 16 | 17 | 32 | 24 | 144 |

Table 6.6: Computation time on fictive instances (ms)

| instance | enhanced model SFCND | enhanced model ESFCND | dynamic programming |
|----------|----------------------|-----------------------|---------------------|
| Fi 10 | 205 | 166 | 322 |
| Fi 11 | 327 | 77 | 652 |
| Fi 12 | 993 | 332 | 1409 |
| Fi 13 | 1130 | 120 | 3800 |
| Fi 14 | 1369 | 347 | 12 403 |
| Fi 15 | 1450 | 98 | 39 654 |
| Fi 16 | 2691 | 280 | 164 243 |

As expected, the dynamic programming algorithm was very sensitive to the

node degree, the computation time growing exponentially. The enhanced MIP formulation for ESFCND was able to solve all instances in less than one second, with an average of 200 ms. This is the opposite of the results obtained on real-life instances, where the dynamic programming was more efficient. As for SFCND, the MIP formulation proved to be efficient, with an average computation time of 900 ms. Although the instances with a higher degree are harder to solve, it stays tractable in practice. One should favor a MIP based approach, regardless of the problem, when dealing with high degree nodes.

6.4.3 Operational considerations

We compared the optimal solutions of both problems. Results are displayed in table 6.7, the column labeled "arcs with rule broken" denotes the number of arcs where the maintenance rule (illustrated in figure 2.8) is broken.

Remark 6.4.1 *The SFCND problem has already been solved numerically in Chapter 4 as the restriction of FCNDA without any tapping. Since some instances are similar, one can see that the results of the third column of Table 6.7 below can already be found in the fourth column of Table 4.8 of Chapter 4.*

Table 6.7: Optimal solution costs and characteristics

| instance | Solution ESFCND | Solution SFCND | arcs with rule broken |
|----------|--------------------|-------------------|--------------------------|
| Ar 1 | 6156.6 | 6087.3 | 6 |
| Ar 2 | 10 357.3 | 9870.0 | 8 |
| Ar 3 | 6546.2 | 6125.8 | 14 |
| Ar 4 | 6720.8 | 6461.9 | 14 |
| Ar 5 | 5081.8 | 5081.8 | 0 |
| Ar 6 | 6546.5 | 6544.2 | 1 |
| Ar 7 | 9348.0 | 8638.6 | 18 |
| Ar 8 | 12 328.3 | 12 248.4 | 4 |
| Ar 9 | 25 619.1 | 24 422.8 | 15 |

The optimal solution of ESFCND is in average 3.7 % more expensive than the optimal solution of SFCND. This can be seen as an acceptable loss in capital expenditure if it is compensated by an easier maintenance, depending on the importance accorded to it.

The maintenance rule is broken in almost every real-life instance we tried (8 out of 9). In average, it is not respected in 6.2 % of the arcs, which is significant. This suggests that the optimal solutions of SFCND will be much harder to repair in case of failure on one of the arcs. These elements can be taken into account to establish a strategy in case of node failure.

6.5 Conclusion

This chapter tackles two fiber cables network design problems, one unconstrained by maintenance consideration (SFCND) and the other one constrained (ESFCND). Regarding the unconstrained problem, one integer programming based solving algorithm was proposed. Associated valid inequalities make it more tractable in practice. We proposed two solution methods for the constrained problem. These methods are complementary, as they prove efficient in different contexts: the dynamic programming approach is generally faster in graphs where nodes have a small degree, whereas the mixed integer programming, embedding efficient valid inequalities, is generally faster otherwise.

On a theoretical level, the unconstrained problem seems much more complex to solve than the constrained problem. Fixing some parameters makes the constrained problem approximable or solvable in polynomial time, depending on the parameters, while the unconstrained problems remains NP-hard and unapproximable within a ratio of $\frac{3}{2}$. Our numerical experiments confirmed this tendency on real-life instances.

As for the operational side, the maintenance rule can be considered as a reasonable compromise between capital expenditure for the network deployment and maintenance costs. Its implementation only increases the optimal solution cost by 3.7 % on our test instances.

Chapter 7

Synthesis

Theoretical Results

Resolution

This thesis introduces several decision problems related to the design of fiber cable networks. We provided at least one resolution method for each of these problems:

- FCND has an integer programming based resolution method.
- FCNDA has two different integer programming based resolution methods. One of them is based on the one ILP used for FCND.
- BFCND has one integer programming based resolution method.
- SFCND has one integer programming based resolution method; it is inspired by the work done for FCNDA.
- ESFCND has one specific mixed integer programming based method as well as a dynamic programming resolution algorithm.

Valid inequalities

The basic formulations of each problem was not so easy to solve on real-life instances. Each was enhanced by several families of valid inequalities, which reduced significantly the computation times. These valid inequalities were especially necessary in the cases where the continuous relaxation of the basic formulation was weak.

The coefficients of the valid inequalities are in most cases easy to determine. However, in some cases we used inequalities with coefficient that are NP-hard to compute (see (3.21) and (6.24), were coefficients are determined respectively by a Steiner Tree problem and a Knapsack problem). Those were easy to handle,

since the problems we solved to determine those coefficients were much smaller in size than the fiber cables problem considered (FCND or SFCND).

Complexity

We showed in this thesis that all problems considered are NP-hard. However, there are still differences between their complexities.

This can be seen in the approximation ratios. It was shown (Section 3.2) that FCND cannot be approximated within a ratio of 2, while we showed that FCNDA, BFCND and SFCND cannot be approximated within a ratio of $\frac{3}{2}$ (respectively in Sections 4.1, 5.2 and 6.3). On the other hand, a FPTAS under some conditions was provided for ESFCND, showing that it can be approximated as close as desired to the optimal solution.

We showed that the FCNDA problem could not be easily modeled with a certain set of variables in Proposition 4.1.2. This result is also valid for the FCND, BFCND and SFCND problems, it can be shown using the same reduction. However, the ESFCND problem, which considers an additional maintenance constraint, has a MIP formulation with less variables that was used to solve it in Section 6.1.2.

Practical results

The results obtained differ in their possible use. All solutions provided are feasible network designs. However, for a decision regarding the entire network, one should focus more on Chapters 3 and 4. Indeed, the introduction of tapping enables more cost savings.

FCND and FCNDA

The design of Fiber Cables Networks in the FTTH is a significant issue for decision makers. By studying the solutions, it was shown in Chapters 3 and 4 that the introduction of the cable separation techniques (splicing and tapping) enables significant cost savings. This justifies our focus on fiber cables.

On a computational side, those savings are not easy to obtain. Regarding FCND, the integer programming solution proposed was able to solve successfully the smallest real-life instances. The valid inequalities introduced significantly decrease the computation times. However, the largest real-life instances stay out of reach. Regarding FCNDA, the integer programming approaches proposed here were efficient enough to solve real-life instances. The enhancements introduced sensitively improve the computational performances. Furthermore, the two integer approaches proposed are efficient in different contexts. In case

of an implementation, one could choose the one more suitable to the cases encountered.

Backfeed

The backfeed technique is practiced in real-life deployments by some Orange affiliates. The introduction of backfeed into the cable network optimisation process enables significant savings at the expense of harder computations.

Operations Administration and Maintenance considerations

We introduced in Chapter 6 a constraint meant to ease Administration and Maintenance Operations, especially repairs in case of failures. We compared two problems which differ by this constraint. After providing several resolution methods, we were able to compare the solutions.

The maintenance rule introduced can be considered as a reasonable compromise between capital expenditure for the network deployment and maintenance costs. Its implementation only increases the optimal solution cost by 3.7 % on our test instances.

We think this rule still needs to be developed for a more general problem in order to be implemented. Indeed, we compared it without the possibility to do any tapping.

We proposed two solution methods for the constrained problem. These methods are complementary, as they prove efficient in different contexts: the dynamic programming approach is generally faster in graphs where nodes have a small degree, whereas the mixed integer programming, embedding efficient valid inequalities, is generally faster otherwise.

Implementation

The work done in his thesis was implemented in a decision aid tool for Orange. The implementation uses a civil engineering arborescence, given by previous decision making. Then, demand nodes are gathered in Points A (PA) for management reasons. Then, on this arborescence, the FCNDA problem is solved, using the path-based model from Chapter 4. Finally, the PA are connected to their respective demand nodes, eventually using the backfeed technique.

This is a significant improvement over the older version of the decision aid tool, which used to tackle these decisions in a sub optimal fashion, using a dynamic programming algorithm.

Another advantage of this approach is the flexibility of integer programming. Indeed, it allows to formulate easily constraints specific to an instance, which is not the case with dynamic programming.

Perspectives

Arborescence selection

Chapter 3 tackles a more general problem than Chapter 4. It enables to make more important savings. On a purely financial level, it would be interesting to be able to tackle FCND. However, the studies of Chapter 3 showed that it is not yet within our computing power.

A different approach for solving FCND consists in selecting an arborescence first and then solve the FCNDA problem on it. This provides a sub-optimal solution for FCND. It is the solution chosen by Orange. A question that arises in this approach is: "which arborescence should we pick in order to have a good sub-optimal solution for FCND?".

Some elements of answer to this question are given by Section 3.3. Indeed, this arborescence is simpler to select in two asymptotic cases: one with negligible cable deployment costs, and the other one with negligible cable separation costs (welds and boxes). In further research, one could investigate more thoroughly this aspect of the problem.

Signal quality

In order to properly function, the network has to ensure a certain level of signal amplitude. While going through an FTTH network, the amplitude of a signal is attenuated. Some of this attenuation is due to the splitters. It is not our concern here, since we consider the splitting ratio of the network as well as the splitter locations to be fixed.

However, two other sources of signal intensity losses are due to the Fiber Cables Network Design:

- The distance on which the signal goes through fiber optics.
- The number and quality of the fiber welds the signal has to go through. Indeed, at each fiber weld, some losses occur.

Since some of this amplitude depends on the FCND solution chosen, it would be interesting to model the signal quality constraints. We leave this problem for future research.

Duct capacity

As specified in Section 2.1, duct capacities should in most real-life cases not be an issue, due to the relatively small size of fiber cables compared to ducts. Furthermore, we showed how to adapt the models of Chapter 4 (Remarks 4.2.3 and 4.3.2) in case capacity constraints are needed.

However, capacity constraints can still occur, and can be harder to model than the adaptation proposed. Indeed, it can be related to a circle packing problem. A possible direction of future research could be to investigate the different ways to express capacity constraints in order to implement them.

Client uncertainty

An aspect that can vary in practical applications is the number of clients. Indeed, it is usually not fixed before the network deployment. It would be interesting to formulate the problem with client uncertainty rather than a fixed demand. However, in order to have a good estimation of the client behavior, one needs preliminary studies.

Bibliography

- [1] Vincent Angilella, Matthieu Chardy, and Walid Ben-Ameur. Cables network design optimisation for the fiber to the home. *Design of Reliable Communication Networks*, 2016.
- [2] Vincent Angilella, Matthieu Chardy, and Walid Ben-Ameur. Design of fiber cable tree networks for the fiber to the home. *International Networks Optimisation Conference, Lisboa, Portugal*, February 2017.
- [3] Vincent Angilella, Matthieu Chardy, and Walid Ben-Ameur. Fiber cable network design in tree networks. *European Journal of Operations Research*, 2018.
- [4] Vincent Angilella, Matthieu Chardy, and Walid Ben-Ameur. Fiber cable network design with operations administration & maintenance constraints. *International Conference on Operations Research and Enterprise Systems, Funchal, Madeira, Portugal*, January 2018.
- [5] ARCEP. Acces au genie civil de france telecom, 2009. http://www.arcep.fr/uploads/tx_gspublication/consult-acces-genie-civil-ft-171209.pdf.
- [6] ARCEP. Electronic communications and postal regulatory authority decision no. 2010-1312 of 14 december 2010, 2010. <http://www.arcep.fr/fileadmin/reprise/dossiers/fibre/2010-1312-arcep-optical-fibre-decision-en.pdf>.
- [7] ARCEP. Observatoire des marches des communications electroniques. service fixes haut et tres haut debit, november 2017. <https://www.arcep.fr/index.php?id=13785>.
- [8] Sanjeev Arora and Boaz Barak. *Computational Complexity A modern Approach*. Cambridge University Press, 2009.
- [9] Alper Atamtürk and Oktay Günlük. Mingling: mixed-integer rounding with bounds. *Mathematical Programming*, pages 315–338, 2010.
- [10] M. Didi Biha, H. Kerivin, and A.R. Mahjoub. Steiner trees and polyhedra. *Discrete Applied Mathematics*, 112(1-3):101–120, september 2001.

- [11] Andreas Bley, Ivana Ljubić, and Olaf Maurer. Lagrangian decompositions for the two-level ftx network design problem. *European Journal of Computational Optimisation*, 1(3):221–252, november 2013.
- [12] Koen Casier, Bart Lannoo, Jan Van Ooteghem, Sophie Verbrugge, Didier Colle, Mario Pickavet, and Piet Demeester. Game-theoretic optimization of a fiber-to-the-home municipality network rollout. *Journal of Optical Communications Networks*, 1(1):30–42, june 2009.
- [13] L. Cen, K. F. Poon, Z. L. Yu, and A. Ouali. Automatic planning of gpon/ftth networks based on lagrangian heuristic optimization. In *2013 IEEE International Conference on Industrial Engineering and Engineering Management*, pages 335–339, Dec 2013.
- [14] M. Chardy, M-C Costa, A Faye, and M. Trampont. Optimising splitter and fiber location in a multilevel optical ftth network. *European Journal on Operational Research*, 222(3):430–440, 2013.
- [15] Sunil Chopra and M.R. Rao. The steiner tree problem i: Formulations, compositions and extension of facets. *Mathematical Programming*, 64(1):209–229, march 1994.
- [16] Sunil Chopra and M.R. Rao. The steiner tree problem ii: Properties and classes of facets. *Mathematical Programming*, 64(1):231–246, march 1994.
- [17] Ivan Contreras and Elena Fernandez. General network design: A unified view of combined location and network design problems. *European Journal of Operational Research*, 219(3):680–697, 2012.
- [18] Marilet de Andrade, Anna Buttaboni, Massimo Tornatore, Pierpaolo Boffi, Paolo Martelli, and Achille Pattavina. Optimisation of long-reach tdm/wdm passive optical network. *Optical Switching and Networking*, 16:36–45, april 2015.
- [19] Dorra Dhouib and Matthieu Chardy. Heuristic solution methods for the fiber to the home cabling problem. *Lecture Notes in Management Science*, 6:198–206, 2014.
- [20] Antonio Eira, Joao Pedro, and Joao Pires. Optimised design of multistage passive optical networks. *Journal of Optimisation in Communication Networks*, 4(5):402–411, may 2012.
- [21] Matteo Fischetti. Facets of two steiner arborescence polyhedra. *Mathematical Programming*, 51(1):401–419, july 1991.
- [22] Yoann Fouquet, Dritan Nace, Michał Pióro, and Michael Poss. An optimization framework for traffic restoration in optical wireless networks with partial link failures. *Optical Switching and Networking*, 23(Part 2):108 – 117, 2017. Design and modeling of Resilient optical networks RNDM 2015.

- [23] General-Cable. Fiber optic cable for voice and data transmission, 2017. <http://www.generalcable.com/na/us-can/products-solutions/communications/fiber-optic-cable/blown-optical-fiber-cable>.
- [24] Fred Glover. Improved linear integer programming formulations of nonlinear integer problems. *Management science*, 22(4):455–460, december 1975.
- [25] Michel X. Goemans. The steiner tree polytope and related polyhedra. *Mathematical Programming*, 63:157–182, 1994.
- [26] Michel X. Goemans and Young soo Myung. A catalog of steiner tree formulations. *Networks*, 23:19–28, 1993.
- [27] Elif Ilke Gokce and Wilbert E. Wilhelm. Valid inequalities for the multi-dimensional multiple-choice 0-1 knapsack problem. *Discrete Optimization*, 17:25–54, august 2015.
- [28] Stefan Gollowitzer, Luis Gouveia, and Ivana Ljubić. 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, march 2013.
- [29] Stefan Gollowitzer and Ivana Ljubić. Mip models for connected facility location: a theoretical and computational study. *Computers and Operations Research*, 38(2):435–449, february 2011.
- [30] Martin Grötschel, Christian Raack, and Axel Werner. Towards optimising the deployment of optical access networks. *European Journal on Computational Optimisation*, 2(1-2):17–53, april 2013.
- [31] Cedric Hervet. *Optimisation of optical network deployment. Considerations on demand uncertainty*. PhD thesis, Université, 2013.
- [32] Cedric Hervet, Alain Faye, Marie-Christine Costa, Matthieu Chardy, and Stanislas Francfort. Solving the two-stage robust ftth network design problem under demand uncertainty. *Electronic Notes in Discrete Mathematics*, 41:335–342, may 2012.
- [33] Hitachi. Premise & fiber optic cable catalog, 2017. <http://www.hca.hitachicable.com/products/hca/catalog/pdfs/hitachi-catalog.pdf>.
- [34] Richard M. Karp. *Reducibility among Combinatorial Problems*, pages 85–103. Springer US, Boston, MA, 1972.
- [35] Kyeong Soo Kim. On the evolution of pon-based ftth solutions. *Information Sciences*, 149:21–30, 2003.
- [36] Youngjin Kim, Youngho Lee, and Junghee Han. A splitter location-allocation problem in designing fiber optic access networks. *European Journal of Operational Research*, 120(2):425–435, april 2011.

- [37] A. Kokangul and A. Ari. Optimisation of passive optical network planning. *Applied Mathematic Modeling*, 35(7):3345–3354, july 2011.
- [38] Bernhard Korte and Jens Vygen. *Combinatorial optimization*. Springer, 2002.
- [39] Markus Leitner, Ivana Ljubić, Markus Sinnl, and Axel Werner. Ilp heuristics and a new exact method for bi-objective 0/1 ilps: Application to ftx-network design. *Computers & Operations Research*, 72:128 – 146, 03 2016.
- [40] Ivana Ljubić, Peter Putz, and Juan-José Salazar-González. A mip-based approach to solve the prize-collecting local access network design problem. *European Journal of Operational Research*, 235(3):727 – 739, 2014.
- [41] MJ Magazine and Osman Oguz. A fully polynomial approximation algorithm for the 0–1 knapsack problem. *European Journal of Operational Research*, 8(3):270–273, 1981.
- [42] Geraldo R. Mateus, Henrique P.L. Luna, and Adriana B. Sirihal. Heuristics for distribution network design in telecommunication. *Journal of Heuristics*, 6:131–148, 2000.
- [43] A. Mitsenkov, P. Bakos, G. Paksy, and T. Cinkler. Technology-independent topology design heuristics for point-to-multipoint optical access networks. In *2013 17th International Conference on Optical Network-ing Design and Modeling (ONDM)*, pages 298–303, April 2013.
- [44] Mariusz Mycek, Michał Pióro, and Mateusz Żotkiewicz. Mip model for efficient dimentioning of real-world ftth trees. *Telecommunication Systems*, September 2017.
- [45] Konstantinos Papaefthimiou, Yonas Tefera, Dimitar Mihaylov, Jose M. Guitierrez, and Michael Jensen. Algorithmic pon/p2p ftth access network design for capex minimization. *21st Telecommunications forum TELFOR*, 2013.
- [46] Gigyoung Park, Youngho Lee, and Junghee Han. A two-level location–allocation problem in designing local access fiber optic networks. *Computers & Operations Research*, 51:52–63, 2014.
- [47] Michał Pióro, Dritan Nace, and Yoann Fouquet. An optimization model for communication networks with partial multiple link failures. In *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2013 5th International Congress on*, pages 171–178. IEEE, 2013.
- [48] R. Sommerhalder and S.C. van Westrhenen. *The theory of computability: programs, machines, effectiveness, and feasibility*. International computer science series. Addison-Wesley Pub. Co., 1988.

- [49] Mathieu Tahon, Jan Van Ooteghem, Koen Casier, Sofie Verbrugge, Didier Colle, Mario Pickavet, and Piet Demeester. Improving the ftth business case—a joint telco-utility network rollout model. *Telecommunications Policy*, 38(5):426 – 437, 2014. Special issue on : Selected papers from the 10th Conference in Telecommunications, Media and Internet Techno-economics Special issue on : The development of public e-services: Empirical analysis and policy issues.
- [50] L. Tiantong and C. Saivichit. A proposed network economic strategic model for ftth access network. In *2013 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, pages 1–6, May 2013.
- [51] Artur Tomaszewski, Michał Pióro, and Mateusz Żotkiewicz. On the complexity of resilient network design. *Networks*, 55(2):108–118, 2010.
- [52] Paolo Toth. Dynamic programming algorithms for the zero-one knapsack problem. *Computing*, 25(1):29–45, 1980.
- [53] Mathieu Trampont. *Modélisation et optimisation du déploiement des réseaux de télécommunications : applications aux réseaux d'accès cuivre et optiques*. PhD thesis, Conservatoire National des Arts et Métiers, 2009.
- [54] S.P. van Loggerenberg, M.J. Grobler, and S.E. Terblanche. Solving the passive optical network with fiber duct sharing planning problem using discrete techniques. *Electronic Notes in Discrete Mathematics*, 41:343–350, june 2013.
- [55] Vodafone. Architectures and competitive models in fibre networks, 2010. <http://www.generalcable.com/na/us-can/products-solutions/communications/fiber-optic-cable/blown-optical-fiber-cable>.
- [56] Başak Yazar, Okan Arslan, Oya Ekni Kardeş, and Bahar Y. Kara. Fiber optical network design problems: A case for turkey. *Omega*, 2015.
- [57] M Żotkiewicz and M Mycek. Reducing the costs of ftth networks by optimized splitter and olt card deployment. *IEEE/OSA Journal of Optical Communications and Networking*, 9(5):412–422, 2017.
- [58] Mateusz Żotkiewicz, Walid Ben-Ameur, and Michał Pióro. Finding failure-disjoint paths for path diversity protection in communication networks. *IEEE Communications Letters*, 14(8):776–778, 2010.