



HAL
open science

Planification adaptative des ressources ferroviaires

Rémi Lucas

► **To cite this version:**

Rémi Lucas. Planification adaptative des ressources ferroviaires. Mathématiques [math]. Institut Polytechnique Paris, 2020. Français. NNT: . tel-03009036

HAL Id: tel-03009036

<https://theses.hal.science/tel-03009036v1>

Submitted on 17 Nov 2020

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.



INSTITUT
POLYTECHNIQUE
DE PARIS



NNT : 2020IPPAE006

Thèse de doctorat

Planification adaptative des ressources ferroviaires

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à École Nationale Supérieure de Techniques Avancées (ENSTA Paris)

École doctorale n°574 École Doctorale de Mathématiques Hadamard (EDMH)
Spécialité de doctorat : Mathématiques appliquées et applications des mathématiques

Thèse présentée et soutenue à Palaiseau, le 25 septembre 2020, par

RÉMI LUCAS

Composition du Jury :

Safia Kedad-Sidhoum Professeur, Cnam (CEDRIC)	Présidente du jury
Stéphane Dauzère-Pérès Professeur, École des Mines de St-Étienne (LIMOS)	Rapporteur
Anass Nagih Professeur, Université de Lorraine (LCOMS)	Rapporteur
Olivier Hudry Professeur, Télécom Paris (LTCl)	Examineur
Paola Pellegrini Directrice de recherche, Université Gustave Eiffel (COSYS-LEOST)	Examinatrice
Sourour Elloumi Professeur, ENSTA Paris (UMA)	Directrice de thèse
Zacharie Alès Maître de Conférence, ENSTA Paris (UMA)	Encadrant
Rémy Chevrier Docteur, Chef de projets, SNCF (Innovation & Recherche)	Encadrant
François Ramond Docteur, Responsable d'équipe, SNCF (Innovation & Recherche)	Encadrant

À mon grand-père, Jacques.

Remerciements

Bien que les remerciements soient systématiquement placés au tout début du manuscrit de thèse, il s'agit toujours de la dernière partie que l'on rédige. Et elle est très importante! C'est en quelque sorte l'*incipit* de tout manuscrit, l'accroche qui permet de captiver le lecteur et d'aiguiser sa curiosité, que je sais insatiable dans la mesure où il vient d'achever la troisième ligne sans avoir refermé l'ouvrage¹. Mais les remerciements, ce sont surtout les pages où le docteur peut se permettre de faire un peu de fantaisie et de laisser transparaître sa patte artistique et littéraire, en remerciant chaque personne l'ayant aidé d'une façon ou d'une autre à achever son travail de thèse². Et je ne vais pas m'en priver.

Je tiens tout d'abord à remercier mes encadrants de thèse avec qui j'ai eu l'occasion de travailler ces dernières années, et qui m'ont tous accordé une grande confiance pour mener mes travaux jusqu'à leur terme.

Mes premiers remerciements vont donc comme il se doit à ma directrice de thèse Sourour Elloumi. C'est en partie grâce à toi que j'ai eu l'opportunité de postuler à cette thèse, et ce fut particulièrement agréable de travailler ensemble. J'ai apprécié la flexibilité que tu m'as laissée quant à l'orientation du sujet, ta disponibilité, tes conseils et ta gentillesse. Dans le cadre d'une thèse CIFRE en entreprise, les relations entre encadrants académiques et industriels sont primordiales, et je suis très satisfait de la façon dont tu as mené cette collaboration avec François et Rémy. D'autre part, tu as su t'adapter et trouver un équilibre entre la pertinence industrielle de mon sujet et les enjeux scientifiques de mes travaux, ce qui n'est jamais évident lorsque les objectifs ne sont pas concordants.

Merci à François Ramond, responsable de l'équipe MOD à SNCF Innovation & Recherche, et encadrant principal de ma thèse à la SNCF. Tu m'as laissé une grande autonomie au quotidien pour mener ce sujet exploratoire, et m'a poussé à valoriser mes travaux en interne dès que cela était possible, me redonnant confiance lors des différentes périodes où j'en manquais. Les nombreux échanges que nous avons eus pour construire ce sujet ont été très enrichissants tant sur le plan scientifique que sur le plan humain. Tu as toi aussi géré avec tact la collaboration avec l'ENSTA, me laissant notamment approfondir pendant de longues périodes certains aspects purement scientifiques de mes travaux.

Merci également à Zacharie Alès, ayant rejoint Sourour pour l'encadrement académique peu après le début de mes travaux. Vous vous êtes très bien complétés, et je ne compte plus les services que tu m'as

1. Ou à défaut, son lecteur pdf.

2. Et avant que ce dernier ne l'achève lui-même.

REMERCIEMENTS

rendus pour accélérer la génération de certaines figures de ce manuscrit. Je te remercie aussi pour ton dynamisme et ta réactivité lors de nos échanges, qu'il s'agisse de discussions sur certains modèles mathématiques, de relectures de ce manuscrit, ou de divers sujets annexes comme les vidéos de vulgarisation sur internet.

Merci enfin à Rémy Chevrier, chef de projet chez MOD, ayant complété l'équipe encadrante de cette thèse avec brio en y apportant toute son expertise ferroviaire et sa bonne humeur. C'est toujours un plaisir de discuter avec toi de sujets aussi divers que le symbolisme de la pomme dans nos cultures occidentales ou de langages informatiques ancestraux dont le nom ne m'évoquait pas grand-chose avant de faire ta connaissance!

J'adresse mes remerciements à l'ensemble de mon jury de thèse, pour avoir accepté de se plonger dans le présent manuscrit avec attention³, avoir assisté à ma présentation que je me garderais bien de commenter ici, et m'avoir questionné avec bienveillance pendant... un certain temps! Un merci particulier aux rapporteurs Stéphane Dauzère-Pérès et Anass Nagih pour leurs lectures minutieuses et leurs remarques dans leurs rapports, ayant permis les premières modifications du manuscrit. Merci également aux examinateurs Olivier Hudry et Paola Pellegrini, ainsi qu'à Safia Kedad-Sidhoum pour avoir présidé le jury. Que vous ayez été masqués et physiquement présents ou à visage découvert à distance, cela n'a pour moi fait aucune différence, mais je remercie particulièrement celles et ceux ayant fait le déplacement dans cette lointaine contrée reculée qu'est Palaiseau, qui plus est un vendredi matin.

Merci aux personnes avec qui j'ai eu l'occasion de travailler à la SNCF, que ce soit à la Direction Innovation & Recherche ou dans d'autres entités : Nicolas et Alexandre chez ROAD pour leur réactivité et leur intérêt à chacune de mes interrogations sur l'outil PRESTO, Vincent et Abdellah chez Fret, Timothée chez Voyages. Je remercie également David de Almeida, chef du département MEV à Innovation & Recherche, pour sa confiance au cours de mes travaux de thèse, son suivi régulier et les idées qu'il a pu apporter durant nos différents points d'étape.

Merci à l'ensemble des membres de l'équipe MOD au sein de laquelle j'ai pu évoluer à la SNCF en y étant pleinement intégré en tant que doctorant. Ce fut un plaisir de partager ses moments quotidiens avec vous : les pots d'arrivée qu'il ne fallait surtout pas manquer, le sacro-saint repas du midi avec son départ groupé, les parties de baby-foot endiablées avec des cannes parfaitement huilées et les discussions philosophiques autour de la machine à café. Merci donc aux chefs de projets Sabine, Valentina, Mathieu, Juliette, Xavier et Fabrice. Merci aux désormais docteurs de cette même équipe, ayant brillamment soutenu leurs thèses respectives : Estelle pour ta bonne humeur et ton rire, Lucile pour ton écoute et ta bienveillance, et Sélim, mon ex-voisin de bureau, pour ta curiosité et tous les bons moments qu'on a pu passer à discuter de tout et de rien pendant ces trois dernières années. Tu as su me redonner confiance dans certains moments où j'en avais le plus besoin. Merci également à Marie et Mohamed pour tous nos échanges sur les difficultés rencontrées dans nos thèses respectives : on est tous arrivé au bout! Je souhaite bon courage aux doctorants Hugo C. et Hugo B. pour la relève dans l'équipe MOD. Merci enfin à l'orateur brillant qu'est Nicolas, chef de

3. Comme vous vous apprêtez vous aussi à le faire.

REMERCIEMENTS

projet dont l'admiration sans faille pour les doctorants n'a d'égal que son comique de répétition tout à fait assumé et bienvenu.

Bien que peu présent au laboratoire UMA à l'ENSTA, j'y ai été très bien accueilli et intégré à chacun de mes passages. Je remercie notamment Frédéric Jean, ex-directeur adjoint de l'école doctorale, ayant été particulièrement attentif et réactif face aux innombrables et récurrents problèmes administratifs que j'ai pu rencontrer à chaque (ré)inscription annuelle. Merci à Marie-Christine Costa, qui m'a fait découvrir la recherche opérationnelle à l'ENSTA pendant mon cycle ingénieur, m'a aidé à affiner mon projet de thèse et a suivi l'évolution de mes travaux à chaque fois que nous nous sommes croisés au Cnam. Merci à Corinne, secrétaire du laboratoire UMA, qui a toujours été là pour répondre à mes questions, et fait tout ce qui était en son pouvoir pour accélérer les choses lorsque des difficultés administratives se présentaient. Merci enfin à Maurice, ingénieur et sysadmin du laboratoire, pour sa gentillesse et sa patience, pour les discussions improvisées sur tout ce qui touchait aux langages de programmation, et pour son aide pour résoudre mes soucis informatiques.

Heureusement, il y a une vie en dehors de la thèse, qui n'est censée occuper le doctorant qu'une quarantaine d'heures hebdomadaires. Je remercie donc tous mes amis sur qui j'ai pu compter pour m'aider et m'évader, particulièrement dans les moments de doute.

Merci au groupe des docteurs anonymes autrement connu sous le nom de « FAQ de RO », (pas) toujours là pour réagir à mes diverses remarques et réactions plus ou moins inintéressantes, surtout lorsqu'elles sont ponctuées de fautes de frappe : Hadrien, dont les avis tranchés et éclairés sont une source inépuisable de débats pour me changer les idées; Rémi, dont je ne comprends toujours pas les blagues mais qui fait encore l'effort de me les expliquer; Émilien, qui a le mérite d'exister et de me battre aux échecs. Merci également au dormeur Youssef et au prudent PV, qui gravitent non loin de ce groupe mais n'aiment pas trop le spam, et ils ont bien raison.

Merci à Robin et Basile pour les discussions au cours de nos randonnées estivales, avec plus ou moins de neige sur le chemin. Merci également à Robin pour les dizaines de concerts qu'on a pu aller voir, même ceux annulés.

Merci à mes amis de prépa : Pritish, Dung et Marie pour tous ces déjeuners à ressasser de bons souvenirs; Lina et Ismaïl pour leur écoute malgré la distance; Léo et Raphaël, dont l'amitié m'est précieuse.

Merci aux membres de la chorale Not'en Bulles et à son conseil d'administration, qui m'ont permis de décompresser dans les moments charnières en m'investissant dans de chouettes projets artistiques à leurs côtés : je citerai notamment Pierre, Thien Van et Rouba, pour leur écoute, ainsi que Steph, pour nos débats souvent futiles mais toujours argumentés et bienveillants.

Merci à l'association Suzanne et ma troupe de théâtre des Déjantés, trop nombreux pour être tous cités⁴ : vous êtes tous géniaux, c'est un réel plaisir et un énorme moyen de décompression que de jouer ensemble. Si j'adapte un jour ce manuscrit de thèse en pièce de théâtre, il ne fait aucun doute qu'il y aura un rôle pour chacun d'entre vous⁵!

4. Donc je citerai seulement son président Manu, parce que j'y suis obligé, ainsi que Léo, et seulement en bas de page.

5. C'est la Vérité.

Je remercie enfin ma famille, qui a toujours été présente lorsque j'ai ressenti le besoin de discuter : mes parents Laurence et Philippe qui m'ont toujours soutenu dans mes choix; mon grand-père Jacques, inventeur incompris du barbecue vertical avant l'heure, ce qui vaut au moins un doctorat; mon frère Simon, génie du web promis à un grand avenir; mon oncle Patoun, jedi en quête de la Force. Merci également à mon parrain Thierry et ma tante Nathalie pour avoir fait le déplacement et avoir été présents à ma soutenance.

Note sur la confidentialité

Cette version du manuscrit est la version diffusable sur Internet jusqu'au 31 décembre 2025. Certains chapitres ont été retirés par rapport à la version confidentielle. En conséquence, certaines références internes au manuscrit sont manquantes. Merci de contacter l'auteur pour de plus amples informations.

REMERCIEMENTS

Table des matières

Remerciements	5
Note sur la confidentialité	9
Liste des tableaux	15
Table des figures	17
Introduction générale	19
1 Contexte et motivations	23
1.1 Le système ferroviaire	23
1.1.1 Organisation	23
1.1.2 Train et ressources ferroviaires	24
1.1.3 Plan de transport et planification ferroviaire	24
1.1.3.1 Pourquoi planifier?	24
1.1.3.2 Les différentes phases de la planification ferroviaire	25
1.2 Hyperplanification du plan de transport	27
1.2.1 Efficacité de la planification	27
1.2.2 Vers une planification adaptative?	28
1.2.2.1 Deux approches pour introduire de la flexibilité	28
1.2.2.2 Une flexibilité difficilement envisageable	29
1.2.2.3 Planification et anticipation	29
1.3 Organisation de la suite du manuscrit	30
2 État de l'art	33

TABLE DES MATIÈRES

2.1	Planification ferroviaire	33
2.1.1	Planification séquentielle des ressources	33
2.1.1.1	Les missions	34
2.1.1.2	La grille horaire	34
2.1.1.3	Le Graphique d’Occupation des Voies (GOV)	39
2.1.1.4	Les agents	39
2.1.1.5	Maintenances	40
2.1.2	Planification intégrée des ressources	40
2.1.3	Replanification	41
2.2	Flexibilité et agilité en planification	42
2.2.1	Cas des problèmes d’ordonnancement sur machine	42
2.2.1.1	Description formelle	42
2.2.1.2	Deux types de flexibilité	43
2.2.2	Flexibilité d’une grille horaire	45
2.2.2.1	Grille horaire flexible pour mieux planifier les autres ressources	45
2.2.2.2	Grille horaire flexible avec variantes de sillons	46
2.2.2.3	Grille horaire flexible en temps réel	47
2.2.3	Contextes d’optimisation à plusieurs étapes	48
2.2.3.1	Aperçu	49
2.2.3.2	Applications au domaine ferroviaire	53
3	Génération de roulements matériels nominaux	57
3.1	Planification du matériel roulant	57
3.1.1	Parc matériel et unités simples	58
3.1.2	Roulement matériel	58
3.2	État de l’art pour la planification du matériel roulant	60
3.2.1	Les différentes formulations de la littérature	60
3.2.2	Les modèles utilisés à la SNCF	63
3.2.2.1	PRESTO	63
3.2.2.2	HERMES	64
3.3	Le Problème de Conception d’un Roulement Matériel Nominal (Problème Nominal)	64

TABLE DES MATIÈRES

3.3.1	Instance du Problème Nominal	66
3.3.2	Graphe de précedence	69
3.3.3	Modèle PLNE	70
3.3.3.1	Notations additionnelles	70
3.3.3.2	Variables	70
3.3.3.3	Objectifs	71
3.3.3.4	Contraintes	74
3.3.4	Génération des lignes de roulement	76
3.4	Paramétrisation et raffinements du modèle	76
3.4.1	Durée maximale entre deux trains	76
3.4.2	Ajout de sommets-garages supplémentaires	77
3.4.3	Contraintes de continuités	79
3.4.4	Durée maximale d'attente en gare	79
4	Conclusion et perspectives	81
	Bibliographie	85

TABLE DES MATIÈRES

Liste des tableaux

3.1	Données d'une instance I relatives aux séries matérielles	67
3.2	Données d'une instance I relatives aux gares	67
3.3	Données d'une instance I relatives aux trains	67
3.4	Notations liées au graphe de précédence du Problème Nominal	69

LISTE DES TABLEAUX

Table des figures

1.1	Exemple de GET	24
1.2	Étapes de la planification ferroviaire en France	26
1.3	Organisation du manuscrit et interdépendances des chapitres	31
3.1	Exemple de roulement matériel	59
3.2	Exemple d'un graphe de transition utilisé par Fioole et al. (2006)	61
3.3	Exemple de graphe espace-temps	62
3.4	Comparaison entre les graphes de PRESTO et ceux d'HERMES	65
3.5	Plusieurs décompositions en chemins pour HERMES	66
3.6	Schéma du Problème Nominal	66
3.7	Exemple d'un graphe de précedence avec 4 trains et 3 gares	68
3.8	Fonction de répartition de la durée entre deux trains consécutifs sur une même ligne de roulement, sur une instance Fret.	77
3.9	Exemple d'un graphe de précedence avec 8 trains et 3 gares, illustrant les sommets-garages intermédiaires	78
3.10	Exemple d'utilisation des contraintes de continuité	79

TABLE DES FIGURES

Introduction générale

Contexte et enjeux

La planification ferroviaire consiste à établir en amont des circulations un plan de transport spécifiant l'emploi du temps de chaque ressource : les sillons, qui correspondent aux horaires des trains, les matériels roulants devant assurer leur circulation, ainsi que les agents à bord de ces trains. Compte tenu de la complexité du système ferroviaire, cette planification est effectuée potentiellement plusieurs années à l'avance, où un plan de transport précis est établi.

Or, une fois qu'une ressource a été planifiée, il est parfois nécessaire d'adapter le plan de transport et donc de le changer. Les raisons peuvent être multiples : changement horaire pour causes de travaux, infrastructure partiellement endommagée, demande d'une rame supplémentaire pour un train, avarie sur un type de matériel roulant... Ainsi, le plan de transport est parfois adapté à plusieurs reprises entre sa conception potentiellement plusieurs années à l'avance et son exécution le jour des circulations. Dès lors, il est légitime de se demander si la conception d'un plan de transport précis avec des échelles de temps aussi importantes est raisonnable.

Dans cette thèse, nous envisageons d'ajouter de la flexibilité dans le processus de planification, afin que les efforts déployés pour adapter le plan de transport lorsque cela est nécessaire soient réduits : nous cherchons ainsi à effectuer ce que nous appelons une planification adaptative, et nous nous intéressons particulièrement au cas du matériel roulant.

Organisation du manuscrit et principales contributions

Le Chapitre 1 présente le contexte industriel dans lequel s'inscrivent ces travaux de thèse. Nous y définissons les principales composantes du système ferroviaire, en mettant particulièrement l'accent sur la notion de planification, paradigme historiquement retenu pour concevoir à l'avance le plan de transport. Nous nous interrogeons alors sur l'efficacité de la procédure telle qu'elle existe aujourd'hui, mettant à la fois en évidence les larges échelles de temps de la planification ferroviaire, le manque d'agilité du processus dans son ensemble, et la rigidité du plan de transport obtenu : on parle alors d'hyperplanification, principal constat ayant motivé les travaux de cette thèse. Nous expliquons les raisons pour lesquelles l'introduction de flexibilité dans la planification et dans le plan de transport est délicate à mettre en œuvre, et l'hyper-

planification difficilement dispensable. La planification adaptative telle que nous l'envisageons est ensuite explicitée : bien qu'elle ne remette pas en question l'hyperplanification, elle permet de pallier un de ses principaux inconvénients, à savoir les multiples adaptations et réoptimisations du plan de transport durant les différentes phases de la planification ferroviaire. En anticipant à l'avance tous ces futurs changements du plan de transport dès sa conception, on peut concevoir un plan de transport adaptatif qui s'adaptera facilement et à moindre coûts dès lors qu'une demande de changement ou un aléa intervient.

Le Chapitre 2 présente un état de l'art de la littérature scientifique sur différents aspects liés aux enjeux de ces travaux de thèse exposés dans le Chapitre 1. Il est organisé en deux sections, concernant d'une part les modèles mathématiques de planification ferroviaire, et d'autre part la notion de flexibilité et agilité en planification. Il est à noter que la première section ne contient pas l'état de l'art sur la planification du matériel roulant, qui sera exposée dans le Chapitre 3. La seconde section traite notamment le cas des problèmes d'ordonnancement sur machines, où la notion de flexibilité a été introduite et discutée par plusieurs auteurs. Nous présentons un état de l'art sur la flexibilité des grilles horaires dans le ferroviaire, les sillons semblant être la seule ressource où une flexibilité telle que définie dans le Chapitre 1 a été étudiée. Enfin, nous présentons les contextes de décision en plusieurs étapes, qui permettent d'introduire une certaine flexibilité envisagée comme de la robustesse.

Le Chapitre 3 traite le cas nominal de la planification du matériel roulant. Après quelques notions et définitions usuelles relatives au matériel roulant, nous réalisons un état de l'art de la littérature scientifique concernant cette ressource, et exposons certains outils logiciels utilisés au sein de SNCF pour concevoir des *roulements matériels* optimisés, un roulement matériel correspondant à un planning pour le matériel roulant. Nous présentons ensuite l'un des trois problèmes que nous avons définis au cours de nos travaux. Le Problème de Conception d'un Roulement Matériel Nominal s'appuie sur un graphe de précedence et un modèle de flot en Programmation Linéaire en Nombres Entiers (PLNE). L'ensemble des variables et contraintes du PLNE sont explicitées dans ce chapitre. Les PLNE du Chapitre ?? reprendront massivement les notations et notions introduites ici.

Le Chapitre ?? s'intéresse spécifiquement à l'adaptation du matériel roulant, et aboutit sur la notion d'adaptabilité, cœur de la planification adaptative telle que nous l'avons envisagée. Il constitue les principales contributions de ces travaux de thèse. Après avoir effectué un état de l'art sur l'adaptation du matériel roulant, nous définissons la notion de coûts d'adaptation d'un plan de transport. Nous discutons de la forme de tels coûts pour le matériel roulant, et nous proposons une définition formelle des coûts d'adaptation d'un roulement matériel pouvant être optimisés au moyen de programmes mathématiques. Nous définissons la notion de scénario d'adaptation et les coûts d'adaptation structurels, pouvant se généraliser à d'autres domaines de la recherche opérationnelle.

Nous introduisons dans un premier temps le Problème d'Adaptation du Matériel Roulant qui, étant donné un roulement nominal et un scénario d'adaptation, cherche un nouveau roulement adapté au scénario et dont l'objectif est de minimiser les coûts d'adaptation, en respectant au mieux les demandes formulées

par le Gestionnaire d'Infrastructure et l'Entreprise Ferroviaire. Nous sommes parvenus à modéliser ce problème par un PLNE dont la taille est du même ordre de grandeur que celle du PLNE du Chapitre 3, cherchant un roulement nominal. La résolution de ce problème s'appuie sur deux briques algorithmiques. La première est la résolution du PLNE. La seconde est un algorithme original de post-traitement pour décomposer un flot en chemins, et nous proposons à la fois un algorithme glouton et une formulation par programmation mathématique.

Nous introduisons dans la dernière section de du Chapitre ?? la notion d'adaptabilité d'un roulement matériel, et définissons le Problème de Conception d'un Roulement Matériel Adaptatif. Cette fois, étant donnés tous les scénarios qui peuvent se produire, disons au nombre de k , on souhaite calculer à la fois un roulement nominal, optimal ou presque, et k roulements adaptés à chacun des k scénarios. L'objectif est de minimiser la somme des coûts d'adaptation des k roulements adaptés relativement au roulement nominal. Nous sommes également parvenus à modéliser ce problème par un PLNE dérivé des précédents mais dont la taille est environ $(k + 1)$ fois plus importante que les deux PLNE précédents. La résolution de ce Problème permet de réaliser la planification adaptative telle qu'envisagée au cours de nos travaux.

Le Chapitre ?? présente nos résultats expérimentaux autour des trois problèmes que nous avons définis dans les chapitres précédents. Les instances utilisées sont issues de jeux de données SNCF, et nous discutons des différentes façons de générer des scénarios d'adaptation. Nous introduisons notamment une méthode pour évaluer les gains de la planification adaptative par rapport à l'approche actuellement en vigueur : pour cela, nous utilisons nos trois problèmes. La qualité des résolutions et les gains réalisés sont discutés et mis en perspective, et l'influence de certains paramètres de modélisation est mise en avant.

Enfin, nous concluons l'ensemble des travaux exposés dans cette thèse dans le Chapitre 4. Nous y présentons également différentes perspectives permettant de poursuivre nos travaux, en améliorant et en tirant davantage partie du concept de planification adaptative.

Chapitre 1

Contexte et motivations

Ce chapitre introductif présente brièvement l'organisation du système ferroviaire en France, en introduisant notamment les notions et définitions essentielles qui seront reprises dans le reste du manuscrit. Dans la seconde section, nous introduisons plus spécifiquement le contexte ferroviaire industriel ayant motivé les travaux développés dans ce manuscrit. On y détaille notamment le cheminement intellectuel qui a permis de parvenir au concept de planification adaptative tel que nous l'avons envisagé compte tenu des enjeux initiaux au début de ce travail de thèse. Enfin, la dernière section détaille l'organisation du reste du manuscrit avec les principales contributions de nos travaux.

1.1 Le système ferroviaire

1.1.1 Organisation

En France, le système ferroviaire est organisé autour du principe fondamental de séparation entre le *gestionnaire d'infrastructure* (GI) et les *entreprises ferroviaires* (EF).

Le GI français est *SNCF Réseau*, l'une des sociétés anonymes (SA) de la maison-mère SNCF¹. Il s'agit du propriétaire de l'ensemble du réseau national. Ainsi, le GI est notamment en charge de sa maintenance et de la construction de potentielles nouvelles lignes. Les EF souhaitant faire circuler un train sur le réseau doivent préalablement en faire la demande au GI, et celui-ci facture l'EF en conséquence.

En France, les principales EF sont rassemblées au sein de la SA *SNCF Voyageurs* pour le transport de voyageurs (TGV, Intercités, TER, Transilien)². Il existe aujourd'hui environ une vingtaine d'EF opérant du transport de marchandises en France, et *Fret SNCF* est l'une d'entre elles. Avec l'ouverture du marché à la concurrence, mise en place entre 2003 et 2006 pour le fret et à partir de 2020 pour le transport de voyageurs,

1. La SNCF est devenue une société anonyme à capitaux publics le 1^{er} janvier 2020. Elle était auparavant constituée de 3 EPIC (Établissement Public à Caractère Industriel et Commercial).

2. Il existe également des EF résultant de coopérations internationales comme Eurostar, ainsi que des EF étrangères pouvant circuler sur le réseau ferré national.

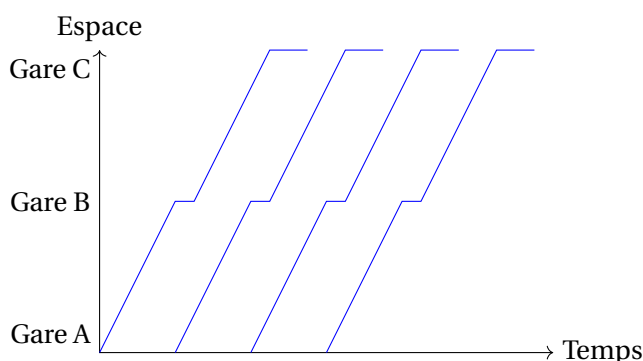


FIGURE 1.1: Exemple de graphique espace-temps (GET) comprenant 4 sillons. Chaque point sur le GET correspond à une position dans l'espace unidimensionnel (en ordonnée) et à un instant précis (en abscisse). Chaque sillon relie la gare A à la gare C avec un arrêt intermédiaire en gare B : deux sillons ne peuvent se croiser s'ils correspondent à des trains circulant sur la même voie. Sur une section donnée, plus la pente est importante, plus la vitesse du train est élevée.

le nombre d'EF françaises est amené à évoluer. Le GI se doit d'être impartial lorsque plusieurs EF formulent des demandes d'accès au réseau incompatibles entre elles : la séparation entre GI et EF est donc essentielle pour l'ouverture du marché à la concurrence.

1.1.2 Train et ressources ferroviaires

Dans la suite de ce manuscrit, on désignera par « train » la rencontre de plusieurs *ressources ferroviaires* de différents types pour la planification d'un service au client final :

- un *sillon*, correspondant à l'itinéraire horairisé effectivement suivi par le train. Lorsque le GI répond favorablement à une EF souhaitant faire circuler un train, on dit alors qu'elle lui *réserve* un sillon. Pour représenter des sillons, on utilise généralement un graphique espace-temps, ou GET (voir Figure 1.1), largement répandu tant dans les entités opérationnelles que dans les publications scientifiques.
- un ou plusieurs matériels roulants, correspondant aux rames physiques qui vont couvrir ce train. On dit alors que ces matériels sont *affectés* au train considéré, ou encore que ces matériels *couvrent* le train;
- des agents, dont notamment le conducteur (appelé *Agent De Conduite*, ou ADC) et les éventuels contrôleurs (appelés *Agents du Service Commercial Train*, ou ASCT).

1.1.3 Plan de transport et planification ferroviaire

1.1.3.1 Pourquoi planifier?

La planification ferroviaire a été introduite et conceptualisée au XIX^e siècle. Elle consiste à *planifier* l'ensemble des ressources ferroviaires à l'avance. La planification ferroviaire permet d'assurer que les ressources

nécessaires à la production de chaque train seront bien disponibles à l'heure et au lieu requis. Le *plan de transport* est le résultat de la planification : il désigne l'ensemble des plans qui ont été établis pour toutes les ressources ferroviaires³.

Selon le cours d'introduction Bouvarel (2012), le système ferroviaire présente deux caractéristiques qui rendent nécessaire la planification :

1. c'est un système à une dimension d'espace;
2. un train sur rails subit un frottement acier sur acier permettant d'atteindre de grandes vitesses mais avec une distance de freinage très importante.

Dès lors, contrairement au cas du trafic routier, il est très difficile pour un ADC d'éviter un obstacle présent sur la voie (comme un autre train) s'il n'a pas été préalablement anticipé. Chaque opérateur du système ferroviaire ne dispose que d'une vision très partielle du système ferroviaire. La *signalisation* ferroviaire permet de pallier ces difficultés et d'assurer la sécurité des circulations. Elle apporte aux ADC une information sur le train situé devant eux, leur indiquant si la voie est *libre*. Nous ne détaillerons pas davantage le fonctionnement et l'intérêt de la signalisation ferroviaire dans ce manuscrit, préférant nous concentrer sur la planification. Ses objectifs sont multiples :

- s'assurer à l'avance que le GI et l'EF sont capables de produire le service promis au client : même si on sait qu'il y aura des aléas en opérationnel, il est essentiel de pouvoir garantir la faisabilité du plan de transport en amont de sa réalisation. Ceci est d'autant plus vrai aujourd'hui, avec la densification du trafic et la complexité croissante du réseau;
- faire en sorte que chaque opérateur du système ferroviaire tente de respecter le plan de transport au mieux : si chacun agit individuellement de telle sorte à recoller le plus possible au plan de transport théorique – dit *nominal* – lors d'aléas opérationnels, les besoins de communication entre les acteurs et les besoins en *régulation* du système ferroviaire sont moins cruciaux;
- rationaliser l'utilisation des ressources dans un souci d'efficacité.

1.1.3.2 Les différentes phases de la planification ferroviaire

La planification ferroviaire en France se déroule généralement en plusieurs étapes essentiellement décrites sur la Figure 1.2. Lors de chaque *phase* de la planification, les ressources ferroviaires sont planifiées avec plus ou moins de précision, et avec un objectif potentiellement différent. En *amont* de la planification, il s'agit surtout de *dimensionner* les ressources : faisabilité *a priori* de faire circuler des trains avec une certaine fréquence sur une ligne donnée, nombre de rames physiques à attribuer à un périmètre géographique pour une certaine période etc. Le plan de transport en aval de la planification, correspondant à *l'opérationnel*, est quant à lui beaucoup plus précis.

3. Certains experts considèrent que le plan de transport désigne uniquement la planification des sillons. La définition que nous avons choisi de retenir dans ce manuscrit est plus générale.

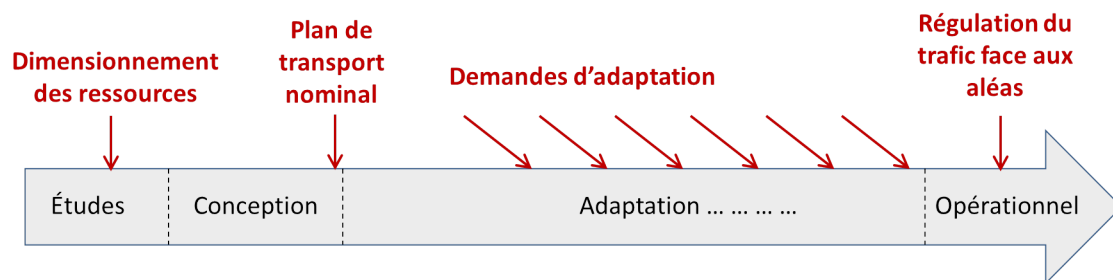


FIGURE 1.2: Étapes de la planification ferroviaire en France

Études et Conception

La planification ferroviaire commence plusieurs années avant la circulation des trains, par la phase d'études. Il peut s'agir d'ouvrir une nouvelle ligne ferroviaire, de restructurer la grille horaire, de jauger les apports d'un nouveau type de matériel roulant etc. Lorsque le plan de transport est globalement reconduit d'une année sur l'autre sans modification importante, la phase d'études est très rapide.

La phase suivante est la *conception*, où un plan de transport est établi pour chaque ressource. Elle commence à une date différente selon la ressource ferroviaire et l'entité considérée.

- La conception de la grille horaire se fait avec l'attribution des sillons par le GI. Cette procédure suit un cycle annuel, où la grille horaire de l'année A est fixée début décembre de l'année $A - 1$. Les EF doivent formuler toutes leurs demandes de sillons pour l'année A dès le mois d'avril de l'année $A - 1$, afin de permettre au GI d'arbitrer entre les différentes demandes de sillons incompatibles et de construire une grille horaire qui soit globalement faisable;
- La conception des plannings pour le matériel pour une semaine de circulation donnée est généralement réalisée par les EF après que les sillons ont été demandés⁴, car l'affectation du matériel roulant dépend de la grille horaire envisagée. Elle commence donc plusieurs mois avant la circulation effective des trains concernés. Généralement, les EF disposent d'un planning hebdomadaire type dont ils s'inspirent pour concevoir ceux spécifiques à chaque semaine;
- La conception des plannings des agents est réalisée elle aussi plusieurs mois à l'avance par rapport au jour de circulation. Chaque agent se voit nominativement attribuer des *journées de service*, avec notamment des horaires précis de prise et fin de service chaque jour du calendrier.

Adaptation

Pour chaque ressource, la phase d'*adaptation* commence une fois qu'un plan a été arrêté durant la phase de conception. Elle se poursuit jusqu'à la phase opérationnelle la veille du jour de circulation.

Durant cette phase, il est parfois nécessaire de modifier le plan de transport préalablement établi en conception. Les raisons peuvent être multiples :

4. et donc avant leur attribution définitive.

- travaux non anticipés au niveau de l'infrastructure;
- avarie sur un type de matériel roulant demandant un allègement du plan de transport sur une zone donnée pendant une période donnée;
- sillons refusés par le GI, ou sillons ajoutés tardivement;
- besoin d'une rame supplémentaire pour un train qui accueillera davantage de voyageurs qu'initialement anticipé.

À chaque demande de modification du plan de transport pour l'une des raisons citées ci-dessus, il faut potentiellement replanifier plusieurs ressources ferroviaires, dans la mesure où leurs planifications ne sont pas indépendantes. Cela peut se produire de nombreuses fois durant toute la phase d'adaptation.

Les demandes d'adaptations sont en réalité de deux types :

1. les demandes *exogènes*, difficilement maîtrisables, dont l'origine est extérieure au système ferroviaire : c'est par exemple le cas d'une avarie matérielle ou de travaux sur l'infrastructure suite à un incident affectant la voie. Il est difficile de prévoir quand un tel événement peut se produire et quel matériel roulant ou quelle partie de l'infrastructure sera concernée;
2. les demandes *endogènes*, que l'on peut davantage anticiper puisqu'elles sont le fruit de requêtes internes : c'est le cas lorsqu'un sillon est refusé ou lorsqu'un train de voyageurs demande une rame supplémentaire car les réservations ont été mal anticipées.

Ces deux types de demandes d'adaptation sont essentiellement traitées de la même façon lorsqu'elles surviennent. Cette distinction permet cependant de déterminer quels sont les axes d'améliorations possibles pour une meilleure gestion de la phase d'adaptation.

Opérationnel

La phase opérationnelle correspond à l'exécution du plan de transport au moment des circulations. En l'absence d'aléa, l'ensemble des opérations est réalisé tel que planifié. Il est cependant fréquent que des aléas surviennent, et que des actions doivent être prises par les acteurs concernés pour limiter et éviter la propagation des retards, afin de *réguler* le trafic.

1.2 Hyperplanification du plan de transport

Dans cette section, nous introduisons la notion d'*hyperplanification* du plan de transport, qui est le principal constat qui a motivé cette thèse. La planification adaptative a été conçue pour pallier les différents inconvénients liés à cette hyperplanification.

1.2.1 Efficacité de la planification

La planification ferroviaire a été introduite pour répondre aux besoins explicités dans la Section 1.1.3.1. Cependant, il convient de se demander à quel point ces raisons historiques sont encore valables aujourd'hui.

d'hui. C'est l'*efficacité* de la procédure qui fait ici question : le plan de transport nominal est très souvent adapté à de nombreuses reprises durant la phase d'adaptation puis la phase opérationnelle, où le plan de transport effectivement réalisé est potentiellement assez éloigné de celui envisagé en conception. Or, les technologies et outils dédiés à la gestion du système ferroviaire en temps réel et à la planification ont beaucoup évolué depuis le XIX^e siècle : pourrait-on alors réduire l'effort en planification et concevoir un plan de transport *à la volée* quelques jours avant la circulation des trains ?

S'il ne fait aucun doute que la construction d'un plan de transport à l'avance est essentielle, on peut néanmoins se demander si un compromis entre une planification très rigide commençant plusieurs années à l'avance et un plan de transport généré à la volée peut être trouvé. C'est donc bien l'*hyperplanification* qui est ici pointée du doigt, et non la planification elle-même : peut-on et doit-on se permettre de « moins planifier » pour obtenir un gain de performance par rapport au processus actuellement en vigueur ?

1.2.2 Vers une planification adaptative ?

1.2.2.1 Deux approches pour introduire de la flexibilité

Originellement, au début de cette thèse, la planification adaptative avait pour objectif de réaliser le compromis précédemment évoqué entre l'hyperplanification et la génération d'un plan de transport à la volée, en rendant la gestion du système ferroviaire plus *flexible* et en gagnant en agilité. Pour cela, on peut envisager deux approches complémentaires :

1. *Flexibilité dans le processus de planification*

Pour chaque ressource ferroviaire, on peut se demander si sa planification n'est pas trop précise lors de chaque phase. Par exemple, le plan de voie en gare doit être planifié à l'avance pour savoir précisément les voies affectées à chaque train, permettant d'obtenir le *Graphique d'Occupation des Voies*, ou GOV. Ce dernier est établi après que les sillons ont été planifiés entre les gares, et avant la phase opérationnelle. Cependant, est-ce véritablement essentiel de le planifier aussi tôt ? Ne pourrait-on pas *retarder* l'établissement du GOV nominal pour éviter d'avoir à l'adapter trop souvent ?

2. *Flexibilité du plan de transport lui-même*

Plutôt que de simplement retarder certains éléments de planification, on peut aussi envisager un plan de transport où les ressources ferroviaires ne sont pas pleinement affectées. Par exemple, plutôt que de définir des sillons avec horaires précis de départ et d'arrivée en chaque gare, on peut définir des intervalles horaires pour chacun de ces événements, qui seraient réduits au fur et à mesure qu'on se rapproche du jour de circulation. De même, plutôt que d'affecter nominativement une rame donnée à un train, on pourrait lui affecter un ensemble de rames qui serait peu à peu réduit peu avant la phase opérationnelle.

1.2.2.2 Une flexibilité difficilement envisageable

Nous avons rencontré plusieurs experts du système ferroviaire au sein de SNCF et de ses filiales, afin de déterminer ce qui était envisageable pour parvenir à gagner en flexibilité dans l'établissement du plan de transport. Nous présentons ici quelques éléments évoqués par ces experts : l'état de l'art sur la flexibilité en planification sera présenté dans le chapitre suivant.

En ce qui concerne la planification des sillons, l'organisation du système ferroviaire avec la séparation du GI et des EF rend la flexibilité très difficile. On ne peut remettre en question le calendrier d'attribution des sillons s'étalant sur huit mois. En effet, le réseau ferré national français est connecté aux autres réseaux européens, et la procédure d'attribution des sillons suit en réalité une réglementation européenne.

La planification des agents est elle aussi très difficile à modifier dans son ensemble, dans la mesure où les agents ont besoin de connaître suffisamment à l'avance leurs horaires de travail. Néanmoins, l'affectation d'un agent (ADC ou ASCT) à un train ou à un autre est envisageable tant que ses horaires de prise et fin de service ne sont pas modifiés en sa défaveur. L'idée est ici de fixer en conception des journées de service avec horaires de prise et fin de service, et d'affecter ces journées à chaque agent. Les trains auxquels chacun sera affecté durant chacune de ses journées de service peuvent alors être définis juste avant la phase opérationnelle, plus tardivement qu'aujourd'hui. Il faudrait alors pouvoir garantir qu'une affectation de l'ensemble des trains aux journées de service soit toujours possible.

Outre ces difficultés, quelle que soit la ressource ferroviaire considérée, la plupart des experts ont considéré que les marges de manœuvre en termes de flexibilité dans la planification et dans le plan de transport étaient faibles. Soucieux de l'applicabilité de nos travaux de recherche au sein de SNCF, nous avons choisi de nous orienter vers une piste différente.

1.2.2.3 Planification et anticipation

Nous nous sommes aperçus que les inconvénients de l'hyperplanification pouvaient être évités sans pour autant modifier la précision du plan de transport ou le processus de planification dans son ensemble. En effet, l'un des principaux défauts de l'hyperplanification réside dans les multiples adaptations et re-planifications de ce plan, de la phase de conception jusqu'à la phase opérationnelle. L'introduction de la

flexibilité avec un plan de transport et une planification plus agiles telle que décrits précédemment permet en effet de limiter les efforts à chaque fois qu'une adaptation est nécessaire.

La planification adaptative, telle que considérée dans la suite de ce manuscrit, procède de façon différente, mais avec le même objectif : réduire les efforts déployés pour adapter et replanifier le plan de transport après la phase de conception. Pour cela, on fait le choix de considérer le champ des possibles en phase de conception en tentant d'*anticiper* les futures adaptations et replanifications du plan de transport, qui surviendront au cours de la phase d'adaptation et la phase opérationnelle. On élabore alors un plan de transport nominal en conception qui est construit *a priori* pour s'adapter « facilement » et « à moindre coût » dans la suite du processus de planification.

L'approche de planification adaptative que nous avons brièvement décrite ici est très générale. Dans ce manuscrit, nous avons choisi de mettre l'accent sur une phase de planification particulière et sur une unique ressource ferroviaire : l'adaptation de la planification du matériel roulant durant la phase d'adaptation. Nous mettons donc de côté les replanifications effectuées durant la phase opérationnelle ainsi que l'adaptation des autres ressources ferroviaires.

1.3 Organisation de la suite du manuscrit

Ce manuscrit est organisé de manière essentiellement linéaire, comme on peut le voir sur la Figure 1.3. Le Chapitre 2 présente un état de l'art sur la *planification ferroviaire* et la *flexibilité* :

- l'état de l'art sur la planification du matériel roulant est réalisé au début du Chapitre 3, et celui sur l'adaptation de cette ressource est présenté au Chapitre ?? . Il est ainsi possible de lire le Chapitre 2 indépendamment des trois chapitres suivants, ces derniers devant cependant être lus dans l'ordre ;
- la flexibilité dont il est ici question est notamment à rapprocher de celles décrites Section 1.2.2.1, que nous n'avons pas retenues pour les raisons exposées Section 1.2.2.2. Cet état de l'art concerne à la fois la flexibilité dans les problèmes de planification en recherche opérationnelle et ses applications au domaine ferroviaire.

Trois Problèmes originaux sont successivement introduits au cours de nos travaux, essentiellement modélisés par des Programmes Linéaires en Nombres Entiers (PLNE) :

- (a) le Problème de Conception d'un Roulement Matériel Nominal (ou *Problème Nominal*), exposé au Chapitre 3. On y définit également les notions et le vocabulaire liés à la ressource matériel roulant. Notamment, un *roulement matériel* correspond à un planning pour le matériel roulant ;
- (b) le Problème d'Adaptation d'un Roulement Matériel (ou *Problème Réactif*), exposé dans les premières sections du Chapitre ?? . On y introduit notamment les coûts d'adaptation spécifiques à la ressource matériel roulant de manière formelle, ainsi que la notion de scénario d'adaptation ;
- (c) le Problème de Conception d'un Roulement Matériel Adaptatif (ou *Problème Proactif*), exposé dans la dernière section du Chapitre ?? , qui permet de proposer des roulements matériels en phase de conception en minimisant les futurs coûts d'adaptation, répondant au concept de planification adaptative.

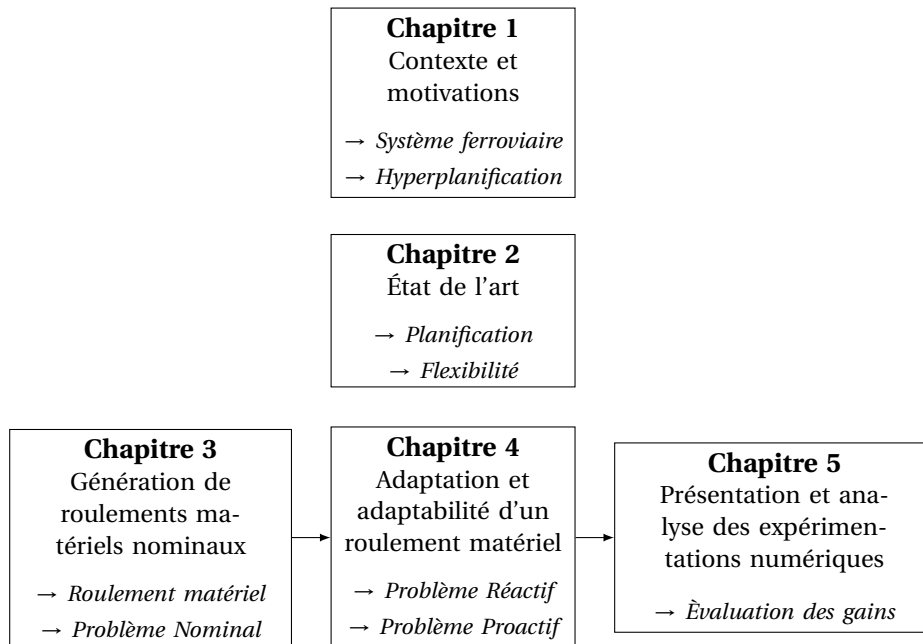


FIGURE 1.3: Organisation du manuscrit et interdépendances des chapitres

Ces trois Problèmes sont testés numériquement sur des instances réelles dans le Chapitre ??, où une méthode originale d'*évaluation* de la planification adaptative est proposée. Elle permet de mettre en évidence les *gains* espérés en se comparant avec l'approche actuelle.

1.3. ORGANISATION DE LA SUITE DU MANUSCRIT

Chapitre 2

État de l'art

Contenu

1.1 Le système ferroviaire	23
1.1.1 Organisation	23
1.1.2 Train et ressources ferroviaires	24
1.1.3 Plan de transport et planification ferroviaire	24
1.2 Hyperplanification du plan de transport	27
1.2.1 Efficacité de la planification	27
1.2.2 Vers une planification adaptative?	28
1.3 Organisation de la suite du manuscrit	30

Dans ce chapitre, nous réalisons un état de l'art de la littérature concernant deux aspects :

1. Les problèmes mathématiques et les modèles issus de la planification ferroviaire, avec un accent particulier sur le *timetabling*. Nous n'avons pas détaillé ici l'état de l'art concernant le matériel roulant : il sera effectué au Chapitre 3 dans la mesure où les modèles de planification adaptative développés dans cette thèse concernent cette ressource ferroviaire.
2. La notion de flexibilité dans les problèmes de planification en recherche opérationnelle, avec un accent particulier sur les cas d'application concernant la planification ferroviaire.

2.1 Planification ferroviaire

2.1.1 Planification séquentielle des ressources

Planifier l'ensemble des ressources est un problème complexe. En effet, les possibilités de planification pour une ressource dépendent assez largement de la planification des autres ressources. Cependant, on peut tout de même observer qu'il est intuitivement nécessaire de déterminer les dessertes d'un train avant d'en fixer les horaires, de fixer les horaires avant de planifier le matériel roulant, et de connaître cette planification avant de déterminer la celle des agents de conduite.

Ainsi, les ressources sont généralement planifiées de manière séquentielle selon un ordre bien établi.

2.1.1.1 Les missions

Le problème du *line planning* est l'un des premiers problèmes de la planification ferroviaire¹. On pourra consulter la revue de Schöbel (2012) pour un aperçu des différents types de problèmes et méthodes de résolution propre au *line planning*, on se contente ici d'en décrire les principales caractéristiques.

Le *line planning* prend en entrée l'infrastructure existante, souvent sous forme de graphe, et vise alors à déterminer les missions (*line* en anglais) des trains, c'est-à-dire :

- le nombre de missions différentes;
- l'itinéraire et les dessertes de chaque mission;
- la fréquence (moyenne) de chaque mission.

On peut distinguer deux grands types de modèles selon le type d'objectif retenu :

- la minimisation des coûts énergétiques et opérationnels

Ces modèles estiment ainsi les coûts qui impactent directement les EF. Pour cela, elles ont besoin d'estimer certaines caractéristiques comme le type et le nombre de voitures par train pour chaque mission, ce qui permet d'estimer :

- les coûts fixes par voiture et par heure comme ceux liés aux maintenances;
- les coûts variables liés à la distance parcourue pour chaque *voiture* comme les coûts liés au nettoyage et certains coûts énergétiques;
- les coûts variables liés à la distance parcourue pour chaque *train* comme les coûts liés aux agents de conduite et d'autres coûts énergétiques.

- la minimisation de l'inconfort des passagers

Ces modèles ont besoin en entrée d'une estimation du flux de voyageurs, généralement sous la forme d'une matrice origine-destination, contenant en position (i, j) le nombre de voyageurs souhaitant se rendre de la gare i vers la gare j . On cherche alors à minimiser le nombre de correspondances nécessaires, le temps d'attente moyen en gare, le temps de parcours moyen, ou à rendre le facteur de charge le plus homogène possible.

2.1.1.2 La grille horaire

Une fois les missions et leurs fréquences déterminées, il faut fixer les horaires et la marche de chaque train, de telle sorte à respecter les fréquences fixées à l'étape précédente.

Dans leur revue, Cacchiani et Toth (2012) distinguent plusieurs problèmes de *timetabling* selon le contexte, la régularité de la grille recherchée, l'entité qui cherche à le résoudre (GI ou EF), la présence ou non de plu-

1. En réalité, le problème du *network design* est résolu avant, afin de déterminer quels sont les emplacements les plus stratégiques pour les gares et l'infrastructure. Ce problème est cependant industriellement considéré comme extérieur à la phase de planification proprement dite : on considèrera toujours que l'infrastructure telle qu'elle existe est une donnée d'entrée.

sieurs lignes interconnectées, la présence de trains de fret... Les données en entrée (et en sortie) sont alors différentes. On présente ici la classification la plus répandue qui distingue les problèmes selon la régularité de la grille obtenue. On notera que Caimi et al. (2017) proposent une classification légèrement différente qui accorde davantage d'importance à l'entité qui est amenée à résoudre le *timetabling* : autant l'EF peut être amenée à planifier les trains de manière périodique ou apériodique, autant le problème résolu par le GI est structurellement différent, puisqu'il dispose en entrée d'un ensemble de grille horaires (périodiques ou non) nécessairement incompatibles et qu'il convient d'ajuster afin de maximiser le contentement des EF. Cependant, ce dernier problème peut parfois être vu comme un *timetabling* apériodique en cherchant à minimiser les écarts entre les grilles idéales demandées par les EF et la grille obtenue en sortie proposée par le GI.

Cas périodique : modèles PESP

Introduit pour la première fois par Serafini et Ukovich (1989), le *Periodic Event Scheduling Problem* (PESP) consiste à attribuer un horaire précis à une liste d'événements se reproduisant périodiquement avec la même période, de telle sorte à respecter des contraintes d'ordonnancement (précédence) entre ces événements.

Plus précisément, étant donné un graphe orienté $G = (X, A)$ et une période T , chaque arc $a \in A$ ayant pour caractéristique un intervalle $I_a = [l_a, u_a]$, il s'agit d'attribuer un horaire $\pi_x \in [0; T]$ à chaque événement $x \in X$, tel que

$$\forall a = [i, j] \in A, \exists q \in \mathbb{Z}, \pi_j - \pi_i + qT \in I_a$$

• Application au domaine ferroviaire

Obtenir des grilles périodiques est souvent d'un grand intérêt dans le transport public, qu'il s'agisse de périodes de l'ordre d'une heure ou d'une journée :

- elles sont beaucoup plus facilement mémorisables et donc attractives pour les passagers ;
- elles permettent de répéter de manière régulière les procédures opérationnelles du côté de l'exploitation et limitent donc les erreurs humaines, parfois responsables d'écarts au plan de transport nominal ;
- elles facilitent la gestion opérationnelle des circulations par la reproduction de certaines décisions opérationnelles qui avaient déjà fait leurs preuves dans des cas de perturbations similaires ;
- elles permettent, sous certaines hypothèses pratiques discutées par Ullrich et al. (2016) dans le cas du tram urbain, de mieux exploiter la capacité de l'infrastructure.

Cependant, la périodicité est une contrainte forte qui réduit considérablement les possibilités pour la grille horaire et les autres ressources, et un compromis est à trouver. Borndörfer et Liebchen (2008) montrent néanmoins par des calculs théoriques que pour un horizon temporel suffisamment grand, on peut toujours

trouver une grille périodique nécessitant le même nombre de rames engagées que la meilleure grille aperiodique.

Le PESP s'adapte très bien dans le cas du *timetabling*. Les événements à planifier correspondent aux départs et arrivées de chaque train en gare, et les arcs de précédence sont essentiellement de quatre types :

- temps de parcours : entre le départ d'un train d'une gare et son arrivée à la gare suivante, ces arcs correspondent au temps nécessaire pour parcourir le tronçon (généralement $l_a = u_a$);
- temps d'attente en gare : entre l'arrivée et le départ d'un même train en une même gare;
- temps de correspondance : entre l'arrivée d'un train et le départ d'un autre à une même gare;
- progression des trains : entre le départ d'un train et l'arrivée d'un autre à une même gare, pour respecter les marges de sécurité entre deux trains. Pour un arc $a = [i, j]$, on a alors $u_a = T - l_a$, ce qui correspond au cas où le train de l'événement i circule devant celui de l'événement j .

Pour parvenir à un tel graphe, il faut au préalable ordonnancer les trains pour obtenir le réticulaire, à partir du résultat du *line planning*². Le PESP obtenu correspond alors à des événements ayant potentiellement des périodicités différentes, selon la fréquence de la ligne à laquelle ils correspondent. Quitte à multiplier l'horizon par le PPCM des périodes et démultiplier les événements, on considère toutefois que toutes les lignes ont la même fréquence.

• Modélisations et résolutions

Le PESP tel qu'introduit par Serafini et Ukovich (1989) est un pur problème d'admissibilité. Le problème de décision associé est prouvé fortement NP-complet par réduction du problème du circuit hamiltonien. Si les variables q sont fixées, ce problème est connu sous le nom de *Feasible Differential Problem* (FDP) et devient polynomial. Il est cependant difficile d'estimer ces variables q , et cela réduit considérablement le nombre de possibilités. Néanmoins, on peut aboutir à des résultats intéressants si on se base sur une grille horaire déjà existante.

Dans le cas du *timetabling*, on adjoint au modèle PESP une fonction objectif. Au premier abord, minimiser les slacks des contraintes de précédence semble le principal critère à ce stade de la planification, bien que cela rende la grille moins robuste : on cherche ainsi à minimiser, pour chaque contrainte $a = [i, j]$, la quantité $\pi_j - \pi_i - l_a + q_{ij}T$. On peut alors attribuer un poids à chacune de ces quantités et minimiser leur somme pondérée. Ce poids peut par exemple être proportionnel au nombre de passagers attendus dans le cas d'une correspondance. On peut de manière générale reprendre les mêmes types de critères que pour le problème du *line planning*.

En notant \mathbf{B} la matrice d'incidence sommets-arcs du graphe G et w le vecteur des poids, on obtient le PLNE classique suivant introduit par Serafini et Ukovich (1989) :

2. Il semble y avoir peu de bibliographie sur ce sujet, et cette étape s'effectue le plus souvent manuellement ou est intégrée dans des problèmes plus globaux.

$$\min w^T \cdot (\mathbf{B}^T \pi + qT)$$

s.c.

$$l \leq \mathbf{B}^T \pi + qT \leq u$$

$$0 \leq \pi \leq T$$

$$q \in \mathbb{Z}^{|A|}$$

On peut alors améliorer sa tractabilité de différentes manières, que ce soit en fixant certaines variables q de manière similaire au FDP ou en reformulant les contraintes avec la matrice d'incidence arc-cycle introduite et étudiée par Liebchen et Peeters (2009).

Le caractère symétrique de la grille horaire est souvent recherché pour certaines lignes particulières (lignes cadencées). Lors d'un itinéraire nécessitant une correspondance entre deux lignes, les grilles symétriques permettent un temps de correspondance identique quel que soit le sens de parcours. Si on veut imposer la symétrie de la grille par rapport à l'instant $s \in [0; \frac{T}{2}]$, il suffit d'ajouter la contrainte

$$\pi_i + \pi_j + q_{ij}T = 2s$$

pour tout arc $[i, j] \in A$.

Inversement, la périodicité est parfois trop contraignante, et on cherche à obtenir des grilles *quasi-périodiques*. Ce processus est souvent réalisé « à la main » en décalant, ajoutant ou supprimant certains sillons de telle sorte à satisfaire certaines contraintes techniques. Les grilles quasi-périodiques peuvent également être utilisées pour planifier la transition entre les heures pleines et les heures creuses de manière plus efficace qu'en supprimant manuellement certains trains et en gardant les mêmes horaires pour les autres. Caimi et al. (2011b) ont développé une approche algorithmique pour parvenir à de telles grilles sans avoir besoin de post-traitement manuel de la grille périodique. Ils considèrent en entrée une *intention périodique de service*, ce qui correspond à un ordonnancement périodique des trains où chaque événement doit avoir lieu dans un certain intervalle de la période (le même quel que soit la période considérée) et uniquement pour certaines périodes de la journée. Ils proposent une projection du problème en une instance de PESP, en introduisant des classes d'équivalence de trains. La projection leur permet alors, avec une probabilité plus ou moins forte qu'il convient de maîtriser, d'obtenir une grille quasi-périodique faisable pour le problème initial respectant l'intention de service.

Cas apériodiques

Dans le cas plus général où l'on s'autorise une grille apériodique, il n'existe pas de formulation mathématique comme le PESP qui s'adapte facilement à tous les cas de figure et dont tous les modèles dérivent. Il existe ainsi des formulations très variées suivant l'objectif et le contexte que l'on se fixe, ce qui aboutit à

des modèles différents, bien qu'ils s'expriment le plus souvent sous la forme de PLNE. On se contente ici d'exhiber quelques-uns de ces modèles.

Une manière naturelle est de partir d'un ensemble de trains, et de leur attribuer des horaires individuellement de telle sorte à ce que la grille obtenue soit macroscopiquement faisable. Plus précisément, on divise l'infrastructure en *blocks* et on discrétise le temps, en n'autorisant à chaque instant la présence d'un seul train par *block*. Pour un train r , la variable binaire x_{bt}^r vaut 1 lorsque le train r utilise le *block* b entre les instants t et $t+1$. L'ensemble T^r contient les sillons x^r possibles pour le train r , un sillon étant ici vu comme un ensemble cohérent de couples (b, t) où le train se trouve sur le *block* b à l'instant t . L'objectif est alors de maximiser la somme pondérée des trains, auxquels on affecte une valeur plus ou moins importante selon leur attractivité. C'est le modèle PLNE initialement proposé par Brännlund et al. (1998), résolu par une relaxation lagrangienne.

$$\begin{aligned} & \max \sum_r f(x^r) \\ \text{s.t.} & \\ & \sum_r x_{bt}^r \leq 1 \quad \forall b, t \\ & x^r \in T^r \quad \forall r \\ & x_{bt}^r \in \{0; 1\} \quad \forall b, t, r \end{aligned}$$

La relaxation lagrangienne de la contrainte couplante ne permet cependant pas de résoudre des instances de taille réelle, car la combinatoire du problème est conséquente. De plus, le profit de chaque train est une quantité difficile à estimer en pratique, même en programmation non linéaire.

Des modèles permettant une plus grande flexibilité sur le choix des trains ont néanmoins été développés. Par exemple, Carey et Lockwood (1995) introduisent un PLNE où il est possible de choisir les horaires de départ et d'arrivée en gare comme dans le PESP, ainsi que des variables binaires pour contrôler l'ordre des trains. Il permet de modéliser le dépassement des trains par des voies d'évitement en gare, et l'objectif choisi est la minimisation des écarts par rapport aux grilles idéales des EF.

De façon similaire, le PLNE de Caprara et al. (2002) dispose en entrée d'un ensemble de sillons incompatibles, et permet de les décaler dans le temps et les étirer/contracter (augmentation/diminution du temps de parcours ou du temps de stationnement) de telle sorte à obtenir une grille sans conflit, en minimisant l'écart par rapport aux grilles idéales demandées par les EF. Il s'appuie sur un graphe représentant le diagramme espace-temps des circulations ainsi planifiées, et est résolu par relaxation lagrangienne.

Ordonnancer les trains de manière aperiodique peut aussi permettre de mieux s'adapter à la demande voyageur. Partant d'une liste de trains avec la même mission et d'une matrice Origine-Destination, Barrena et al. (2014) proposent par exemple trois formulations MILP pour planifier les trains de manière irrégulière en minimisant le temps de parcours et le temps d'attente moyen. Les modèles proposés se restreignent à

un corridor et les MILP font tous intervenir des variables de flots qui permettent de rester linéaire. Ces trois modèles sont résolus par des méthodes de Branch&Cut.

2.1.1.3 Le Graphique d'Occupation des Voies (GOV)

Le problème du *routing* consiste à établir le plan de voies en une gare du réseau, en déterminant pour chaque train arrivant en (*resp.* partant de) cette gare un itinéraire tracé sur le plan des voies de cette gare. Notamment, il s'agit de déterminer le quai associé à chaque train, ce qui constitue le problème du *platforming*.

Le problème de *routing* peut être résolu par un PLNE. On détermine au préalable un ensemble de *patrons*, chaque patron correspondant à un itinéraire d'arrivée, un quai, et un itinéraire de départ de la gare. On recherche alors un ensemble de cliques $\mathcal{K} \subset K$ pour détecter les affectations incompatibles entre elles, et on obtient le PLNE suivant issu de Caprara et al. (2007).

$$\left\{ \begin{array}{l} \min \sum_t \sum_{P \in P_t} c_{t,P} x_{t,P} + \sum_t M_t s_t \\ \text{s.t.} \\ s_t + \sum_{P \in P_t} x_{t,P} = 1, \quad t \in T \quad (2.1) \\ \sum_{(t,P) \in K} x_{t,P} \leq 1, \quad K \in \mathcal{K} \quad (2.2) \\ x, s \in \{0, 1\} \end{array} \right.$$

La variable s indique si un train t est supprimé. $x_{t,P}$ vaut 1 si le train t est affecté au patron P . M_t est une pénalité assez grande et $c_{t,P}$ est le coût d'affectation de t à P . Les contraintes (2.1) assurent que soit un train est supprimé soit il est affecté à un patron. Les contraintes (2.2) permettent d'éviter l'incompatibilité des affectations de trains.

2.1.1.4 Les agents

Les problèmes du *crew scheduling* et *crew rostering* concernent la planification des agents à bord des trains. Les modèles mathématiques associés considèrent généralement le cas des agents de conduite, celui des ASCT étant similaire et rarement considéré explicitement dans la littérature. Industriellement, l'objectif est de créer des *journées de service* pour chaque agent, où chaque journée de service correspond à une tranche horaire de travail avec un certain nombre de trains à conduire.

Le *crew scheduling* consiste à établir des journées de service génériques sur une période donnée (généralement jour ou semaine), sans attribuer spécifiquement chaque journée de service à un agent donné : c'est précisément ce que l'on cherche à résoudre dans le problème du *crew rostering*, où il convient de tenir compte des spécificités de chaque agent. Par exemple, ils ne sont pas tous habilités à conduire certains

types de matériels roulants. Heil et al. (2020) ont réalisé une revue de la littérature des deux dernières décennies concernant le *crew scheduling*. Les modèles couramment utilisés sont des modèles de flots assez semblables à ceux relatifs à la planification du matériel roulant. En effet, ces deux ressources (matériels et agents) peuvent en première approximation se planifier de façon similaire : il s'agit d'attribuer un emploi du temps à chaque matériel roulant (*resp. agent*) disponible en s'assurant que tous les trains soient couverts (*resp. conduits*) par un nombre minimum de rames (*resp. agents*). Les différences résident alors dans les contraintes opérationnelles associées à ces deux ressources.

2.1.1.5 Maintenances

Les maintenances sont planifiées tout au long du processus de planification, que ce soit très en amont pour des maintenances lourdes concernant l'infrastructure, ou proche de la phase opérationnelle lorsqu'il s'agit d'un imprévu ou d'une maintenance plus légère. La revue de Lidén (2015) retrace les différents types de maintenances pouvant intervenir, et les classe en plusieurs catégories selon les ressources mobilisées et la phase de planification où elles apparaissent.

2.1.2 Planification intégrée des ressources

Bešinović et al. (2016) proposent une procédure visant à intégrer l'étape de *timetabling* au niveau macroscopique et au niveau microscopique. En effet, certaines solutions faisables à un niveau de description macroscopique ne le sont pas avec un niveau de description plus précis, où l'on calcule notamment finement les temps de parcours et où l'infrastructure est plus détaillée. Les auteurs fournissent une procédure itérative avec des algorithmes capables de passer de l'un à l'autre niveau de description, et détaillent un modèle MILP renvoyant une grille macroscopiquement faisable à partir d'une grille microscopique dont les conflits ont été détectés. Comme ce modèle possède un nombre exponentiel de contraintes, les auteurs introduisent une heuristique *ad hoc* fondée sur des méthodes de Monte-Carlo pour planifier successivement chaque ligne dans un ordre aléatoire, en choisissant le meilleur chemin possible à chaque fois et en maintenant la faisabilité macroscopique de la grille. Leur algorithme permet également d'évaluer la robustesse de la grille obtenue, dont le coût est considéré conjointement à celui des chemins sélectionnés pour chaque ligne.

Dans le même esprit, dans le cadre du réseau belge très dense avec plusieurs gares complexes, Dewilde et al. (2013) ont développé une procédure algorithmique pour parvenir à une planification intégrée robuste de la grille horaire (*timetabling*), des itinéraires de chaque train en gare (*routing*), et de l'affectation des quais (*platforming*). La robustesse est ici évaluée du point de vue voyageur de manière pragmatique au moyen de la simulation comme la différence entre le temps de transport réel et le temps de transport nominal. L'objectif choisi en conception est la maximisation des marges entre chaque train³. La procédure

3. Plus précisément, les auteurs cherchent à minimiser la somme des inverses des marges entre deux trains consécutifs, ce qui favorise l'homogénéisation des marges.

proposée utilise la recherche locale pour modifier la planification d'une ou plusieurs ressources de manière itérative, et s'appuie sur des modules algorithmiques appliqués successivement. Le *routing* est résolu (à grille et quais fixes) par un MILP publié dans un autre article. La planification des sillons est traitée par une recherche tabou avec trois structures de voisinage :

- décalage d'un sillon
- décalage de plusieurs sillons
- changement de l'ordre entre deux trains

L'affectation des trains aux quais peut être modifiée en inversant deux quais dans la solution courante, ce qui nécessite alors de résoudre à nouveau le problème de *routing* et de *timetabling*. Les résultats obtenus sont détaillés pour les gares de la zone de Bruxelles et celle d'Antwerp, qui sont deux zones structurellement très différentes au sens où il existe beaucoup de possibilités de *routing* pour un même quai. dans les deux cas, les auteurs montrent que leur méthode permet une amélioration de la robustesse de 4 à 8 points par rapport à la procédure actuellement en vigueur.

À la SNCF, plusieurs thèses ont été menées autour de la planification intégrée des ressources. Les travaux de Benhizia (2012) ont notamment permis d'intégrer la planification du matériel roulant et celle des agents au moyen de PLNE résolus par relaxation lagrangienne. Plus récemment, Zehrouni (2016) s'est intéressée aux gains de la planification intégrée sur la robustesse du plan de transport, en considérant à la fois l'infrastructure ferroviaire, le matériel roulant et les agents. La robustesse du plan de transport proposé est alors évaluée et validée grâce à des outils de simulation. Enfin, les travaux de Bréthomé (2018) ont notamment permis de proposer un modèle intégrant la planification des missions et la conception de grilles horaires.

2.1.3 Replanification

Une fois l'ensemble des ressources planifiées, le plan de transport est souvent réadapté pour faire face aux changements et aléas de nature multiple intervenant durant le processus de planification. Industriellement, on parle de *replanification* lorsqu'il s'agit de changer la planification des ressources en temps réel lors de la phase opérationnelle, et on parle d'*adaptation* si cette replanification s'effectue en amont, entre la publication du plan de transport et le début de la phase opérationnelle. Dans la littérature scientifique, ces deux notions sont parfois confondues, dans la mesure où les principales différences résident dans la nature des aléas et le temps de calcul disponible pour replanifier. L'adaptation de la planification du matériel roulant sera décrite dans le Chapitre ??, nous nous intéressons brièvement ici au cas de la replanification en temps réel.

L'un des problèmes les plus étudiés est la détection et résolution des conflits, *Conflict Detection and Resolution*, *Train Timetabling Rescheduling* ou de manière équivalente *Traffic Management Problem*. Il s'agit de replanifier les sillons en temps réel lorsque les retards de certains trains rendent la grille infaisable. On pourra consulter la revue de littérature de Cacchiani et al. (2014) pour une présentation des différents modèles et algorithmes concernant ce problème. À la SNCF, les travaux de Altazin et al. (2017) ont permis d'introduire la notion de *stabilité* du plan de transport en temps réel et de proposer des algorithmes de

replanification en zones denses, sur des portions du réseau où l'infrastructure est sollicitée à une fréquence particulièrement importante, et où la gestion des retards en temps réel est essentielle.

Lorsque certains trains sont en retard, certaines correspondances ne peuvent être assurées. Le *Delay Management Problem* consiste alors à prendre la décision de maintenir certaines correspondances en retardant volontairement le départ de certains trains, ou au contraire de ne pas maintenir ces correspondances, une solution alternative devant alors être trouvée par la suite pour satisfaire les voyageurs souhaitant se rendre à destination. Les objectifs du problème sont alors assez similaires à ceux du *timetabling*. Kanai et al. (2011) introduisent également la satisfaction des voyageurs dans leurs critères, en proposant une modélisation de l'insatisfaction des voyageurs par une fonction affine par morceaux selon le taux de remplissage du train, autrement appelé *facteur de charge*. On pourra par exemple consulter Heilporn et al. (2008) pour un aperçu des différents algorithmes et modèles de *Delay Management* issus de la littérature.

2.2 Flexibilité et agilité en planification

La rigidité de la planification ferroviaire est le constat initial motivant le concept de planification adaptative. Bien que l'introduction de flexibilité dans la planification ferroviaire et dans le plan de transport ait finalement été écartée, nous présentons dans cette section un état de l'art sur la flexibilité. Nous nous concentrons d'abord sur la flexibilité dans les problèmes d'ordonnancement, puis présentons les différentes approches issues de la planification ferroviaire où la notion de flexibilité intervient explicitement : à notre connaissance, elles concernent toutes la ressource sillons. Enfin, nous nous intéressons au cas de contextes d'optimisation à plusieurs étapes, qui peuvent être rapprochés de la notion de flexibilité et sont au cœur de la planification adaptative telle que nous l'avons finalement envisagée.

2.2.1 Cas des problèmes d'ordonnancement sur machine

La notion de flexibilité en planification apparaît de manière naturelle pour les problèmes d'ordonnancement de tâches sur machine(s).

En ordonnancement, flexibilité et robustesse sont deux notions très liées : Billaut et al. (2008) définit ainsi la flexibilité comme la donnée de solutions partielles pouvant facilement s'affiner et se préciser en temps réel pour mieux s'adapter à la conjoncture. La robustesse est alors vue comme la performance d'un algorithme en cas de situation perturbée, et un algorithme flexible permet donc *a priori* d'augmenter la robustesse de la méthode choisie.

2.2.1.1 Description formelle

Il existe de très nombreux types de problèmes d'ordonnancement. On décrit ici une forme relativement générale du *job-shop scheduling problem*. Les données sont :

- Un ensemble \mathcal{M} de M machines.

- Un ensemble \mathcal{J} de J jobs, chaque job $j \in \mathcal{J}$ devant s'effectuer successivement sur chaque machine de \mathcal{M} . On appelle alors *opération* le fait d'effectuer le job j sur la machine m , et on regroupe les opérations dans l'ensemble \mathcal{S} .
- Un triplet (r_i, p_i, d_i) pour chaque opération $i \in \mathcal{S}$: cette opération nécessite une durée p_i , doit commencer après l'instant r_i et terminer avant l'instant d_i .
- Un ensemble de contraintes de précédence entre des couples d'opérations i_1 et i_2 , spécifiant que l'opération i_2 ne peut pas commencer avant la date de fin de l'opération i_1 .

Il s'agit alors de déterminer, pour chaque opération $i \in \mathcal{S}$, une date t_i où commencera le traitement de cette opération sur la machine $m_i \in \mathcal{M}$, de telle sorte à minimiser le *makespan*, c'est-à-dire l'instant à partir duquel toutes les opérations sont terminées.

La solution obtenue peut aussi se voir sous la forme d'un *diagramme de Gantt*, où l'on représente une machine par ligne avec un axe des temps horizontal, une opération $i \in \mathcal{S}$ affectée à la machine $m \in \mathcal{M}$ entre les instants t et $t + p_i$ étant représentée par un rectangle sur la m -ième ligne entre ces deux instants.

2.2.1.2 Deux types de flexibilité

Flexibilité temporelle et graphes disjonctifs

Roy et Sussmann (1964) proposent de représenter une instance d'un problème d'ordonnancement par un *graphe disjonctif*.

- Chaque sommet correspond à une opération, et on ajoute un sommet source s et un sommet puits ω .
- Les arcs de précédence entre deux opérations assujetties à une contrainte de précédence dans l'instance initiale. Un arc de précédence (i_1, i_2) est alors valué par p_{i_1} et représente le fait que l'opération i_2 ne peut être effectuée qu'après la fin de l'opération i_1 .
- On ajoute pour chaque opération i un arc de précédence depuis la source s valué par r_i , ainsi qu'un arc de précédence vers le puits ω valué par $p_i - d_i$.
- Les arcs disjonctifs se présentent par paires et relient deux opérations i_1 et i_2 s'effectuant sur la même machine $m \in \mathcal{M}$; le premier arc relie i_1 à i_2 et le second i_2 à i_1 . Ces arcs ne sont pas valués.

Avec cette formulation, les auteurs montrent que toute solution faisable d'une instance revient à *sélectionner*, pour chaque arc disjonctif, un des deux arcs de la paire, de telle sorte à ce que le graphe obtenu soit sans cycle. On peut alors facilement calculer le *makespan* comme le plus court chemin de s à ω , et on peut retrouver les dates de début de chaque opération en procédant « au plus tôt ».

Remarquons que les graphes disjonctifs ont été généralisés par Mascis et Pacciarelli (2002) sous la forme de *graphes alternatifs*, où une paire d'arcs alternatifs se présente sous la forme $[(n_1, n_2), (n_3, n_4)]$, signifiant qu'il faut sélectionner (n_1, n_2) ou (n_3, n_4) . Dans le cas des graphes disjonctifs, on a toujours $n_1 = n_4$ et $n_2 = n_3$. Une telle formulation a pu être exploitée dans le domaine ferroviaire dans le cas de la replanification en temps réel, dans le cas du *delay management* (voir par exemple Samà et al. (2017)).

Avec la formulation sous la forme de graphes disjonctifs, une solution n'est plus vue comme la détermination d'une date de début pour chaque opération, mais comme un ordonnancement des opérations sur chaque machine. Comme noté par Billaut et al. (2008), cela apporte de la flexibilité *temporelle* à la planification des tâches. En effet, en cas d'aléas (par exemple un temps de réalisation plus long pour une opération), la structure de la solution reste inchangée et on peut se contenter de « simplement » décaler les dates des opérations suivantes sur cette machine, tant que cela est nécessaire.

Cette flexibilité temporelle est systématiquement laissée à l'ordonnanceur en temps réel, les durées des opérations étant souvent connues avec une précision limitée avant leur exécution.

Flexibilité séquentielle et sélection partielle

La formulation avec les graphes disjonctifs permet d'introduire un autre type de flexibilité : en laissant à l'ordonnanceur le choix de sélectionner l'un ou l'autre arc de certaines paires d'arcs disjonctif, celui-ci dispose d'une flexibilité *séquentielle* quant à l'ordre des opérations sur chaque machine.

Artigues et al. (2005) proposent de maximiser cette flexibilité séquentielle afin d'obtenir une solution plus robuste, tout en garantissant une certaine qualité de la solution en imposant un *makespan* correct dans les contraintes⁴. Plus précisément, l'objectif est de déterminer non plus un ordre total sur chacune des opérations, mais un ordre partiel. Ils se restreignent cependant au cas des groupes d'opérations : ils partitionnent l'ensemble des opérations et imposent un ordre total sur ces classes (on préfère parler de « groupes »). Ainsi, ils ne donnent pas l'ordre dans lequel effectuer les opérations *au sein d'un groupe*, mais imposent d'avoir terminé *toutes* les opérations d'un groupe avant de commencer celles d'un autre groupe.

Une façon naturelle d'évaluer la flexibilité d'un tel ordre partiel est de compter le nombre de plannings complets que l'on peut obtenir à partir de cette solution partielle. En effet, plus ce nombre est important, meilleure sera la flexibilité laissée à l'ordonnanceur. Cependant, pour des raisons de combinatoire, ce nombre est difficile à évaluer : la fonction de comptage est montrée #P-complet par Brightwell et Winkler (1991), et les auteurs préfèrent maximiser le nombre de groupes, ce qui est intuitivement corrélé au nombre de plannings complets (mais pas toujours).

Dans le cas où il n'y a pas de dates de début et de fin pour les opérations, ils proposent un algorithme polynomial pour calculer de manière exacte le plus grand *makespan* associé à un ordre partiel de la forme étudiée, en utilisant notamment l'algorithme de Bellman pour calculer un plus long chemin. Ils peuvent ainsi proposer une borne supérieure des *makespan* parmi tous les plannings possibles respectant l'ordre partiel.

Dans un article plus récent, Aloulou et Artigues (2010) s'intéressent au cas d'un ordre partiel plus général, sans qu'il ne corresponde nécessairement à des groupes d'opérations. Une *solution flexible* est alors définie comme un ordre partiel sur l'ensemble des opérations, c'est-à-dire une *sélection partielle* parmi les

4. Il s'agit de la technique d' ϵ -contrainte, méthode usuelle en optimisation multicritères.

paires d'arcs disjonctifs, de telle sorte à ce que tous les plannings respectant cette sélection soient faisables. Le problème consiste alors à trouver une sélection partielle en optimisant deux critères *a priori* contradictoires :

- Minimiser le pire *makespan* parmi toutes les plannings complets respectant la sélection partielle, comme ce qui avait été fait par Artigues et al. (2005) dans le cas des groupes d'opérations.
- Maximiser la flexibilité : dans le même esprit que la maximisation du nombre de groupes par Artigues et al. (2005), il s'agit de maximiser le nombre de paires d'arcs disjonctifs non sélectionnés dans la sélection partielle.

2.2.2 Flexibilité d'une grille horaire

On présente ici trois contextes de planification se rapprochant de la notion de planification adaptative en introduisant de la flexibilité dans la grille horaire lors de sa conception.

2.2.2.1 Grille horaire flexible pour mieux planifier les autres ressources

Caimi et al. (2011a) ont proposé une adaptation du modèle PESP capable d'introduire de la flexibilité horaire sur chaque événement. Le *Flexible Periodic Event Scheduling Problem* consiste ainsi à attribuer non pas un unique horaire par événement mais un intervalle de temps dont la longueur caractérise la *flexibilité* de cet événement. Bien que les auteurs mentionnent les avantages qu'une telle flexibilité peut apporter pour la gestion opérationnelle (voir les travaux de D'Ariano et al. section 2.2.2.3), l'objectif visé ici est d'obtenir une grille macroscopique moins rigide à partir de laquelle une grille microscopique faisable sans conflit pourra être élaborée. Ils poursuivent ainsi le même objectif que Bešinović et al. (2016), sans toutefois fournir de modèle pour établir une telle grille microscopique à partir d'une grille macroscopique flexible.

Le modèle FPESP possède des variables correspondant, pour chaque événement i , à son horaire *au plus tôt* $\underline{\pi}_i$ et à sa flexibilité δ_i , de telle sorte à ce que quel que soit l'horaire définitif $\pi_i \in [\underline{\pi}_i; \underline{\pi}_i + \delta_i]$ attribué à *chaque* événement i , la grille soit faisable. Il faut bien noter que les différents intervalles de flexibilité sont indépendants : changer un ou plusieurs horaires d'événements dans leur intervalle de flexibilité sans modifier les autres doit mener à une grille faisable. Pour qu'une telle flexibilité existe, il est nécessaire que la durée entre deux événements (temps d'attente en gare, parcours d'un tronçon) soit variable. Ainsi, dans le cas du départ d'un train en gare à l'instant $\pi_i = \underline{\pi}_i + \delta_i$, ce train doit pouvoir augmenter sa vitesse pour être capable d'arriver à la gare suivante en $\pi_{i+1} \in [\underline{\pi}_{i+1}; \underline{\pi}_{i+1} + \delta_{i+1}]$, et donc en particulier en $\pi_{i+1} = \underline{\pi}_{i+1}$.

Une conséquence importante de l'indépendance des intervalles de flexibilité est que deux événements i et j adjacents dans le graphe FPESP vont avoir des flexibilités qui vont dépendre l'une de l'autre, au sens où plus la flexibilité de i est importante, moins celle de j le sera. Plus précisément, si la variabilité de la durée entre i et j est γ_{ij} , c'est-à-dire que $\pi_j - \pi_i$ peut être dans un intervalle de longueur γ_{ij} , les auteurs montrent que $\delta_i + \delta_j \leq \gamma_{ij}$.

Caimi et al. cherchent alors à résoudre le modèle FPESP *en maximisant la flexibilité* conjointement aux

objectifs usuels (minimisation des marges pour minimiser le temps de parcours sur une ligne). Ils proposent ainsi plusieurs fonctions objectifs combinant les deux critères par des méthodes classiques d'optimisation bi-objectif. En particulier, les auteurs notent qu'en imposant une qualité de la grille égale à celle du modèle PESP sans aucune flexibilité, il est possible de trouver une solution avec flexibilité non nulle : les solutions rigides du PESP possèdent donc potentiellement de la flexibilité puisqu'on peut les transformer en solutions flexibles du FPESP sans détériorer la performance de la grille ainsi obtenue.

Les auteurs proposent en fin d'article une extension du concept de flexibilité⁵. Une *flexbox* F avec une certaine flexibilité associée δ_F est un ensemble d'événements (connexes dans le graphe) tel que si tous ces événements sont *simultanément* décalés d'une *même durée* d'au plus δ_F par rapport à leur horaire au plus tôt, la grille reste faisable. Les *flexboxes* permettent ainsi de briser le caractère indépendant de certains événements et apportent davantage de possibilités pour la flexibilité. La flexibilité associée à un seul événement peut se voir comme une *flexbox* associée à cet unique événement. Cette extension est motivée par le fait que certains événements comme par exemple l'arrivée et le départ d'un même train en gare ne sont pas indépendants dans la réalité : si le temps d'arrêt comporte potentiellement très peu de flexibilité, on peut toutefois décaler de plusieurs minutes l'ensemble de ces deux événements dans la grille tout en maintenant sa faisabilité.

2.2.2.2 Grille horaire flexible avec variantes de sillons

En Suède, Forsgren et al. (2012) ont proposé une nouvelle approche pour l'attribution des sillons par le gestionnaire d'infrastructure et pour les négociations avec les EF. Ils font le constat de la rigidité du processus actuellement en vigueur, qui impose une grille horaire quotidienne sans conflit répétée durant les 364 jours d'exploitation de l'année de service. Or, chaque train ne circule pas chaque jour de l'année, et certains sillons ont des régimes disjoints : résoudre tout conflit entre de tels sillons est alors inutile, et cela mène à des pertes de capacité. De plus, la rigidité de la planification des sillons ne permet pas de les décaler en cas d'annulation d'un sillon et rend donc trop souvent impossible la planification d'un autre train (type Fret).

Ainsi, les auteurs proposent d'établir en amont une planification où chaque train possède plusieurs sillons possibles, qui sont autant de variantes rendant l'exploitation plus agile. La décision de l'utilisation de l'une ou l'autre des variantes est reportée dans le temps, et peut même changer lors de la phase pré-opérationnelle en cas d'insertion d'un train dans le graphique espace-temps au dernier moment. Pour garantir la viabilité d'une telle approche et donner suffisamment d'informations aux EF sur la réservation des sillons, il faut revoir la façon dont les sillons sont alloués par le GI. Le processus, nommé *allocation incrémentale de la capacité*, se déroule de la façon suivante.

1. Chaque EF demande au GI, pour chaque train :
 - des horaires de départ/arrivée souhaités à chaque gare

5. Il est à noter que l'article existe sous une autre version avec le même titre, en tant que *conference paper* pour ATMOS 2007, mais cette version est moins détaillée. En particulier, elle ne présente pas cette extension du concept de flexibilité

- un compromis entre peu de sillons ne respectant pas les horaires demandés mais *a contrario* un temps de parcours globalement sous-optimal, et inversement beaucoup de sillons ne respectant pas l'horaire demandé mais avec un temps de parcours proche de l'optimum
2. Le GI trace les sillons, où chaque train possède plusieurs variantes selon les jours, en faisant en sorte :
 - de respecter au mieux les horaires demandés par les EF pour un maximum de variantes
 - de respecter au mieux le compromis formulé par les EF pour chaque train
 - de limiter le nombre de variantesCela est fait au moyen d'algorithmes, comme par exemple le PLNE présenté par Gestrelus et al. (2015) s'inspirant de modèles d'allocation déjà existants et résolu pour chaque jour de l'année avec un horizon roulant.
 3. Établissement d'une *grille horaire de contrôle* sur un jour, où sont reportées toutes les variantes de tous les sillons. L'horaire d'arrivée (*resp.* départ) d'un train à une gare correspond alors au maximum (*resp.* minimum) de l'horaire d'arrivée (*resp.* départ) parmi toutes les variantes
 4. Communication de la grille horaire de contrôle aux EF. Pour certains trains, les horaires obtenus diffèrent des horaires demandés, car il existe des variantes qui ne les respectent pas.
 5. Négociations itératives selon un procédé non encore fixé par les auteurs.
 6. Publication de la grille horaire annuelle, avec un seul jeu d'horaires d'arrivée/départ pour chaque train en chaque gare. L'EF peut communiquer ces horaires aux clients car elles seront respectées par toutes les variantes, quel que soit le jour de circulation du train.

Bien que chaque variante soit *a priori* affectée à un régime particulier, cela peut changer lors de la phase d'adaptation ou la phase pré-opérationnelle pour mieux réarranger le diagramme espace-temps en cas d'aléa, ce qui confère à ce processus une plus grande flexibilité.

2.2.2.3 Grille horaire flexible en temps réel

D'Ariano et al. (2008) ont proposé d'introduire de la flexibilité dans la grille horaire y compris lors de la phase opérationnelle. Cette flexibilité n'est appliquée que pour les zones denses et permet davantage de libertés en opérationnel, où certaines décisions habituellement arrêtées plus tôt dans la phase de planification sont prises. Cela permet ainsi une meilleure utilisation des marges et donc un débit plus important sans pour autant faire baisser la robustesse et la ponctualité de la grille.

Concrètement, les auteurs proposent de retarder trois types de décisions :

1. L'assignation des quais en gare (en ligne)
2. Les horaires précis de départs/arrivées de chaque train, qui sont remplacés par des intervalles de départs/arrivées au plus tôt/tard
3. L'ordre de convergence, cisaillement de couples de trains, ainsi que les possibilités de dépassement *via* les voies d'évitement en gare

Les deux dernières décisions sont celles étudiées par les auteurs. En partant d'une grille horaire rigide, introduire de la flexibilité δ consiste alors à :

- transformer l'instant rigide d'arrivée a en un intervalle $[a - \delta; a]$
- transformer l'instant rigide de départ d en l'intervalle $[d - \delta; d]$

Cela impose notamment d'être transparent vis à vis du client en lui annonçant la date d'arrivée au plus tard et la date de départ au plus tôt. Ainsi, du point de vue du client, cette transparence impose de communiquer une arrivée en a et un départ en $d - \delta$. Les durées nominales de stationnement vont donc diminuer, et les temps de parcours nominaux annoncés vont augmenter par rapport à ceux de la grille rigide, bien que cela permette *in fine* en temps réel une meilleure gestion opérationnelle des circulations et un temps de parcours moyen inférieur compte tenu des multiples perturbations plus facilement maîtrisables. Il est alors possible de diminuer les marges habituelles associées à chaque train en augmentant la flexibilité sur les horaires. Trouver le compromis entre marges et flexibilité est un enjeu important partiellement traité dans les résultats expérimentaux mais pouvant encore nécessiter des travaux plus poussés tant théoriques que pratiques.

Le problème résolu en temps réel est celui de la *résolution des conflits*, modélisé avec les graphes alternatifs de Mascis et Pacciarelli (2002), où l'on cherche à obtenir une grille sans conflit entre les sillons. Trois heuristiques simples s'apparentant à des stratégies sont successivement appliquées, ainsi qu'un Branch & Bound, initialisé avec la meilleure solution trouvée.

2.2.3 Contextes d'optimisation à plusieurs étapes

Dans le domaine ferroviaire la flexibilité telle que définie pour les problèmes d'ordonnement est rarement envisagée comme un critère à part entière, et les articles faisant explicitement référence à la notion de flexibilité ou agilité en planification, comme ceux de la section précédente, sont relativement peu nombreux, et semblent uniquement concerner la grille horaire. Néanmoins, comme pointé par Billaut et al. (2008), flexibilité et robustesse sont deux notions intimement liées. Ainsi, certains articles traitant de la robustesse dans les différents problèmes de la planification ferroviaire proposent des méthodes de résolution agiles, et c'est cette agilité qui permet de garantir la robustesse de la planification.

On présente donc dans cette section des contextes d'optimisation où les variables de décision sont fixées à différentes étapes, avec une connaissance plus ou moins fine des données du problème à chaque étape. De manière générale, l'objectif est d'améliorer la robustesse de la solution choisie en amont en s'autorisant une certaine flexibilité sur le choix de certaines variables, et on trouve quelques applications pour le domaine ferroviaire.

2.2.3.1 Aperçu

Optimisation stochastique biniveau

L'optimisation stochastique est un cadre formel qui permet de modéliser l'incertitude sur certains paramètres au moyen de leur distribution de probabilité, supposée connue à l'avance. On pourra consulter le tutoriel de Shapiro et Philpott (2007) pour un aperçu de l'optimisation stochastique, et notamment la section 2 en ce qui concerne l'optimisation stochastique biniveau. On se contente ici de présenter le contexte décisionnel associé, dont s'inspirent assez largement tous les contextes décisionnels à plusieurs étapes, bien qu'ils ne fassent pas nécessairement intervenir des distributions de probabilité.

On considère deux types de variables :

- on regroupe dans x les variables fixées *en première étape*
- on regroupe dans y les variables fixées *en seconde étape*

Le contexte décisionnel est le suivant. En toute connaissance des distributions de probabilité des différents paramètres apparaissant dans le modèle :

- on choisit x
- on « révèle » le scénario, c'est-à-dire qu'on fixe une réalisation $\omega \in \Omega$, de telle sorte à ce que chaque paramètre incertain a prenne désormais la valeur $a(\omega)$
- on choisit y

L'objectif est alors de minimiser (ou maximiser) le gain moyen, c'est-à-dire l'espérance mathématique d'une certaine fonction objectif fixée dépendant de x , de y et de paramètres éventuellement incertains.

Dans le cas linéaire, on peut écrire le modèle suivant :

$$\min_x c^T x + \mathbb{E}_\omega \left[\min_y \left[q^T(\omega) y \right] \right]$$

$$\left[T(\omega)x + W(\omega)y \leq h(\omega) \right]$$

On peut réécrire le modèle sous la forme suivante, en considérant de manière équivalente que \tilde{y} est une variable aléatoire à déterminer qui à une réalisation ω de la distribution de probabilité associée une valeur de la variable y :

$$\min_{x, \tilde{y}} c^T x + \mathbb{E}_\omega [q^T(\omega) \tilde{y}(\omega)]$$

s. c. $T(\omega)x + W(\omega) \tilde{y}(\omega) \leq h(\omega) \quad \forall \omega \in \Omega$

Optimisation robuste ajustable et adaptative

L'optimisation stochastique n'est pas toujours une approche envisageable : en effet, dans de nombreuses applications, la distribution de probabilité associée aux événements est très difficile à estimer. Il est alors

illusoire de tenter de résoudre un modèle d'optimisation stochastique. De plus, la plupart des modèles sont difficiles à résoudre, même dans le cas d'une seule étape.

Suite à ces deux constats, Bertsimas et Sim (2004) ont développé l'optimisation robuste, contexte décisionnel où l'on remplace les distributions de probabilité par des *scénarios*, chaque scénario $s \in S$ correspondant à un jeu de valeurs pour l'ensemble des paramètres. Dans sa version classique avec un seul niveau de décision, l'objectif est de trouver une solution x qui soit faisable quel que soit le scénario considéré, et qui garantisse une certaine valeur de la fonction objectif pour chacun des scénarios. On cherche alors à trouver un x qui minimise cette valeur, correspondant au *pire* scénario qui pourrait arriver après avoir choisi la solution x .

En suivant la même approche que pour l'optimisation stochastique, Ben-Tal et al. (2004) introduisent alors l'optimisation robuste biniveau, également appelée *optimisation ajustable* ou *optimisation adaptative*. De manière similaire à l'optimisation stochastique biniveau, ce contexte décisionnel considère lui aussi deux types de variables x et y correspondant aux décisions à fixer avant et après révélation du scénario, et l'objectif est de trouver x tel que, quel que soit le scénario $s \in S$, il existe y^s rendant (x, y^s) admissible.

Les auteurs montrent que ce problème d'admissibilité pure est intractable dans le cas général. Parmi toutes les solutions x' respectant ce premier critère d'admissibilité, un problème plus intéressant est de choisir l'une de celles qui minimisent une fonction objectif dans le pire *scénario* possible, cette fonction objectif pouvant éventuellement dépendre de la variable y^s fixée en seconde étape. Formellement, on peut réécrire ce problème de la façon suivante dans le cas linéaire :

$$\min_x c^T x + \max_s \min_y (q^s)^T y$$

$$T^s x + W^s y \leq h^s$$

En considérant de manière équivalente qu'une solution du problème consiste en la variable x et une collection de variables $(y^s)_{s \in S}$ à appliquer pour chacun des scénarios, ce problème peut se réécrire de la façon suivante :

$$\min_{x, (y^s)_{s \in S}} c^T x + \max_s (q^s)^T y^s$$

$$\text{s. c. } T^s x + W^s y^s \leq h^s \quad \forall s \in S$$

Des résultats théoriques sont démontrés. En particulier, dans le cas où l'ensemble des scénarios peut être décrit par un polyèdre convexe paramétré par ξ et où la matrice de recours W^s est nulle, les auteurs montrent qu'il existe une solution optimale $(x, (y(\xi))_\xi)$ avec $y(\xi) = k + K\xi$ dépendant affinement de l'incertitude, et montrent alors que le problème est équivalent à un programme linéaire qui se résout polynomialement. Cette notion de *stratégie de retour affine* directement inspirée du retour linéaire d'état en automatique est reprise et améliorée par Bertsimas et Goyal (2012) : le cas d'une stratégie affine est opti-

mal si l'ensemble d'incertitudes est un simplexe⁶. Dans les cas d'un ensemble d'incertitudes convexe plus général, les auteurs montrent des bornes inférieures pour une stratégie affine par rapport à l'optimum.

Robustesse avec recours

Dans l'optimisation adaptative et l'optimisation stochastique biniveau, la variable de seconde étape y se comprend comme une variable du modèle initial dont la valeur est fixée plus tardivement, une fois que la conjoncture (réalisation ω ou scénario s) s'est précisée. Cependant, on peut également considérer que y est une modification de x . Le terme $c^T x$ dans la fonction objectif est alors le coût associé à la solution initialement choisie, et le terme $q^T y$ correspond au coût dû à la transformation de x en y . On parle alors de variable de *recours* pour y : elle permet de réparer la solution après révélation du scénario.

L'optimisation robuste avec recours est donc bien un cas particulier de l'optimisation adaptative. La contrainte de faisabilité concernant x et y est alors remplacé par :

- une contrainte assurant l'admissibilité de x dans le scénario nominal
- une contrainte assurant l'admissibilité de y_s pour chaque scénario s

Introduit initialement par Liebchen et al. (2009), la robustesse avec recours peut également se voir comme l'application d'un *algorithme* de recours $A \in \mathcal{A}$. Cet algorithme prend en entrée la solution initialement choisie x ainsi que le scénario s , et renvoie en sortie la nouvelle solution $y = A(x, s)$. Les auteurs décrivent une formulation générique d'un problème d'optimisation robuste avec recours, dont une solution consiste en un couple (x, A) , appelée *précaution*. Si P et P_s désignent respectivement l'ensemble des solutions admissibles pour le scénario initial et le scénario s , on aboutit à la formulation suivante.

$$\begin{aligned} & \inf_{(x,A) \in P \times \mathcal{A}} f(x) \\ \text{s. c. } & A(x, s) \in P_s \quad \forall s \end{aligned}$$

Le coût d'application de l'algorithme est supposé nul sans perte de généralité, car il est possible d'introduire un coût dans les contraintes *via* une variable de budget associée. Ainsi, ce coût associé au recours permet de retrouver le formalisme de l'optimisation robuste ajustable, dont l'optimisation robuste avec recours n'est qu'un cas particulier avec une formulation différente.

Il existe plusieurs classes d'algorithmes \mathcal{A} . On propose ici une classification inspirée par celle de Cicerone et al. (2009) :

- Le cas de l'optimisation robuste statique de Bertsimas et Sim (2004) correspond à l'application d'aucun recours : $A(x, s) = x$
- Inversement, le cas de la ré-optimisation sans contrainte correspond à celui où les recours sont possibles : $A(x, s)$ est une solution du problème où les paramètres ont été fixés à leur valeur dans le scénario s

6. Simplexe : enveloppe convexe de $m + 1$ points affinement indépendants dans \mathbb{R}^m

- On peut parvenir à un compromis entre l'optimisation robuste statique et la ré-optimisation sans contrainte en introduisant des restrictions pour l'algorithme :
 - L'algorithme de recours doit s'effectuer en un temps raisonnable (borné par une constante)
 - L'algorithme doit fournir une solution $y = A(x, s)$ qui est « proche » de x pour une certaine norme : $\|A(x, s) - x\| \leq K$, où K est une constante fixée initialement. Cela correspond à l'optimisation robuste avec *recours incrémental* étudiée notamment par Nasrabadi et Orlin (2013).

Ancrage et solution stable

Récemment, Bendotti et al. (2017) ont introduit la notion d'*ancrage d'une solution* d'un problème d'optimisation avec incertitudes. Partant du constat que modifier une solution suite à un changement la rendant non admissible est une opération très coûteuse, ils cherchent des solutions initiales proactives qui s'adaptent facilement à tous les aléas, de manière similaire à l'optimisation avec recours. Ils définissent ainsi le *niveau d'ancrage* entre deux solutions comme la similitude structurelle entre ces solutions, définie par exemple par le nombre de variables de décision identiques. Ils définissent alors deux types de problèmes :

- Le problème réactif à niveau d'ancrage maximal intervient après révélation du scénario. Il s'agit de déterminer une solution y avec un niveau d'ancrage maximal vis à vis de la solution initialement choisie x , supposée non admissible pour le scénario considéré. Ce problème peut se voir avec le formalisme de l'optimisation robuste avec recours comme le choix d'un algorithme A à coût minimal, c'est-à-dire garantissant la faisabilité de la solution avec un niveau d'ancrage maximal.
- Le problème proactif du *choix d'une solution initiale avec niveau d'ancrage garanti*. Il s'agit du problème résolu à la première étape avant révélation du scénario. Une telle solution est dite *ancrée* lorsque, quel que soit le scénario considéré, il existe une solution avec laquelle elle a un niveau d'ancrage élevé.

De manière similaire, Akkan et al. (2016) ont étudié la robustesse d'une solution dans le cas particulier de l'élaboration d'emploi du temps pour des projets d'étudiants. Ils cherchent ainsi à produire des emplois du temps qui soient facilement adaptables en cas d'empêchement majeur d'un intervenant ou retard important, cette adaptabilité se traduisant par l'existence d'un autre emploi du temps admissible qui soit *peu différent*. On parle alors de *stabilité* de la solution trouvée. Les auteurs construisent un graphe dont chaque sommet représente une solution admissible pour un ou plusieurs scénarios donnés, et où une arête relie deux emplois du temps jugés similaires. Il s'agit alors de choisir un sommet qui soit relié à le plus de sommets possibles pour un maximum de scénarios. Les auteurs proposent ainsi plusieurs distances heuristiques sur ce graphe pour caractériser au mieux la stabilité d'une solution vis à vis d'un ensemble de scénarios.

Recours systématique vers l'optimalité

Goerigk et Schöbel (2014) ont introduit un concept d'optimisation sous incertitude en deux étapes nommé *recovery-to-optimality*. Les principales différences avec la robustesse avec recours sont les suivantes :

- La solution x choisie initialement n'est pas nécessairement admissible pour le scénario nominal considéré
- Le recours consiste à choisir $y = A(x, s)$ comme une solution optimale (non nécessairement unique) pour le scénario révélé s

L'objectif est alors de minimiser la moyenne des coûts de la transformation de x en y , ou bien le pire coût, *ie.* respectivement la somme des distances ou le maximum de la distance de x à la solution optimale y_s d'un scénario s pour tous les scénarios. On retrouve ici la définition du niveau d'ancrage entre deux solutions, qu'on cherche à maximiser en choisissant une solution initiale fortement ancrée.

Ce contexte peut donc être vu comme un cas particulier de la robustesse avec recours et de la recherche de solution ancrée, à l'exception de la condition de faisabilité de la solution initiale qui peut ici être omise. Ce contexte peut également se rapprocher de la robustesse *min-max regret*, qui consiste à choisir une solution initiale qui minimise le manque-à-gagner dans le pire scénario, c'est-à-dire la différence entre la fonction objectif de la solution choisie pour ce scénario et une solution optimale pour ce scénario. En percevant cette différence de fonction objectif comme une distance, la robustesse *min-max regret* peut ainsi se voir comme un cas particulier de *recovery-to-optimality*, où la solution n'est toutefois pas modifiée en seconde étape et où l'on impose sa faisabilité dans le scénario initial et dans chacun des scénarios envisagés.

2.2.3.2 Applications au domaine ferroviaire

Le *timetabling* avec recours

Dans leur article fondateur sur la robustesse avec recours, Liebchen et al. (2009) appliquent ce concept au problème du *timetabling*. Ils s'appuient sur une modélisation de type FDP avec un graphe des événements. Les scénarios considérés correspondent à des temps minimaux plus grands entre certaines paires d'événements. Dans le cas robuste classique, le problème consiste alors à augmenter tous les temps minimaux d'une certaine valeur fixée, ce qui est très conservatif. Les auteurs proposent alors trois algorithmes pour la robustesse avec recours, sans applications numériques.

1. Le premier algorithme interdit la suppression d'un train, et consiste simplement à définir un nouvel horaire π_i^s pour chaque événement i et scénario s , tel que $\pi_i^s \geq \pi_i$, π faisable et π_i^s faisable pour le scénario s . Ainsi, le problème de robustesse avec recours peut être modélisé avec un PL. De plus, le prix de la robustesse avec recours est contrôlé au moyen de pénalisations dans la fonction objectif. Le modèle obtenu n'est cependant pas tractable compte tenu du nombre de variables et contraintes proportionnels au nombre de scénarios.

2. Le deuxième algorithme est une variante également intractable où on peut supprimer des correspondances, en supprimant des arcs du graphe dans l'algorithme de recours.
3. Le troisième algorithme est tractable, et correspond à une résolution avec système de règles : en pratique, on autorise un train à attendre au plus un certain temps fixé par rapport à son horaire nominal, après quoi il est autorisé à repartir en supprimant la correspondance. Cet algorithme est non-anticipatif : même si le conducteur sait qu'il lui faudra attendre au moins ce temps, il n'a pas le droit de faire partir le train avant.

Cicerone et al. (2009) proposent également d'appliquer la robustesse avec recours au problème du *timetabling*. Ici, chaque scénario correspond à l'accroissement d'une unique durée minimale dans le graphe des événements, et ne peut dépasser une certaine durée α . La vision des auteurs est cependant différente de celle de Liebchen et al. (2009), puisque la robustesse avec recours est vue comme l'application successive de deux algorithmes indépendants :

1. dans un premier temps l'algorithme robuste : attribution des horaires avec définition d'une certaine *marge de robustesse* pour chaque arête du graphe des événements, ce qui permet de définir une solution robuste admissible pour le *timetabling*;
2. dans un second temps l'algorithme de recours, après révélation du scénario : l'horaires d'au plus Δ événements peut être modifié. L'algorithme de recours est donc un algorithme de *timetabling* auquel on ajoute la contrainte sur le nombre maximum d'événements dont l'horaire peut être modifié.

Les auteurs s'intéressent alors au prix de la robustesse et démontrent des résultats théoriques en fonction de la structure du graphe. En particulier, si la fonction objectif est une combinaison linéaire des durées entre chaque paire d'événements, ils développent un algorithme pseudo-polynomial pour le problème robuste avec recours, qui est prouvé optimal. De plus, le prix de la robustesse est évalué au moyen d'une borne supérieure dépendant de α et Δ , correspondant à l'écart de performance maximal entre une solution nominale optimale et une solution optimale robuste avec recours.

Enfin, Goerigk et Schöbel (2014) ont appliqué le principe *recovery-to-optimality* au *timetabling*, en se comparant notamment avec la robustesse stricte et la robustesse avec recours de Liebchen et al. (2009). Ils considèrent essentiellement un ensemble d'incertitude uniforme, où les durées minimales entre deux événements sont incertaines. Si on considère la norme L^1 pour évaluer la distance entre deux solutions lors du recours, les auteurs remarquent que la relaxation de la contrainte de faisabilité de la solution initiale permet d'améliorer les résultats obtenus, et leur algorithme est montré tractable.

Choix d'un itinéraire robuste avec recours

Goerigk et al. (2013) décrivent le problème du choix d'un itinéraire pour un voyageur se déplaçant sur un réseau de transport. Ils s'intéressent notamment à la version robuste avec recours, où il est possible de changer d'itinéraire en temps réel en cas de perturbation sur une partie du réseau. Un scénario correspond alors à un ensemble de nouveaux horaires pour chaque sommet du graphe des événements : en pra-

2.2. FLEXIBILITÉ ET AGILITÉ EN PLANIFICATION

tique les instances correspondent à la propagation d'un retard avec des règles très simples. L'objectif est de trouver un chemin qui soit relativement rapide mais également robuste avec recours, au sens où si le temps de parcours est plus long sur une partie du graphe, le voyageur a un itinéraire alternatif de bonne qualité à partir du premier sommet perturbé. Les auteurs démontrent l'optimalité d'un processus simple et surtout facilement implémentable pour calculer un bon chemin avec stratégies de recours efficaces. Des résultats numériques sont présentés sur des instances réelles, où des courbes de Pareto permettent de confronter la robustesse de la solution et le temps de parcours dans le cas nominal.

Chapitre 3

Génération de roulements matériels nominaux

Contenu

2.1 Planification ferroviaire	33
2.1.1 Planification séquentielle des ressources	33
2.1.2 Planification intégrée des ressources	40
2.1.3 Replanification	41
2.2 Flexibilité et agilité en planification	42
2.2.1 Cas des problèmes d'ordonnancement sur machine	42
2.2.2 Flexibilité d'une grille horaire	45
2.2.3 Contextes d'optimisation à plusieurs étapes	48

Dans ce chapitre, nous traitons le cas nominal de la planification du matériel roulant. Après avoir présenté les caractéristiques de cette ressource et défini le Problème Nominal que nous cherchons à résoudre, nous effectuons un état de l'art bibliographique concernant ce problème tant dans la littérature scientifique qu'au sein de la SNCF. Nous détaillons alors notre modèle dans les deux dernières sections de ce chapitre en définissant le « Problème Nominal ».

3.1 Planification du matériel roulant

Comme évoqué dans la première partie, la planification du matériel roulant intervient généralement une fois les horaires des trains fixés, après résolution du *timetabling*. Cette planification est résolue par chaque EF sur chacun de ses périmètres. Il s'agit de déterminer l'emploi du temps de chaque élément du parc matériel sur une période donnée, en respectant les horaires et caractéristiques des trains. Des caractéristiques essentielles sont par exemple la puissance de traction nécessaire, le nombre de places demandées et/ou le type des éléments pouvant être affectés à ce train.

3.1.1 Parc matériel et unités simples

Le parc matériel d'une EF est constitué de plusieurs éléments de nature différente. Dans le problème de planification du matériel roulant, on s'intéresse aux *unités simples*, chacune étant indépendante et éventuellement constituée de différents éléments insécables. On parle aussi d'*engins moteurs*, ou plus simplement d'*engins*.

Par exemple, une unité simple sur le périmètre TGV correspond à une rame automotrice indépendante, disposant de ses propres moyens de traction, et constituée d'un certain nombre d'éléments (ou voitures) indissociables. Sur le périmètre Fret, une unité simple correspond à une locomotive avec une puissance de traction donnée. En général, la planification des *wagons* est réalisée en amont de la planification des locomotives, ce qui permet de définir les puissances de traction nécessaires pour chaque train. On considère donc qu'il ne s'agit pas d'unités simples et on ne s'intéressera pas à leur planification.

Une unité simple possède un certain type, qui permet à lui seul de préciser ses caractéristiques. On parle alors de sa *série matérielle*.

Il est possible de *coupler* plusieurs unités simples de séries matérielles identiques ou compatibles pour former une unité multiple. Par exemple, un TGV peut être couvert par une à deux rames, un TER par une à trois rames, et un train de fret jusqu'à quatre locomotives. Lorsqu'on réalise le couplage d'unités simples, on dit que l'on effectue des *accroches*. Inversement, lorsque ces unités sont découplées, on parle alors de *coupes*. Bien qu'elles nécessitent un certain temps d'exécution, les coupes et les accroches sont des manœuvres très couramment réalisées et apportent davantage de possibilités dans la planification.

Une *composition* est un n -uplet de séries matérielles, et indique la liste et l'ordre des unités simples au sein des trains. Cet ordre est important en opérationnel, car il conditionne notamment quelles unités simples peuvent être couplées ou découplées en fonction du temps disponible pour réaliser cette manœuvre et des restrictions liées au plan de voie en gare. Cependant, l'ordre des unités simples au sein d'une composition n'a pas été pris en compte dans nos travaux, qui concernent des phases de planification en amont de la phase opérationnelle.

Dans la suite de ce manuscrit, on désignera une unité simple par le terme générique « unité », par souci de simplification. Un train circulant avec une seule unité sera alors qualifié de « train en unité simple », ou train US, tandis qu'un train circulant avec plusieurs unités sera appelé « train unité multiple », ou train UM. On peut aussi parler de train UM2 ou train UM3 pour préciser la taille de l'unité multiple considérée.

3.1.2 Roulement matériel

L'emploi du temps d'une unité peut être représenté par une succession d'états. À chaque instant, une unité est dans l'un des états suivants :

- couverture d'un train ;
- maintenance (en un certain centre de maintenance) ;

3.1. PLANIFICATION DU MATÉRIEL ROULANT

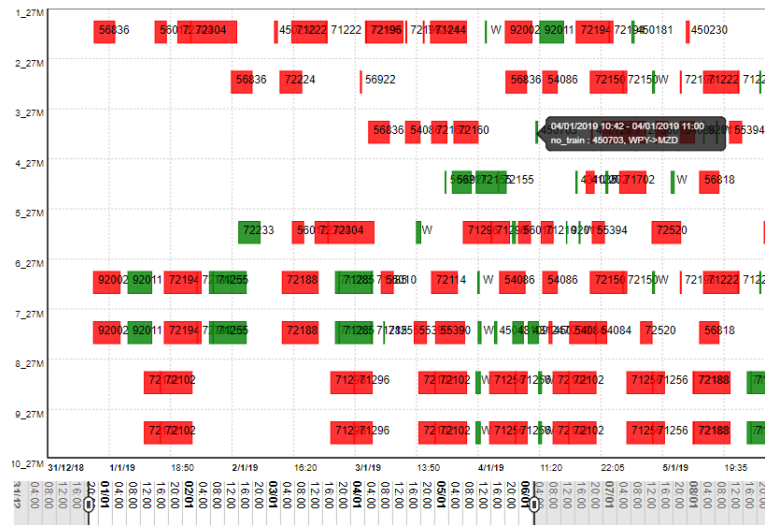


FIGURE 3.1: Exemple de roulement matériel, sous la forme d'un diagramme de Gantt. L'axe des abscisses correspond au temps, et chaque ligne de roulement de l'axe des ordonnées correspond à l'emploi du temps d'une unité. Chaque rectangle représente alors un train à couvrir ou une maintenance ou opération à effectuer.

- réalisation d'une manœuvre telles les coupes et accroches;
- en attente (en un lieu donné).

L'emploi du temps d'un ensemble d'unités est appelé *roulement matériel*¹ (voir Figure 3.1), et celui d'une unité donnée est appelé *ligne de roulement*. Certains experts et opérateurs du système ferroviaire désignent par « roulement » l'ensemble des lignes de roulement correspondant aux unités d'une même série matérielle. Nous ne retiendrons pas cette définition dans le cadre de ce manuscrit : un roulement peut donc être constitué d'unités de séries matérielles différentes.

Dans la suite de nos travaux, nous avons choisi de ne pas considérer les contraintes de maintenance, qui sont très spécifiques à chaque entité, afin de ne pas surcharger les modèles. Cependant, nos modèles peuvent aisément être complétés en intégrant de telles contraintes, et il est souvent possible de planifier les maintenances en post-traitement du modèle PLNE que nous proposons.

De la même façon, nous avons considéré que la durée d'une coupe et d'une accroche était nulle, et nous avons négligé toutes les manœuvres autres que les coupes et accroches. En particulier, on ne considère pas les mouvements d'une unité pour rejoindre un dépôt dans les environs d'une gare, ou pour acheminer cette unité jusqu'à un quai peu avant son départ : mis à part les mouvements pour acheminer une unité jusqu'au centre de maintenance, ces décisions relèvent généralement d'un problème plus microscopique à l'échelle

1. Le terme de « roulement » n'a aucun lien avec le caractère « roulant » du matériel. Il est ici inspiré du domaine de la logistique, où les individus se *remplacent* alternativement pour effectuer certaines tâches. Ce n'est pas toujours le cas, mais un roulement matériel est souvent conçu sur une semaine et de telle sorte à ce que les lignes de roulement puissent *s'enchaîner* selon un ordre défini à l'avance. Ainsi, l'unité affectée à la ligne de roulement 1 sera par exemple affectée à la ligne de roulement 2 la semaine suivante, et on définit de tels *enchaînements* pour chacune des lignes. Ceci impose notamment des restrictions sur les gares à atteindre en début et en fin d'horizon temporel, comme nous le verrons plus tard avec la cyclicité des roulements.

d'une gare, appelé problème de *shunting*.

3.2 État de l'art pour la planification du matériel roulant

3.2.1 Les différentes formulations de la littérature

Dans cette section, on effectue une revue de la littérature des problèmes de planification du matériel roulant. On ne s'intéressera ici qu'aux modèles traitant uniquement de cette ressource, ce qui exclue les modèles de planification intégrée, évoqués dans le Chapitre 2.

Fioole et al. (2006) ont défini le Problème de Circulation du Matériel Roulant (*Rolling Stock Circulation Problem*). Par la suite, il a été étendu et très largement utilisé dans la littérature, notamment par Nielsen et al. (2012); Cacchiani et al. (2012); Wagenaar et al. (2017). On décrit ici le modèle d'origine de Fioole et al., les mêmes auteurs ayant proposé un algorithme de Branch & Price pour accélérer sa résolution dans Peeters et Kroon (2008).

L'originalité de leur modèle par rapport à la littérature existante est de pouvoir considérer les opérations de coupes et accroches, ainsi que de pouvoir déterminer la position d'une unité dans une composition. Ainsi, un train circulant en UM2 avec deux unités de séries matérielles différentes A et B peut se voir affecté la composition AB (A à l'avant, B à l'arrière) ou BA (B à l'avant, A à l'arrière). En entrée, ils considèrent une grille horaire avec un ensemble de trains à planifier, ainsi qu'une estimation du nombre de voyageurs pour chaque train. De plus, ils considèrent que les *successeurs* d'un train sont connus à l'avance. Ainsi, une unité effectuant le train i devra nécessairement être affecté à un train $j \in N(i)$ une fois le train i arrivé : elle devra soit être affectée à un de ces trains, soit être découplée puis attendre au dépôt de la gare pendant une durée suffisamment grande, avant d'être affectée à un autre train. L'ensemble $N(i)$ est toujours de taille 1 ou 2, le second cas correspondant à un train se *séparant* en deux autres trains. Enfin, l'ensemble des compositions admissibles pour chaque train est considérée en entrée du problème. Les auteurs proposent alors un modèle PLNE minimisant les trois critères suivants :

1. le nombre total de kilomètres parcourus pour chaque série matérielle;
2. le nombre de voyageurs n'ayant pas de place assise;
3. le nombre de coupes et accroches.

On dit qu'un train UM sans voyageurs debout est en *surcomposition* s'il est possible de réduire sa composition en supprimant une unité, et tel que tous les voyageurs aient une place assise. Dans ce cas, on peut se permettre de ne pas ouvrir cette unité à la clientèle. Les surcompositions sont fréquentes dans la mesure où elles permettent de *repositionner* une unité en une autre gare, afin qu'elle puisse être affecté à un train. Plus il y a de surcompositions, plus le nombre de kilomètres parcourus est important : ainsi, le premier objectif est une façon équivalente de minimiser ces surcompositions.

Le modèle PLNE s'appuie sur un *graphe de transition* (voir Figure 3.2), qui permet de représenter les changements de compositions entre deux trains, et donc de calculer finement les coupes et accroches. Trois

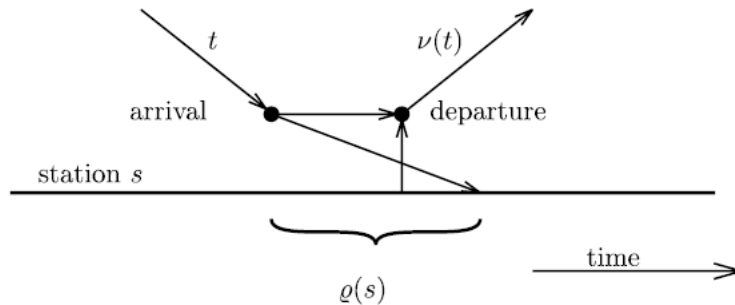


FIGURE 3.2: Exemple d'un graphe de transition utilisé par Fioole et al. (2006)

types de variables sont définies :

- les variables binaires $Z_{tpp'}$ indiquent si la composition p est affectée au train p et si la composition p' est affectée au train $v(t)$, dans le cas où t n'a qu'un seul successeur $v(t)$;
- les variables binaires $Z_{tpp_1p_2}$ indiquent si la composition p est affectée au train p , la composition p_1 au premier successeur de t et la composition p_2 au second successeur de t , dans le cas où t a deux successeurs;
- les variables binaires de travail X_{tp} indiquent si la composition p est affectée au train t ;

Les contraintes liant ces variables entre elles se déduisent aisément de leur définition. Les auteurs introduisent également des contraintes et variables additionnelles permettant de dénombrer les unités présentes en chaque gare à chaque instant, ce qui permet notamment de considérer un *inventaire* imposé au début de l'horizon temporel. Le modèle est résolu avec le solveur IBM ILOG Cplex, une méthode de branchements à base de priorités est proposée, et les auteurs décrivent une heuristique utilisant de la recherche locale pour résoudre leur problème.

Ahuja et al. (2005) ont étudié un problème similaire à celui de Fioole et al. (2006) adapté au transport de marchandises, et ont proposé une modélisation différente. Leur modèle PLNE consiste à calculer un multiflot sur un graphe d'événements s'appuyant sur le graphique espace-temps, appelé *graphe espace-temps*. La Figure 3.3 présente un exemple d'un tel graphe. Chaque flot correspond à une série matérielle. Les sommets correspondent essentiellement au départ et à l'arrivée d'un train. Si on considère un flot d'un certain type, un arc (d, a) de flux $\phi_{da} \in \mathbb{N}$ entre un sommet-départ d et un sommet-arrivée a correspond alors à la circulation d'un train couvert par ϕ_{da} unités de la série matérielle correspondante. Les autres arcs, qui sont tous horizontaux sur la Figure 3.3b, correspondent à l'attente d'unités en une gare sans qu'elles ne couvrent de train.

Cacchiani et al. (2010) ont étudié le problème d'Affectation des Unités de Train, ou *Train Unit Assignment Problem*. Le graphe peut être assimilé au graphe dual de celui de Ahuja et al. (2005) : chaque sommet correspond à un train, et un arc (i, j) entre deux sommets avec un flux $\phi_{ij} \in \mathbb{N}$ correspond à la couverture du train i puis du train j par ϕ_{ij} unités. Contrairement au modèle de Fioole et al. (2006), ce modèle ne contraint

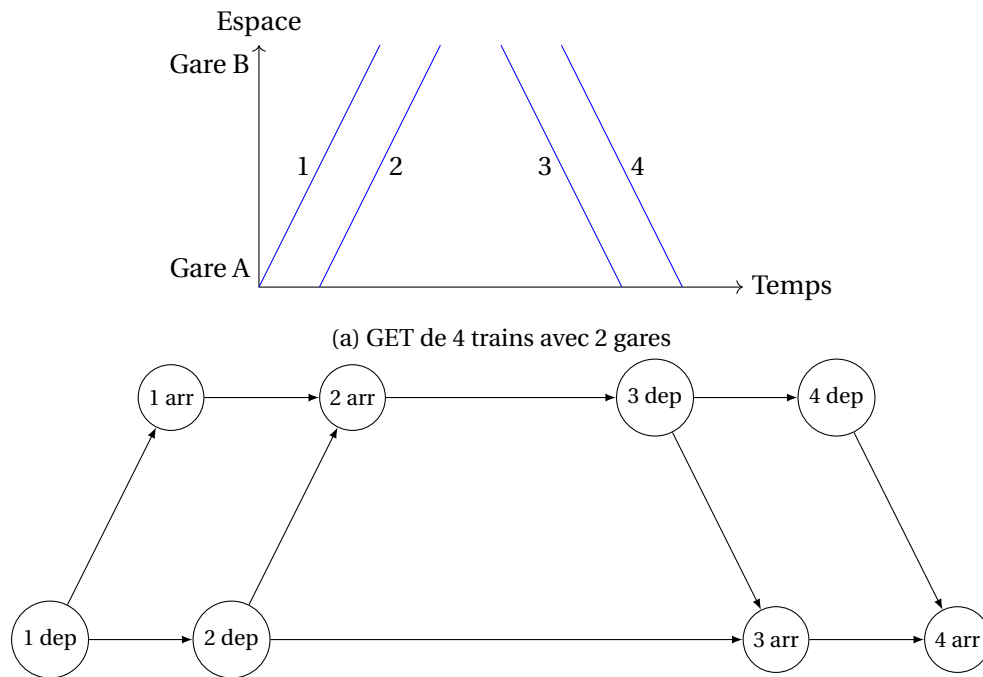


FIGURE 3.3: Exemple de graphe espace-temps

pas les successeurs de chaque train en entrée. Un PLNE s'appuyant sur ce graphe est proposé pour calculer un multiflot. Les auteurs introduisent une formulation par arcs, où chaque variable entière x_{ij} indique le nombre d'unités affectées au train i puis au train j . Une reformulation par chemins est proposée, qui est ensuite résolue heuristiquement.

Giacco et al. (2014) ont proposé un modèle intégrant la planification des rames et de leurs maintenances. Ils rendent explicitement possible les trains non commerciaux correspondant à des trains sans voyageurs ni marchandises. On parle ainsi de « sillons de repositionnement », ou de manière équivalente de *sillons W* , pour « Vide Voyageur ». Ces deux termes seront utilisés tout au long de ce manuscrit. Le modèle proposé consiste à trouver un ensemble de circuits hamiltoniens. Il est formulé par un PLNE, et résolu par un solveur commercial. Il s'agit d'une formulation par chemins, et l'objectif est la minimisation du nombre de rames.

Plus récemment, (Borndörfer et al., 2011; Borndörfer et al., 2016) ont introduit une nouvelle approche pour calculer des roulements matériels, développée plus en détail dans la thèse de Reuther et al. (2012) : les modèles d'hypergraphes. Ils résolvent là encore un problème similaire, appelé Problème de Rotation du Matériel Roulant, ou *Rolling Stock Rotation Problem*. Un hypergraphe est constitué de sommets, d'arcs, et d'*hyperarcs* reliant un ensemble de sommets à un autre. On associe à chaque train un sommet départ et un sommet arrivée pour chacune de ses compositions admissibles. Un hyperarc orienté relie alors l'ensemble

des sommets arrivée d'un train i à l'ensemble des sommets départs d'un train j , et indique la possibilité de couvrir ces trains avec les mêmes unités. Un roulement matériel s'obtient alors en calculant un multiflot sur ce graphe. Une formulation MILP par chemins est proposée par les auteurs, et résolue avec un algorithme de générations de colonne.

3.2.2 Les modèles utilisés à la SNCF

À la SNCF, il existe plusieurs outils permettant d'optimiser la planification du matériel roulant. Nous présentons dans cette section deux d'entre eux : PRESTO et HERMES. Ils ont des données d'entrée similaires et permettent d'établir des roulements matériels durant différentes phases de planification. Cependant, ces outils s'appuient sur des modèles mathématiques différents.

3.2.2.1 PRESTO

L'outil PRESTO est issu des travaux de thèse de Nicolas Marcos chez SNCF Innovation & Recherche (voir Marcos et al. (2005); Marcos (2006)). Il est utilisé en phase d'études (1 à 2 ans avant le jour de circulation des trains), notamment lors de la conception de l'offre de transport afin de dimensionner les ressources matérielles à affecter à un périmètre géographique. Ces roulements matériels optimisés seront adaptés et retouchés par le chargé d'études afin de prendre en compte des contraintes spécifiques.

Le modèle est une extension de celui de Ahuja et al. (2005). Une solution de multiflot admissible sur le graphe doit respecter des contraintes de demande minimale et capacité maximale d'unités pour chaque train, ainsi que certaines contraintes comme les mouvements d'unités associés. L'ordre des unités au sein d'une composition n'est pas pris en compte, et les maintenances sont planifiées en post-traitement de la résolution du modèle PLNE. Il est possible de réaliser des sillons W , mais ceux-ci doivent préalablement avoir été anticipés en pré-traitement, par l'ajout de trains supplémentaires de demande nulle. Le problème principal est alors un problème multiobjectif où on minimise différents critères de performance, qui sont, par ordre d'importance : le nombre de trains non couverts, le nombre d'unités engagées, les kilomètres parcourus en sillon W ainsi que ceux parcourus en surcomposition. En pratique, un unique PLNE est résolu, où les différents objectifs sont agrégés avec une pondération telle que cela revient à optimiser ces objectifs dans l'ordre lexicographique.

Le modèle PLNE est une formulation par arcs, et PRESTO doit donc effectuer en post-traitement une décomposition du flot en chemins. Dans un souci de robustesse, suite aux travaux de (Tréfond, 2014; Tréfond et al., 2017), le choix consiste à équilibrer les lignes de roulements le mieux possible tout en restant cohérent avec la solution renvoyée par le PLNE. PRESTO cherche à homogénéiser les *temps de retournement* des unités, c'est-à-dire les durées d'attente en gare pour un engin entre deux trains consécutifs, et il est souhaité d'avoir un temps de retournement plus long afin d'absorber un retard potentiel. Ainsi, bien que PRESTO ne permette pas *a priori* de se prémunir contre des aléas au cours de la phase d'adaptation, il intègre une dimension de robustesse vis à vis des aléas opérationnels.

3.3. LE PROBLÈME DE CONCEPTION D'UN ROULEMENT MATÉRIEL NOMINAL (PROBLÈME NOMINAL)

3.2.2.2 HERMES

L'outil HERMES est développé depuis 2016 à la Direction Générale de l'Exploitation au sein de SNCF Réseau et Fret. Utilisé à l'origine par diverses entités, son moteur d'optimisation est aujourd'hui à la base des modèles de planification de Fret SNCF pour le matériel roulant, tant en phase de conception qu'en phase d'adaptation.

Le modèle sous-jacent s'appuie sur un graphe similaire à celui de Cacchiani et al. (2010), donc correspondant au graphe dual de celui de PRESTO.

Comme pour PRESTO, une solution de flot sur ce graphe est calculée au moyen d'une modélisation PLNE par arcs. En revanche, la donnée d'une telle solution laisse beaucoup moins de marge de manœuvre pour construire un roulement matériel complet en post-traitement de la résolution du PLNE. En effet, une solution de flot sur ce graphe définit l'ensemble des successeurs de chaque train (voir exemple Figure 3.4). L'ensemble des décisions à prendre en post-traitement de la résolution du PLNE est donc plus limité. La Figure 3.5 illustre un cas élémentaire où un algorithme de post-traitement mène à des roulements différents.

Les objectifs du PLNE sont essentiellement les mêmes que pour PRESTO mais l'agrégation n'est pas tout à fait équivalente à un ordre lexicographique. Chez Fret SNCF, par exemple, tous ces objectifs ont pu être évalués et valorisés en euros.

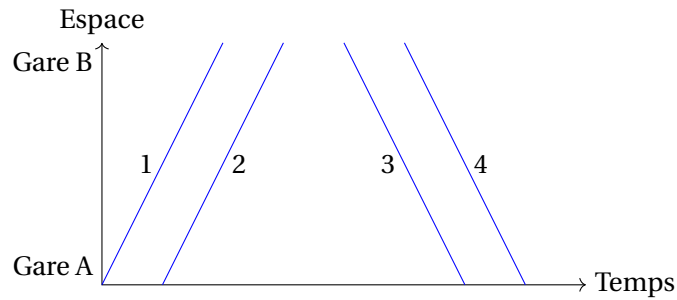
3.3 Le Problème de Conception d'un Roulement Matériel Nominal (Problème Nominal)

Comme nous le verrons dans le chapitre suivant, la notion d'adaptabilité d'un roulement matériel est très liée aux successions entre trains consécutifs sur une même ligne de roulement. Ainsi, le choix de la modélisation doit pouvoir intégrer les successions en amont du processus de décision. Un modèle s'appuyant sur le graphe tel que celui de PRESTO n'est donc pas adapté, car on ne pourra évaluer l'adaptabilité d'un roulement matériel dans le PLNE dans le cas d'une formulation par arcs. De même, le modèle de Fioole et al. (2006) ne peut convenir, au sens où le successeur d'un train est généralement unique et doit être défini en amont du modèle.

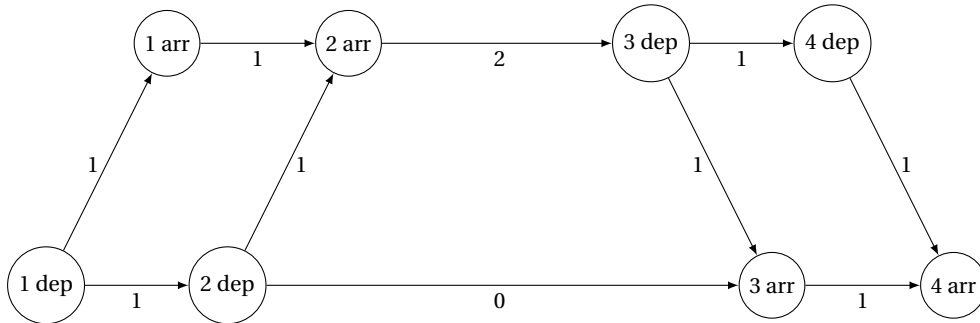
Nous avons donc choisi d'utiliser un modèle par graphe de précedence, proche de celui de Cacchiani et al. (2010).

Pour faciliter la lecture, le Problème de Conception d'un Roulement Matériel Nominal sera désigné par « Problème Nominal » dans la suite de ce manuscrit. La Figure 3.6 présente schématiquement les entrées et sorties de ce problème. Une instance I étant donnée telle qu'explicitée Section 3.3.1, on cherche à construire un roulement matériel \mathcal{R} au moyen de deux briques algorithmiques successivement appliquées : un modèle PLNE fournissant une solution intermédiaire R ainsi qu'une brique de post-traitement.

3.3. LE PROBLÈME DE CONCEPTION D'UN ROULEMENT MATÉRIEL NOMINAL
(PROBLÈME NOMINAL)



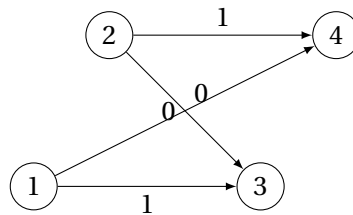
(a) GET de 4 trains avec 2 gares



(b) Sous-graphe associé par PRESTO avec la solution de flot renvoyée. Remarquons que la source et le puits n'ont pas été représentés, pour faciliter la lecture.

US1 : 1 3 et US1 : 1 4
US2 : 2 4 US2 : 2 3

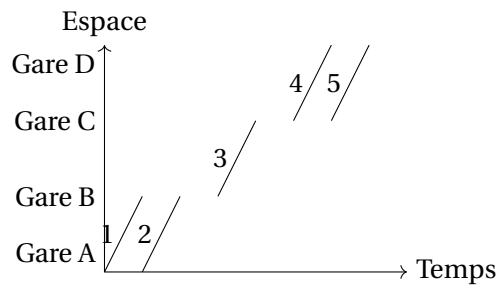
(c) Deux roulements matériels correspondant à la solution de flot de PRESTO



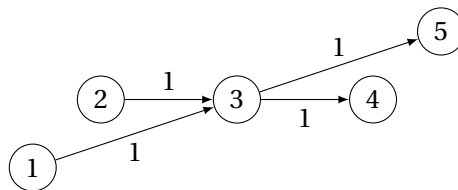
(d) Sous-graphe associé par HERMES avec la solution de flot renvoyée. Le roulement obtenu en post-traitement est ici unique, c'est le premier de la Figure 3.4c.

FIGURE 3.4: Comparaison entre les graphes de PRESTO et ceux d'HERMES. Dans cet exemple, une solution de flot sur le graphe de PRESTO correspond à deux roulements matériels différents, alors qu'une solution de flot sur le graphe d'HERMES détermine entièrement le roulement matériel obtenu en sortie.

3.3. LE PROBLÈME DE CONCEPTION D'UN ROULEMENT MATÉRIEL NOMINAL (PROBLÈME NOMINAL)



(a) GET de 5 trains avec 4 gares. On suppose que le train 3 requiert une UM2, les autres pouvant circuler en US.



(b) Sous-graphe associé par HERMES avec sa solution de flot

US1 : 1 3 5 et US1 : 1 3 4
 US2 : 2 3 4 US2 : 2 3 5

(c) Deux roulements matériels différents correspondant à la solution de flot d'HERMES

FIGURE 3.5: Plusieurs décompositions en chemins pour HERMES. Dans cet exemple élémentaire, il existe deux possibilités pour établir un roulement matériel en post-traitement à partir d'une solution de flot.

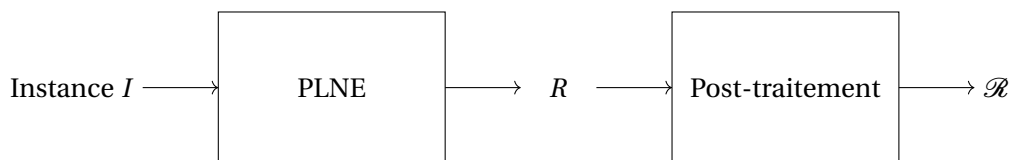


FIGURE 3.6: Schéma du Problème Nominal

3.3.1 Instance du Problème Nominal

Une *instance I* du Problème Nominal est définie par :

1. un ensemble de séries matérielles \mathcal{M} dont les caractéristiques sont représentées sur le Tableau 3.1 ;
2. un ensemble de gares \mathcal{G} dont les caractéristiques sont représentées sur le Tableau 3.2 ;
3. un ensemble de trains T^{train} à planifier, dont les caractéristiques sont représentées sur le Tableau 3.3.

3.3. LE PROBLÈME DE CONCEPTION D'UN ROULEMENT MATÉRIEL NOMINAL
(PROBLÈME NOMINAL)

$K_m \in \mathbb{N}$	Nombre d'unités de série matérielle $m \in \mathcal{M}$ à disposition
$C_m \in \mathbb{R}$	Coût d'engagement d'une unité de série matérielle $m \in \mathcal{M}$
$P_m \in \mathbb{N}$	Nombre de places dans une unité de série matérielle $m \in \mathcal{M}$, ou sa puissance de traction dans le cas du transport de marchandises
$Comp : \mathcal{M} \times \mathcal{M} \rightarrow \{0;1\}$	Fonction booléenne indiquant la compatibilité entre deux séries matérielles données

TABEAU 3.1: Données d'une instance I relatives aux séries matérielles

$Distance : \mathcal{G} \times \mathcal{G} \rightarrow \mathbb{R}$	Fonction indiquant la distance entre deux gares, arbitrairement grande si elles ne sont pas reliées entre elles
$CA : \mathcal{G} \rightarrow \{0;1\}$	Fonction booléenne indiquant la possibilité de réaliser des coupes et accroches en une gare donnée
$t_g^{ret} \in \mathbb{R}$	Temps de retournement minimal en la gare $g \in \mathcal{G}$, c'est-à-dire la durée minimale nécessaire pour qu'une unité arrivant en la gare g puisse manœuvrer pour être affectée à un autre train partant de g

TABEAU 3.2: Données d'une instance I relatives aux gares

$G^{dep} : T^{train} \rightarrow \mathcal{G}$	Fonctions indiquant la gare de départ d'un train donné
$G^{arr} : T^{train} \rightarrow \mathcal{G}$	Fonctions indiquant la gare d'arrivée d'un train donné
$D_i \in \mathbb{N}$	Demande cible du train $i \in T^{train}$ en nombre de places, ou sa puissance minimale dans le cas du transport de marchandises
$E_i^{max} \in \llbracket 0; \max(0, D_i - 1) \rrbracket$	Écart maximal autorisé par rapport à la demande cible si le train $i \in T^{train}$ est effectivement couvert; la demande minimale est alors $D_i - E_i^{max} > 0$, et tout train i est soit non couvert, soit couvert par une composition ayant au moins $D_i - E_i^{max}$ places
$D_i^{max} \in \mathbb{N}$	Capacité maximale du train $i \in T^{train}$ en nombre d'unités, toutes séries matérielles confondues
$D_{im}^{max} \in \mathbb{N}$	Capacité maximale du train $i \in T^{train}$ en nombre d'unités de série matérielle $m \in \mathcal{M}$

TABEAU 3.3: Données d'une instance I relatives aux trains

3.3. LE PROBLÈME DE CONCEPTION D'UN ROULEMENT MATÉRIEL NOMINAL
(PROBLÈME NOMINAL)

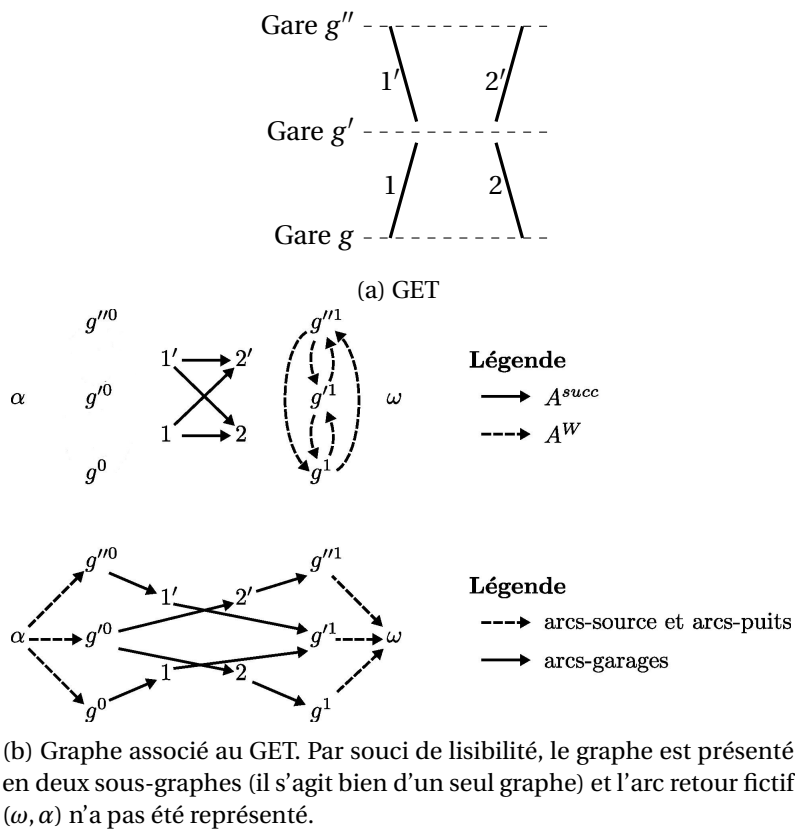


FIGURE 3.7: Exemple d'un graphe de précédence avec 4 trains et 3 gares

3.3. LE PROBLÈME DE CONCEPTION D'UN ROULEMENT MATÉRIEL NOMINAL (PROBLÈME NOMINAL)

T	Ensemble des sommets
$T^{train} \subset T$	Ensemble des sommets-trains, ou ensemble des trains
$T^{garage} \subset T$	Ensemble des sommets-garages
$\alpha \in T$	Sommet-source du graphe, ou source
$\omega \in T$	Sommet-puits du graphe, ou puits
$g^n \in T$	Sommet-garage correspondant à la gare $g \in \mathcal{G}$ en début ($n = 0$) ou fin ($n = 1$) d'horizon temporel
$A \subset T \times T$	Ensembles des arcs
$A^{succ} = A \cap (T^{train} \times T^{train})$	Ensemble des arcs-successeurs
$A^W = A \cap (T^{garage} \times T^{garage})$	Ensemble des arcs-repositionnement, ou arcs-W

TABLEAU 3.4: Notations liées au graphe de précédence du Problème Nominal

3.3.2 Graphe de précédence

Soit $G = (T, A)$ le graphe sur lequel nous nous appuyons pour résoudre le Problème Nominal. Un exemple de graphe est représenté sur la Figure 3.7. On rassemble dans l'ensemble des sommets T les sommets suivants :

- la source α ;
- les sommets-garages T^{garage} , comportant deux sommets par gare g , un sommet *initial* noté g^0 et un sommet *final* noté g^1 , par lequel passent effectivement les unités positionnées en cette gare au début (*resp.* à la fin) de l'horizon temporel ;
- les sommets-trains T^{train} , comportant un sommet par train ;
- le puits ω .

Les arcs de A sont orientés et sont également de différents types :

- les arcs successeurs A^{succ} , reliant deux sommets-trains pouvant être enchaînés par une même unité, selon les conditions expliquées ci-dessous ;
- les arcs-repositionnement ou arcs-W A^W , reliant chaque couple de sommets-garages finaux relatifs à deux gares où un sillon de repositionnement est possible en fin d'horizon temporel ;
- les arcs-source, reliant la source α à chacun des sommets-garage initiaux ;
- les arcs-puits, reliant chacun des sommets-garages finaux au puits ;
- les arcs-garage, reliant d'une part chaque sommet-garage initial aux sommets-trains au départ de la gare correspondante, et reliant d'autre part chaque sommet-train au sommet-garage final correspondant à sa gare d'arrivée ;
- un arc-retour (ω, α) .

Les différentes notations que nous avons introduites sont rappelées sur le Tableau 3.4

Pour que deux trains puissent être enchaînés par une même unité en une gare $g \in \mathcal{G}$, il faut que la durée séparant l'arrivée du premier et le départ du second soit supérieure à t_g^{ret} . Si la gare d'arrivée du premier train n'est pas celle de départ du second, il faut réaliser dans l'intervalle un sillon de repositionnement à

3.3. LE PROBLÈME DE CONCEPTION D'UN ROULEMENT MATÉRIEL NOMINAL (PROBLÈME NOMINAL)

vide, qui nécessite donc d'autant plus de temps, et ce sillon doit pouvoir être réalisable. Toutes ces discussions doivent avoir lieu en amont de la résolution de façon à générer un graphe le plus réaliste possible.

3.3.3 Modèle PLNE

Nous introduisons dans cette section un modèle PLNE s'appuyant sur le graphe décrit précédemment. Il consiste à trouver un multiflot entre la source et le puits, chaque flot correspondant à une série matérielle. Ce multiflot doit alors respecter certaines contraintes spécifiques au problème de planification du matériel roulant tout en optimisant des objectifs de nature différente.

3.3.3.1 Notations additionnelles

Nous introduisons quelques notations supplémentaires pour faciliter la lisibilité de notre modèle PLNE :

- $\mathcal{M}_i \subset \mathcal{M}$, séries matérielles compatibles avec le sommet $i \in T$: si $i \in T^{train}$, il s'agit des séries compatibles avec ce train, sinon $\mathcal{M}_i = \mathcal{M}$;
- $\mathcal{M}_{ij} = \mathcal{M}_i \cap \mathcal{M}_j$, séries matérielles compatibles avec les sommets $i \in T$ et $j \in T$;
- $Gare : T^{gare} \rightarrow \mathcal{G}$, fonction indiquant la gare associée à un sommet-gare.

3.3.3.2 Variables

Notre modèle comporte deux types de variables :

- des variables de décision x ;
- des variables de travail t, e, δ et δ' .

Une fois les variables de décision x fixées, les valeurs des variables de travail en découlent immédiatement. Autrement dit, imposer une valeur pour chaque composante de x détermine entièrement la valeur des autres variables du modèle. En particulier, seules les variables x seront étudiées en post-traitement pour générer un roulement matériel solution de notre problème.

La variable $x_{ijm} \in \mathbb{N}$ indique le flux sur l'éventuel arc $(i, j) \in A$ pour le flot de type $m \in \mathcal{M}$. Si $(i, j) \notin A$, alors $x_{ijm} = 0$. Lorsque i et j sont des trains, cela correspond au nombre d'unités de série matérielle $m \in \mathcal{M}$ affectées au train $i \in T^{train}$ puis immédiatement au train $j \in T^{train}$.

La variable $t_i \in \{0; 1\}$ vaut 1 si le train $i \in T^{train}$ est couvert par au moins une unité, 0 sinon. Couvrir un train n'étant pas une contrainte dure du modèle, la solution $x = 0$ est toujours admissible, et correspond à un roulement matériel vide. Les variables t_i seront maximisées pour inciter la couverture des trains, mais elles permettent de garantir qu'il existe toujours une solution admissible pour notre problème. Si un train $i \in T^{train}$ est couvert mais sa demande cible D_i est non respectée, la variable $e_i \in \mathbb{N}$ indique l'écart à la demande cible D_i en nombre de places. Si ce train n'est pas couvert ou si sa demande est respectée, alors

3.3. LE PROBLÈME DE CONCEPTION D'UN ROULEMENT MATÉRIEL NOMINAL (PROBLÈME NOMINAL)

$e_i = 0$. Ainsi, on peut écrire les relations

$$t_i = \begin{cases} 1 & \text{si } \sum_{j \in T} \sum_{m \in \mathcal{M}} x_{ijm} > 0 \\ 0 & \text{sinon} \end{cases} \quad (3.1)$$

$$e_i = \begin{cases} 0 & \text{si } t_i = 0 \text{ ou si } \sum_{m \in \mathcal{M}} P_m \cdot x_{ijm} \geq D_i \\ D_i - \sum_{m \in \mathcal{M}} P_m \cdot x_{ijm} & \text{sinon} \end{cases} \\ = \max(0, D_i \cdot t_i - \sum_{m \in \mathcal{M}} P_m \cdot x_{ijm}) \quad (3.2)$$

La variable $\delta_{ijm} \in \{0;1\}$ vaut 1 si $x_{ijm} > 0$, 0 sinon. La variable $\delta'_{ij} \in \{0;1\}$ indique s'il existe une unité couvrant le train i puis le train j , et vaut donc 1 s'il existe $m \in \mathcal{M}_{ij}$ tel que $\delta_{ijm} = 1$, 0 sinon. On a donc les relations suivantes :

$$\delta_{ijm} = \begin{cases} 1 & \text{si } x_{ijm} > 0 \\ 0 & \text{sinon} \end{cases} \quad (3.3)$$

$$\delta'_{ij} = \max_{m \in \mathcal{M}} \delta_{ijm} \quad (3.4)$$

On peut également retenir les identités suivantes concernant la variable t_i .

$$t_i = \max_{j \in T} \delta'_{ij} = \max_{j \in T} \delta'_{ji} \quad (3.5)$$

Exprimées sous forme linéaires, elles seront plus fortes que les contraintes linéaires correspondant à (3.1) et permettront donc de renforcer notre formulation PLNE.

Dans la suite de ce manuscrit, on désignera par R l'ensemble des variables du PLNE, c'est-à-dire le quintuplet $(x, t, e, \delta, \delta')$. Par abus de langage, on dira que R est un roulement matériel, bien qu'il soit pour cela nécessaire de décomposer le flot x en chemins.

3.3.3.3 Objectifs

L'objectif du problème de roulement nominal est de maximiser la *performance* du roulement obtenu en sortie. Nous reprenons ici les différents critères utilisés actuellement dans les outils de planification de roulement à la SNCF ainsi que dans la littérature scientifique, par ordre décroissant d'importance a priori.

Critères à optimiser

La performance d'un roulement matériel $R = (x, t, e, \delta, \delta')$ peut donc être décrite au moyen des critères ci-dessous.

• Couverture des trains

Le premier objectif est de couvrir chaque train $i \in T^{train}$ dont la demande D_i est positive par un nombre d'unités contenant au moins $D_i - E_i^{max}$ (≥ 1) places. Il s'agit donc de maximiser

$$Couv(R) := \sum_{\substack{i \in T^{train} \\ D_i > 0}} t_i \quad (3.6)$$

3.3. LE PROBLÈME DE CONCEPTION D'UN ROULEMENT MATÉRIEL NOMINAL (PROBLÈME NOMINAL)

- **Écart à la demande**

Dans un second temps, pour chaque train i couvert, on cherche à respecter au mieux sa demande cible D_i . On va donc chercher à minimiser l'écart à cette demande cible, écart précisément donné par la variable e_i . On cherche donc à minimiser

$$Ecart(R) := \sum_{i \in T^{train}} e_i \quad (3.7)$$

- **Coûts des unités engagées**

Tous les modèles de planification du matériel roulant intègrent *a minima* un objectif pour minimiser le coût des unités engagées dans le roulement sur le périmètre spatio-temporel donné en entrée. En effet, le coût d'une unité est très impactant pour une EF. On cherchera donc à minimiser

$$Unit(R) := \sum_{m \in \mathcal{M}} C_m \cdot x_{\omega \alpha m} \quad (3.8)$$

- **Kilomètres en sillon de repositionnement**

Les sillons de repositionnement, autrement appelés sillons W (« Vide Voyageur ») ou sillons haut-le-pied, sont couramment utilisés par toutes les EF. Ils consomment de l'énergie et nécessitent au moins un agent de conduite, et doivent donc être minimisés. Outre les sillons de repositionnement en fin d'horizon temporel possibles grâce aux arcs de A^W , cela concerne aussi les trains de demande nulle, souvent présents dans les instances réelles, et qui peuvent être assimilés à des sillons de repositionnement. De plus, il faut également considérer les couples de trains couverts par une même unité mais où un sillon de repositionnement doit être effectué dans l'intervalle. On cherche donc à minimiser l'objectif suivant :

$$W(R) := \sum_{m \in \mathcal{M}} \left(\sum_{(i,j) \in A^W} Distance(Gare(i), Gare(j)) \cdot x_{ijm} \right. \\ \left. + \sum_{\substack{i \in T^{train} \\ D_i=0}} \sum_{j \in T} Distance(GDep(i), GArr(i)) \cdot x_{ijm} \right. \\ \left. + \sum_{(i,j) \in A^{Succ}} Distance(GArr(i), GDep(j)) \cdot x_{ijm} \right) \quad (3.9)$$

- **Kilomètres en surcomposition**

Lorsqu'un train circule en UM alors qu'un train US était demandé, la demande est satisfaite, mais on dit que ce train est en surcomposition. Plus généralement, une unité en surcomposition sur un train $i \in T^{train}$ avec $D_i > 0$ est une unité facultative pour ce train, au sens où la demande D_i serait quand même respectée si cette unité n'était pas affectée au train. Une unité en surcomposition est préférable à une unité en sillon W, car elle ne nécessite pas d'agent de conduite supplémentaire, mais l'énergie supplémentaire consommée est néanmoins importante². On cherche donc à minimiser les surcompositions. Comme certains trains

2. Sur un périmètre TGV, une unité en surcomposition dispose de ses propres moyens de traction, telle une rame automotrice.

3.3. LE PROBLÈME DE CONCEPTION D'UN ROULEMENT MATÉRIEL NOMINAL (PROBLÈME NOMINAL)

sont potentiellement couverts par plusieurs unités de séries matérielles différentes couplées entre elles, et que ces unités peuvent avoir un nombre de places assises (ou une puissance de traction) différent(e)s, il est difficile d'exprimer le nombre d'unités en surcomposition de façon linéaire. On préférera donc minimiser, de façon similaire, le nombre de places vides (ou la puissance excédentaire disponible) dans la composition affectée aux trains de demande strictement positives, pondéré par les kilomètres parcourus. On supposera dans la suite qu'on se situe dans le cas du transport de voyageurs, mais le modèle peut très bien s'adapter aux transport de marchandises. La donnée P_m (resp. e_i) ne représente alors plus le nombre de places de la série matérielle $m \in \mathcal{M}$ (resp. le nombre de places manquantes), mais sa puissance de traction (resp. la puissance de traction manquante). On cherche donc à minimiser :

$$Surcompos(R) := \sum_{\substack{i \in T^{train} \\ D_i > 0}} Distance(GDep(i), GArr(i)) \left(\sum_{j \in T} \sum_{m \in \mathcal{M}} P_m \cdot x_{ijm} + e_i - D_i \cdot t_i \right) \quad (3.10)$$

• Coupes et accroches

Le dernier critère que l'on cherche à minimiser est le nombre de coupes et d'accroches dans le roulement matériel obtenu, dans la mesure où elles impliquent des opérations manuelles supplémentaires et ont tendance à rendre la planification moins robuste en opérationnel. Si un train $i \in T^{train}$ est couvert, le nombre de coupes à son arrivée est donné par $\sum_{j \in T} \delta'_{ij} - 1$, et le nombre d'accroches à son départ est donné par $\sum_{j \in T} \delta'_{ji} - 1$. On cherche donc à minimiser :

$$CA(R) := \sum_{i \in T^{train}} \left(\sum_{j \in T} (\delta'_{ij} + \delta'_{ji}) - 2t_i \right) \quad (3.11)$$

Gestion du caractère multiobjectif

Dans le Chapitre ??, nous allons discuter de l'optimisation de ces différents critères de plusieurs façons différentes :

- agrégation multiobjectif, où l'on résout un unique PLNE en optimisant une somme pondérée des critères précédents;
- résolution par ordre lexicographique au moyen de plusieurs PLNE successifs : après avoir optimisé un critère, on adjoint au modèle une contrainte pour imposer la valeur optimale ou quasi-optimale de ce critère, puis on résout à ce nouveau modèle en optimisant le critère suivant;
- compromis entre agrégation et résolution par PLNE successifs : certains critères sont agrégés et on optimise ces nouveaux critères par résolutions successives.

Elle consomme presque autant d'énergie qu'une rame avec voyageurs, car elle doit pouvoir tracter toutes les voitures de la rame, qui sont néanmoins vides de voyageurs. Sur le périmètre Fret, une unité en surcomposition est une locomotive dont la puissance de traction peut ne pas être utilisée, ou seulement partiellement, dans la mesure où celles des locomotives principales peut suffire. On dit alors que la locomotive circule « comme un véhicule », car elle est assimilable à un wagon, sa puissance de traction étant peu ou pas utilisée.

3.3. LE PROBLÈME DE CONCEPTION D'UN ROULEMENT MATÉRIEL NOMINAL (PROBLÈME NOMINAL)

3.3.3.4 Contraintes

Les contraintes du modèle PLNE sont précisées sur la page ci-après, avec la notation supplémentaire

$$X_{ijm}^{max} = \begin{cases} 0 & \text{si } (i, j) \notin A \text{ ou } m \notin \mathcal{M}_i \text{ ou } m \notin \mathcal{M}_j \\ \min(D_i^{max}, D_j^{max}, D_{im}^{max}, D_{jm}^{max}, K_m) & \text{sinon} \end{cases}$$

Le premier bloc de contraintes (3.12)–(3.17) permet de définir la structure globale du multiflot donné par les variables x et de lier ces variables aux variables t . Les contraintes (3.12) assurent la conservation du flot en chaque nœud. Les contraintes (3.13) assurent d'une part que $t_i = 1$ si et seulement si le train i est effectivement couvert par une rame, conformément à l'identité (3.1), et d'autre part que la capacité globale D_i^{max} est respectée. Si le train i est couvert, les contraintes (3.14) imposent une capacité maximale pour chaque type de série matérielle et les contraintes (3.15) imposent le respect de la demande D_i en nombre de places avec une tolérance d'au plus E_i^{max} . Les contraintes (3.16) limitent le nombre d'unités à disposition pour chaque série matérielle. Les contraintes (3.17) imposent la cyclicité du roulement matériel obtenu, où l'inventaire de chaque série matérielle doit être identique en chaque gare entre le début et la fin de l'horizon temporel. Ces contraintes ne sont pas systématiquement ajoutées, car la cyclicité d'un roulement matériel n'est pas forcément souhaitée par certaines entités lors de certaines phases de la planification.

Le second bloc de contraintes (3.18)–(3.25) impose les liens entre les différentes variables du modèle, de telle sorte à fixer les composantes de e , δ et δ' en fonction de celles de x et t . Les contraintes (3.18) assurent que e_i corresponde à l'écart à la demande tel qu'il a été défini en (3.2)³. Les contraintes (3.19) fixent les variables δ conformément à l'identité (3.3), et les contraintes (3.20) – (3.21) fixent δ' conformément à l'identité (3.4). Enfin, les contraintes (3.22) – (3.25) renforcent la formulation, en précisant les liens entre t et δ' exhibés en (3.5). Elles ne sont pas indispensables pour résoudre le modèle PLNE, mais nous les avons inclus dans tous nos résultats expérimentaux Chapitre ?? pour obtenir de meilleures relaxations à la racine des arbres de recherche lors de la résolution des PLNE.

Les séries de contraintes du troisième bloc (3.26)–(3.29) correspondent à des restrictions opérationnelles. Les contraintes (3.26) et (3.27) tiennent compte de l'incompatibilité de certaines séries matérielles entre elles, qui ne peuvent être simultanément utilisées pour un même train. Les contraintes (3.28) et (3.29) interdisent respectivement les coupes et les accroches en certaines gares.

Le quatrième bloc de contraintes (3.30)–(3.34) décrit le domaine d'admissibilité de chacune des variables du modèle. Il est à noter que si $X_{ijm}^{max} = 0$, alors $x_{ijm} = 0$ par (3.30), et (3.19) impose immédiatement $\delta_{ijm} = 0$. De même, si $(i, j) \notin A$, $X_{ijm}^{max} = 0$ pour tout $m \in \mathcal{M}$, (3.19) et (3.21) imposent respectivement $\delta_{ijm} = 0$ pour tout $m \in \mathcal{M}$ et $\delta'_{ij} = 0$.

Dans les instances industrielles, on se trouve généralement dans l'un des cas suivant :

3. Dans la mesure où toutes ces variables e_i sont minimisées dans la fonction objectif, il n'est pas nécessaire d'ajouter les contraintes $e_i \leq E_i^{max} \cdot t_i$.

3.3. LE PROBLÈME DE CONCEPTION D'UN ROULEMENT MATÉRIEL NOMINAL
(PROBLÈME NOMINAL)

$$\sum_{i \in T} x_{ihm} = \sum_{j \in T} x_{hjm} \quad h \in T, m \in \mathcal{M} \quad (3.12)$$

$$t_i \leq \sum_{j \in T} \sum_{m \in \mathcal{M}} x_{ijm} \leq D_i^{max} \cdot t_i \quad i \in T^{train} \quad (3.13)$$

$$\sum_{j \in T} x_{ijm} \leq D_{im}^{max} \cdot t_i \quad i \in T^{train}, m \in \mathcal{M} \quad (3.14)$$

$$\sum_{j \in T} \sum_{m \in \mathcal{M}} P_m \cdot x_{ijm} \geq (D_i - E_i^{max}) \cdot t_i \quad i \in T^{train} \quad (3.15)$$

$$x_{\omega\alpha m} \leq K_m \quad m \in \mathcal{M} \quad (3.16)$$

$$x_{\alpha g^0 m} = x_{g^1 \omega m} \quad g \in \mathcal{G}, m \in \mathcal{M}, \quad (3.17)$$

$$e_i \geq D_i \cdot t_i - \sum_{j \in T} \sum_{m \in \mathcal{M}} P_m \cdot x_{ijm} \quad i \in T^{train} \quad (3.18)$$

$$\delta_{ijm} \leq x_{ijm} \leq X_{ijm}^{max} \cdot \delta_{ijm} \quad i \in T, j \in T, m \in \mathcal{M} \quad (3.19)$$

$$\delta'_{ij} \geq \delta_{ijm} \quad i \in T, j \in T, m \in \mathcal{M} \quad (3.20)$$

$$\delta'_{ij} \leq \sum_{m \in \mathcal{M}} \delta_{ijm} \quad i \in T, j \in T \quad (3.21)$$

$$t_i \geq \delta'_{ij} \quad i \in T^{train}, j \in T \quad (3.22)$$

$$t_i \geq \delta'_{ji} \quad i \in T^{train}, j \in T \quad (3.23)$$

$$t_i \leq \sum_{j \in T} \delta'_{ij} \quad i \in T^{train} \quad (3.24)$$

$$t_i \leq \sum_{j \in T} \delta'_{ji} \quad i \in T^{train} \quad (3.25)$$

$$\delta_{ijm} + \delta_{ij'm'} \leq 1 + Comp(m, m') \quad m, m' \in \mathcal{M}, i \in T^{train}, j \in T, j' \in T \quad (3.26)$$

$$\delta_{jim} + \delta_{j'im'} \leq 1 + Comp(m, m') \quad m, m' \in \mathcal{M}, i \in T^{train}, j \in T, j' \in T \quad (3.27)$$

$$\sum_{j \in T} \delta'_{ij} = t_i \quad i \in T^{train} : CA(G^{arr}(i)) = 0 \quad (3.28)$$

$$\sum_{j \in T} \delta'_{ji} = t_i \quad i \in T^{train} : CA(G^{dep}(i)) = 0 \quad (3.29)$$

$$x_{ijm} \in \llbracket 0; X_{ijm}^{max} \rrbracket \quad i \in T, j \in T, m \in \mathcal{M} \quad (3.30)$$

$$\delta_{ijm} \in \{0; 1\} \quad i \in T, j \in T, m \in \mathcal{M} \quad (3.31)$$

$$\delta'_{ij} \in \{0; 1\} \quad i \in T, j \in T \quad (3.32)$$

$$e_i \in \llbracket 0; E_i^{max} \rrbracket \quad i \in T^{train} \quad (3.33)$$

$$t_i \in \{0; 1\} \quad i \in T^{train} \quad (3.34)$$

3.4. PARAMÉTRISATION ET RAFFINEMENTS DU MODÈLE

- Les capacités globales D_i^{max} pour chaque train $i \in T$ ne sont pas définies, ou de manière équivalente valent $+\infty$, mais les capacités D_{im}^{max} propres à chaque série matérielle $m \in \mathcal{M}_i$ sont définies. On peut alors poser :

$$D_i^{max} := \sum_{m \in \mathcal{M}_i} D_{im}^{max}$$

- Les capacités D_{im}^{max} pour chaque train $i \in T$ et chaque série matérielle $m \in \mathcal{M}_i$ ne sont pas définies, ou de manière équivalente valent $+\infty$, mais les capacités globales D_i^{max} sont définies. On peut alors poser :

$$D_{im}^{max} := D_i^{max}$$

3.3.4 Génération des lignes de roulement

Le programme linéaire précédent permet de trouver une solution de flot x pour le graphe en attribuant un flux pour chaque arc. Pour parvenir à un roulement matériel complet \mathcal{R} , comme pour PRESTO et HERMES, il faut pouvoir décomposer ce flot en chemins, et chaque chemin correspond alors à une ligne de roulement.

La décomposition en chemins s'effectue pour l'instant de manière arbitraire, car les objectifs à atteindre diffèrent assez largement selon les entités : si certaines cherchent à homogénéiser les lignes de roulements en terme de nombre de trains, d'autres préfèrent au contraire avoir des lignes déséquilibrées. Nous ne nous sommes donc pas intéressés à ce problème afin que notre modèle puisse s'adapter à différents périmètres métiers.

3.4 Paramétrisation et raffinements du modèle

Nous introduisons dans cette section certains paramètres régissant le modèle, qui permettent notamment de mieux maîtriser le nombre d'arcs dans le graphe.

3.4.1 Durée maximale entre deux trains

Pour qu'un arc de A^{succ} relie deux sommets-trains, il est nécessaire que la gare d'arrivée du premier soit la gare de départ du second, ou qu'un sillon de repositionnement soit réalisable entre ces deux gares dans l'intervalle. Ces conditions sont *a priori* suffisantes pour caractériser tous les arcs possibles. Cependant, sur des solutions réelles, il est très rare que les arcs reliant deux sommets-trains séparés par plus d'un jour soient effectivement utilisés : par exemple, sur les roulements Fret, 97 % des arcs utilisés concernent deux trains circulant avec moins de 24h de différence, comme on peut le voir sur la Figure 3.8.

Nous avons donc choisi de n'introduire qu'une partie des arcs de A^{succ} dans notre modèle, en paramétrant la durée maximale entre deux trains. La valeur de 24 heures est choisie par défaut.

3.4. PARAMÉTRISATION ET RAFFINEMENTS DU MODÈLE

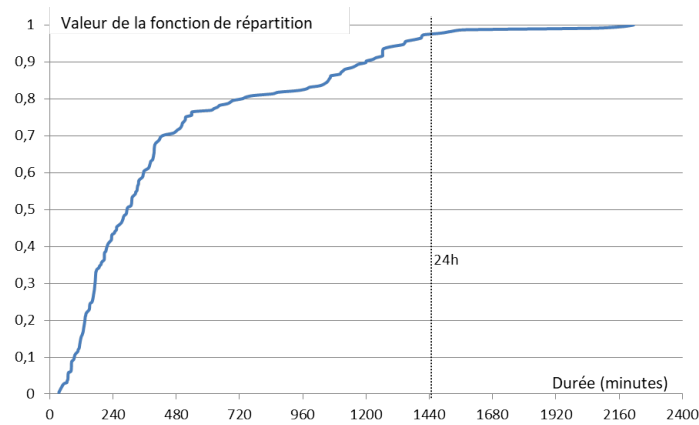


FIGURE 3.8: Fonction de répartition de la durée entre deux trains consécutifs sur une même ligne de roulement, sur une instance Fret.

3.4.2 Ajout de sommets-garages supplémentaires

Pour les entités transportant des voyageurs, hors Intercités de nuit et situations de perturbations majeures, il existe toujours un créneau horaire en milieu de nuit où aucun train ne circule. Il s'agit généralement de la tranche horaire entre 3 et 5 heures du matin.

Afin de réduire le nombre d'arcs dans le graphe, on peut introduire des sommets-garages supplémentaires chaque nuit entre deux jours consécutifs, pour chaque gare, un exemple étant représenté sur la Figure 3.9.

Ainsi, un sommet-train arrivant à la gare $g \in \mathcal{G}$ le lundi à 23h59 ne sera pas relié au sommet-train partant de $g' \in \mathcal{G}$ le mardi à 6h00 : en revanche, il sera relié au sommet-garage correspondant à la gare g la nuit du lundi au mardi, et le sommet-garage correspondant à g' sera relié au second train partant à 6h00. De plus, si $g \neq g'$, il faut pouvoir modéliser le sillon de repositionnement entre ces deux gares, ce qui se fait aisément au moyen d'un arc reliant les deux sommets-garages précédemment évoqués. L'ensemble des sommets-garages correspondant à chaque nuit forme donc une clique dans le graphe, dans le cas où tous les sillons de repositionnement sont autorisés et réalisables dans l'intervalle de temps de la nuit. Un tel choix de modélisation a deux conséquences.

1. Diminution du nombre d'arcs dans le graphe G .

Si n trains circulent chaque jour, la modélisation avec sommets-garages intermédiaires implique un nombre d'arcs de l'ordre de $2n + |\mathcal{G}|^2$ pour représenter les enchaînements de trains entre deux jours consécutifs sur une même ligne de roulement. Dans la modélisation précédente, en supposant que tous les sommets-trains d'un jour donné sont reliés aux sommets-trains du jour suivant (ce qui correspond à une durée maximale entre deux trains de l'ordre de 48 heures), on a un nombre d'arcs de l'ordre de n^2 . Comme n est grand devant $|\mathcal{G}|$, le nombre d'arcs diminue nettement avec les sommets-garages supplémentaires, et le nombre de variables et contraintes du modèle est donc significative-

3.4. PARAMÉTRISATION ET RAFFINEMENTS DU MODÈLE

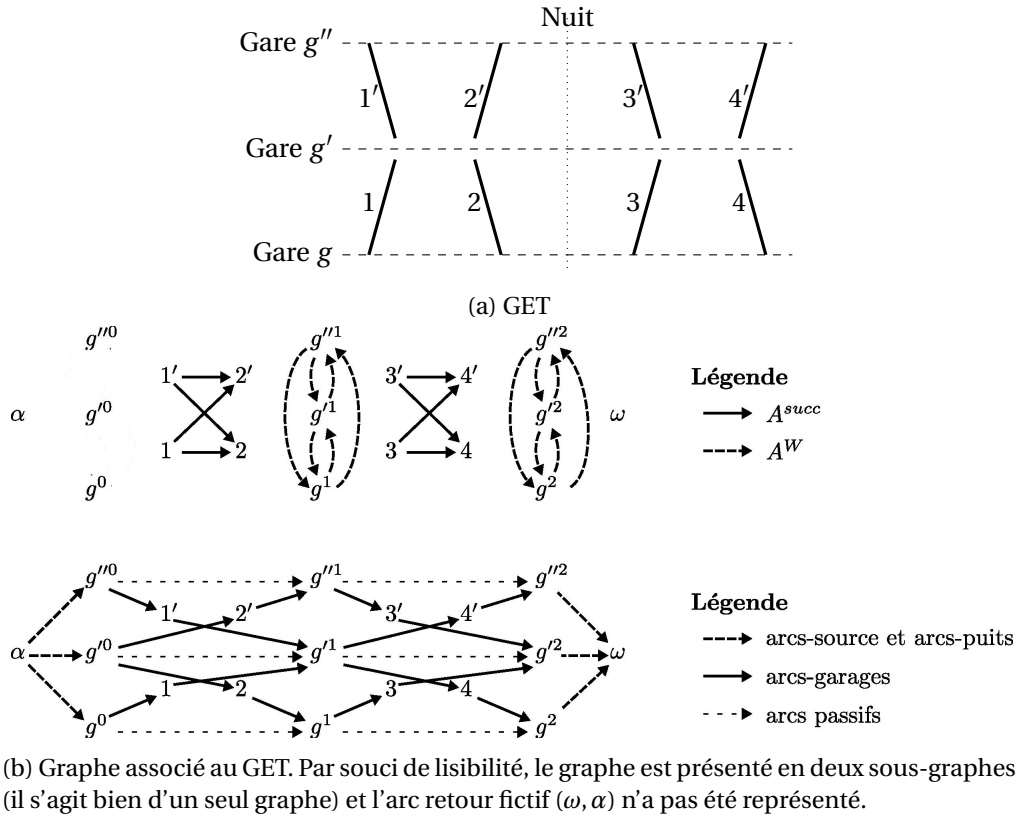


FIGURE 3.9: Exemple d'un graphe de précédence avec 8 trains et 3 gares, illustrant les sommets-garages intermédiaires

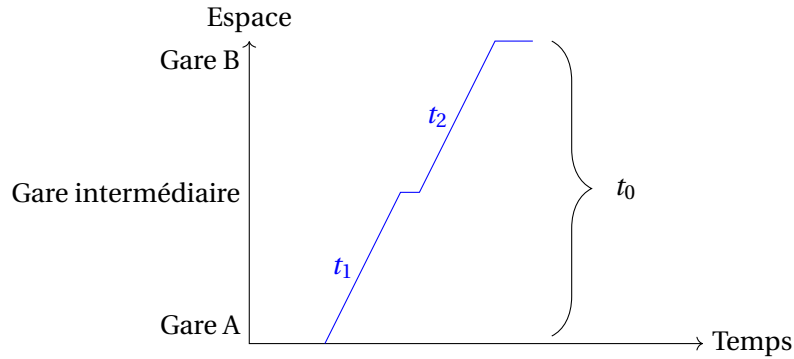
ment plus faible.

2. Solution de flot du PLNE moins précise, au sens où elle correspond à davantage de décompositions en chemins et donc davantage de roulements matériels possibles.

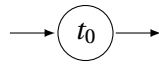
Elle ne permet plus de savoir quels sont les prochains trains que vont couvrir des unités présentes en une même gare $g \in \mathcal{G}$ à la fin d'une journée. Cette décision est alors reportée en post-traitement, au moment de la génération des lignes de roulement.

Avec ces sommets-garages supplémentaires, un arc (i, j) reliant deux trains i et j n'existera que si i et j circulent le même jour.

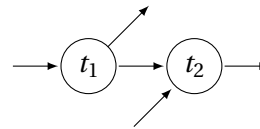
Pour modéliser la couverture successive de deux trains ayant des jours de circulation consécutifs par une même unité, il convient de passer par le sommet-garage exposé précédemment. Si ces jours de circulation ne sont pas consécutifs, le modèle peut être étendu pour modéliser la couverture successive de ces trains par une même unité : il suffit de créer des arcs supplémentaires appelés *arcs passifs* reliant chaque sommet-garage de la gare $g \in \mathcal{G}$ à la fin du jour n au sommet-garage de la même gare g à la fin du jour $n + 1$. Le flux sur cet arc indique alors le nombre d'unités qui restent en la gare g durant le jour $n + 1$, sans couvrir aucun train : un exemple est représenté sur la Figure 3.9.



(a) Graphique Espace-Temps d'un train t_0 composé de deux sous-trains t_1 et t_2



(b) Sous-graphe correspondant sans coupe/accroche possible en la gare intermédiaire



(c) Sous-graphe correspondant avec coupe/accroches possibles en la gare intermédiaire, où des contraintes de continuité doivent être ajoutées entre t_1 et t_2

FIGURE 3.10: Exemple d'utilisation des contraintes de continuité

3.4.3 Contraintes de continuités

Pour certaines entités, certains couples de trains doivent être couverts par au moins une unité en commun. C'est par exemple le cas lorsqu'on veut pouvoir anticiper la planification des agents, ou lorsqu'un train peut nécessiter des coupes ou accroches sur une gare intermédiaire de son parcours. Dans ce dernier cas, il convient alors de définir les gares intermédiaires où telle opération est possible, et de remplacer le train initial par un ensemble de sous-trains liés par une contrainte de continuité (voir Figure 3.10).

En notant $Continuites \subset T^{train} \times T^{train}$ les couples de trains devant être liés par une contrainte de continuité, ces contraintes peuvent facilement être ajoutées au modèle PLNE de la façon suivante :

$$\delta'_{ij} = t_i = t_j \quad (i, j) \in Continuites \quad (3.35)$$

3.4.4 Durée maximale d'attente en gare

Certaines instances Fret ont un paramètre indiquant la durée maximale d'attente d'une unité en chaque gare, qui est généralement de l'ordre de quelques heures. Si cette durée est dépassée, les unités doivent être acheminées vers un lieu de dépôt.

1. Si la gare $g \in \mathcal{G}$ en question est elle-même un dépôt, cela implique une manœuvre supplémentaire en gare et on considère qu'elle est réalisable.

2. Sinon, les unités en question doivent être entreposées dans le dépôt d'une autre gare $g' \in \mathcal{G}$. Il convient alors de générer un premier sillon W de g vers g' , puis un second sillon W de g' vers g .

Pour modéliser ces situations, on peut déterminer en pré-traitement les arcs $(i, j) \in A$ concernés. On peut alors ajouter un objectif pour minimiser les manœuvres à réaliser en gare (cas 1), et modifier le critère $W(R)$ pour prendre en compte le cas 2. Dans les expérimentations du Chapitre ??, nous n'avons pas introduit de coût pour réaliser les manœuvres du cas 1, et avons seulement modifié notre modèle pour prendre en compte le cas 2 : cela se fait simplement en modifiant la fonction objectif associée aux sillons de repositionnement ($W(R)$) : si une unité ayant effectué le train i de destination g doit passer par la gare de dépôt g^D lorsqu'elle réalise ensuite le train j , on ajoute dans la fonction objectif W l'expression

$$\sum_{m \in \mathcal{M}} 2Distance(g, g^D) \cdot x_{ijm}. \quad (3.36)$$

Remarquons que ces considérations peuvent s'étendre aux arcs $(i, j) \in A$ entre deux trains i et j nécessitant un sillon W entre l'arrivée de i en la gare g et le départ de j en la gare $g' \neq g$. En effet, la durée maximale d'attente dans l'une de ces deux gares peut être dépassée. Il convient alors de trouver une gare de dépôt g^D , puis de générer un premier sillon W de g vers g^D puis un second de g^D vers g' . On modifie alors à nouveau le critère $W(R)$ en ajoutant l'expression

$$\sum_{m \in \mathcal{M}} (Distance(g, g^D) + Distance(g^D, g')) \cdot x_{ijm}. \quad (3.37)$$

Dans le cas où on ajoute des sommets-garages supplémentaires, seules les gares ayant un dépôt doivent être considérées.

Chapitre 4

Conclusion et perspectives

Cette thèse a permis de mener des travaux exploratoires autour de ce que pourrait être une planification adaptative plus agile que celle actuellement en vigueur dans le système ferroviaire. Menés à la Direction Innovation & Recherche au sein de SNCF, ils ont abouti à l'introduction de plusieurs concepts permettant de corriger certains défauts de l'hyperplanification du plan de transport ferroviaire.

Au cours du Chapitre 1, nous avons introduit les principaux éléments du système ferroviaire en France, notamment le concept de planification. Le constat initial de nos travaux repose sur l'apparente hyperplanification du plan de transport : le processus de planification tel qu'il existe actuellement est jugé trop rigide, et on se demande s'il est possible d'adopter un nouveau paradigme rendant la planification plus agile et le plan de transport plus flexible. L'introduction de flexibilité dans le processus s'est révélée industriellement complexe à mettre en œuvre, et nous avons su élargir la problématique initiale pour nous concentrer plus spécifiquement sur l'anticipation des besoins d'adaptation. C'est le sens que nous avons donné à la planification adaptative au cours de nos travaux.

Le Chapitre 2 a permis de réaliser un état de l'art de la recherche scientifique sur la planification ferroviaire. Les différents modèles et problèmes d'optimisation qui interviennent sont de nature variée et permettent d'entrevoir la difficulté industrielle pour optimiser un plan de transport, tant ils sont difficiles à résoudre et dépendent les uns des autres. Par ailleurs, ce chapitre a également été l'occasion de discuter de la notion de flexibilité en recherche opérationnelle au travers de différents problèmes de planification non nécessairement appliqués au domaine ferroviaire. Nous avons notamment pu exhiber certains articles introduisant explicitement de la flexibilité dans la planification des sillons, et nous nous sommes intéressés aux contextes décisionnels faisant intervenir plusieurs étapes et niveaux de décisions. Tout modèle de planification rentre en effet dans cette catégorie, puisqu'on doit planifier en conception, puis adapter ou réoptimiser le plan de transport ainsi conçu durant les phases de la planification ferroviaire.

À partir du Chapitre 3, nous nous sommes essentiellement concentrés sur la planification du matériel roulant. Nous avons réalisé un état de l'art de la littérature scientifique et des outils industriels à la SNCF concernant la planification de cette ressource. Le Problème de Conception d'un Roulement Matériel Nomi-

nal a ensuite été introduit, essentiellement constitué d'un PLNE s'appuyant sur un graphe d'adjacence que nous avons défini. Un roulement matériel peut alors être assimilé à un flot décomposé en chemins sur ce graphe. Nous avons montré que ce modèle pouvait facilement s'étendre pour prendre en compte différentes spécificités et contraintes industrielles liées à la planification du matériel roulant.

C'est au cours du Chapitre ?? que nous nous sommes intéressés à l'adaptation du matériel roulant et que nous avons défini formellement la planification adaptative. Après un état de l'art sur l'adaptation du matériel roulant, nous avons introduit la notion de coûts d'adaptation d'un plan de transport, qui permet de juger de la qualité d'une adaptation. Plus particulièrement et plus formellement, nous avons défini des coûts d'adaptation pour le matériel roulant. Nous avons alors introduit le Problème d'Adaptation d'un Roulement Matériel, qui consiste à adapter un roulement matériel existant à un scénario d'adaptation en minimisant les coûts d'adaptation. Un PLNE de taille comparable à celui du Chapitre 3 permet de trouver un tel roulement matériel sous la forme d'un flot, et nous proposons deux méthodes pour le décomposer en chemins : un algorithme glouton ainsi qu'un programme mathématique avec objectif quadratique. Enfin, nous avons défini le Problème de Conception d'un Roulement Matériel Adaptatif, qui permet de réaliser la planification adaptative du matériel roulant en phase de conception. Le PLNE que nous avons introduit est de taille plus importante et permet d'anticiper différents scénarios d'adaptation dès la phase de conception, en minimisant les futurs coûts d'adaptation attendus.

Le Chapitre ?? est consacré aux expérimentations numériques et à leur analyse. Nous y avons présenté sept instances réelles constituant notre jeu de tests, et nous avons discuté des différentes façons d'obtenir des scénarios pertinents pour l'adaptation d'un roulement matériel. Les trois problèmes introduits précédemment ont été successivement appliqués sur chacune de nos instances, et nous avons analysé nos résultats tant d'un point de vue de la performance de la résolution que de celui de la performance et des caractéristiques des solutions. Nous avons notamment décrit une procédure pour évaluer l'intérêt de la planification adaptative par rapport au processus actuellement en vigueur, en mettant en évidence les gains de cette nouvelle façon de planifier. Bien que les temps de résolution de l'approche par planification adaptative soient plus importants, nos modèles ont permis de montrer que certains roulements matériels actuellement utilisés en conception pourraient être améliorés. Sans aucune concession sur leur performance en termes d'utilisation de ressources, il est parfois possible de les rendre plus adaptatifs aux demandes de changements survenant en phase d'adaptation.

Pour poursuivre nos travaux, plusieurs pistes peuvent être envisagées.

1. *Cas d'usage industriel spécifique*

La planification adaptative telle que nous l'avons définie a permis d'engager la discussion avec certains opérateurs métiers. Ils ont pu prendre conscience de l'intérêt des concepts que nous avons introduits pour mieux rationaliser l'utilisation de leurs ressources. C'est notamment le cas de SNCF Fret, pour qui la phase d'adaptation présente des échelles de temps et des contraintes très différentes de celles relatives au transport de voyageurs. Nos modèles actuels pour la planification du matériel rou-

lant sont de nature assez générique et peuvent *a priori* s'adapter à plusieurs entreprises ferroviaires : il est nécessaire de prendre en compte certaines de leurs spécificités pour que les roulements matériels que nous proposons avec la planification adaptative puissent être intégrés dans les outils industriels. Outre ces contraintes spécifiques, les changements dans les successions entre trains ne sont peut-être pas les seuls coûts d'adaptation structurels pour le matériel roulant. Par exemple, le maintien de l'affectation de certaines rames à certains trains importants doit peut-être être considéré. De façon similaire, le maintien de certaines successions stratégiques ou de certains sillons W commandés à l'avance peuvent être des critères à considérer pour modéliser plus finement les coûts d'adaptation structurels. La pertinence de telles extensions dépend fondamentalement de l'entité ferroviaire considérée, et doit être préalablement discutée de manière approfondie dans le cadre d'une collaboration industrielle.

Une réflexion reste notamment à mener sur la façon la plus pertinente de générer des scénarios d'adaptation. La génération de scénarios aléatoires n'est pas satisfaisante sur le plan industriel, et il est difficile d'obtenir des scénarios d'adaptations réalistes en discutant avec les opérateurs métiers. Pour y parvenir, il faut pouvoir engager une discussion plus durable sur le long-terme avec ces entités métiers.

2. Généralisations à d'autres ressources ou domaines

Au cours de nos travaux, nous avons élaboré des modèles spécifiquement dédiés à la planification et l'adaptation du matériel roulant. Cependant, il s'agit bien là d'un exemple de ce que nous proposons avec la démarche de planification adaptative, qui relève d'un concept plus global et peut être étendue à d'autres ressources durant d'autres phases de planification. Par exemple, il est tout à fait envisageable d'appliquer la planification adaptative à la planification des sillons ou des agents : il faut notamment pouvoir décrire les coûts d'adaptation correspondant à ces deux ressources. De même, on peut envisager des scénarios ne concernant non pas la phase d'adaptation mais la phase opérationnelle, où les aléas sont de nature différente.

Par ailleurs, les coûts d'adaptation structurels que nous avons introduits peuvent se généraliser indépendamment du contexte ferroviaire. Maintenir la structure d'un graphe peut être pertinent pour d'autres cas d'application. On peut notamment penser au domaine des télécommunications, ou plus généralement dès lors que l'on cherche à optimiser dans des réseaux.

3. Accélération de la résolution

Notre dernier modèle ne permet pas de résoudre de trop grosses instances industrielles, dès lors que le nombre de scénarios d'adaptation envisagés est important. Or, notre modèle PLNE semble bien s'adapter aux méthodes de décompositions-coordination. Entre autres, une relaxation lagrangienne des quelques contraintes couplantes du modèle pourrait permettre d'accélérer sa résolution. Nous avons tenté d'utiliser la décomposition de Benders pour améliorer la résolution de la relaxation continue, mais les résultats étaient mitigés, et nous n'avons pas pu expliquer pourquoi cela ralentissait certaines instances tandis que d'autres se résolvent plus rapidement. Une étude plus approfondie

des spécificités des instances pourrait permettre de mieux comprendre ces différences de comportement, et par la suite de développer des méthodes de décomposition *ad hoc*.

De plus, notre graphe d'adjacence rend possible de très nombreux enchaînements de trains, créant autant d'arcs et des variables et contraintes dans le modèle. En discutant avec les opérateurs métiers et en étudiant certaines instances, il serait peut-être possible de diminuer ce nombre d'arcs, ce qui permet de réduire significativement la taille de nos modèles et donc d'accélérer la résolution.

Bibliographie

- AHUJA, R. K., LIN, J., ORLIN, J. B., SHARMA, D., ET SHUGHART, L. 2005. Solving real-life locomotive-scheduling problems. *Transportation Science* 39 :503 – 517.
- AKKAN, C., KÜLÜNK, M. E., ET KOÇAŞ, C. 2016. Finding robust timetables for project presentations of student teams. *European Journal of Operational Research* 249 :560 – 576.
- ALLOULOU, M. A. ET ARTIGUES, C. 2010. Flexible solutions in disjunctive scheduling : General formulation and study of the flow-shop case. *Computers & Operations Research* 37 :890 – 898. Disruption Management.
- ALTAZIN, E., DAUZÈRE-PÉRÈS, S., RAMOND, F., ET TRÉFOND, S. 2017. Rescheduling through stop-skipping in dense railway systems. *Transportation Research Part C : Emerging Technologies* 79 :73 – 84.
- ARTIGUES, C., BILLAUT, J.-C., ET ESSWEIN, C. 2005. Maximization of solution flexibility for robust shop scheduling. *European Journal of Operational Research* 165 :314–328.
- BARRENA, E., CANCA, D., COELHO, L. C., ET LAPORTE, G. 2014. Exact formulations and algorithm for the train timetabling problem with dynamic demand. *Computers & Operations Research* 44 :66 – 74.
- BEŠINOVIĆ, N., GOVERDE, R. M. P., QUAGLIETTA, E., ET ROBERTI, R. 2016. An integrated micro–macro approach to robust railway timetabling. *Transportation Research Part B : Methodological* 87 :14 – 32.
- BEN-TAL, A., GORYASHKO, A., GUSLITZER, E., ET NEMIROVSKI, A. 2004. Adjustable Robust Solutions of Uncertain Linear Programs. *Mathematical Programming* 99 :351.
- BENDOTTI, P., CHRÉTIENNE, P., FOUILHOX, P., ET QUILLIOT, A. 2017. Anchored reactive and proactive solutions to the CPM-scheduling problem. *European Journal of Operational Research* 261 :67 – 74.
- BENHIZIA, F. 2012. Optimisation du plan de transport par planification intégrée des ressources. Thèse de doctorat, École Nationale Supérieure des Mines de Saint-Étienne et SNCF.
- BERTSIMAS, D. ET GOYAL, V. 2012. On the Power and Limitations of Affine Policies in Two-stage Adaptive Optimization. *Mathematical Programming* 134 :491–531.
- BERTSIMAS, D. ET SIM, M. 2004. The Price of Robustness. *Operations Research* 52 :35–53.

BIBLIOGRAPHIE

- BILLAUT, J.-C., MOUKRIM, A., ET SANLAVILLE, É. 2008. Flexibility and Robustness in Scheduling, chapitre Introduction. Wiley-ISTE.
- BORNDÖRFER, R. ET LIEBCHEN, C. 2008. When periodic timetables are suboptimal. *Dans* J. Kalcsics et S. Nickel (eds.), Operations Research Proceedings 2007 : Selected Papers of the Annual International Conference of the German Operations Research Society (GOR) Saarbrücken, September 5–7, 2007, pages 449–454, Berlin, Heidelberg. Springer Berlin Heidelberg.
- BORNDÖRFER, R., REUTHER, M., SCHLECHTE, T., ET WEIDER, S. 2011. A Hypergraph Model for Railway Vehicle Rotation Planning. *Dans* A. Caprara et S. Kontogiannis (eds.), 11th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems, volume 20 de *OpenAccess Series in Informatics (OASIs)*, pages 146–155, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- BORNDÖRFER, R., REUTHER, M., SCHLECHTE, T., WAAS, K., ET WEIDER, S. 2016. Integrated Optimization of Rolling Stock Rotations for Intercity Railways. *Transportation Science* 50 :863–877.
- BOUVAREL, P. 2011 - 2012. Cours de chemin de fer : présentation générale de l'exploitation. *École des Ponts*.
- BRIGHTWELL, G. ET WINKLER, P. 1991. Counting Linear Extensions is #P-complete. *Dans* Proceedings of the Twenty-third Annual ACM Symposium on Theory of Computing, STOC '91, pages 175–181, New York, NY, USA. ACM.
- BRÄNNLUND, U., LINDBERG, P. O., NÖU, A., ET NILSSON, J.-E. 1998. Railway timetabling using lagrangian relaxation. *Transportation Science* 32 :358–369.
- BRÉTHOMÉ, L. 2018. Passenger-oriented modelling and optimization of the railway transportation plan in a mass transit system. Thèse de doctorat, École Centrale de Lille et SNCF.
- CACCHIANI, V., CAPRARA, A., GALLI, L., KROON, L., MARÓTI, G., ET TOTH, P. 2012. Railway rolling stock planning : Robustness against large disruptions. *Transportation Science* 46 :217–232.
- CACCHIANI, V., CAPRARA, A., ET TOTH, P. 2010. Solving a real-world train-unit assignment problem. *Mathematical Programming* 124 :207–231.
- CACCHIANI, V., HUISMAN, D., KIDD, M., KROON, L., TOTH, P., VEELANTURE, L., ET WAGENAAR, J. 2014. An overview of recovery models and algorithms for real-time railway rescheduling. *Transportation Research Part B : Methodological* 63 :15–37.
- CACCHIANI, V. ET TOTH, P. 2012. Nominal and robust train timetabling problems. *European Journal of Operational Research* 219 :727 – 737. Feature Clusters.
- CAIMI, G., FUCHSBERGER, M., LAUMANN, M., ET SCHÜPBACH, K. 2011a. Periodic railway timetabling with event flexibility. *Networks* 57 :3–18.

BIBLIOGRAPHIE

- CAIMI, G., KROON, L., ET LIEBCHEN, C. 2017. Models for railway timetable optimization : Applicability and applications in practice. *Journal of Rail Transport Planning & Management* 6 :285 – 312.
- CAIMI, G., LAUMANN, M., SCHÜPBACH, K., WÖRNER, S., ET FUCHSBERGER, M. 2011b. The periodic service intention as a conceptual framework for generating timetables with partial periodicity. *Transportation Planning and Technology* 34 :323–339.
- CAPRARA, A., FISCHETTI, M., ET TOTH, P. 2002. Modeling and Solving the Train Timetabling Problem. *Operations Research* 50 :851–861.
- CAPRARA, A., GALLI, L., ET TOTH, P. 2007. Solution of the Train Platforming Problem. Dans C. Liebchen, R. K. Ahuja, et J. A. Mesa (eds.), 7th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS'07), volume 7 de *OpenAccess Series in Informatics (OASISs)*, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- CAREY, M. ET LOCKWOOD, D. 1995. A Model, Algorithms and Strategy for Train Pathing. *Journal of the Operational Research Society* 46 :988–1005.
- CICERONE, S., D'ANGELO, G., DI STEFANO, G., FRIGIONI, D., NAVARRA, A., SCHACHTEBECK, M., ET SCHÖBEL, A. 2009. Recoverable Robustness in Shunting and Timetabling, pages 28–60. Dans R. K. Ahuja, R. H. Möhring, et C. D. Zaroliagis (eds.), *Robust and Online Large-Scale Optimization*, volume 5868 de *Lecture Notes in Computer Science*. Springer.
- D'ARIANO, A., PACCIARELLI, D., ET PRANZO, M. 2008. Assessment of flexible timetables in real-time traffic management of a railway bottleneck. *Transportation Research Part C : Emerging Technologies* 16 :232 – 245.
- DEWILDE, T., SELS, P., CATTRYSSSE, D., ET VANSTEENWEGEN, P. 2013. Robust railway station planning : An interaction between routing, timetabling and platforming. *Journal of Rail Transport Planning & Management* 3 :68 – 77. Robust Rescheduling and Capacity Use.
- FIOOLE, P.-J., KROON, L., MARÓTI, G., ET SCHRIJVER, A. 2006. A rolling stock circulation model for combining and splitting of passenger trains. *European Journal of Operational Research* 174 :1281–1297.
- FORSGREN, M., ARONSSON, M., GESTRELIUS, S., ET DAHLBERG, H. 2012. Opportunities and challenges with new railway planning approach in Sweden. Technical Report T2012 :11, Swedish Institute of Computer Science (SICS).
- GESTRELIUS, S., ARONSSON, M., ET BOHLIN, M. 2015. On the uniqueness of operation days and delivery commitment generation for train timetables. Dans RailTokyo.
- GIACCO, G. L., D'ARIANO, A., ET PACCIARELLI, D. 2014. Rolling stock rostering optimization under maintenance constraints. *Journal of Intelligent Transportation Systems* 18 :95–105.

BIBLIOGRAPHIE

- GOERIGK, M., HESSE, S., MÜLLER-HANNEMANN, M., SCHMIDT, M., ET SCHÖBEL, A. 2013. Recoverable Robust Timetable Information. Dans D. Frigioni et S. Stiller (eds.), ATMOS - 13th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems - 2013, volume 33 de *OpenAccess Series in Informatics (OASICs)*, pages 1–14, Sophia Antipolis, France. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- GOERIGK, M. ET SCHÖBEL, A. 2014. Recovery-to-optimality : A new two-stage approach to robustness with an application to aperiodic timetabling. *Computers & Operations Research* 52 :1 – 15.
- HEIL, J., HOFFMANN, K., ET BUSCHER, U. 2020. Railway crew scheduling : Models, methods and applications. *European Journal of Operational Research* 283 :405 – 425.
- HEILPORN, G., DE GIOVANNI, L., ET LABBÉ, M. 2008. Optimization models for the single delay management problem in public transportation. *European Journal of Operational Research* 189 :762–774. SCOPUS : ar.j.
- KANAI, S., SHIINA, K., HARADA, S., ET TOMII, N. 2011. An optimal delay management algorithm from passengers' viewpoints considering the whole railway network. *Journal of Rail Transport Planning & Management* 1 :25 – 37. Robust Modelling of Capacity, Delays and Rescheduling in Regional Networks.
- LIDÉN, T. 2015. Railway Infrastructure Maintenance - A Survey of Planning Problems and Conducted Research. *Transportation Research Procedia* 10 :574 – 583. 18th Euro Working Group on Transportation, EWGT 2015, 14-16 July 2015, Delft, The Netherlands.
- LIEBCHEN, C., LÜBBECKE, M., MÖHRING, R., ET STILLER, S. 2009. The concept of recoverable robustness, linear programming recovery, and railway applications, pages 1–27. Dans R. K. Ahuja, R. H. Möhring, et C. D. Zaroliagis (eds.), *Robust and Online Large-Scale Optimization : Models and Techniques for Transportation Systems*, chapitre 1. Springer Berlin Heidelberg, Berlin, Heidelberg.
- LIEBCHEN, C. ET PEETERS, L. 2009. Integral cycle bases for cyclic timetabling. *Discrete Optimization* 6 :98 – 109.
- MARCOS, N. 2006. Modélisation et optimisation de la gestion du matériel roulant à la SNCF. Thèse de doctorat, Université Paris 13 et SNCF.
- MARCOS, N., DE ALMEIDA, D., GAUYACQ, D., ALFANDARI, L., NAGIH, A., ET PLATEAU, G. 2005. Une approche de modélisation générique pour la gestion des locomotives fret à la SNCF. Dans ROADEF 2005, pages –, Tours, France.
- MASCIS, A. ET PACCIARELLI, D. 2002. Job-shop scheduling with blocking and no-wait constraints. *European Journal of Operational Research* 143 :498 – 517.
- NASRABADI, E. ET ORLIN, J. B. 2013. Robust optimization with incremental recourse. *CoRR* abs/1312.4075.

BIBLIOGRAPHIE

- NIELSEN, L. K., KROON, L., ET GÁBOR, M. 2012. A rolling horizon approach for disruption management of railway rolling stock. *European Journal of Operational Research* 220 :496 – 509.
- PEETERS, M. ET KROON, L. 2008. Circulation of railway rolling stock : a branch-and-price approach. *Computers Operations Research* 35 :538 – 556. Part Special Issue : Location Modeling Dedicated to the memory of Charles S. ReVelle.
- REUTHER, M., BORNDORFER, R., SCHLECHTE, T., ET WEIDER, S. 2012. Integrated optimization of rolling stock rotations for intercity railways. Technical Report 12-39, ZIB, Takustr. 7, 14195 Berlin.
- ROY, B. ET SUSSMANN, B. 1964. Les problèmes d’ordonnancement avec contraintes disjonctives. *Note ds* 9.
- SAMÀ, M., D’ARIANO, A., CORMAN, F., ET PACCIARELLI, D. 2017. A variable neighbourhood search for fast train scheduling and routing during disturbed railway traffic situations. *Computers & Operations Research* 78 :480 – 499.
- SCHÖBEL, A. 2012. Line planning in public transportation : models and methods. *OR Spectrum* 34 :491 – 510.
- SERAFINI, P. ET UKOVICH, W. 1989. A Mathematical Model for Periodic Scheduling Problems. *SIAM J. Discret. Math.* 2 :550–581.
- SHAPIRO, A. ET PHILPOTT, A. 2007. A Tutorial on Stochastic Programming.
- TRÉFOND, S. 2014. Planification robuste des roulements d’engins dans le domaine ferroviaire. Thèse de doctorat, Cnam et SNCF.
- TRÉFOND, S., BILLIONNET, A., ELLOUMI, S., DJELLAB, H., ET GUYON, O. 2017. Optimization and simulation for robust railway rolling-stock planning. *Journal of Rail Transport Planning & Management* 7 :33–49.
- ULLRICH, O., LÜCKERATH, D., ET SPECKENMEYER, E. 2016. Do regular timetables help to reduce delays in tram networks? It depends! *Public Transport* 8 :39–56.
- WAGENAAR, J., KROON, L., ET FRAGKOS, I. 2017. Rolling stock rescheduling in passenger railway transportation using dead-heading trips and adjusted passenger demand. *Transportation Research Part B : Methodological* 101 :140–161.
- ZEHROUNI, A. 2016. Planification intégrée et robuste des ressources ferroviaires. Thèse de doctorat, École Nationale Supérieure des Mines de Saint-Étienne et SNCF.

Titre : Planification adaptative des ressources ferroviaires

Mots clés : Planification ferroviaire, Optimisation, Matériel roulant, Recherche Opérationnelle, Programmation Linéaire en Nombres Entiers

Résumé : La planification ferroviaire consiste à établir en amont des circulations un plan de transport spécifiant l'emploi du temps de chaque ressource : les sillons, qui correspondent aux horaires des trains, les matériels roulants devant assurer leur circulation, ainsi que les agents à bord de ces trains. Compte tenu de la complexité du système ferroviaire, cette planification est effectuée potentiellement plusieurs années à l'avance, où un plan de transport précis est établi. Or, une fois qu'une ressource a été planifiée, il est parfois nécessaire d'adapter le plan de transport et donc de le changer. Les raisons peuvent être multiples : changement horaire pour causes de travaux, infrastructure partiellement endommagée, demande d'une rame supplémentaire pour un train, avarie sur un type de matériel roulant. . . Ainsi, le plan de transport est parfois adapté à plusieurs reprises entre sa conception potentiellement plusieurs années à l'avance et son exécution le jour des circulations. Dès lors, il est légitime de se demander si la conception d'un plan de transport précis avec des échelles de temps aussi importantes est raisonnable.

Dans cette thèse, nous envisageons d'ajouter de la flexibilité dans le processus de planification, afin que les efforts déployés pour adapter le plan de transport lorsque cela est nécessaire soient réduits. Nous introduisons la notion de coûts d'adaptation du plan de transport, et nous proposons d'explicitier ces coûts d'adaptation pour le matériel roulant. Nous précisons notamment les coûts d'adaptation structurels, qui permettent de quantifier les similitudes entre deux planifications différentes du matériel roulant. La planification adaptative telle que nous l'envisageons consiste alors à anticiper dès la phase de conception les futures adaptations possibles du plan de transport. Plusieurs modèles de Programmation Linéaire en Nombres Entiers sont introduits et des résultats expérimentaux sont proposés sur des instances réelles issues de SNCF. L'approche de planification adaptative a pu être comparée à l'approche actuelle, et on montre que les coûts d'adaptation du plan de transport peuvent être effectivement réduits en adoptant cette nouvelle façon de planifier.

Title : Adaptive planning of railway resources

Keywords : Railway planning, Optimization, Rolling stock, Operations Research, Mixed Integer Programming

Abstract : Railway planning consists of drawing up a transportation plan before operational time, specifying a planning for each resource: the train paths, which correspond to the schedules, the rolling stock that will be used to move the trains, and the crew on board the trains. Given the complexity of the railway system, this transportation plan is carried out potentially several years in advance, when a precise transportation plan is drawn up. However, once a resource has been planned, it is sometimes necessary to adapt the transportation plan and therefore change it. There may be many reasons for this: schedule changes due to work, partially damaged infrastructure, request for an additional train set for a train, damage to a type of rolling stock, etc. Thus, the transportation plan may sometimes be adapted several times between its design, potentially several years in advance, and its execution on the day of traffic. It is therefore legitimate to ask whether the

design of a precise transport plan with such large time scales is reasonable. In this thesis, we consider the adding of flexibility to the planning process so that the effort to adapt the transportation plan when necessary is reduced. We introduce the notion of transportation plan adaptation costs, and we propose to make these adaptation costs explicit for the rolling stock resource. In particular, we specify the structural adaptation costs, which make it possible to quantify the similarities between two different rolling stock planings. Adaptive planning consists in anticipating possible future adaptations of the transport plan as early as the design phase. Several Integer Linear Programs are introduced and experimental results are proposed on real SNCF instances. The adaptive planning approach can be compared with the current approach, and it is shown that the adaptation costs can be effectively reduced by adopting this new way of planning.