



HAL
open science

Modélisation et optimisation de la gestion opérationnelle des circulations en cas d'aléas.

Laurent Gély

► **To cite this version:**

Laurent Gély. Modélisation et optimisation de la gestion opérationnelle des circulations en cas d'aléas..
Mathématiques [math]. Université Sciences et Technologies - Bordeaux I, 2010. Français. NNT : .
tel-00551419

HAL Id: tel-00551419

<https://theses.hal.science/tel-00551419>

Submitted on 7 Mar 2011

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.

N ° d'ordre : 2.1.0



THÈSE

présentée à



l'Université de Bordeaux
École doctorale de Mathématiques et Informatique

par
GÉLY Laurent

pour obtenir le grade de
Docteur

SPÉCIALITÉ MATHÉMATIQUES APPLIQUÉES

*Modélisation et Optimisation
de la Gestion Opérationnelle du Trafic Ferroviaire
en cas d'Aléas.*

Directeurs de recherche : PESNEAU Pierre et VANDERBECK François.

Soutenue le 14 décembre 2010 à l'IMB
devant la commission d'examen formée de :

M. MONGEAU Marcel	Professeur d'université	Université Paul Sabatier
M. FEILLET Dominique	Professeur d'université	École des Mines de Saint-Étienne
M. DESSAGNE Gilles	Docteur	Réseau Ferré de France
M. SOURD Francis	Docteur	Société Nationale des Chemins de Fer
M. PESNEAU Pierre	Maître de Conférence	Université Bordeaux I
M. VANDERBECK François	Professeur d'université	Université Bordeaux I
M. MILLER Andrew	Professeur d'université	Université Bordeaux I
M. PÊCHER Arnaud	Professeur d'université	IUT Bordeaux I

TABLE DES MATIÈRES

1	Preambule	7
2	Introduction	13
3	Problématique	17
3.1	Contexte	17
3.1.1	La production ferroviaire	17
3.1.2	La gestion opérationnelle des circulations	18
3.1.3	Notions ferroviaires élémentaires	20
3.2	Définition du projet	30
3.2.1	Cadre du projet	30
3.2.2	La modélisation du système ferroviaire	31
3.2.3	Formulations mathématiques et optimisation	33
3.2.4	Vers la mise en œuvre industrielle	34
4	État de l'art	37
4.1	Modélisation du système ferroviaire	38
4.1.1	Représentation métier du réseau ferré	38
4.1.2	Abstraction d'un modèle ferroviaire	38
4.2	Modèles mathématiques	40
4.2.1	Notations générales	41
4.2.2	Formulations en temps continu	41
4.2.3	Formulation en temps discret	45
4.3	Méthodes et algorithmes de résolution	51
4.3.1	Systèmes experts, systèmes de règles	51
4.3.2	Optimisation par des méthodes exactes	52
4.3.3	Optimisation par des méthodes heuristiques	54
4.4	Analyse	55
4.4.1	Modèles mathématiques	55
4.4.2	Algorithmes et méthodes de résolution	56
4.4.3	Mises en œuvre, domaine d'étude	56

4.4.4	Conclusion	59
5	Modélisation	61
5.1	Représentations du système ferroviaire	62
5.1.1	Représentation métier	62
5.1.2	Représentations macroscopiques et microscopiques :	63
5.1.3	Une nouvelle représentation de l'infrastructure	66
5.2	Leviers d'action	70
5.3	Objectifs métier	73
5.3.1	Les outils de contrôle du système ferroviaire	73
5.3.2	Les principaux objectifs métiers de la régulation	73
5.3.3	Options techniques concernant la mise en œuvre	74
5.3.4	Analyse	75
5.4	Contraintes métiers	75
5.4.1	Contraintes d'exploitation	76
5.4.2	Contraintes opérationnelles	78
5.4.3	Contraintes commerciales	79
5.4.4	Classes de contraintes mathématiques	79
5.5	Données	81
5.5.1	Données requises	81
5.5.2	Solution	81
6	Formulations	83
6.1	Notations	83
6.1.1	Ensembles	83
6.1.2	Relations	84
6.1.3	Données	84
6.1.4	Notations mathématiques	84
6.2	Formulation mathématique en temps continu	85
6.2.1	Variables de décision	85
6.2.2	Contraintes	86
6.2.3	Fonction objectif	90
6.2.4	Synthèse du modèle	91
6.2.5	Extensions du modèle	92
6.3	Formulation mathématique en temps discret	95
6.3.1	Discrétisation du temps	95
6.3.2	Variables de décision	95
6.3.3	Contraintes	99
6.3.4	Fonction objectif	102
6.3.5	Synthèse du modèle	103
6.3.6	Mise en œuvre	104
6.4	Formulation mixte	107
6.4.1	Variables de décision	107
6.4.2	Contraintes	109
6.4.3	Fonction objectif	111

6.4.4	Synthèse du modèle	112
6.4.5	Extensions du modèle	113
6.5	Mise en œuvre pratique des modèles mathématiques	115
6.5.1	Formulations et résolution	116
7	Mise en œuvre industrielle et intégration système	119
7.1	Intégration en phase d'étude	120
7.1.1	Les modules en interaction	120
7.1.2	La gestion des circulations	121
7.1.3	Intégration	122
7.1.4	Conclusion	122
7.2	Vers un outil opérationnel	124
7.2.1	Concernant les données : analyse des flux de données nécessaires	124
7.2.2	EIFEL et le modèle de données	125
7.2.3	Les autres principaux éléments du système logiciel	128
7.2.4	Fonctionnement du système logiciel	130
8	Résultats et analyses	135
8.1	Enseignements de la planification	136
8.1.1	Paramétrage	136
8.1.2	Analyse sur la recherche de solutions	138
8.2	Prototype de replanification	140
8.2.1	Explosion des retards induits	140
8.2.2	Exemples de taille d'études traitées	141
8.2.3	Résultats du prototype	141
8.3	Une première version d'EIFEL	143
8.3.1	Taille des formulations	143
8.3.2	Tests préliminaires	144
8.3.3	Conclusion	146
9	Perspectives et conclusions	147
9.1	Perspectives	147
9.1.1	Définition des objectifs	147
9.1.2	De la gestion efficace des interactions entre les CCR	148
9.1.3	Couplage avec les autres problématiques ferroviaires	148
9.1.4	Architecture technique et méthodes de résolution	148
9.1.5	Le Système d'Information	149
9.2	Conclusions	149
A	Annexes	151
B	Glossaire	153

Venant de l'aérospatial, qu'est ce qui vous a le plus surpris en entrant dans « l'atmosphère ferroviaire » ?

L'extraordinaire complexité de l'organisation des circulations ferroviaires, qui doit gérer chaque jour, le guidage de 13.000 trains et 200.000 wagons.

Et il ne faut pas perdre de vue que nous sommes une activité de service. On ne stocke pas la production. Le client utilise donc le service en même temps qu'il se fabrique et subit en temps réel tous les dysfonctionnements de l'entreprise. Il vit avec nous. Ce qui donne une relation très particulière, très affective, entre nos clients, la SNCF et ses personnels.

(interview pour Science&Vie, hors série juin 2003, « Les trains du XXI^e siècle »)
Louis Galois (président de la SNCF de 1996 à 2006)

CHAPITRE 1

PREAMBULE

REMERCIEMENTS

Ce travail est le résultat d'une réunion d'acteurs qui m'ont permis de le réaliser. C'est pourquoi je tiens à les remercier ici.

Avant toute chose, je remercie l'Association Nationale de la Recherche et de la Technologie ([ANRT](#)) pour avoir rendu possible cette thèse CIFRE, collaboration entre la Direction de l'Innovation et de la Recherche de la SNCF ([DIR-SNCF](#)), l'Institut de Mathématiques de Bordeaux, Université Bordeaux I ([IMB](#)) et le groupe de recherche *Reformulation and algorithms for Combinatorial Optimization* de l'INRIA sud-ouest ([INRIA-RealOpt](#)).

Je remercie également toutes celles et ceux qui ont rendu ce travail possible et ont permis de l'enrichir au sein de la Direction de l'Innovation et de la Recherche, de la Direction de l'Ingénierie, la Direction de l'Infrastructure, ainsi que les membres du groupe transverse Gestion des Circulations et l'ensemble des acteurs de la régulation, en particulier nos interlocuteurs au Centre National des Opérations.

Je remercie ensuite mes collègues d'Eurodécision pour avoir partagé leur expertise sur la mise en œuvre des techniques d'optimisation et mes collègues étrangers pour les échanges toujours enrichissant concernant l'optimisation de l'exploitation ferroviaire.

Enfin, je remercie les membres du jury et les rapporteurs pour le temps qu'il m'ont accordé, leurs critiques et leurs conseils.

Mais ce sont bien sûr tout particulièrement François, Pierre et Gilles que je remercie : François et Pierre, qui ont toujours veillé à développer le recul critique et la rigueur nécessaire à ce travail.

Gilles, car il est l'initiateur visionnaire d'une solution industrielle globale.

Merci à tous pour votre confiance et votre soutien !

RÉSUMÉ-ABSTRACT

Résumé

La régulation ferroviaire sur de vastes zones est un problème complexe. Elle intervient dans la phase opérationnelle de la production. Son rôle consiste à reconstruire un plan de transport rendu irréalisable suite à l'apparition d'un incident. La contribution de ce travail s'organise autour de trois axes. Le premier consiste à définir une formalisation exhaustive du système ferroviaire, associé à une représentation plus cohérente (modèle multiniveau). Le deuxième axe s'articule autour de l'étude des modèles mathématiques pour la régulation du trafic ferroviaire : évolutions d'un modèle en temps continu complet (espacements dynamiques), proposition d'un modèle innovant à temps discret et d'un modèle mixte (continu-discret) adossé au modèle multiniveau. Enfin, le dernier axe traite de la mise en œuvre concrète au niveau industriel, en particulier des gains attendus du couplage avec un outil de simulation.

Abstract

On-line re-scheduling of trains aims to find accurate solutions after an incident has occurred on the railway network. When we consider a large area, solutions consist in calculating new timetables, sequences and routes of trains. This challenging problem asks for decision support tools based on operational research techniques. This thesis develops solutions toward an operational tool, it articulates around three main parts.

The first part defines an exhaustive model of the railway system operations and a multiscalable description of the infrastructure. The second part presents three mathematical models of the rescheduling problem associated with this description. We present an exhaustive continuous time based model including headways depending on speeds of trains, plus a discrete time based model and an hybrid model combining both formulations. The last part describes first implementations and the global framework we need to develop in order to solve the real world problem. It emphasizes on expected synergies, specially between simulation and optimization techniques.

INTRODUCTION

-Bonjour, dit le petit prince.

-Bonjour, dit l'aiguilleur.

-Que fais-tu ici ? dit le petit prince.

-Je trie les voyageurs, par paquets de mille, dit l'aiguilleur. J'expédie les trains qui les emportent, tantôt vers la droite, tantôt vers la gauche.

Et un rapide illuminé, grondant comme le tonnerre, fit trembler la cabine d'aiguillage.

-Ils sont bien pressés, dit le petit prince. Que cherchent-ils ?

-L'homme de la locomotive l'ignore lui-même, dit l'aiguilleur.

Et gronda, en sens inverse, un second rapide illuminé.

-Ils reviennent déjà ? demanda le petit prince...

-Ce ne sont pas les mêmes, dit l'aiguilleur. C'est un échange.

-Ils n'étaient pas contents, là où ils étaient ?

-On n'est jamais content là où on est, dit l'aiguilleur.

Et gronda le tonnerre d'un troisième rapide illuminé.

-Ils poursuivent les premiers voyageurs demanda le petit prince.

-Ils ne poursuivent rien du tout, dit l'aiguilleur. Ils dorment là-dedans, ou bien ils bâillent. Les enfants seuls écrasent leur nez contre les vitres.

-Les enfants seuls savent ce qu'ils cherchent, fit le petit prince. Ils perdent du temps pour une poupée de chiffons, et elle devient très importante, et si on la leur enlève, ils pleurent...

-Ils ont de la chance, dit l'aiguilleur.

Le Petit Prince, CHAPITRE XXII
Antoine de Saint Exupéry

La régulation ferroviaire lorsqu'elle porte sur un périmètre important est un problème complexe. Elle intervient dans la phase opérationnelle de la production. Son rôle consiste à trouver de nouvelles solutions en termes de planification des mouvements de trains suite à l'apparition d'un incident empêchant la réalisation normale du plan de transport tel que préétabli lors des phases amonts de production.

La contribution de ce travail s'organise autour de trois axes : d'abord la modélisation du système ferroviaire, puis les modèles mathématiques pour l'optimisation de la régulation du trafic ferroviaire et pour finir la définition d'un cadre de mise en œuvre industrielle.

Le premier axe consiste à définir une représentation abstraite du réseau ferroviaire et formaliser l'ensemble du fonctionnement du système ferroviaire. Ce modèle couvre en particulier une description de l'infrastructure qui permet plusieurs niveaux de représentation et inclut ses différents modes de fonctionnement (contraintes d'exploitation). Ce travail de formalisation comprend aussi une analyse des contraintes métiers (opérationnelles et commerciales) ainsi que des objectifs possibles de la régulation. Ce modèle servira dans l'avenir de socle commun à l'ensemble des applications liées à la production ferroviaire et en particulier à la chaîne horaire : depuis la phase de construction des horaires, jusqu'à la gestion en temps réel du système ferroviaire.

Le deuxième axe s'articule autour de l'étude des modèles mathématiques pour la régulation du trafic ferroviaire associés à cette nouvelle représentation. Nous y présentons les évolutions d'un modèle en temps continu complet (espacements dynamiques), puis nous proposons un modèle innovant à temps discret et un modèle mixte (continu-discret) adossés au modèle multiniveau.

Enfin, l'objectif industriel consiste à développer des solutions permettant une intégration des outils d'aide à la décision dans le futur système de gestion des circulations. La définition de l'architecture système nécessaire et particulièrement des interactions et synergies avec le module de simulation ont donc aussi constitué une bonne part de nos travaux de recherche et de développement d'un outil opérationnel. Pour finir, une description chronologique du déroulement des opérations de gestion du trafic ferroviaire est présentée ; elle permettra d'embrasser l'architecture logicielle nécessaire pour la mise en œuvre de ce projet et les enchaînements entre ses divers éléments.

Ces trois axes sont exposés en suivant le plan qui suit.

Le second chapitre est consacré à la définition de la problématique de l'exploitation ferroviaire et à définir tant les besoins que les ambitions du projet.

Le troisième passe en revue la bibliographie traitant des techniques et modèles employés sur le sujet de la re-planification mais aussi pour la conception de grille horaire.

Dans le chapitre quatre, nous développons un nouveau modèle abstrait de représentation du système ferroviaire. Ce modèle couvre en particulier une description de l'infrastructure qui permet plusieurs niveaux de représentation et inclut ses différents modes de fonctionnement.

Dans le cinquième chapitre, nous détaillons les modèles mathématiques et les trois formulations que nous avons développées sur la base du modèle abstrait.

La sixième chapitre traite de la mise en œuvre concrète au niveau industriel ; il présente

les différents éléments constitutifs d'un outil d'aide à la décision pour l'opérationnel et les synergies possibles.

Le septième est consacré à la présentation des résultats obtenus au cours des expérimentations que nous avons menées.

Enfin, le dernier chapitre expose nos conclusions et les perspectives.

PROBLÉMATIQUE

...l'essentiel, le plus sûr, pour ne pas dire la seule certitude dans un monde mouvant, c'est notre production, la quantité, la qualité et la performance de notre production : la production de services à nos clients finaux - voyageurs et chargeurs - , et ce, quels que soient nos métiers, dans des chaînes de services, dans des chaînes de valeur et de croissance ...

Anne-Marie Idrac (ancienne présidente de la SNCF de 2006 à 2008)

3.1 Contexte

Le secteur ferroviaire en Europe vit une phase d'accélération de son évolution (expansion des trains à grande vitesse, mises en concurrence, développement des trafics transnationaux, préférence environnementale, etc). En conséquence, les trafics s'intensifient et les méthodes de production se rationalisent. Ainsi certains secteurs de la production ferroviaire gagnent en complexité et nécessitent de nouveaux outils d'aide à la décision. C'est le cas en particulier de la construction des grilles horaires et de la gestion opérationnelle des circulations.

3.1.1 La production ferroviaire

Ainsi que l'évoque Anne-Marie Idrac, le produit final de la production ferroviaire est le résultat d'une chaîne d'opérations industrielles qui pour la plupart restent hors champ du public. Cette chaîne commence très en amont (les grands projets d'infrastructure se conçoivent une dizaine d'années avant leur mise en service) et se termine par la gestion opérationnelle qui consiste à s'assurer du bon déroulement des opérations et gérer en continu les inévitables aléas qui surviennent.

C'est pourquoi on retrouve, tant dans la littérature (par exemple [13, 16]) que dans l'organisation même des équipes chargées d'optimiser les principales phases de la production ferroviaire, une partition en plusieurs catégories de problèmes d'optimisation dont voici les principaux :

- l'**identification de la demande** et l'**optimisation des offres** pouvant y répondre. Ceci représente la phase la plus amont de la production pour une entreprise ferroviaire,
- l'**optimisation des investissements en infrastructure** permet de proposer de nouvelles offres sur la base des infrastructures déjà existantes, et d'étudier les travaux indispensables pour répondre à de nouvelles demandes (exemple : demande d'augmentation du trafic sur des lignes régionales),
- la **planification** correspond à la phase de construction de l'ensemble des horaires des circulations (aussi appelé « grille horaire »),
- le **routage** (*platforming*) consiste à affecter les voies aux différentes circulations,
- la **gestion du matériel** couvre l'établissement des plannings de roulements, l'optimisation des maintenances (retours aux dépôts) du matériel roulant,
- la **gestion des agents** concerne l'affectation des personnels (aux circulations, à la maintenance,...) en tenant compte notamment des compétences et des réglementations,
- la **gestion opérationnelle** des circulations doit répondre en temps réel aux aléas qui rendent la réalisation des horaires théoriques impossible en l'absence d'action de régulation du trafic ferroviaire. On parle aussi de *replanification*.

Bien évidemment, ces problèmes ne sont pas indépendants; les résultats de l'un influencent les autres. Par exemple, le choix des roulements du matériel dépend des horaires planifiés et influence le planning des agents de circulation (en réalité tous ces problèmes se nourrissent mutuellement). Cependant il n'est pas possible aujourd'hui de traiter l'ensemble de ces problèmes globalement. Lorsque nous saurons résoudre de manière pleinement satisfaisante chacun de ces problèmes indépendamment, il est probable que les recherches porteront sur des modèles et des algorithmes permettant de les résoudre de manière plus globale. Par exemple, certaines de ces problématiques commencent déjà à être traitées simultanément [31, 45, 3]. Néanmoins, l'intégration des problèmes de ressources matérielles et humaines dans la phase de replanification concernera une étape ultérieure de la recherche dans le secteur ferroviaire.

Par contre, nous avons pris le parti de ne pas dissocier complètement le routage de la gestion opérationnelle des circulations, dans la mesure où c'est un problème important en pratique.

Le cadre de cette thèse est donc limité à la seule étude de la gestion opérationnelle des circulations, indépendamment des aspects ressource humaine et matérielle en particulier (qui seront donc des *données* éventuelles mais pas les *objets* de l'optimisation).

3.1.2 La gestion opérationnelle des circulations

La densité des trafics approche la saturation de l'infrastructure dans de nombreuses zones. Cette densification est d'ailleurs elle-même rendue possible en partie grâce aux outils d'optimisation de construction des grilles horaires. Cependant cette même densification rend les incidents plus fréquents et surtout leurs conséquences plus importantes ou, au moins, beaucoup plus difficiles à gérer efficacement.

D'autre part on assiste à une prise de conscience du rôle des interactions entre différentes parties du système ferroviaire ; on souhaite donc à présent optimiser la régulation de grandes zones de manière globale (par exemples : les corridors européens, les LGV, ...) et non plus seulement gérer de petites zones localement.

La replanification et la gestion des circulations en temps réel nécessitent donc de plus en plus des outils d'aide à la décision adaptés.

Schématiquement le problème pourrait se résumer de la manière suivante : suite à l'apparition d'un ou plusieurs incidents, les horaires théoriques ne sont plus réalisables. La recherche d'une solution consiste à trouver de nouveaux horaires, de nouveaux routages et de nouveaux ordonnancements pour chaque circulation. Cette recherche doit se limiter à l'espace des solutions réellement possibles qui est régi par un ensemble de contraintes métiers à respecter absolument. Parmi ces solutions, l'objectif consiste généralement à trouver celle qui limitera le plus les retards induits.

Bien sûr, cette définition succincte pourrait finalement assez bien s'appliquer à chaque étape qui concourt à la conception de la chaîne horaire, c'est à dire le processus allant de la phase de planification initiale à la gestion opérationnelle des circulations, en passant par l'adaptation qui consiste à anticiper des demandes exceptionnelles de circulation ou de travaux rendus nécessaires. C'est pourquoi nous allons commencer par clarifier ce qui distingue ces deux principaux problèmes.

Liens avec le problème de la construction de grilles horaires (planification)

Le problème de la gestion opérationnelle des circulations ressemble donc beaucoup en principe à celui de la planification.

Du point de vue *mathématique*, les similitudes concernent les deux points suivants. D'une part **les mêmes catégories de contraintes métiers** décrivent le domaine des solutions (contraintes d'exploitations, opérationnelles et commerciales décrites en détails au chapitre 5.4). D'autre part, **des objectifs très semblables** régissent la recherche de solutions, en particulier la minimisation des pénalités liées à des écarts horaires (respectivement entre l'horaire commercial souhaité et l'horaire obtenu dans le cas de la planification, et entre l'horaire théorique et l'horaire finalement réalisé dans le cas de la replanification). Néanmoins, la hiérarchie entre plusieurs objectifs peut aussi différer sensiblement.

Il est ensuite nécessaire de souligner les points clefs qui rendent ces deux problèmes assez différents en pratique. Ainsi, la principale spécificité est liée à l'aspect *temps réel* de la gestion des circulations. Ceci amène plusieurs conséquences.

D'abord **le temps de calcul** devient évidemment une limite opérationnelle forte. Alors que la planification met en jeu des problèmes dont les solutions seront appliquées au plus tôt dans quelques jours, dans le cas de la gestion opérationnelle, le délai atteint à peine quelques minutes. Les méthodes de résolution ne sont donc pas nécessairement semblables (à défaut d'être *optimale*, il faut a minima une *bonne* solution réalisable, rapidement).

Ensuite, le niveau de **précision** requis constitue une différence fondamentale. En phase de planification, aujourd'hui l'outil d'aide à la décision peut servir de support à l'expert pour construire la grille horaire ; il le guide dans la construction du scénario global, mais souvent, l'objectif ne consiste pas à construire directement une grille horaire opérationnelle (c'est à dire effectivement réalisable en l'état sur le terrain, avec des consignes de vitesse très précises et une prise en compte détaillée des aspects de sécurité et de signalisation). De ce fait, le niveau d'exigence en termes de précision des solutions s'avère plus faible qu'avec les problèmes opérationnels qui nous concernent, où l'on doit s'assurer de la parfaite réalisabilité de la solution proposée sans compromettre son efficacité. En effet, en phase de planification, on choisit par exemple d'ajouter des *marges de régularité* le long du parcours (afin d'atténuer les effets d'imprévus mineurs). De plus ces marges de sécurité sont elles aussi volontairement surévaluées. A contrario, la gestion des circulations en temps réel va (entre autres actions) exploiter ces marges introduites en phase de conception au plus juste, afin de rattraper efficacement les retards. Cette précision accrue a évidemment des conséquences en termes de taille des problèmes mathématiques traités et de difficulté de résolution, mais c'est un aspect incontournable du problème de replanification.

Enfin, il faut souligner qu'on peut disposer **d'une solution initiale** plus facilement dans le cadre de la replanification. Lors de la planification on doit pouvoir construire toute (ou une partie de) la grille ex-nihilo. Dans le cas de la phase de replanification, l'incident (et ses conséquences) qui vient de survenir rend la grille (partiellement) irréalisable. On peut donc mettre à profit l'information émanant de la grille théorique (par exemple, la plupart des ordonnancements seront préservés) afin de construire une solution initiale ; la phase d'optimisation à proprement parler peut ensuite consister à améliorer cette solution.

Ainsi, si la planification des circulations et la gestion opérationnelle des circulations sont des problèmes qui semblent proches en pratique, il serait dangereux de négliger les différences intrinsèques entre ces deux problèmes. Néanmoins, le problème de planification des circulations fournit malgré tout une bonne porte d'entrée pour l'étude théorique de notre problème, d'autant plus qu'elle a fait l'objet de davantage d'études à ce jour comme nous le verrons au chapitre 4.

Pour approfondir l'étude du problème (métier) de la planification , on se référera à [23, 28].

3.1.3 Notions ferroviaires élémentaires

Avant de continuer plus avant, il paraît nécessaire de développer quelques concepts spécifiques au monde ferroviaire et au problème métier qui nous intéresse. En effet, bien qu'au cœur de la société industrielle que nous connaissons aujourd'hui et qui s'est développée il y a un peu plus d'un siècle, le chemin de fer demeure généralement mal connu. Or on ne peut pas faire abstraction sans dommage pour la qualité de notre réflexion de ses spécificités.

Histoire des chemins de fer et répartition des rôles

Historiquement, la première expansion des réseaux ferroviaires (à partir de 1840) relevait de compagnies ferroviaires privées et nationales construisant leurs propres infrastructures afin d’y faire circuler (uniquement) leurs propres trains.

Puis vint en Europe l’époque de la rationalisation des réseaux nationaux, souvent jugée stratégique par les états, par le biais de nationalisations (les réseaux de chaque compagnie n’étant pas nécessairement interopérables au sein du même pays). En France, la fusion des différentes compagnies ferroviaires et de leurs réseaux respectifs fut adoptée en 1938 avec la création de la SNCF. Cependant bon nombre de contraintes techniques nous restent de ce premier siècle de l’histoire ferroviaire qui structurèrent l’environnement que nous lui connaissons aujourd’hui. À titre d’exemple, on exploite toujours en France deux systèmes électriques différents selon les régions (courant continu 1,5 kV et courant alternatif 25 kV 50 Hz).

Un demi-siècle plus tard, dans une démarche visant à permettre une libre concurrence entre opérateurs (à terme intra-européenne), l’Union Européenne s’est imposée de dissocier à l’échelle nationale le gestionnaire d’infrastructure qui possède, développe, entretien et gère l’utilisation de l’infrastructure, des entreprises ferroviaires qui opèrent des circulations sur ce réseau. Pour ce qui concerne tant la planification que de la gestion opérationnelle des circulations, c’est donc le gestionnaire d’infrastructure qui est chargé d’arbitrer les requêtes des entreprises ferroviaires et d’allouer les sillons selon la capacité que peut supporter le réseau.

En France, la charge de gestionnaire d’infrastructure revient à Réseau Ferré de France (RFF). La SNCF au travers de ses propres activités Voyage France Europe, SNCF Proximités et Fret est donc une entreprise ferroviaire. Cependant son activité Infrastructure reste en charge de missions d’ingénierie ferroviaire, d’entretiens et de conceptions. De plus, cette dernière activité a aussi pour mission d’aider à la conception des horaires (via les bureaux horaires) ainsi que la charge de l’exploitation opérationnelle (régulation) du réseau en tant que Gestionnaire d’Infrastructure Délégué pour le compte de RFF.

Enfin, l’Europe incite à adopter un nouveau standard commun de système de signalisation (ERTMS) afin de limiter le nombre de systèmes de signalisation avec lesquels les matériels roulants doivent être compatibles pour opérer des circulations transnationales. Pour illustration, actuellement les rames Thalys de type PBKA (pour Paris-Brussels-Köln-Amsterdam) sont quadricourant et compatibles avec pas moins de sept systèmes de signalisation différents.

Le système ferroviaire est donc l’héritier d’un siècle et demi d’histoire industrielle, technique et politique (car paradoxalement, les incompatibilités entre réseaux furent longtemps perçues comme stratégiques ; l’écartement spécifique des voies en Espagne en demeure l’exemple le plus emblématique en Europe). Beaucoup d’aspects techniques rendent son exploitation plus contraignante que ce qu’il serait permis de supposer, d’autre part la répartition des rôles entre différents acteurs ne permet pas non plus toutes les latitudes en matière de prise de décision. Par exemple en ce qui concerne la phase de définition d’objectifs pour la régulation ferroviaire, le rôle de la SNCF se limite à être force de proposition.

Sur ce thème, le dossier dédié au thème de la répartition des rôles entre RFF et la SNCF dans [23] pourra compléter cette courte présentation. Cependant, il est nécessaire de noter que l'organisation actuelle ne semble pas satisfaire pleinement les objectifs et que sa pérennité n'est pas absolument garantie [5, 29, 19].

Définitions élémentaires

Ce premier paragraphe propose quelques définitions indispensables à la bonne compréhension du sujet. D'autre part, l'annexe B présente un glossaire qui récapitule les termes les plus utiles.

Le plan de transport regroupe l'ensemble des dispositions destinées à organiser les ressources (humaines, matérielles, ...) mises en place pour transporter des biens ou des personnes (ici par voie ferrée). Parmi celles-ci **les circulations** désignent les mobiles horairisés, c'est-à-dire des ressources associant un matériel roulant dont les mouvements sont planifiés et une **mission** ou **desserte** constituée d'une succession de points d'arrêts commerciaux ou techniques (arrêts de service, ...) utiles à l'entreprise ferroviaire. La **marche** d'une circulation désigne l'ensemble des temps de parcours intermédiaires entre deux points du réseau ferré. Enfin, l'**infrastructure** est constituée de ressources fixes, qui comprennent la signalisation.

Les **aléas** ou **incidents** recouvrent toute perturbation imprévue. Sans intervention de régulation, ces incidents initiaux génèrent en général un lot de perturbations secondaires en cascade. Néanmoins, celles-ci sont prévisibles dès lors que la perturbation initiale est connue. En revanche de nouveaux aléas peuvent s'ajouter à tout instant à la situation courante ; on parle alors de **sur-incident**.

Représentations

La représentation la plus classique est le **Graphique Espace-Temps** (GET) ou **Graphique des Circulations** qui sert à représenter sur une période de temps donnée l'évolution des circulations le long d'une portion de réseau (voir la figure 3.1). L'abscisse représente l'axe du temps, l'ordonnée l'espace (une portion nécessairement linéaire du réseau) où on représente une succession de points remarquables, appelée *manchette*. Chaque courbe représente le parcours en fonction du temps d'une circulation. La pente correspond à la vitesse de la circulation, mais l'infrastructure pouvant être parcourue dans les deux sens, il existe des pentes négatives et positives. Ainsi, un palier horizontal indique un arrêt, alors qu'un croisement entre deux circulations de sens opposés (sur des voies séparées) est immédiatement repérable car deux courbes de pentes opposées se croisent au point remarquable considéré.

L'utilisation du GET est tellement ancrée dans le métier d'horairiste ou de régulateur qu'elle a généré sa propre sémantique. Par exemples :

- *graphicage* : opération d'établir -dessiner- une grille horaire,
- *coucher* la grille horaire : dans la représentation française (temps en abscisse, espace en ordonnées), ralentir les circulations s'interprète graphiquement par une diminu-

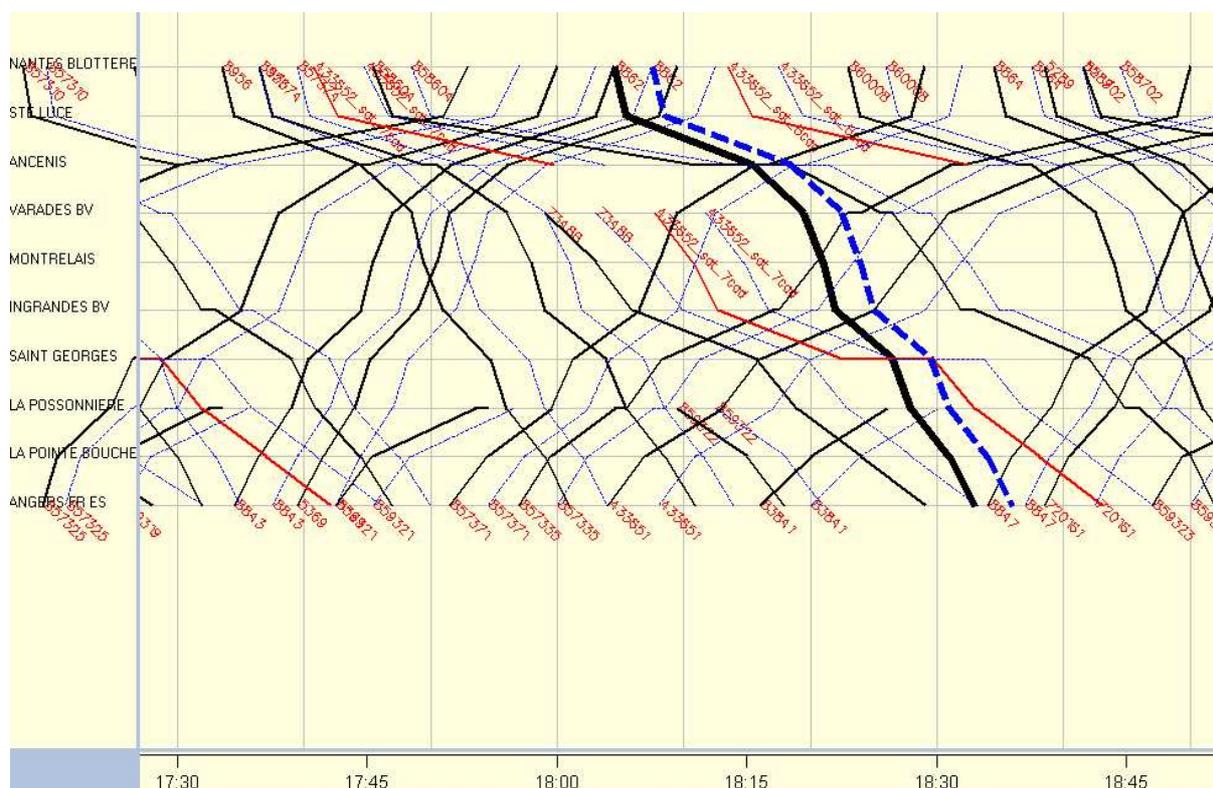


FIGURE 3.1 – Graphique Espace Temps

tion générale des pentes ; donc, visuellement, à *coucher la grille*.

Pourtant le GET n'est pas dénué de défauts : il est difficile de représenter des portions de réseau qui ne sont pas strictement linéaires (par exemple : un réseau en étoile, ou plus simplement une bifurcation) ; on peut alors empiler sur le même schéma chaque graphique correspondant à une branche du réseau, mais il ne peut y avoir de raccord entre elles. D'autre part, on ne peut pas non plus visualiser l'occupation précise des voies (dans une gare par exemple) ; d'où une forme complémentaire de représentation : le **graphique d'occupation des voies** (GOV) qui permet de visualiser l'occupation de chaque ressource (en ordonnées) au cours du temps (en abscisse) (voir la figure 3.2).

Chacune des deux représentations est bien adaptée pour un aspect particulier, mais aucune ne permet d'avoir une vue globale du système ferroviaire. D'ailleurs ces limitations techniques correspondent aussi à une séparation du point de vue du métier puisqu'il est de coutume de séparer la gestion *en ligne* (où le GET est bien adapté) de la gestion des *nœuds ferroviaires* ou *en gare* (où le GOV s'avère plus pratique). Néanmoins, comme nous le soulignons dans l'introduction, les deux problèmes (schématiquement celui des horaires en ligne et celui du routage en gare) ne sont pas indépendants et toute dissociation peut être préjudiciable à l'efficacité de la solution qui devrait prendre en compte un point de vue global.

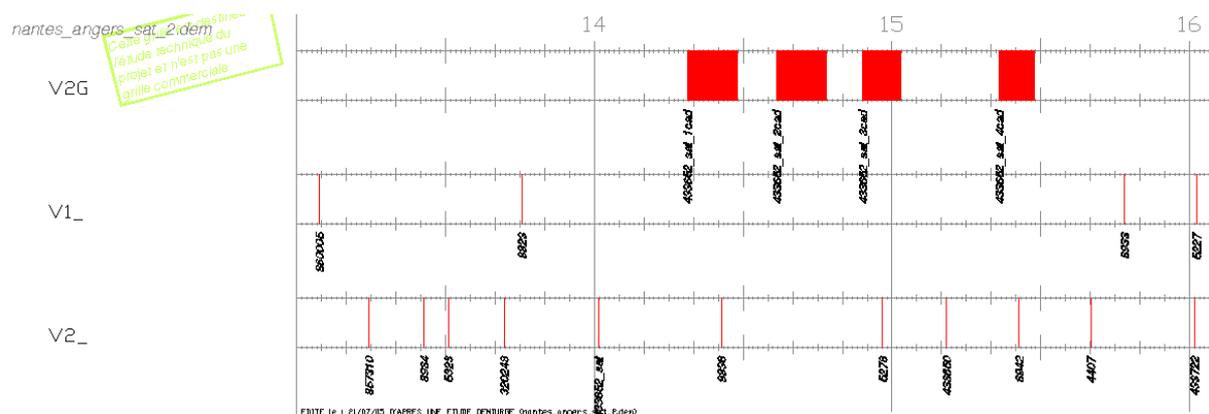


FIGURE 3.2 – Graphique d'Occupation des Voies

Espacements de sécurité et systèmes de signalisation : principes

L'avantage du chemin de fer réside dans la faible résistance à l'avancement du contact roue-rail; ceci permet un très bon rendement et permet d'acheminer de très lourdes charges moyennant un faible effort de traction (sur une section plane). Néanmoins cette faible résistance à l'avancement a pour revers une mauvaise adhérence en cas de décélération importante.

Les distances de freinages en cas d'urgence dépassent donc souvent de beaucoup celles de la visibilité maximum du conducteur¹. Ainsi se sont mis en place très tôt dans l'histoire ferroviaire des systèmes de sécurité garantissant un espacement suffisant entre circulations. Ces systèmes sont mis en œuvre au moyen de **systèmes de signalisation**. Quel que soit le degré de perfectionnement ou la technique utilisée (automatisme, contrôle, affichage, ...), tous reposent sur le principe de cantonnement, c'est à dire une zone clairement délimitée dans laquelle une seule circulation est admise. La présence de la circulation déclenche la signalisation qui interdit à toute autre circulation de pénétrer dans la zone de sécurité, comme l'illustre la figure 3.3. (Plus exactement, elle garantit même qu'un certain nombre de cantons vides séparent deux circulations qui se succèdent, permettant ainsi le freinage en toute sécurité.)

Malheureusement les systèmes de signalisation, héritage de l'histoire ferroviaire, sont multiples au sein d'un même pays et varient d'un pays à l'autre. Bien qu'ils soient tous basés sur ce même principe, cela engendre de nombreuses incompatibilités techniques entre matériels (surtout pour les circulations transnationales). Ainsi, un nouveau système de signalisation ERTMS est voué à devenir un nouveau standard commun en Europe. Sa version la plus évoluée repose sur une variante du principe précédent : le canton n'est plus fixe, deux trains doivent se succéder en respectant en permanence (et non en un nombre fini de points du réseau) un espacement permettant le freinage sûr du train suiveur en cas d'incident (valeur calculée en permanence en fonction des vitesses des deux circulations). C'est le concept de canton (ressource) flottant(e).

Nous verrons que cette dernière déclinaison apporte une difficulté supplémentaire en

1. pour un TGV la distance de freinage en situation normale est de l'ordre de 1000m à 160km/h

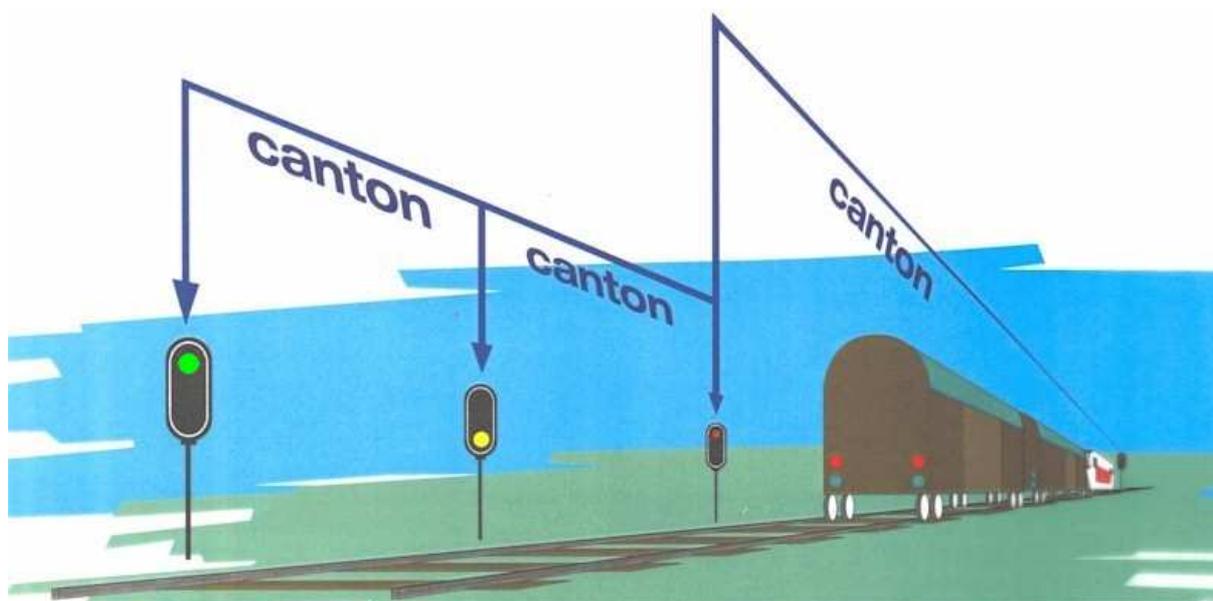


FIGURE 3.3 – Principe du cantonnement

termes de modélisation.

Pour les zones plus complexes (comme l'avant-gare illustré par la figure 3.4) ce sont généralement des câblages matériels et des règles logiques appelées *consignes roses* qui régissent la signalisation. Ces règles logiques visent à interdire une utilisation conflictuelle de ressource d'infrastructure, en particulier afin d'interdire l'utilisation simultanée d'itinéraires physiquement incompatibles (par exemple un couple d'itinéraires qui se croiseraient comme sur la figure 3.5).



FIGURE 3.4 – Aiguillages en entrée de gare

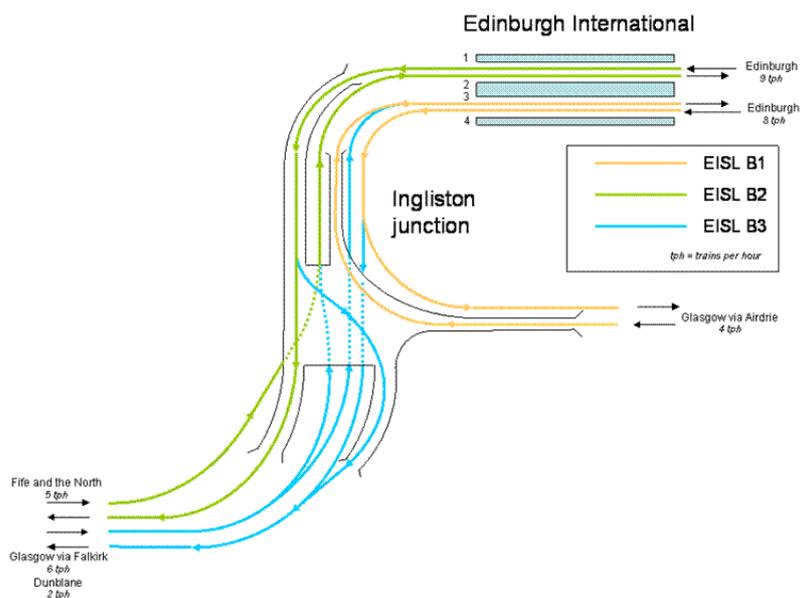


FIGURE 3.5 – Itinéraires incompatibles

Pour plus de détails on pourra se référer à [48, 57] pour les systèmes français qui nous intéressent en premier lieu. Concernant l'ERTMS, on se tournera vers le site du consortium industriel en charge de son développement [24].

Les situations à éviter

Le rôle combiné de la régulation ferroviaire et des systèmes de sécurité est d'éviter que ne survienne un accident entre une circulation et un élément d'infrastructure, ou entre plusieurs circulations.

En matière d'accidents entre circulations, il existe trois catégories :

L'**affrontement** ou le **nez à nez** : deux circulations se dirigent en sens opposé l'une vers l'autre jusqu'à une collision frontale.

Le **rattrapage** : deux circulations qui se succèdent rentrent en collision par l'arrière.

La **prise en écharpe** : les itinéraires de deux circulations s'entrecroisent (cas du **cisaillement**) ou se rejoignent (cas de la **convergence**) pour aboutir à une collision latérale.

Ces situations sont gérées grâce aux systèmes de sécurité et de signalisation qui garantissent des espacements de sécurité (cas du rattrapage) ou de n'emprunter une portion d'infrastructure (une voie unique ou un couple d'itinéraires incompatibles par exemple) seulement une circulation après l'autre.

Le champ des accidents liés à l'interaction entre une circulation et l'infrastructure est plus vaste, même si ceux-ci peuvent paraître plus triviaux. Nous illustrerons seulement les principales interactions qui entrent en jeu.

En premier lieu, le système doit s'assurer de guider la circulation sur le bon itinéraire (via les commandes d'aiguillages), c'est à dire sur des ressources d'infrastructures compatibles avec les caractéristiques de la circulation (systèmes de traction, systèmes de sécurité, charges, gabarits, ...) et libres de tout danger, tel un obstacle entravant la voie, un rail sectionné, ou encore une caténaire arrachée.

Réciproquement, le système doit garantir que la circulation s'adapte à l'infrastructure, c'est à dire laisse aux appareils de voie (aiguillages) le temps nécessaire pour se mettre en position ou que la circulation respecte les limitations temporaires de vitesse par exemple.

Nous retiendrons que les circulations doivent respecter des contraintes d'exploitation pour garantir la sécurité en toutes circonstances : en particulier, elles doivent respecter des espacements de sécurité et ne peuvent emprunter n'importe quelle portion d'infrastructure librement. Nous détaillerons ces contraintes métiers au chapitre 5.4. Le document [6] explique de manière détaillée comment les divers systèmes de sécurité et de signalisation permettent de s'assurer sur le terrain que ces accidents ne se produisent pas.

De l'hétérogénéité des circulations et ses conséquences

Ce paragraphe présente la typologie des circulations selon les réseaux et les conséquences inhérentes en termes de gestion du trafic.

Le trafic en Amérique du Nord se concentre sur des flux de fret, les convois y sont beaucoup plus imposants et le réseau desservi beaucoup plus vaste. Le trafic est très peu dense, les problématiques de régulation opérationnelle y sont moins prégnantes et les défis de l'optimisation de la chaîne horaire concernent plutôt la phase de planification des dessertes entre hubs de fret.

En Europe², et particulièrement en France, le trafic voyageur est très fortement majoritaire. L'apparition de trains à grande vitesse capables de concurrencer efficacement l'avion sur l'échelle favorable des voyages intra-européens a joué un grand rôle dans le développement de cette activité « grandes lignes ». Mais la structure de l'habitat plus dense a aussi rendu plus intéressante la desserte urbaine. Enfin, bien qu'ayant eu à subir une forte concurrence de la route ces dernières décennies, le fret ferroviaire connaît un regain d'intérêt. Étant à nouveau perçu comme un enjeu de société, son trafic est donc voué à s'accroître lui aussi à moyen et long termes.

Tout ceci concourt à une grande hétérogénéité des trafics partageant la même infrastructure et une forte densité des circulations qui exploitent pour beaucoup les capacités des réseaux actuels en France.

D'autre part cela rend la construction des grilles horaires plus compliquée pour faire coexister sur le même réseau des circulations aux vitesses différentes. De même, en régulation, les interactions sont potentiellement plus variées et les conséquences plus complexes à gérer sans les outils adéquats. Cet aspect est la contrepartie naturelle de la coexistence efficace entre circulations (types, matériels et missions) hétérogènes sur le même réseau.

2. comme au Japon

Dans le paragraphe qui suit, nous expliquerons comment sont parfois structurés ces flux de circulations.

De la structuration des grilles horaires

Le plan de transport regroupe l'ensemble des dispositions destinées à organiser les ressources mises en place pour transporter des biens ou des personnes par voie ferrée. Parmi ces dispositions, l'établissement de « grilles horaires » (en référence à la représentation graphique des mouvements et horaires des circulations) est celle qui nous concerne en premier lieu (puisque nous excluons l'organisation des ressources matérielles et humaines du cadre de notre étude).

Or l'organisation des grilles horaires n'est pas partout identique. En effet, certaines grilles horaires ont une structure bien spécifique qui doit être respectée, en particulier : les **grilles cycliques** (ou *cadencement*) désignent une organisation en intervalles réguliers de mêmes schémas (structures) de dessertes (heures de départ, arrêts en cours de route, heures d'arrivée), comme l'illustrent les figures 3.6 et 3.7.

Parmi ces grilles cycliques, il existe plusieurs variantes du concept ; par exemple le cadencement au départ est utilisé pour les TGV afin de définir des horaires de départs réguliers. Le cadencement en réseau suppose que les circulations utilisent les mêmes schémas (complets) de dessertes. Enfin le cadencement symétrique correspond à un schéma d'organisation où d'une part l'avancement des circulations est symétrique dans les deux sens de parcours de chaque ligne et où d'autre part, un maximum de circulations se trouvent en même temps aux gares joignant plusieurs lignes afin de permettre de meilleures correspondances.

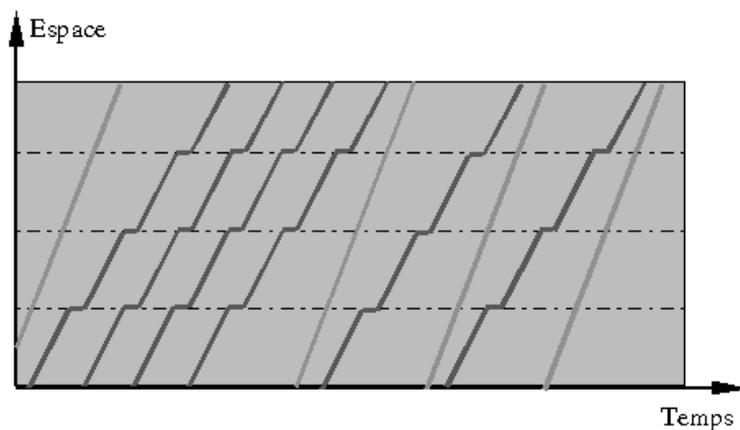


FIGURE 3.6 – Principe du cadencement

On trouvera un bref rappel historique dans [43] décrivant la diffusion de ces concepts, tout d'abord dans les métros de plusieurs capitales, puis l'extension aux chemins de fer en général. Les dessertes de bus sont aussi souvent organisées de la sorte.

Aujourd'hui plusieurs pays d'Europe fonctionnent selon ces principes pour organiser

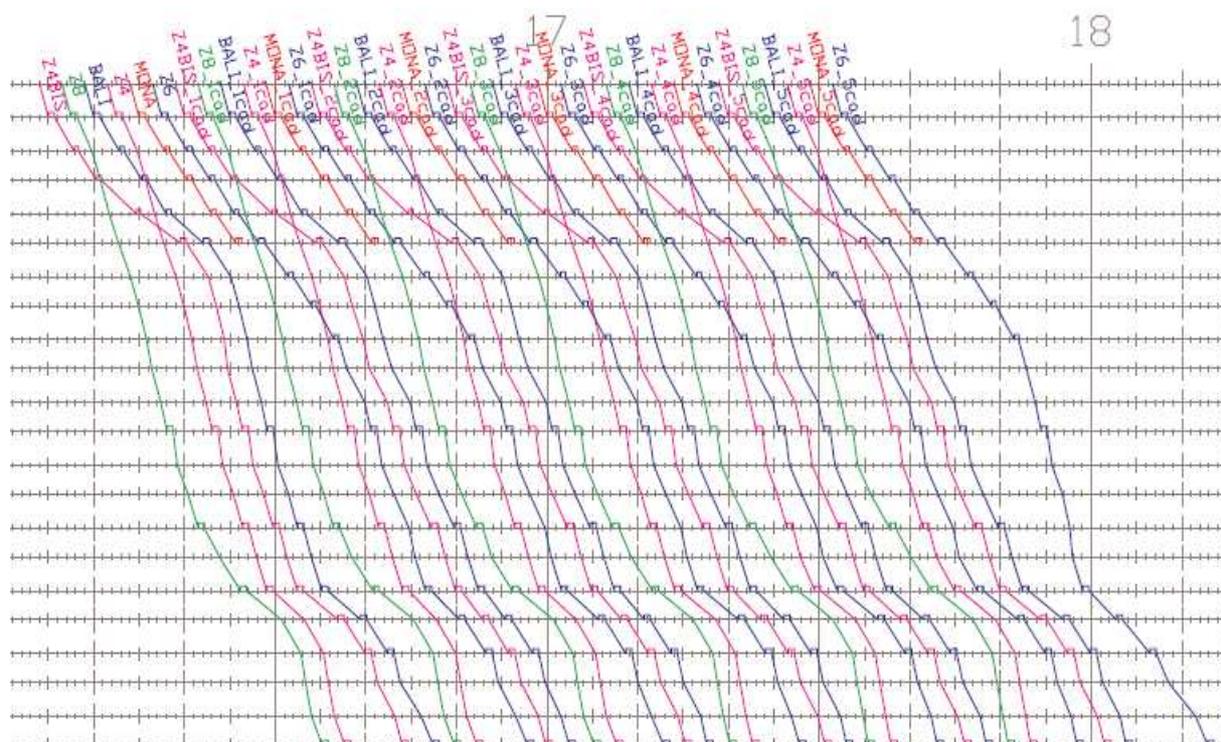


FIGURE 3.7 – Détail d’une grille horaire cadencée (RER ligne C)

l’ensemble de leur système ferroviaire : Pays-Bas, Suisse, Allemagne et Belgique notamment.

Jusqu’ici, la France n’avait pas choisi de construire son réseau et l’ensemble de ses grilles horaires autour du concept de grilles cycliques. Le concept trouve pourtant un regain d’intérêt récent, surtout pour le transport régional. Cependant il n’existe pas encore de grille horaire strictement basée sur ce concept pour les réseaux ferroviaires français (sauf à isoler certaines portions de lignes), mais plutôt des trafics mixtes (cadencés et non cadencés), voire s’appuyant sur des concepts plus souples que le cadencement strict tel que généralement défini à l’étranger, comme le cadencement au départ utilisé pour les TGV, ou le cadencement souple développé dans l’outil DEMIURGE [53, 22] pour les trains de banlieue.

Enfin, le problème des grilles cadencées peut être vu comme un sous-ensemble particulier de grilles horaires ; en contraignant davantage le problème, on perd notamment en capacité d’adaptation à la demande, ce qu’on gagne en simplicité de gestion (en phase de planification comme en opérationnel).

On trouvera notamment dans [40] des techniques d’optimisation spécifiques à ces problèmes (PESP : *Periodic Event Scheduling Problems*).

3.2 Définition du projet

Dans cette section, nous traiterons en premier lieu du cadre et des étapes principales du projet. Nous aborderons ensuite successivement le contenu des trois axes qui ont composé notre réflexion : tout d'abord l'aspect modélisation du système ferroviaire, puis l'optimisation et les modèles mathématiques, et enfin la mise en œuvre de l'outil d'aide à la décision.

3.2.1 Cadre du projet

Ce travail de thèse s'appuie sur trois expériences successives au cours desquelles notre vision a bien sûr évolué. Ainsi, les trois axes tels que nous les présenterons dans les sections suivantes correspondent à la version la plus aboutie que nous avons développée, c'est à dire la définition de l'objectif à atteindre pour la réalisation d'un outil opérationnel. Nous ne détaillerons pas ici l'architecture de ces différents projets (nous l'aborderons au chapitre 7.2.4), mais nous en évoquerons les grandes lignes afin de mieux comprendre comment s'est construite notre réflexion, ce qui relève de la conception, du développement ou de l'intégration d'outils existants.

Planification des circulations

Avant même le début de ces travaux, nous avons développé une expertise sur le problème de planification au travers du projet DEMIURGE. Ce projet arrivait à son terme, notre intervention a consisté en premier lieu à améliorer les performances du moteur d'optimisation et la formulation mathématique existante : réglage du logiciel *Cplex*, ajouts de coupes, de priorités . . . Cet aspect nous a permis de tester l'efficacité d'un grand nombre d'options techniques pour l'optimisation de l'exploitation ferroviaire.

D'autre part, nous avons apporté des évolutions fonctionnelles en lien avec les premiers utilisateurs de l'outil. Cette dernière expérience s'est révélée riche en enseignements pour la compréhension des réels enjeux pour les utilisateurs, de l'usage pratique des outils d'aides à la décision, et enfin pour développer notre expertise de l'exploitation ferroviaire (pour la partie conception d'horaires).

Un premier prototype de replanification, l'environnement LIPARI

La deuxième étape s'est déroulée au sein de l'environnement LIPARI (décrit à la section 7.1).

L'architecture logicielle (un logiciel d'optimisation pour la régulation ferroviaire, un simulateur du système ferroviaire et un outil statistique de gestion du plan d'expérience) était déjà en place. Nous avons d'ailleurs pu récupérer de nombreux jeux d'études déjà développés et les faire évoluer.

Notre intervention s'est inscrite dans la partie concernant le logiciel d'optimisation. Elle a consisté principalement à la ré-écriture du modèle mathématique et une classification plus homogène des contraintes (métiers et mathématiques), mais aussi à l'intégration des réglages issus du projet DEMIURGE. Les nombreux jeux de données ont aussi permis de tester et comprendre les enjeux spécifiques de la régulation ferroviaire.

Par ailleurs, dans le cadre de ce système logiciel, nous avons pu nous confronter au problème de fluidification (projet *Fluidification du nœud de Baudrecourt*), décrit dans la section suivante où nous expliquerons aussi les spécificités du problème de détournement. Cette expérience nous a permis de développer la formulation intégrant la prise en compte des vitesses dans le calcul des valeurs d'espacement.

Par la suite, d'autres équipes ont développé un prototype (spécifique, en dehors du cadre LIPARI) pour gérer le problème du détournement (projet *MOGADOR*). Nos nombreux échanges nous ont mené à la conclusion qu'une base commune serait souhaitable pour toutes ces catégories de problèmes.

Enfin, si l'environnement technique a présenté des limites que nous avons clairement identifiées, il a aussi permis d'entrevoir des perspectives intéressantes en termes de coopérations entre plusieurs outils logiciels et des synergies qui pourraient rendre un outil d'aide à la décision plus efficace. Il a aussi permis de prendre conscience des différents écueils auxquels se heurteraient un logiciel d'optimisation seul pour répondre correctement au problème métier dans un environnement opérationnel (en particulier vis à vis du problème des données).

Le système EIFEL

La dernière étape s'inscrit dans le cadre de la conception et du développement du système EIFEL (décrit à la section 7.2.2).

En premier lieu, notre action a porté sur la conception d'un modèle métier qui s'appuie sur une description exhaustive du système ferroviaire et une nouvelle représentation qui permet une approche multiniveau. Le développement de la base de données EIFEL a ensuite été confié à d'autres intervenants, comme le développement de l'architecture logicielle associée.

En effet, le système qui s'agrège autour de la base de donnée EIFEL est constitué de services logiciels, dont la partie optimisation. Ici, nous avons conçu et développé complètement les modèles et formulations mathématiques. Pour les autres éléments (services), en particulier la simulation à évènement discret et les outils statistiques notre rôle a consisté à souligner l'intérêt de ces projets de recherche et les synergies attendues pour notre projet.

Ainsi, nous avons pris part à la conception d'une vision globale dans le projet *Régulation du futur* afin de proposer une vision cohérente pour l'intégration de l'outil d'aide à la décision dans le cadre de la réalité industrielle de la régulation ferroviaire. Cette étape a surtout permis de concevoir des solutions aux différents écueils que nous avons identifiés au cours du projet LIPARI.

3.2.2 La modélisation du système ferroviaire

Nous allons à présent détailler le contenu du premier axe de réflexion qui a porté sur la modélisation. Cette description correspond au cahier des charges pour un outil opérationnel. C'est celui qui a guidé la conception du système EIFEL.

Description du système ferroviaire

La description exhaustive du système ferroviaire (notamment de l'infrastructure existante sur le terrain) est le point de départ de toute réflexion sur la modélisation du système ferroviaire si nous souhaitons définir un outil d'aide à la décision capable de s'adapter à toutes les zones de régulation.

Pour la partie **infrastructure**, il doit être capable de gérer une topologie de réseau quelconque et non uniquement une zone dédiée à un usage spécifique : qu'elle soit composée d'un corridor, d'une partie de réseau urbain ou d'un nœud ferroviaire complexe. D'ailleurs, l'étude de zones où des portions d'infrastructure hétérogènes se rejoignent (lignes à grande vitesse, lignes internationales et structures de desserte régionale) est généralement très instructive pour tester nos outils. Ensuite, il doit pouvoir gérer des topologies d'infrastructure potentiellement complexes : gestion des IPCS (voies uniques) et des itinéraires incompatibles (notamment en gares) en particulier.

Pour l'aspect **matériel roulant**, il est nécessaire de savoir différencier le matériel hétérogène aux vitesses variables (et personnalisées¹). Ensuite, il faut pouvoir s'adapter à tout type de structure des **grilles horaires** (comme par exemple une grille purement cyclique, mais surtout une grille normale ou un cas mixte).

Représentation abstraite du système ferroviaire

À partir du réseau réel, en fonction des modes d'utilisation de l'outil, des caractéristiques du réseau ou de l'étude considérée, il est nécessaire d'extraire une représentation abstraite (souvent simplifiée) du réseau et du fonctionnement du système ferroviaire. Cette représentation doit tenir compte de l'ensemble des contraintes métiers et de la formalisation de l'objectif que la régulation doit suivre.

Commençons par l'**expression des besoins en termes de représentation** ; il est d'abord nécessaire de pouvoir gérer et utiliser tous les niveaux possibles de représentation (du plus détaillé au plus synthétique, de manière hétérogène et combinée) ceci à discrétion du descripteur de l'infrastructure ou définis automatiquement par le système logiciel. En particulier, on doit pouvoir gérer l'infrastructure finement si cela est nécessaire localement (en particulier dans les zones critiques).

En effet, l'objectif est de définir de la manière la plus réaliste possible le fonctionnement d'une zone de régulation en prenant en compte les différents systèmes de sécurité. En particulier il doit rester compatible avec n'importe quel système de signalisation. C'est à dire être capable de reproduire aussi bien le concept le plus simple du cantonnement (BAL) que les concepts de ressources flottantes utilisées pour les systèmes les plus récents et complexes (ERTMS 3). Un point important concerne aussi les itinéraires incompatibles (couples d'itinéraires ne pouvant être alloués en même temps -souvent au sein d'une gare). Enfin la gestion de la voie unique ou à double sens (exemple : IPCS) est aussi indispensable.

Ensuite, cette représentation abstraite doit permettre de modéliser la composante dynamique du système, c'est à dire des espacements de sécurité réalistes en fonction des

1. calculées par la simulation, en tenant compte du profil de l'infrastructure et des capacités de traction

catégories de trains ainsi que de leurs vitesses (et du système de signalisation). De même, la prise en compte simplifiée des temps d'accélération (approximations linéaires ou valeurs forfaitaires) serait souhaitable.

Enfin le modèle doit pouvoir servir de socle commun aux futurs outils de régulation et de conception, il doit permettre en particulier une coopération optimale avec d'autres composants au sein d'un système logiciel complexe (par nature) et évolutif.

L'analyse de l'ensemble des **contraintes métiers** régissant le fonctionnement du système ferroviaire est la deuxième étape incontournable de la formalisation des processus. Nous nous assurerons par la suite que l'ensemble de ces contraintes puisse trouver une traduction dans les modèles mathématiques proposés.

Enfin, il n'est pas du ressort de cette thèse de définir la **fonction objectif métier** de la régulation ferroviaire. Comme expliqué dans le préambule sur les notions ferroviaires, ce rôle n'appartient d'ailleurs pas non plus à la SNCF. C'est pourquoi nous avons essayé d'analyser les différents objectifs métiers couverts par la régulation en pratique. Ensuite, nous aurons à cœur de proposer une modélisation mathématique adaptable à chacune de ces fonctions objectifs et d'engager une réflexion sur le sujet.

3.2.3 Formulations mathématiques et optimisation

Formulations mathématiques

Il s'agit de définir des modèles mathématiques correspondant à la représentation abstraite précédente, et capables de reproduire le fonctionnement du système ferroviaire.

Plusieurs modèles mathématiques peuvent coexister. Nous souhaitons analyser les atouts de chacun afin de combiner leur utilisation de manière efficace.

Schéma de résolution

Plusieurs schémas de résolution pourraient être adoptés : de type distribué, local, indépendant, ou coopératif.

Néanmoins, l'objectif industriel est de proposer un outil d'aide à la décision qui s'insère dans le cadre de la réorganisation des systèmes de régulation en France. En particulier l'objectif est de fournir un outil utilisable pour les Commandes Centralisées du Réseau. Chacune d'elles aura en charge une grande zone de régulation indépendamment (au moins dans le cadre de la définition de leurs rôles). Les interactions entre les différentes zones ne sont pas abordées dans le cadre de ce projet. Le schéma de résolution retenu ici est donc de type indépendant.

Nous soumettons cependant cette problématique dans la partie perspectives et rappelons qu'elle doit être, à minima, prise en compte dans la partie définition de la fonction objectif.

Méthodes de résolution

L'existence de méthodes exactes permet de mesurer la qualité d'une solution (et l'intérêt des méthodes qui seront testées). C'est une garantie de performance (de la solution proposée) utile en pratique pour l'aide à la décision en opérationnel. Nous souhaitons proposer un schéma efficace l'intégrant.

Les méthodes heuristiques ne sont pourtant pas à écarter du champ d'investigation car elles pourraient servir à traiter un sous-problème ou fournir une première solution. Néanmoins, nous n'aborderons pas directement cet aspect.

3.2.4 Vers la mise en œuvre industrielle

Spectre d'incidents à traiter

La gestion des circulations ferroviaires englobe un large spectre de problèmes. À titre d'exemple voici trois problèmes types, de complexités et de tailles croissantes (ces exemples ont été présentés en détail dans [27]) :

- Le problème de **fluidification** consiste à gérer préventivement les vitesses des circulations avec la plus grande précision en amont d'un point délicat du réseau afin d'éviter tout conflit (un ralentissement conséquent, voire un arrêt imposé par le système de sécurité) et ainsi maximiser le débit. C'est un cas courant sur certaines parties du réseau. Le dessin 3.8 illustre la situation : il représente le GET de deux circulations qui se succèdent et l'état des signalisations associées, comparativement avec et sans action préventive. Il faut se représenter qu'un train de fret peut mettre longtemps avant d'atteindre à nouveau sa vitesse théorique si il est ralenti voire arrêté. Il est alors probable que cela génère une cascade de ralentissements en aval.
- Le problème de **régulation** correspond à la survenue d'un aléa qui rend la grille initialement prévue irréalisable en l'état, le régulateur doit jouer des leviers d'action à sa disposition afin de limiter les conséquences de l'incident. C'est le cœur de mission de la gestion opérationnelle des circulations.
- Le problème du **détournement** ou de **coordination d'axe** apparaît suite à un incident majeur (inondation ou incendie sur un axe important, par exemple), lorsque tout ou partie des circulations doivent être détournées selon des itinéraires alternatifs (généralement parmi un ensemble prédéfini). L'illustration 3.9 présente le réseau sur lequel s'opère un détournement classique qui consiste à maintenir le trafic TGV entre la région parisienne (au nord-ouest) d'un côté, la vallée du Rhône (au sud), Lyon, Chambéry et Genève, de l'autre, suite à un incident survenu sur la LGV au nord de Lyon. Leurs itinéraires sont alors détournés sur les voies classiques qui passent par Dijon. Insérer au mieux ce nouveau flux dans la grille existante (sur le réseau classique) est une tâche ardue en pratique.

Notre outil doit pouvoir traiter un large spectre de problèmes. Tous les leviers d'action réels doivent être envisageables.

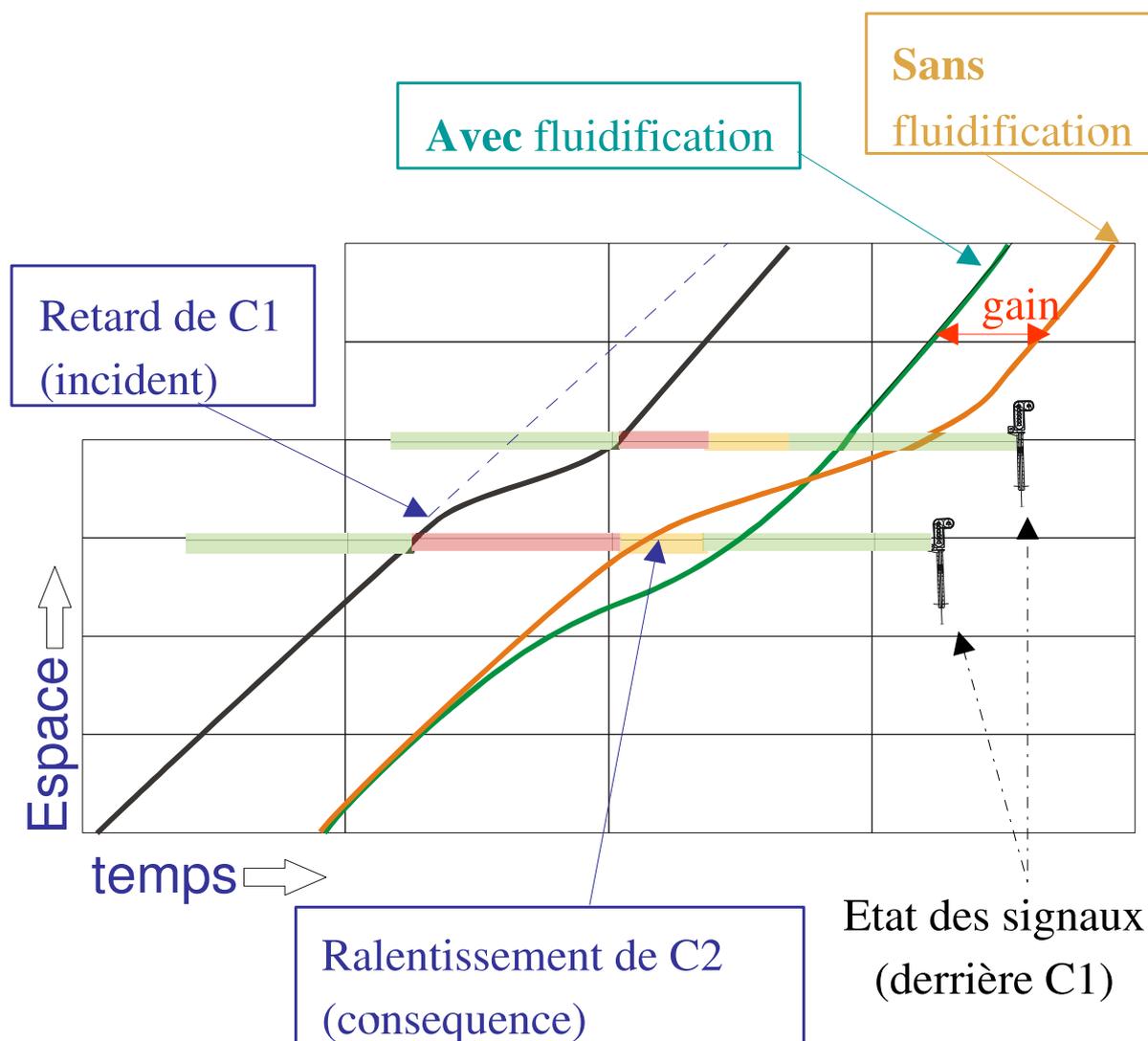


FIGURE 3.8 – Principe de la fluidification

Mode de fonctionnement

L'ambition de ce projet est de pouvoir développer un outil d'aide à la décision qui propose une solution complète réalisable. Il ne s'agit pas de développer une interface informatisée basée sur le principe interactif d'essai-erreur (tâche déjà dévolue aux outils de projection-simulation). En effet ce dernier type d'outil ne permettra pas de répondre rapidement et précisément à un problème de régulation compliqué (par exemple avec beaucoup de circulations en interaction). Or c'est précisément pour ces cas difficiles qu'un outil d'aide à la décision peut être le plus utile. Il s'agit de proposer une solution optimisée globale, bien que les paramètres métiers soient modifiables par l'opérateur.

En conséquence, le temps de calcul acceptable est celui de la prise de décision correspondante en opérationnel ; cela dépend de la gravité du problème auquel l'opérateur doit faire face. Concrètement, si on reprend les exemples précédents :

- Le problème de fluidification requiert une réponse immédiate du système (de l'ordre



FIGURE 3.9 – Exemple : LGV Sud-Est : détournement par Dijon

de la dizaine de secondes).

- Pour un problème de taille intermédiaire, plusieurs centaines de secondes représentent un délai raisonnable en général.
- Dans le cas d’une interruption grave du trafic nécessitant un détournement important des circulations, le délai de réaction peut dépasser l’heure (notamment à cause des procédures de concertation à engager). Dans ce cas il semble raisonnable de fixer une limite de temps de calcul à 5 minutes (car au delà se pose un problème majeur d’ergonomie).

Intégration industrielle

Il va de soi qu’un outil d’aide à la décision repose sur d’autres éléments que le seul cœur de calcul (le logiciel d’optimisation), de plus il doit s’insérer dans une réalité industrielle.

Il est donc nécessaire de penser à l’avance les interactions possibles avec les systèmes d’information présent et à venir. Plus important encore, il est utile d’envisager les interactions avec d’autres outils logiciels afin de rendre le système d’aide à la décision plus efficace en pratique.

ÉTAT DE L'ART

Whenever a theory appears to you as the only possible one, take this as a sign that you have neither understood the theory nor the problem which it was intended to solve.

Objective Knowledge : An Evolutionary Approach
Karl Popper

Nous allons ici passer en revue de manière synthétique les différents modèles ferroviaires et mathématiques ainsi que les techniques de résolution et les outils qui ont déjà été employés ou sont utilisés pour les problèmes de construction de grilles horaires et de replanification ferroviaire.

Les études portant sur les problématiques ferroviaires apparaissent dans les années 1970, le problème de planification est traité plus spécifiquement depuis les années 1990 et la replanification est un sujet d'intérêt relativement récent. L'évolution des techniques employées au cours de la dernière décennie apparaît particulièrement dans les revues (*surveys*) suivantes : [61] et [13, 16, 8] . On pourra aussi se référer à la revue [2] pour les études antérieures.

On remarquera que cette bibliographie se concentre principalement sur des références européennes. Cela s'explique aisément ; comme précisé dans le chapitre précédent, la problématique ferroviaire actuelle en Amérique du Nord est généralement assez différente de celle qu'on retrouve en Europe ou au Japon. Quant à ce dernier, les références accessibles sont malheureusement assez peu nombreuses. Enfin, l'Asie du Sud-Est (Chine, Inde, Corée du Sud), bien qu'en fort développement, ne présentent pas encore une littérature abondante sur les problématiques qui nous concernent.

Nous allons d'abord considérer les modèles (abstrait) du système ferroviaire, puis les principaux modèles mathématiques associés. Ensuite, nous passerons en revue les algorithmes et les techniques de résolution qui ont été appliqués aux problèmes qui nous occupent. Enfin une dernière partie propose une analyse qui servira de revue critique générale et qui nous mène à la conclusion qu'une réponse globale et complète à notre problème tel que nous l'avons défini au chapitre précédent manque encore.

4.1 Modélisation du système ferroviaire

4.1.1 Représentation métier du réseau ferré

A l'exception de quelques exercices intéressants de descriptions formelles [4], les représentations du réseau ferré réel sont rarement évoquées. Les acteurs industriels le décrivent naturellement sous la forme d'un graphe avec des arcs et des arêtes qui représentent des éléments de voies et les sommets, leurs intersections (et points remarquables).

4.1.2 Abstraction d'un modèle ferroviaire

Cette représentation abstraite s'appuie systématiquement sur le modèle précédent. En effet, on retrouve dans les représentations utilisées couramment en opérationnel des versions simplifiées de celle-ci où ne sont plus représentés que les points ferroviaires importants en pratique et les voies (ou axes) qui les relient : cf. la manchette (la partie correspondant à l'axe espace) d'un graphique espace-temps qui ne présente que les points remarquables essentiels, ou un Tableau de Contrôle Optique (tableau synoptique représentant schématiquement une zone complexe) qui peut devenir très synthétique si la zone qu'il couvre est vaste.

C'est donc aussi ce type de graphe qui est retenu dans toutes les études pour la représentation abstraite du modèle ferroviaire (cf. [55]), ceci avec un niveau plus ou moins détaillé de l'infrastructure (par exemple : [43]) préalablement définis pour chaque étude.

Choix d'une partie implicite dans la description abstraite

Le parcours d'une circulation suit une succession de ressources d'infrastructures adjacentes. La date de sortie de la ressource d'infrastructure précédente correspond donc à la date d'entrée dans la ressource qui lui succède. De même son parcours dans le graphe représentant le modèle abstrait suit une succession alternée de nœuds et de segments. Il n'est pas nécessaire a priori de décrire les horaires associés à chacun d'eux ; on choisit généralement de ne décrire explicitement qu'une seule de ces catégories (cf. [61, 55]).

Une formulation courante consiste donc à décrire l'ordre des circulations sur un segment donné à l'aide de variables binaires et les horaires d'entrée et de sortie (des segments) sous forme de variables continues. Le modèle abstrait est alors composé explicitement de segments (les sommets apparaissent de manière implicite). A l'inverse d'autres formulations (plus rares) décrivent l'ordre et les horaires d'entrée et de sortie aux nœuds, ce sont alors les segments qui apparaissent de manière implicite.

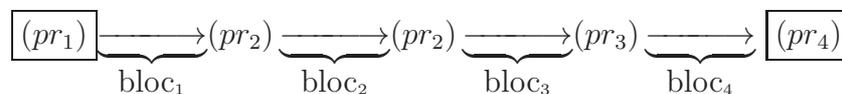
Nota bene :

- Avec le premier choix, l'écriture de contraintes différentes (par exemple durées de parcours minimales différentes) portant sur plusieurs voies parallèles (arcs ou segments non explicitement désignés dans cette description) peut être difficile, puisqu'on ne pourra manipuler que les horaires d'entrée et de sortie aux mêmes nœuds à leurs extrémités .
- Dans le deuxième cas, la représentation du fonctionnement de zones d'aiguillages ou de gares connectées à plusieurs axes de circulation est mal aisée (ce qui est courant dès que l'on considère non plus une simple ligne mais un réseau), puisqu'un segment n'a que deux extrémités.

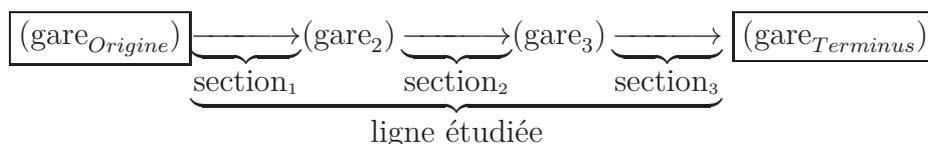
Niveau de détails

Le niveau de description retenu correspond à la catégorie d'études envisagées. Par exemple, pour une étude de type affectation des circulations à une voie (*platforming*) le graphe correspond à une description relativement détaillée d'un plan de gare, alors que pour l'étude d'un axe ferroviaire on se contente d'une vue très simplifiée. Ces différents niveaux de description sont incompatibles entre eux (et donc, souvent avec une autre catégorie d'études).

Ainsi dans quelques exemples, l'infrastructure est détaillée jusqu'au niveau (de sécurité) élémentaire -c'est à dire les blocs séparés par des points remarquables (notés *pr* dans l'illustration qui suit)- comme dans [7].



Mais, elle est aussi très souvent [12, 14] décrite à un niveau beaucoup plus simplifié. Généralement l'étude s'appuie sur un modèle composé d'un axe plus ou moins important reliant deux gares principales et possédant quelques gares intermédiaires :



Néanmoins, on assiste au développement d'approches combinant deux niveaux distincts de descriptions. Par exemple, [18] peut ainsi utiliser deux passes successives, ou [11] décrit un réseau de manière différente selon les zones (critiques ou pas).

Leviers d'actions

De même que pour le niveau de représentation, la plupart des études n'envisagent pas tous les leviers d'actions possibles, mais seulement ceux qui correspondent au cadre particulier de l'étude (par exemple le choix de voies pour le *platforming*).

Il est rare que soient réunis les choix d'horaires et de durées (i.e. de vitesse) de parcours, les choix d'ordonnancement, les choix de routage ou de détournement voire les

suppressions de trains.

Contraintes métiers

Comme le choix de représentation abstraite permet difficilement de concilier deux niveaux de représentation, la plupart des études se focalisent encore une fois soit sur la gestion des gares soit sur celle des lignes (problématique qu'on retrouve d'ailleurs au niveau métier, où -comme nous l'avons vu- les outils permettant de visualiser ces deux problèmes en même temps n'existent pas).

Les travaux présentés dans [49] permettent maintenant une approche à la fois plus générale et réaliste, car ils exposent une bonne formalisation des principes de sécurité et d'espacement (*Blocking Time Theory*).

Objectifs

Les horaires de passage (départ/arrivée) des circulations de chaque élément de l'infrastructure correspondent chacun à des variables. La recherche d'une solution consiste alors à déterminer les valeurs optimales des horaires de passage pour toutes les circulations suivant l'objectif global défini. Cependant, les variables de décision peuvent aussi concerner le choix de voie, le choix d'ordonnancement auxquelles peuvent être affectées des valeurs dans la fonction objectif.

Pour la planification horaire on retrouve couramment une fonction objectif composée, pour chaque circulation c , d'une valeur (constante) nominale π_c de la desserte à l'heure, d'une pénalité $\phi_c()$ pour l'allongement de la durée de parcours (détente horaire) ν_c , et enfin d'une pénalité $\gamma_c()$ pour le retard par rapport à l'horaire de référence μ_c (translation de l'horaire) :

$$f = \sum_c \pi_c - \phi_c(\nu_c) - \gamma_c(\mu_c)$$

4.2 Modèles mathématiques

Nous allons ici passer en revue de manière synthétique les principaux modèles mathématiques et les techniques de résolution qui ont déjà été employés pour les problèmes de construction de grilles et de replanification.

Nous avons choisi de les séparer en deux catégories :

- ceux **à temps continu** : les variables horaires sont des variables réelles, les variables d'ordonnements des variables binaires (ces dernières sont nécessaires pour traiter les contraintes disjonctives).
- ceux **à temps discret** : ici, on ne considère que des période de temps (de tailles prédéfinies et constantes) durant lesquelles des ressources d'infrastructures sont utilisées (cas des problèmes de type *job shop*) ou bien durant lesquelles un départ ou une arrivée de circulation peut se produire. Les variables horaires sont des variables binaires associées aux triplets (circulation, ressource d'infrastructure, période). Le nombre de variables est donc inversement proportionnel (pseudo polynomial) à la

taille de la période de temps de référence et proportionnel à la durée de la plage d'étude.

4.2.1 Notations générales

Si nous prenons la convention de décrire l'ordre et les horaires d'entrée et de sortie aux (seuls) nœuds, on note :

C : ensemble des circulations, notées c ,

N : ensemble des nœuds, notés n ,

$N(c)$: ensemble des nœuds pouvant être empruntés par la circulation $c \in C$,

$C(n)$: ensemble des circulations pouvant utiliser le nœud $n \in N$

$\text{succ}(c, n)$: nœud succédant immédiatement à $n \in N(c)$, dans le sens de circulation de $c \in C$.

$C^{\text{corres}}(r)$: ensemble des couples de circulations en correspondance au nœud $n \in N$,

$c_1 \prec c_2$: relation de succession entre c_1 et c_2 (c_1 est avant c_2),

4.2.2 Formulations en temps continu

Les travaux [64] et [22] sont basés sur des formulations en temps continu. Nous présentons ici celle utilisée dans [22].

Variables

D'une part ce type de formulation présente des variables horaires associées à chaque évènement.

$$\forall c \in C, \forall n \in N(c)$$

$a_{c,n} \in \mathbb{R}^+$: horaire d'arrivée de la circulation c au nœud n ,

$d_{c,n} \in \mathbb{R}^+$: horaire de départ de la circulation c au nœud n .

D'autre part, pour chaque nœud n des variables booléennes permettent de gérer l'ordre entre évènements pour chaque couple de circulations $(c_i, c_j) \in C(n)$, donc entre les variables horaires associées (e_i, e_j) où $e_i \in \{a_{c_i,r}, d_{c_i,r}\}$, et $e_j \in \{a_{c_j,r}, d_{c_j,r}\}$.

$$\forall n \in N, \forall c_i \in C(n), \forall c_j \in C(n)$$

$$o_{(e_i \leq e_j)} = \begin{cases} 1 & , \text{si l'évènement lié à } e_i \text{ se produit avant l'évènement lié à } e_j \\ 0 & , \text{sinon} \end{cases}$$

(i.e. $o_{(e_i \leq e_j)} = 1 \Leftrightarrow e_i \leq e_j$)

Exemple : la variable d'ordre entre l'arrivée de c_1 et le départ de c_2 du nœud $n \in N(c_1) \cap N(c_2)$ est définie par :

$$o_{(a_{c_1,n} \leq d_{c_2,n})} = \begin{cases} 1 & , \text{ si } c_1 \text{ arrive avant le départ de } c_2 \text{ au nœud } n \\ 0 & , \text{ sinon} \end{cases}$$

i.e. $o_{(a_{c_1,n} \leq d_{c_2,n})} = 1 \Leftrightarrow a_{c_1,n} \leq d_{c_2,n}$

Données

$p(a_{c,n})$: (fonction retournant la) valeur correspondant dans la fonction objectif à l'arrivée de la circulation c au nœud n à la date $a_{c,n}$,

ε_{min} : valeur minimale de durée d'espacement pour chaque contrainte (par abus de notation, afin de simplifier l'exposé, nous garderons la même notation pour toutes les contraintes, bien que cette valeur puisse être en réalité différente pour chaque catégorie de contrainte, de circulation, d'infrastructure et éventuellement chaque ordre de passage),

M : valeur arbitrairement grande (nécessaire pour les choix disjonctifs).

Contraintes du modèle à temps continu

Contraintes de durée minimale d'arrêt :

$$\forall c \in C, \forall n \in N(c)$$

$$d_{c,n} \geq a_{c,n} + \varepsilon_{min} \quad (4.1)$$

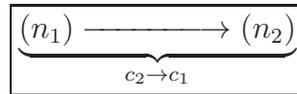
Contraintes de durée minimale de parcours dans un nœud :

$$\forall c \in C, \forall n_i \in N(c), n_{i+1} = succ(c, n_i)$$

$$a_{c,n_{i+1}} \geq d_{c,n_i} + \varepsilon_{min} \quad (4.2)$$

Contraintes de sécurité et d'espacement entre deux nœud : deux situations peuvent se présenter.

Dans le premier cas (**succession**), deux circulations c_1 et c_2 partagent la même infrastructure délimitée par n_1 et n_2 et se suivent :



Alors, les contraintes de sécurité consistent à assurer que les espacements seront respectés au départ comme à l'arrivée (hypothèse importante : les vitesses sont homogènes sur chaque portion d'infrastructure) :

$$a_{c_2,n_2} \geq a_{c_1,n_2} + \varepsilon_{min} - M(1 - o_{(a_{c_1,n_2} \leq a_{c_2,n_2})}) \quad (4.3)$$

$$d_{c_2,n_1} \geq d_{c_1,n_1} + \varepsilon_{min} - M(1 - o_{(d_{c_1,n_1} \leq d_{c_2,n_1})}) \quad (4.4)$$

et

$$a_{c_1,n_2} \geq a_{c_2,n_2} + \varepsilon_{min} - M(o_{(a_{c_1,n_2} \leq a_{c_2,n_2})}) \quad (4.5)$$

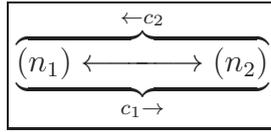
$$d_{c_1,n_1} \geq d_{c_2,n_1} + \varepsilon_{min} - M(o_{(d_{c_1,n_1} \leq d_{c_2,n_1})}) \quad (4.6)$$

et que l'ordre au départ et à l'arrivée sera le même :

$$O(d_{c_1, n_1} \leq d_{c_2, n_1}) = O(a_{c_1, n_2} \leq a_{c_2, n_2}) \quad (4.7)$$

N.B. la capacité de l'infrastructure est gérée de manière implicite grâce aux contraintes d'espacements.

Dans le deuxième cas (**conflit**), les circulations ne peuvent utiliser l'infrastructure que l'une après l'autre; par exemple une voie unique délimitée par n_1 et n_2 utilisée par deux circulations c_1 et c_2 en sens contraire :



Alors, le départ de la seconde circulation ne peut avoir lieu qu'après l'arrivée de la première :

$$d_{c_2, n_2} \geq a_{c_1, n_2} + \varepsilon_{min} - M \cdot (1 - o_{(a_{c_1, n_2} \leq d_{c_2, n_2})}) \quad (4.8)$$

et

$$d_{c_1, n_1} \geq a_{c_2, n_1} + \varepsilon_{min} - M \cdot o_{(a_{c_1, n_2} \leq d_{c_2, n_2})} \quad (4.9)$$

Contraintes de durée minimale de correspondance : La durée minimale de correspondance permet aux voyageurs en provenance de c_1 d'embarquer dans c_2 avant son départ.

$$\forall n \in N, \forall (c_1, c_2) \in C^{corres}(n)$$

$$d(c_2, n) \geq a(c_1, n) + \varepsilon_{min} \quad (4.10)$$

Synthèse de la formulation à temps continu

$$\begin{aligned}
& \min && \sum_{c,n} p(a_{c,n}) \\
& \text{s.c.} && \\
& \text{Durée d'arrêt :} && \forall c \in C, \forall n \in N(c) \\
& && d_{c,n} \geq a_{c,n} + \varepsilon_{min} \\
& \text{Durée de parcours :} && \forall c \in C, \forall n_i \in N(c), n_{i+1} = succ(c, n_i) \\
& && a_{c,n_{i+1}} \geq d_{c,n_i} + \varepsilon_{min} \\
& \text{Espacements de type Succession :} && \\
& c_1, c_2 \in C, n_1, n_2 \in N(c_1) \cap N(c_2), && n_2 = succ(c_1, n_1) \text{ et } n_2 = succ(c_2, n_1) \\
& && a_{c_2,n_2} \geq a_{c_1,n_2} + \varepsilon_{min} - M(1 - o_{(a_{c_1,n_2} \leq a_{c_2,n_2})}) \\
& && d_{c_2,n_1} \geq d_{c_1,n_1} + \varepsilon_{min} - M(1 - o_{(d_{c_1,n_1} \leq d_{c_2,n_1})}) \\
& && \text{et} \\
& && a_{c_1,n_2} \geq a_{c_2,n_2} + \varepsilon_{min} - M(o_{(a_{c_1,n_2} \leq a_{c_2,n_2})}) \\
& && d_{c_1,n_1} \geq d_{c_2,n_1} + \varepsilon_{min} - M(o_{(d_{c_1,n_1} \leq d_{c_2,n_1})}) \\
& && \text{avec} \\
& && o_{(d_{c_1,n_1} \leq d_{c_2,n_1})} = o_{(a_{c_1,n_2} \leq a_{c_2,n_2})} \\
& \text{Espacements de type Conflit :} && \\
& c_1, c_2 \in C, n_1, n_2 \in N(c_1) \cap N(c_2), && n_2 = succ(c_1, n_1) \text{ et } n_1 = succ(c_2, n_2) \\
& && d_{c_2,n_2} \geq a_{c_1,n_2} + \varepsilon_{min} - M.(1 - o_{(a_{c_1,n_2} \leq d_{c_2,n_2})}) \\
& && \text{et} \\
& && d_{c_1,n_1} \geq a_{c_2,n_1} + \varepsilon_{min} - M.o_{(a_{c_1,n_2} \leq d_{c_2,n_2})} \\
& \text{Correspondance :} && \forall n \in N, \forall (c_1, c_2) \in C^{corres}(n) \\
& && d(c_2, n) \geq a(c_1, n) + \varepsilon_{min} \\
& && \\
& \text{Avec} && \\
& \forall c \in C, \forall n \in N(c) && \\
& && a_{c,n} \geq 0 \\
& && d_{c,n} \geq 0 \\
& \forall n \in N, \forall c_1, c_2 \in C(n) && e_1 \in \{a_{c_1,n}, d_{c_1,n}\} \text{ et } e_2 \in \{a_{c_2,n}, d_{c_2,n}\} \\
& && o_{(e_1 \leq e_2)} \in \{0, 1\}
\end{aligned}$$

Une formulation de type programmation par contraintes

La gestion en temps réel d'un nœud ferroviaire est étudiée dans [58] qui propose un modèle combinant un problème d'affectation de ressources (routage) et un problème de planning.

Le **problème d'affectation** consiste à affecter aux circulations une route composée

d'une séquence d'utilisation de ressources d'infrastructure r . Ces occupations de ressources sont appelées activités. Les contraintes associées sont :

- contraintes du nombre maximum d'activités fixé (on ajoute des activités factices pour les routes alternatives qui en comporteraient moins),
- contraintes d'affectation : choix parmi des tuples de valeurs possibles (cette suite d'activités défini une route possible),
- contraintes de capacité : chaque ressource est de capacité unitaire. Comme les activités ne peuvent être préemptées, en étendant la notation pour les nœuds aux ressources $r \in R$, on a la contrainte disjonctive (analogue à 4.8 et 4.9) : $\forall r \in R, \forall c_i, c_j \in C(r)$

$$\{d_{c_i,r} > a_{c_j,r}\} \vee \{d_{c_j,r} > a_{c_i,r}\} \quad (4.11)$$

Le problème de **planning de tâches** gère les contraintes de durée de parcours.

Enfin, il est intéressant de noter que la problématique des durées d'accélération est étudiée (ajout d'activités fictives, linéarisation de la fonction d'accélération par rapport à la masse de la circulation et la valeur -durée- de son ralentissement) : retard = attente + durée d'accélération.

4.2.3 Formulation en temps discret

On retrouve cette catégorie de formulation tant en conception de grille [7, 12] ; que pour des études de replanification [59]. Elle consiste à diviser la plage d'étude en périodes t et fait intervenir des variables binaires pour associer un (ou plusieurs) évènement(s) à une (ou plusieurs) période(s) horaire.

Problème de type programmation entière

On retrouve cette formulation dans [7]. L'infrastructure y est divisée en cantons, la plage d'étude est subdivisée en périodes d'une minute. Si on utilise à nouveau la notation étendue aux ressources $r \in R$ qui désignent ici les cantons :

$$\forall c \in C, \forall r \in R(c),$$

$$x_{c,r,t} = \begin{cases} 1 & , \text{ si la circulation } c \text{ occupe la ressource } r \text{ durant la période } t \\ 0 & , \text{ sinon} \end{cases}$$

Soit T^c l'ensemble des grilles horaires possibles pour chaque circulation c prise indépendamment ; c'est à dire comme si chaque circulation était la seule à utiliser l'ensemble de l'infrastructure. Chacune de ces grilles horaires est définie par le vecteur $x^c = (x_{c,r,t})_{r,t}$, de valeur $v^c(x^c)$.

Alors la formulation du problème est la suivante :

$$\max \sum_{c \in C} v^c(x^c) \quad (4.12)$$

s.c.

$$\sum_{c \in C} (x_{c,r,t}) \leq 1, \forall r \in R, \forall t \quad (4.13)$$

$$x^c \in T^c, \quad \forall c \in C, \quad (4.14)$$

où 4.13 modélise la capacité unitaire d'un canton r .

Problème de type multiflot dans un domaine espace-temps

Cette formulation, utilisée dans [12, 14], est basée sur le graphe acyclique $G = (V, A)$ tel que :

- $V = \{(n, t)\}_{n,t}$ ensemble des nœuds $v = (n, t)$ représentant des dates possibles (période t) d'arrivée et de départ aux gares (infrastructure n). Ainsi $V = U \cup W$ où U est l'ensemble des nœuds du graphe associé aux arrivées et W l'ensemble de ceux associés aux départs. En particulier W^n et U^n désignent spécifiquement les sous-ensembles associés à la gare n et V^c celui associé à la circulation c . De plus, ρ et τ désignent respectivement les nœuds fictifs puits et source.
- $A = \{a = (c, n_1, n_2, t_1, t_2)\}$, ensemble des arcs a décrivant un parcours d'une circulation c , entre deux dates (période de départ, période d'arrivée) (t_1, t_2) entre deux gares (n_1, n_2) ou un arrêt en gare (si $n_1 = n_2$). De plus A^c représente la partition correspondant aux arcs associés à la circulation c .

Le problème consiste à trouver des chemins dans ce graphe espace-temps.

La variable x_a représente le choix d'un arc dans la solution :

$$x_a = \begin{cases} 1 & , \text{ si l'arc } a \in A^c \text{ est sélectionné dans la solution} \\ 0 & , \text{ sinon.} \end{cases}$$

Ainsi la fonction objectif est composée de profits p_a associés aux choix des arcs $a \in A$:

$$f = \sum_{a \in A} p_a \cdot x_a$$

Les contraintes *de capacité* sont portées par la définition de l'ensemble K des couples d'arcs (a_i, a_j) non compatibles. Deux arcs sont considérés comme n'étant *pas compatibles* deux à deux dans les cas suivants :

- les départs sont trop rapprochés,
- les arrivées sont trop rapprochées,
- ils correspondent à des itinéraires matériellement incompatibles (néanmoins, l'étude envisage uniquement l'interdiction de dépassement sur une même voie).

L'ensemble K est clairement de taille pseudo polynomiale (mais pas exponentielle puisque $|K| < |A|^2$).

Dans [12], les auteurs proposent d'abord la formulation multiflot suivante où pour chaque circulation $c \in C$, $\delta^+(c, n, t) \subset A^c$ et $\delta^-(c, n, t) \subset A^c$ désignent l'ensemble des arcs entrant et sortant du nœud (n, t) :

$$\begin{aligned} \max \sum_{a \in A} p_a \cdot x_a \\ \text{s.c.} \\ \sum_{a \in \delta^+(c, \rho)} x_a \leq 1 \quad , \forall c \in C \end{aligned} \quad (4.15)$$

$$\sum_{a \in \delta^-(c, n, t)} x_a = \sum_{a \in \delta^+(c, n, t)} x_a \quad , \forall c \in C, \forall (n, t) \in V \setminus \{\rho, \tau\} \quad (4.16)$$

$$\begin{aligned} x_i + x_j &\leq 1 \quad , \forall (x_i, x_j) \in K \\ x_a &\in \{0, 1\} \quad , \forall a \in A \end{aligned} \quad (4.17)$$

Une évolution de cette formulation multiflot est proposée dans [14]. Celle-ci étend la formulation dans l'espace de variables des nœuds (n, t) :

Soient $\forall (n, t) \in V$,

$$y_{n,t} = \begin{cases} 1 & , \text{si il existe un chemin comprenant le nœud } (n, t) \\ 0 & , \text{sinon} \end{cases}$$

et $\forall c \in C, \forall (n, t) \in V^c$,

$$x_{c,(n,t)} = \begin{cases} 1 & , \text{si le chemin associé à la circulation } c \text{ passe par le nœud } (n, t) \\ 0 & , \text{sinon} \end{cases}$$

Alors,

$$\begin{aligned} x_{c,n,t} &= \sum_{a \in \delta^+(c, n, t)} x_a \quad , \forall c \in C, \forall (n, t) \in V^c \\ y_{n,t} &= \sum_c x_{c,n,t} \quad , \forall (n, t) \in V^c \end{aligned}$$

et la contrainte (4.17) peut s'exprimer en la développant de la manière suivante (où ε représente la valeur minimale d'espacement) :

Conflit à l'arrivée :

$$\sum_{\bar{t} \in [t, t+\varepsilon[} y_{(n, \bar{t})} \leq 1 \quad , \forall n \in N, \forall (n, t) \in U^n \quad (4.18)$$

Conflit au départ :

$$\sum_{\bar{t} \in [t, t+\varepsilon[} y_{(n, \bar{t})} \leq 1 \quad , \forall n \in N, \forall (n, t) \in W^n \quad (4.19)$$

Interdiction de dépassement :

$$\begin{aligned} \forall c_1, c_2 \in C, \\ \forall n_1, n_2 \in N(c_1) \cap N(c_2), \quad n_2 = \text{succ}(c_1, n_1) &= \text{succ}(c_2, n_1) \\ \forall t_1, t_2, t_3, t_4 \quad \text{tels que} & t_1 < t_2 < t_3 < t_4 \\ x_{c_1, (n_1, t_1)} + x_{c_2, (n_1, t_2)} + x_{c_1, (n_2, t_4)} + x_{c_2, (n_2, t_3)} &\leq 3 \end{aligned} \quad (4.20)$$

Ainsi, la relaxation lagrangienne de ces contraintes permet d'associer les pénalités *aux nœuds* (et non plus aux arcs comme pour la relaxation lagrangienne des cliques 4.17 dans la première formulation). En pratique, la représentation utilisée dans [12] stocke uniquement les nœuds du graphe et utilise une technique de type sous-gradient et un sous problème de plus long chemin pour résoudre le programme linéaire ci-dessus de manière approchée.

D'autre part, [12] propose une évolution de la contrainte d'interdiction de dépassement (4.20), $\forall c_1, c_2 \in C, \forall n_1, n_2 \in N(c_1) \cap N(c_2), n_2 = succ(c_1, n_1) = succ(c_2, n_1)$:

$$\sum_{t \leq t_1} x_{c_1, n_1, t} + \sum_{t \geq t_1} x_{c_2, n_1, t} + \sum_{t \geq t_2} x_{c_1, n_2, t} + \sum_{t \leq t_2} x_{c_2, n_2, t} \leq 3 \quad (4.21)$$

$$\forall t_1, t_2 : t_1 \leq t_2$$

qui permet de générer moins de contraintes dans le modèle (de l'ordre de $O(c^2, n, t^2)$ versus $O(c^2, n, t^4)$ auparavant). Néanmoins, leur grand nombre nécessite encore de faire appel à un algorithme de génération dynamique des contraintes.

Une troisième formulation, adaptée aux techniques de génération de colonnes, est proposée par [9]. Elle est une évolution de la précédente ([12]) qui prend pour hypothèse supplémentaire une durée de parcours fixée pour chaque circulation (i.e. les vitesses sont constantes et fixées).

Soient P l'ensemble des chemins p du graphe espace-temps, P^c est le sous ensemble associée à la circulation c et v_p la valeur associée au chemin p . On définit la variable

$$x_{c,p} = \begin{cases} 1 & , \text{si } c \text{ utilise le chemin } p \text{ dans le graphe espace-temps} \\ 0 & , \text{sinon.} \end{cases}$$

De plus ε_n^d et ε_n^a correspondent aux valeurs minimales d'espacement entre deux départs et deux arrivées au nœud n , et $r_{(n_i, n_j)}^{c_i}$ à la durée de parcours (fixée) de la circulation c_i pour l'infrastructure (n_i, n_j) . Ainsi,

$$t_2 + r_{(n_i, n_j)}^{c_2} \geq t_1 + r_{(n_i, n_j)}^{c_1} + \varepsilon_{n_j}^a \Leftrightarrow t_2 - t_1 \geq \varepsilon_{n_j}^a + r_{(n_i, n_j)}^{c_1} - r_{(n_i, n_j)}^{c_2}$$

On définit alors la valeur d'espacement suivante :

$$\varepsilon_{(n_i, n_j)}^{1,2} = \max\{\varepsilon_{n_i}^d, \varepsilon_{n_j}^a + r_{(n_i, n_j)}^{c_1} - r_{(n_i, n_j)}^{c_2}\}$$

Alors la nouvelle formulation devient :

$$\begin{aligned} & \min && \sum_{c,p} v_p \cdot x_{c,p} \\ & \text{s.c.} && \\ \text{Unicité du chemin : } & \sum_p x_{c,p} &\leq 1 &, \forall c \in C \end{aligned} \quad (4.22)$$

Espacement aux départs :

$$\sum_{(c,p): p \ni (n, \bar{t}), \bar{t} \in [t, t + \varepsilon_n^d[} x_{c,p} \leq 1, \forall n \in N, \forall (n, t) \in W^n \quad (4.23)$$

Espacement aux arrivées :

$$\sum_{(c,p): p \ni (n, \bar{t}), \bar{t} \in [t, t + \varepsilon_n^a[} x_{c,p} \leq 1, \forall n \in N, \forall (n, t) \in U^n \quad (4.24)$$

Interdiction de dépassement :

$$\begin{aligned} & \forall c_1, c_2 \in C, \\ & \forall n_1, n_2 \in N(c_1) \cap N(c_2) : n_2 = \text{succ}(c_1, n_1) = \text{succ}(c_2, n_1) \\ & \quad \forall (t_1, t_2) : t_1 \leq t_2 < t_1 + \varepsilon_{(n_1, n_2)}^{1,2} \\ & \quad \sum_{p: p \ni (n_1, t): t_1 \leq t < t_2} x_{c_1, p} + \sum_{p: p \ni (n_1, t): t_2 \leq t < t_2 + \varepsilon_{(n_1, n_2)}^{2,1}} x_{c_1, p} \\ & + \sum_{p: p \ni (n_1, t): t_2 \leq t < t_1 + \varepsilon_{(n_1, n_2)}^{1,2}} x_{c_2, p} \leq 1 \end{aligned} \quad (4.25)$$

Avec

$$x_{c,p} \in \{0, 1\}, \forall c \in C, \forall p \in P$$

N.B. : la contrainte 4.25 est une évolution renforcée de 4.20, puisqu'elle prend aussi en compte les espacements.

Formulation de type job shop

Le partage d'une ligne à voie unique est par exemple abordée dans [47] comme un problème de type *job shop*. L'infrastructure y est constituée de ressources de deux catégories :

- les voies : ressources de capacité unitaire,
- les points de dépassement (gares) : ressources de capacité supérieure à 1 .

Les circulations ne peuvent donc se doubler ou se croiser qu'aux points de dépassement. Un conflit apparaît lorsque le nombre de circulations devant utiliser une ressource durant une même période dépasse sa capacité.

Soient :

$R = \{r_1, \dots, r_n\}$: ensemble (ordonné) des ressources r_i ,

$\text{conflicts}(r_i, t)$: nombre de conflits pour la ressources r_i à la période t ,

J : ensemble des missions,

$J_c = \{t_{c,1}, \dots, t_{c,n}\}$: mission associée à la circulation c composée d'une succession de tâches c'est à dire d'occupations de ressources,

$t_{c,i}$: occupation de la $i^{\text{ème}}$ ressource de la mission associée à la circulation c ,

$d(t_{c,i})$: horaire de départ théorique de c de la $i^{\text{ème}}$ ressource parcourue,

$\overline{d(t_{c,i})}$: horaire de départ effectif de c de la $i^{\text{ème}}$ ressource parcourue,

$p(t_{c,i})$: durée d'occupation (fixe) de la $i^{\text{ème}}$ ressource par c ,

δ_c : retard de c (translation de son horaire théorique).

Ainsi

$$\begin{aligned} d(t_{c,i+1}) &= d(t_{c,i}) + p(t_{c,i}) \\ \text{et} \\ \delta_c &= \overline{d(t_{c,0})} - d(t_{c,0}) \end{aligned}$$

L'objectif est d'affecter les tâches aux ressources, afin de faire disparaître les conflits en minimisant la somme des retards et en respectant les contraintes additionnelles.

$$\min D = \sum_{J_c \in J} \delta_c \quad (4.26)$$

La méthode suivie consiste à construire une solution réalisable selon l'algorithme suivant, où

$$horizon = \sum_{J_c \in J} \sum_{t_{c,j} \in J_c} p(t_{c,j})$$

-
- 1: **while** $\exists r_i : \text{conflits}(r_i, t) \neq 0, t \in [0..horizon]$ **do**
 - 2: considérer r_i correspondant au premier conflit
 - 3: sélectionner une tâche parmi celles en conflit pour r_i selon l'algorithme *SPT*
 - 4: l'insérer à la suite de celles qui sont déjà programmées
 - 5: **if** l'insertion est impossible dans cet ordre **then**
 - 6: revenir au choix d'une autre tâche {backtracking}
 - 7: **end if**
 - 8: **end while**
-

où l'algorithme *SPT* (*shortest processing time*) choisit une tâche parmi deux (notées t_A et t_B) comme suit :

-
- 1: *SPT*(t_A, t_B) :
 - 2: **if** $d(t_A) + p(t_A) - d(t_B) < d(t_B) + p(t_B) - d(t_A)$ **then**
 - 3: return t_A
 - 4: **else**
 - 5: return t_B
 - 6: **end if**
-

Dès qu'une solution réalisable est trouvée, on ajoute la contrainte :

$$\sum_{J_c \in J} \sum_{t_{c,j} \in J_c} \overline{d(t_{c,j})} - d(t_{c,j}) < D, \forall J_c \in J \quad (4.27)$$

où D est la valeur de la fonction objectif pour cette solution.

Variante avec des contraintes de type *no wait* et *blocking* :

Dans le cadre du projet COMBINE, [44] traite un cas plus évolué de *job shop* où :

- la machine est bloquée tant que la tâche (même achevée) ne la libère pas (*Blocking constraint*),
- les tâches doivent s'enchaîner immédiatement entre les machines (*No wait constraint*).

Variante avec des durées variables selon l'ordonnement :

L'article [38] présente une formulation avec une durée de tâche variant selon l'ordonnement choisi (deux durées; une selon chaque ordre).

4.3 Méthodes et algorithmes de résolution

Dans cette partie nous commencerons par expliquer les méthodes utilisées sur le terrain, puis nous verrons les algorithmes de résolution proposés dans les études dont les modèles ont été exposés précédemment.

4.3.1 Systèmes experts, systèmes de règles

Dans la quasi totalité des outils ou méthodes existants pour résoudre les problèmes de régulation ferroviaire en opérationnel, on retrouve des systèmes experts ou à base de règles.

Principes des règles de priorité

Une règle de priorité traite un conflit (entre deux circulations); on part donc du principe que régler les conflits deux à deux indépendamment (séquentiellement ou parallèlement) va aboutir à une solution satisfaisante pour la gestion de toutes les circulations.

La plupart des règles de priorité reposent par exemple sur les principes suivants :

- priorité chronologique (*F.I.F.O.* : premier arrivé, premier servi),
- priorité donnée à une catégorie de circulation (exemple : les rapides/voyageurs face aux lents/frets),
- priorité donnée aux circulations à l'heure (ou au contraire à celles en retard).

Ces principes de base sont souvent combinés et agencés par niveau de priorité pour former un **système de règles de priorité hiérarchisé** :

Si la règle du niveau considéré n'arrive pas à déterminer une priorité -égalité entre circulations selon le critère considéré-, on considère une règle de niveau inférieur, jusqu'à épuisement des niveaux de règle.

Panorama des systèmes de règles en Europe

Il existe une présentation des travaux du groupe Rail Net Europe [60] qui propose un comparatif des règles de priorité utilisées en opérationnel dans divers pays européens.

On y retrouve les principes de base précédents, déclinés sous plusieurs variantes et combinaisons. On remarquera que selon le pays, les choix de règles de priorité sont parfois totalement antagonistes (priorité donnée aux trains en retard versus à l'heure, par exemple). Il est probable que ces choix reposent avant tout sur des approches de principe (à noter que la structure des grilles -cycliques- à l'étranger peut introduire un biais dans l'interprétation de ces choix) et non sur une efficacité éprouvée.

Limites

Un système de règles présente plusieurs avantages :

- les règles sont relativement simples à formuler (individuellement),
- elles fournissent au régulateur un ensemble de conduites simples à respecter par défaut,
- le système de règles est aisément automatisable (ce qui est pratique pour un automate gérant une simple bifurcation par exemple ; c'est le cas d'outils tels [41, 54]).

Néanmoins, un système de règles de priorité, seul ne peut gérer une situation globale de manière optimale. Celui-ci fonctionne selon un mode programmé qui présuppose que tous les cas possibles sont envisagés et décrits. Comme ceci est bien évidemment impossible, tout les cas qui ne sont pas explicitement prévus risquent d'être traités de manière inadéquate (voire totalement contre-productive).

Autrement dit, on peut certainement construire pour chaque système de règle, une base de contre exemples qui le mettra en défaut (vis à vis de l'efficacité de la régulation).

Cette approche est clairement antagoniste à la gestion par objectif c'est à dire l'optimisation globale qui nous intéresse. Cette dernière consiste à aborder une situation sans préjuger du moyen systématique à adopter mais d'y répondre au mieux au regard de la stratégie générale retenue (le *critère* ou *l'objectif*) en prenant en compte la globalité du problème ; c'est à dire l'ensemble des interactions entre circulations.

4.3.2 Optimisation par des méthodes exactes

Ici nous proposons un rapide panorama des méthodes exactes employées pour résoudre les problèmes mathématiques de planification et de régulation.

Résolution par des logiciels d'optimisations

Parmi les possibilités pour résoudre les problèmes mathématiques formulés au début de ce chapitre, on retrouve des références qui font directement appel à des outils spécialisés dans la résolution de problèmes mathématiques d'optimisation. Ces outils ne sont absolument pas dédiés aux seuls problèmes ferroviaires, mais font preuve d'une efficacité reconnue pour résoudre des problèmes industriels (notamment grâce à leur capacité d'adaptation ou de paramétrage).

Dans cette catégorie on retrouve ILOG CPLEX qui est utilisé directement pour résoudre les modèles de type temps continu vus à la section 4.2.2, présentés dans [22, 64] par exemple.

Le modèle de type *job shop* de la section 4.2.3 présenté est implémenté avec ILOG Scheduler puis résolu via ILOG Solver dans [47]. Il en est de même pour le modèle de programmation par contrainte proposé par [58], vu à la section 4.2.2.

De type relaxation lagrangienne

Ces méthodes sont utilisées notamment dans [7, 12, 14]. Dans [7] par exemple, la relaxation de la formulation de type programmation entière présenté en section 4.2.3 porte sur la contrainte de capacité unitaire (4.13). Ainsi le problème devient :

$$\begin{aligned}
 (P_\lambda) : \\
 \max \sum_{c \in C} v^c(x^c) + \sum_{r,t} \lambda_{r,t} \cdot (1 - \sum_{c \in C} (x_{c,r,t})) \\
 \text{s.c.} \\
 x^c \in T^c, \forall c
 \end{aligned} \tag{4.28}$$

Celui-ci peut être décomposé en sous-problèmes indépendants (un pour chaque circulation c) :

$$\begin{aligned}
 (P_\lambda^c) : \\
 \max v^c(x^c) - \sum_{r,t} \lambda_{r,t} \cdot (x_{c,r,t}) \\
 \text{s.c.} \\
 x^c \in T^c, \forall c
 \end{aligned} \tag{4.29}$$

Ce dernier problème consiste alors à rechercher un chemin de profit maximum dans le graphe (acyclique) espace-temps qui définit les mouvements possibles de la (seule) circulation c .

La résolution du problème initial s'effectue en suivant deux étapes.

Dans un premier temps la méthode retenue consiste à construire une première solution (grille horaire complète) admissible. Cette première solution est obtenue en recherchant successivement un chemin pour chaque train (insertion dans la grille horaire) en suivant un ordre déterminé par des règles de priorité et en utilisant les ressources laissées libres. Au sein de chaque niveau de priorité l'ordre est défini par ordre de coût.

Dans la deuxième phase, l'amélioration de cette solution initiale repose sur la résolution du problème dual via des techniques de sous-gradient.

De type Programmation Par Contrainte

[58] propose le schéma de résolution suivant pour le modèle présenté à la section 4.2.2 :

Labelling :

- ordre suivi pour le choix des variables à tester :
 1. d'abord les variables d'affectation (aux ressources), en fonction de la taille du domaine des valeurs admissibles (ordre statique),
 2. ensuite les variables horaires (horaires de début d'occupation d'abord, les horaires de fin d'occupation sont traités ensuite en suivant le même ordre fixé précédemment), selon la valeur au plus tôt possible -*earliest due date scheduling*- (ordre dynamique)
- la première valeur testée est la valeur par défaut (planifiée initialement), il n'y a pas d'ordre défini ensuite.

Cohérence et propagation :

1. propagation des bornes horaires,
2. *arc-B-consistency* pour la contrainte disjonctive 4.11,
3. *edge finding* pour déterminer si une activité se place avant ou après un lot d'activités.

Backtracking : de type antéchronologique (quand toutes les valeurs d'une variable ont été explorées sans succès on revient sur le choix de la valeur de la dernière variable fixée).

4.3.3 Optimisation par des méthodes heuristiques

Cette dernière catégorie de méthodes s'appuie sur des spécificités du problème : structure particulière du modèle mathématique, ou savoir spécifique lié au fonctionnement du système ferroviaire utilisé pour améliorer la résolution du problème.

Hybrides

L'étude [59] est basée sur un modèle de type multiflot dans un graphe espace-temps semblable à celui de la section 4.2.3, où les variables sont les couples (période, point remarquable)

$$N = \{ \text{nœuds : (lieu, période de temps)} \} (+ \{ \text{puits} \} \text{ et } \{ \text{sources} \})$$

$$A = \{ \text{arcs : d'arrivée, de départ, de stationnement (attente)} \}$$

Elle propose trois heuristiques :

- **heuristique combinée à Cplex** : basée directement sur le solveur commercial Cplex, elle consiste à diminuer progressivement la valeur maximale admissible du retard pour chaque train, chaque fois qu'une solution réalisable est trouvée.
- **heuristique constructive basée sur la simulation à événements discrets** : l'algorithme résout les conflits dans leur ordre d'apparition chronologique en projetant dans le temps (par la simulation) les conséquences des décisions précédentes

jusqu'au prochain conflit. Quand il se trouve dans une situation de blocage (solution infaisable) il revient sur sa décision et change d'option. Plusieurs critères de sélection du train à décaler peuvent s'appliquer, dont une relaxation de la formulation précédente où on compare les résultats associés à chacune des options de choix.

- **heuristique d'énumération gloutonne** : la procédure précédente est maintenant envisagée sous l'angle d'un arbre de décision (chaque conflit est un nœud, chaque option une branche).

De type constructive

Dans [12] on trouve un algorithme glouton pour la formulation de type génération de colonnes : on fixe les horaires d'une circulation à la fois (dans l'ordre décroissant du profit lagrangien) de manière à rester compatible avec la solution partielle précédente.

De type recherche locale, voisinages

L'article [12] propose une réoptimisation séquentielle des horaires d'une seule circulation (les autres étant fixés) : chaque recherche est alors un problème de plus long chemin dans le modèle de la section 4.2.3.

Ce même article propose une autre méthode qui consiste à supprimer k circulations (celles de plus faibles profits lagrangien) puis les réinsérer de manière gloutonne.

Dans [9], k circulations (celles de plus forts profits lagrangien) sont fixées et les auteurs appliquent une heuristique constructive au problème restant.

De type Métaheuristiques

[63] présente une méthode de recherche tabou et une méthode de recuit simulé pour résoudre le problème de régulation.

4.4 Analyse

Cette dernière section présente une synthèse critique de l'état de l'art. En effet, une analyse numérique comparative de ces différentes formulations et méthodes (taille en nombre de variables et de contraintes, en qualité de formulation, ...) se révèle impossible car elles ne s'appliquent ni aux mêmes problèmes métiers, ni ne s'appuient sur les mêmes représentations.

4.4.1 Modèles mathématiques

Lors de cette revue de l'art, il manquait une formalisation exhaustive du système ferroviaire et des modèles à la fois solides et suffisamment flexibles pour s'adapter à tous les types de problèmes métiers que recouvre la replanification. Les modèles mathématiques développés reposaient donc sur des bases intrinsèquement incomparables.

La parution récente de [49] et nos travaux où sont posées en partie ces bases permettront de meilleures comparaisons dans l'avenir.

4.4.2 Algorithmes et méthodes de résolution

La synthèse précédente montre le vaste champs d'algorithmes et de méthodes adoptés pour la résolution des problèmes de planification et de replanification.

Malheureusement, comme le souligne [62] leurs mises en œuvre et leurs domaines d'études diffèrent grandement. Effectivement, il manque dans le domaine ferroviaire, une large base commune de problèmes ferroviaires de type Miplib [46] qui permettrait une comparaison de l'efficacité de chaque méthode proposée sur des problèmes reconnus (cependant, comme chaque système ferroviaire propose ses particularités, on conçoit qu'une telle librairie soit difficile à mettre en œuvre et convienne à tous).

Par ailleurs, le manque d'uniformisation des descriptions et modèles ferroviaires pose aussi un problème de comparabilité, même entre études traitant du même problème métier comme le montre [65].

Pourtant des pistes intéressantes sont menées dans le sens de la standardisation, notamment en termes de modèle de données comme dans le cas de l'initiative RailML [56]. Le présent travail entend aussi proposer une vision exhaustive (et flexible) de la description du modèle ferroviaire ainsi qu'un modèle de données associé (EIFEL).

4.4.3 Mises en œuvre, domaine d'étude

Conception et replanification

Dans le cadre de la conception de grilles horaires, la programmation linéaire associée à des *solvers* commerciaux ([64] ou DEMIURGE comme nous le verrons plus loin) peut répondre convenablement aux attentes actuelles, bien que la formulation elle-même du problème nécessite pour cela une attention particulière. En effet, on obtient notamment une précision horaire qui n'est même pas atteignable en pratique par le système actuel de contrôle-commande¹, tout en prenant en compte l'ensemble des contraintes ferroviaires. C'est pourtant sur ce sujet que portent la plupart des études.

Précision

La replanification, comme nous l'avons déjà évoqué, ne peut se contenter de la vision très simplifiée du réseau (approche souvent qualifiée de *capacitaire*) qui est généralement étudiée car les problèmes apparaissent souvent localement sur des zones ferroviaires complexes, ou des gares qui représentent les *goulets d'étranglement* du réseau. Une description détaillée y est nécessaire pour rendre compte de la complexité de leur fonctionnement.

De même il est nécessaire de conserver une bonne précision en termes de pas de temps (les horaires théoriques peuvent afficher aujourd'hui une précision de l'ordre de la décaseconde dans certaines zones). En effet, le trafic voyageur est très dense en Europe. Ainsi un simple ralentissement d'une circulation par la signalisation peut induire des problèmes de régulation (à cause de l'inertie des circulations qui ne retrouveront pas

1. l'objet du présent travail, étant bien de définir de **nouveaux** outils pour les **futurs systèmes** de contrôle-commande

leurs horaires théoriques avant que cet incident mineur ne se propage) comme exposé au paragraphe 3.8 ou détaillé dans [27].

C'est la raison de l'intérêt croissant en phase de construction pour les études concernant la robustesse des grilles face à l'aléa (comme le projet européen ARRIVAL² [1]).

De nouvelles réponses émergent d'ailleurs qui prennent en compte cette problématique : couvrir une zone relativement vaste mais gérer aussi certaines parties dans le détail. En ce sens les derniers travaux [18, 11] sont assez représentatifs. L'un [18] propose deux passes ; la première englobant un vaste périmètre la seconde assurant une précision suffisamment détaillée. L'autre [11] part sur un autre postulat : diviser le réseau en zones de compressions et de compensations : c'est à dire qui gère dans le détail les zones critiques (zones ferroviaires complexes, gares) et de manière plus simplifiée les zones moins critiques (lignes). Pourtant les différents niveaux de description sont prédéfinis, de plus l'interface entre les deux descriptions pourrait poser des problèmes (retour au niveau global pour le premier, précision au passage entre deux zones pour le second)

enfin, il est intéressant de constater que les problématiques de l'ajustement des espaces en fonction des vitesses dans l'un et des profils dans l'autre sont maintenant abordés.

Topologie de réseau

Les études réalisées sont souvent basées sur une simplification assez réductrice : l'étude d'une ligne indépendamment du reste du réseau ferroviaire (voire sans voie partagée). Or, ce sont précisément les interactions du réseau ferroviaire qui sont difficiles à gérer et plus précisément les zones dépendant d'intersections entre infrastructures aux caractéristiques ou aux topologies différentes.

On peut, par exemple, mentionner le cas des zones où une LGV, une ligne classique (fret et grandes lignes), et des lignes régionales (TER) se rejoignent.

Bien sûr, l'étude de grandes gares (comme dans [18] qui traite d'un nœud ferroviaire important du réseau hollandais) est un pas important dans cette démarche vers des études de plus en plus réalistes du point de vue métier.

Homogénéité des circulations

L'étude du seul trafic de fret (problématique typiquement nord américaine), d'un trafic de banlieue, voire d'une ligne dédiée uniquement à la grande vitesse (LGV) induit un biais. En effet les circulations se comportent de manière extrêmement homogène (vitesses très semblables) : la combinatoire des conflits potentiels est moins élevée que dans le cas d'un trafic mixte impliquant par exemple des rattrapages et des dépassements.

Or c'est évidemment dans le cas difficile des trafics mixtes que les outils d'aide à la décision peuvent être les plus intéressants.

N.B. : en ce sens, les grilles cycliques ont aussi un effet d'homogénéisation des circulations, car même si le matériel employé est différent selon l'affluence, les durées de parcours (donc les vitesses) restent identiques.

2. Algorithms for Robust and online Railway optimization : Improving the Validity and reliability of Large scale systems.

Structure des grilles horaires

- L'étude des grilles purement cycliques [42, 52, 39] diffère relativement de notre sujet :
- elles introduisent une réduction drastique de l'ensemble des solutions admissibles (on combine des séries de trains et non plus l'ensemble des couples de trains entre eux).
 - les valeurs arbitrairement grandes (M), utilisées pour décrire des contraintes disjonctives du modèle continu peuvent aussi être plus facilement bornées (par la valeur correspondant à la durée du cycle), ce qui renforce d'autant la formulation.

En pratique, une règle de régulation appliquée consiste d'ailleurs à supprimer une circulation dont le retard dépasse un (faible) seuil qui compromettrait la desserte prévue (ce seuil est donc très inférieur à la durée d'un cycle).

Malheureusement, les travaux portant sur ce postulat sont nombreux, ce qui est compréhensible puisque de nombreux systèmes ferroviaires étrangers fonctionnent selon ces principes industriels.

Vitesses

Considérer que les vitesses sont fixées reviendrait à s'interdire le principal levier d'action sur le terrain, tel qu'on le conçoit jusqu'ici en France.

Mais c'est aussi introduire une simplification drastique du problème mathématique : dans le cas du temps discrétisé, pour joindre le nœud A au nœud B , supposons qu'on ait n_A créneaux horaires de départ et n_B créneaux horaires d'arrivées possibles. Si les vitesses ne sont pas fixées on doit choisir parmi $(n_A \times n_B)$ arcs possibles, contre seulement $\min(n_A, n_B)$ dans le cas contraire. En pratique, le passage de l'un à l'autre des problèmes relève donc plus que d'une simple subtilité dans la mesure où les problèmes mathématiques associés n'ont pas du tout la même taille. Néanmoins, il faut reconnaître que dans le cas d'un système cadencé, un objectif légitime de la régulation peut consister à limiter le nombre de circulations dont le cycle sera perturbé (donc à préserver au maximum des durées de parcours -i.e. des vitesses- identiques au plan de déplacement initial), cette simplification est alors légitime.

Nous avons déjà dit que dans un souci de précision, pour des études locales au moins, la problématique des espacements dépendant des vitesses et celle des profils de vitesse (d'accélération) réalistes sont des aspects qu'il est bon de pouvoir gérer. A ce jour, seuls [58, 11] semblent avoir pris en compte ce problème.

Données

Le secteur ferroviaire souffre du manque de données accessibles sous forme numérique et plus généralement de l'accès à des systèmes d'informations modernes. Dans ce cadre, il est d'autant plus raisonnable d'envisager -à court terme- des outils qui ne reposent pas trop sur des données exhaustives sauf à faire le pari de la conjonction d'une refonte des systèmes d'information avec le déploiement rapide de nouveaux systèmes de contrôle-commande.

L'accès à des données statiques décrivant l'infrastructure et les matériels roulant semble un postulat minimum, cependant indisponible à ce jour au format numérique.

4.4.4 Conclusion

On trouvait peu de références qui correspondent à la problématique du cas général au début de ce projet. De plus, ce sont le plus souvent des approches pour la conception de grille qui ont été étudiées. Pourtant notre thématique commence à se faire jour (voir les projets européens ARRIVAL [1] et IAROR [32] notamment) et il est intéressant de constater que les problèmes soulevés ici sont aussi l'objet de recherches similaires à l'étranger.

Bien qu'aucune étude ne rassemble toutes les conditions que nous avons évoquées au chapitre précédent, ou encore que beaucoup concernent des problématiques spécifiques comme celles des grilles cadencées pour le trafic voyageur en Europe, il faut noter néanmoins une volonté forte de faire correspondre les solutions à la réalité du problème métier qui se pose. Notamment grâce aux travaux de revue critique de J. Tornquist [62], l'effort de formalisme et de synthèse des travaux récents de l'équipe réunie autour des professeurs I.A. Hansen et J. Pärchl [50], mais aussi particulièrement du projet Roma et des travaux de A. D'ariano [20] et F. Corman [17] d'une part, ainsi que ceux de G. Caimi [10] et M. Fuchsberger de l'autre.

MODÉLISATION

Science may be described as the art of systematic over-simplification - the art of discerning what we may with advantage omit.

The Open Universe : An Argument for Indeterminism Karl Popper,
(1992), p. 44

5.1 Représentations du système ferroviaire

5.1.1 Représentation métier

Au niveau le plus élémentaire, l'infrastructure du réseau ferroviaire est composée de sections de voies, par exemple des cantons. Ces voies se rejoignent en des points remarquables (aiguillages, joints de zone, ...). Le réseau ferroviaire physique est donc naturellement représenté sous la forme d'un graphe non orienté¹.

Ainsi la représentation naturelle (métier) du réseau ferroviaire correspond à un graphe où les nœuds représentent les points remarquables du réseau et les arcs les sections de voies :

$$G_{\text{métier}} = (Pr, Sv) \quad (5.1)$$

avec $Pr = \{\text{points remarquables}\}$
et $Sv = \{\text{sections de voies}\}$

Par exemple la figure 5.1 présente la représentation d'une partie de l'infrastructure correspondant à la gare TGV Lyon St Exupéry.

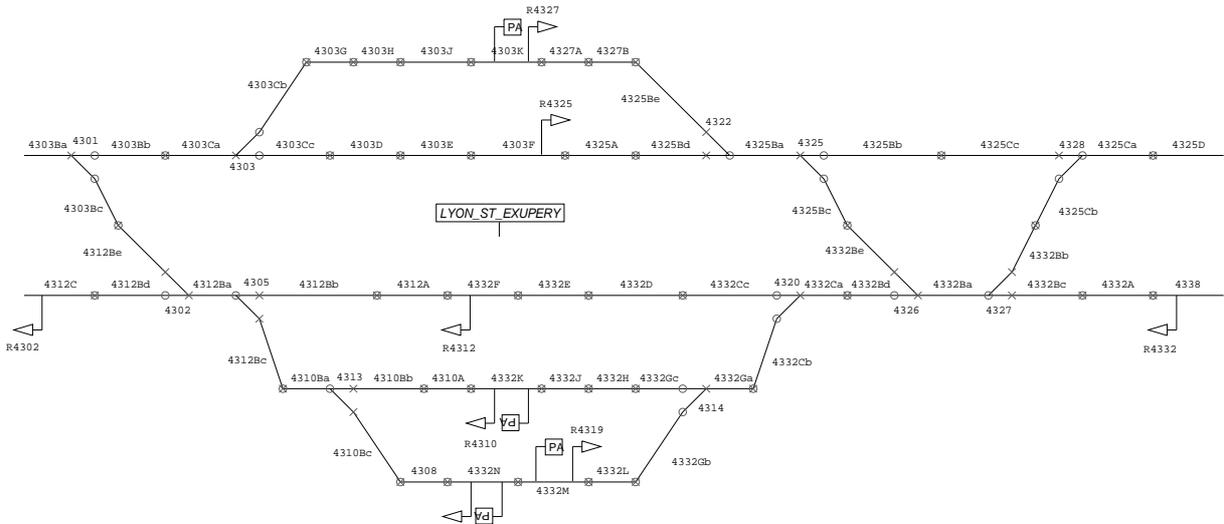


FIGURE 5.1 – Représentation métier de la gare TGV Lyon St Exupéry

Cependant, cette infrastructure ferroviaire peut être décrite plus ou moins finement selon les données disponibles et l'utilité escomptée. En effet, certaines zones critiques doivent être décrites exhaustivement ; par exemple les nœuds ferroviaires complexes dont le fonctionnement régit l'ensemble du sous réseau attenant. Inversement, décrire exhaustivement l'ensemble du réseau n'est pas systématiquement nécessaire pour l'obtention de solutions intéressantes.

Ce travail de description abstraite de l'infrastructure est effectué par un expert qui se base sur son analyse des conflits probables et de la topologie du sous réseau considéré.

1. l'orientation est portée par les itinéraires et la signalisation

Les données nécessaires concernant l'infrastructure sont renseignées pendant cette même étape. Les contraintes métiers vont donc découler de cette phase de modélisation. Elles seront ensuite traduites en contraintes mathématiques dans les formulations que nous allons définir par la suite.

5.1.2 Représentations macroscopiques et microscopiques :

Définitions

La **représentation microscopique** du réseau nécessite une description des ressources élémentaires. Ces ressources ont une capacité unitaire : elles ne peuvent être occupées que par une circulation à la fois.

Exemple : une ressource représente un canton.

La **représentation macroscopique** correspond à une description où les éléments de base sont des agrégations. Les ressources ont une capacité finie (non nécessairement unitaire). Par exemple, une ressource représente une succession de sections de voies, ou de cantons, voire un nœud ferroviaire complexe.

Remarques :

- la description du réseau n'est pas nécessairement homogène ; certaines parties peuvent être représentées sous forme macroscopique, d'autres sous forme microscopique (typiquement la pleine voie pour la première, les nœuds ferroviaires complexes et autres points bloquant du réseau pour la seconde).
- la représentation macroscopique recouvre tous les niveaux de représentation possibles qui sont supérieurs au niveau élémentaire (microscopique). Ainsi une ressource de la représentation macroscopique peut être encore agrégée ou bien décomposée.

Critères de choix

Selon le type d'étude, des niveaux de description différents peuvent être adoptés. En effet, plus la description de l'infrastructure est précise, plus la dimension (périmètre géographique et plage temporelle) de l'étude envisagée doit être réduite.

Par exemple, les projets suivants proposent nécessairement une précision inversement proportionnelle à l'étendue des zones considérées (cf. [27] pour une présentation détaillée) :

1. La fluidification d'un nœud ferroviaire complexe nécessite la gestion fine des vitesses dans un nœud ferroviaire.
2. La re-planification opérationnelle requiert un niveau de précision intermédiaire et une bonne vision des interactions entre différentes zones.
3. Le re-routage de circulations sur de nouveaux itinéraires se contente de description macroscopique étant données les vastes zones considérées (grandes lignes).

De plus, comme nous l'avons déjà souligné, une étude de planification horaire peut se contenter d'un niveau de description assez macroscopique (la solution sert de guide à la réalisation précise par un expert) alors que la replanification nécessite un niveau de

précision plus proche du modèle microscopique, car la solution doit être immédiatement (et effectivement) réalisable.

D'autre part, au sein d'une même étude, plusieurs niveaux de descriptions doivent pouvoir coexister. En effet, il peut s'avérer inutile de décrire très finement certaines sous-parties du réseau étudié, ce sont surtout les zones couvrant des goulets d'étranglement qui doivent être décrites en détails (cf. [11]).

On peut aussi imaginer une description dynamique dont la finesse de représentation puisse varier pour une étude de re-planification, car on ne connaît pas nécessairement le niveau idéal *a priori* pour tout type d'incident. En cela, le couplage à la simulation² en temps réel de nos solutions pourrait permettre d'identifier automatiquement les zones où les changements de représentation sont nécessaires. Nous verrons l'intérêt de ces synergies à la section 7.2.4.

Enfin, le choix du niveau de représentation peut être guidé par l'intérêt mathématique résultant de la formulation associée. Ainsi, dans une zone d'aiguillages très complexe (comme une arrivée en gare), l'utilisation d'une description microscopique liée à utilisation de ressources unitaires représente une combinatoire beaucoup moins importante que la gestion de l'ensemble des itinéraires incompatibles (telle qu'elle était utilisée avec une représentation macroscopique).

En effet, dans le modèle microscopique où les ressources ont une capacité unitaire, la formulation des contraintes peut être simplifiée, comme nous le verrons plus en détail dans la partie 6.5.

Les limites de la représentation naturelle

Certains systèmes de signalisation ne reposent pas expressément sur des cantons fixes mais sur le concept métier de **ressources flottantes** ou *moving blocs* : les circulations doivent maintenir *en permanence* des espacements de sécurité ; c'est le cas de la future norme ERTMS3. Ceux-ci ne sont pas modélisables avec des ressources de capacité unitaire comme pour les cantons.

On modélise ainsi le fonctionnement des LGV sur le principe d'une description macroscopique (ressource partagée) avec des contraintes de succession et de respect de l'ordre.

On notera que la représentation naturelle de l'infrastructure ferroviaire ne permet pas d'exploiter **plusieurs niveaux de représentation**. En effet, dans la représentation naturelle, on mélange généralement des objets de natures différentes. La cohérence n'étant pas respectée, les opérations d'agrégation ou de désagrégation ne sont plus possibles dans le cas général.

La figure 5.2 illustre différentes représentations du trajet du TGV 9835. Ainsi, le nœud rouge dans le premier graphe n'est pas de la même nature que les nœuds gris (ressources gares qui sont des agrégations de voies). C'est un point remarquable du réseau : la bifurcation **des Angles** (où se sépare la ligne vers Marseille ou Perpignan) , représentée plus en détail dans le second graphe. De même le dernier graphe représente ce même

2. par essence, nécessairement liée à une description microscopique

parcours divisé en deux parties : d'une part la LGV entre Bruxelles et Nîmes, et d'autre part la ligne classique entre Nîmes et Perpignan ; l'opération correspondrait à l'agrégation des nœuds et des arcs qui composaient la partie LGV du premier graphe.

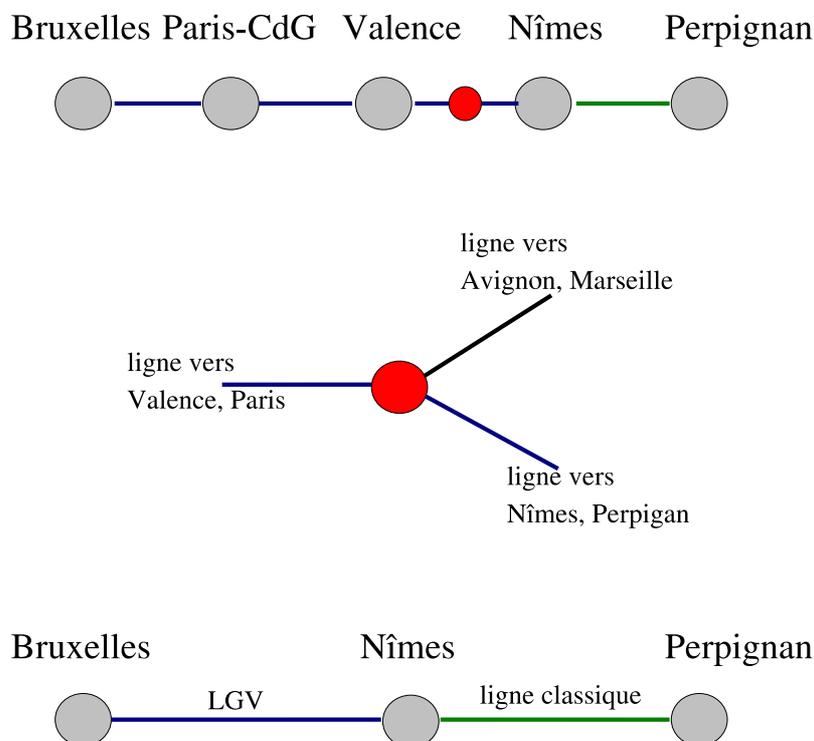


FIGURE 5.2 – Différents niveaux de représentations pour le trajet du TGV 9835

D'autre part, si la représentation naturelle (dans laquelle les arcs sont des ressources voies, les nœuds sont les points remarquables les liant) est adaptée à la description microscopique du réseau (cf. la représentation de la gare de Lyon Saint Exupéry sur la figure 5.1) elle ne permet pas de gérer la représentation agrégée de certaines ressources. Comment représenter sous forme agrégée une gare avec plusieurs axes de circulation ? La figure 5.3 qui représente la gare de Bordeaux avec les axes Paris, Toulouse et Hendaye illustre ce problème.

Dans un modèle macroscopique la gare est une agrégation de ressources de type voies (quais, etc), et non une agrégation de points remarquables, car elle n'est pas ponctuelle (des contraintes de durée de parcours s'y appliquent). De plus, une gare peut avoir plusieurs liaisons. Pour rester cohérent entre les différents niveaux de représentation, les ressources *section de voie* et *gare* doivent donc être représentées de la même manière. Or autant une voie n'est généralement liée qu'à deux autres ressources (aux extrémités), autant une agrégation de voies (comme une gare) peut être liée à plusieurs ressources simultanément. Il est donc cohérent de choisir une représentation sous forme de nœuds pour les ressources de type voie ou gare, c'est à dire les ressources d'infrastructure requises par les circulations.

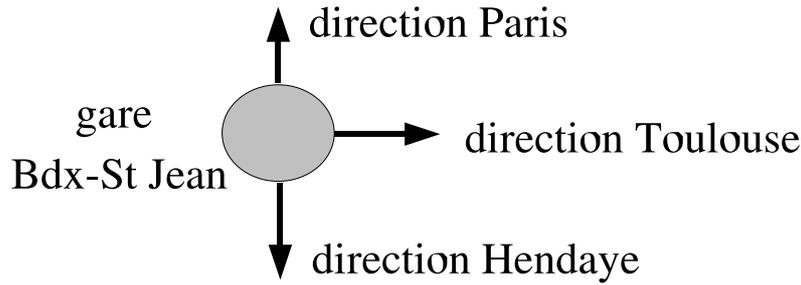


FIGURE 5.3 – Axes de circulation en gare de Bordeaux - Saint Jean

5.1.3 Une nouvelle représentation de l'infrastructure

Nous allons à présent présenter les choix retenus pour représenter le modèle ferroviaire.

Choix de représentation

Dans sa représentation abstraite (sous forme de graphe), l'infrastructure est représentée à l'aide des **objets élémentaires génériques** suivants :

1. les sommets représentent chaque ressource d'infrastructure,
2. les arcs lient les sommets; elles représentent une relation de connexité entre les ressources.

Ainsi l'abstraction du réseau ferroviaire se représente sous la forme du graphe de connexité comme illustré sur la figure 5.4 :

$$G_{\text{connexité}} = (R, Pm) \quad , \quad (5.2)$$

où $R = \{\text{sommets}\} = \{\text{ressources d'infrastructure}\}$
 et $Pm = \{\text{arcs}\} = \{\text{relation de connexité}\}$

NB : à chaque arc, on peut associer un point de mesure (pm) qui permet de mesurer la date de passage de la circulation entre deux ressources.

Un tel formalisme permet de s'adapter aux différents niveaux de description. Par exemple, il permet de traiter un réseau représenté à l'échelle du canton (description microscopique) comme sous forme agrégée (description macroscopique). La représentation choisie peut prendre toutes les formes intermédiaires et permet d'intégrer une description hétérogène du réseau.

Agrégations

Ce paragraphe traite quelques exemples d'agrégations et principalement l'évolution des propriétés des ressources considérées. Nous nous intéresserons en particulier au débit (quantité maximale de circulations pouvant traverser la ressource par unité de temps), noté $Debit(r)$ et à la capacité (quantité maximale de circulations présentes en même temps dans la ressource) notée $Capacite(r)$. NB : ici la relation de connexité (arc) entre deux ressources r_1 et r_2 est notée $Fr(r_1, r_2)$.

Pour commencer, nous allons considérer les **agrégations en série**, de type itinéraire :

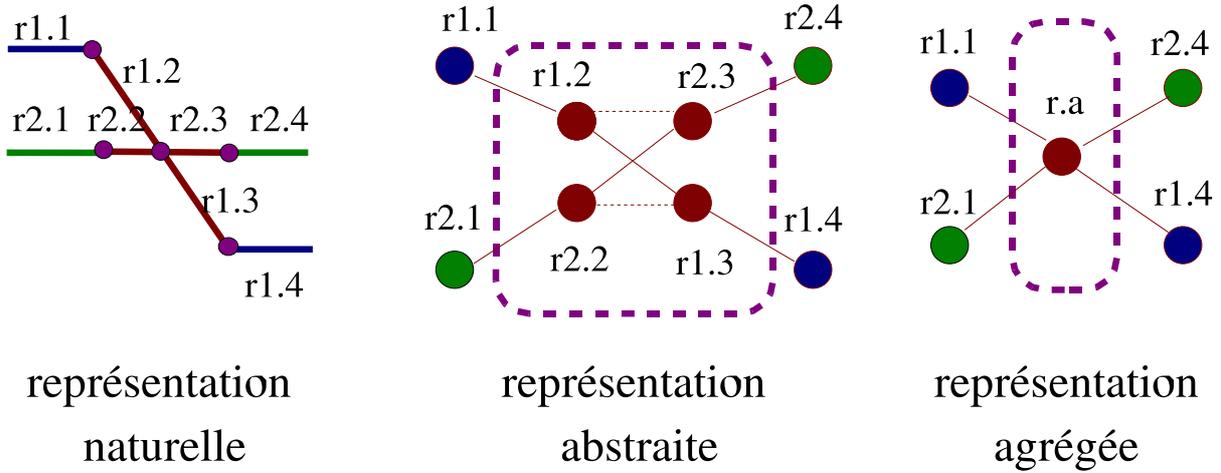


FIGURE 5.4 – Illustration du graphe de connectivité

1. agrégation des ressources d'infrastructure : l'agrégation en série correspond au concept d'itinéraire ; c'est une liste ordonnée de ressources d'infrastructure. Ainsi, sur l'exemple de la figure 5.5, on a : $r_a = \cup_i r_{2.i} = \langle r_{2.1}, r_{2.2}, r_{2.3} \rangle$
2. agrégation des points de mesure : les points de mesure reliant la ressource agrégée aux autres sont les mêmes que ceux de chacune des ressources frontières qui constituent l'agrégation. Exemple : ici, $Fr(r_1, r_a) = Fr(r_1, r_{2.1})$ et $Fr(r_a, r_3) = Fr(r_{2.3}, r_3)$.
3. Hypothèses :
 - (a) $Debit(r_a) = \min_i Debit(r_{2.i})$
 - (b) $Capacite(r_a) = \sum_i Capacite(r_{2.i})$
 - (c) les espacements sont de type succession,
 - (d) il n'y a pas de changements d'ordre possibles.

Nous allons à présent voir le cas des **agrégations en parallèle** :

1. agrégation des ressources d'infrastructure : l'agrégation en parallèle correspond à l'agrégation de plusieurs voies parallèles.

Exemple de la figure 5.6 : $r_a = \cup_i r_{2.i} = \begin{cases} r_{2.1} \\ r_{2.2} \\ r_{2.3} \end{cases}$

2. agrégation des points de mesure : les points de mesure doivent fusionner en un seul. Exemple : ici, $Fr(r_1, r_a) = \cup_i Fr(r_1, r_{2.i})$ et $Fr(r_a, r_3) = \cup_i Fr(r_{2.i}, r_3)$.
3. Hypothèses :
 - (a) $Debit(r_a) = \sum_i Debit(r_{2.i})$
 - (b) $Capacite(r_a) = \sum_i Capacite(r_{2.i})$
 - (c) les espacements sont de type succession,
 - (d) les changements d'ordre sont possibles.

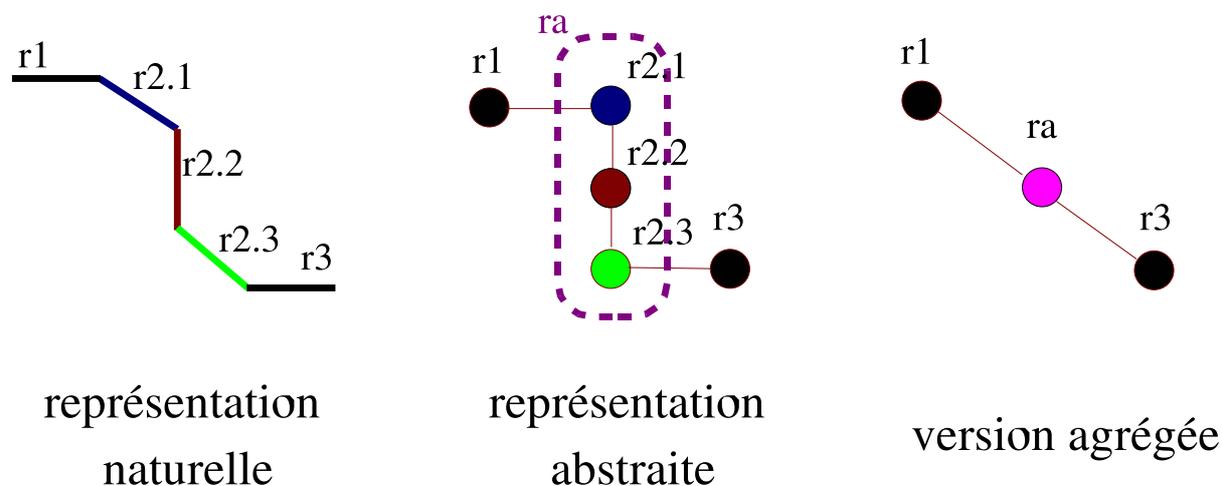


FIGURE 5.5 – Agrégation en série

Convention pour la mesure des horaires

Chaque arête est associée à un unique point de mesure qui correspond à l'horaire de passage de la tête des circulations au point limite entre les deux ressources.

Ainsi, pour décrire les mouvements des circulations, un horaire est associé au passage de chaque circulation à chaque point de mesure et un planning peut donc être transcrit sur le graphe G_r .

Remarques :

- Comme les ressources d'infrastructure se succèdent dans le parcours d'une circulation, implicitement, l'horaire de sortie de la première ressource correspond à l'horaire d'entrée dans la seconde ressource (dans le sens de parcours de la circulation considérée).
- Deux ressources sont reliées par un unique point de mesure. (Un point de mesure peut aussi être une agrégation de points de mesure : voir le cas d'une agrégation en parallèle.)

Passage au point de mesure et occupation des ressources

D'une part, il ne faut pas négliger la longueur d'une circulation. On mesure la date de passage de la tête d'une circulation; selon la vitesse et la longueur de la circulation, l'ensemble de la circulation peut quitter effectivement la ressource bien après la mesure du passage de la tête (cf. convois de fret).

D'autre part, chaque circulation continue à utiliser la ressource qu'elle occupe pendant une durée dite de **libération** (dégagement) *après* son horaire effectif de *sortie* de la ressource. De même, la ressource est réservée pendant une durée dite de **réservation** (visibilité) *avant l'entrée* effective de la circulation dans la ressource. L'espacement entre deux circulations est donc décomposé en deux parties, fonctions de chacune des circula-

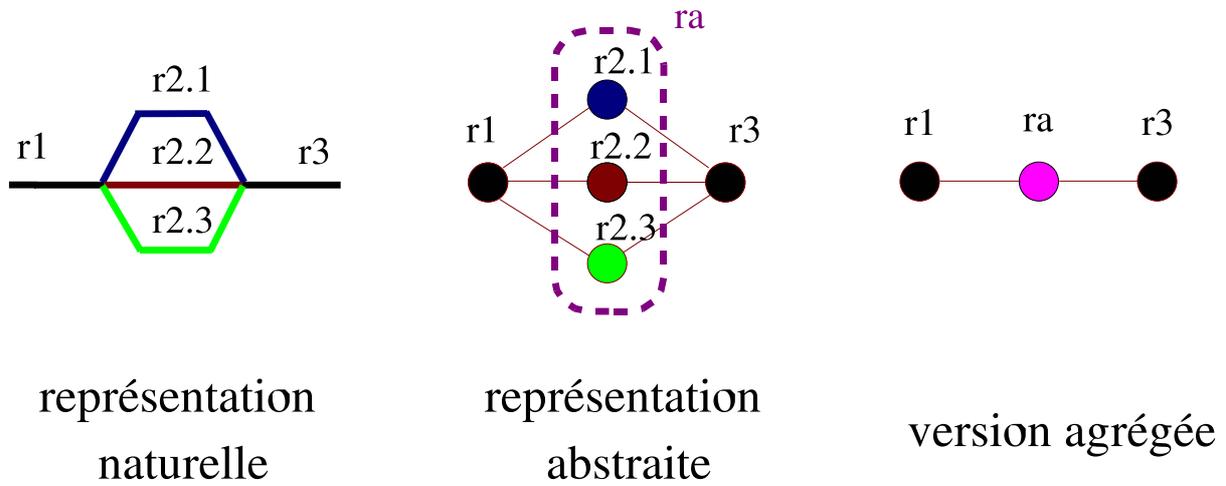
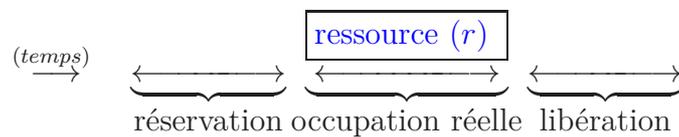


FIGURE 5.6 – Agrégation en parallèle

tions indépendamment.

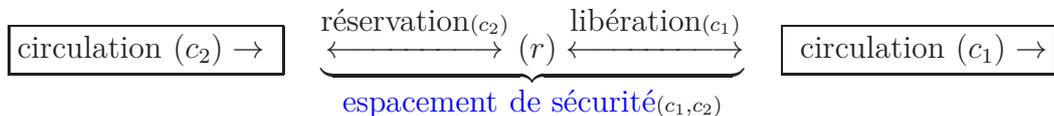
Ainsi, pour chaque circulation occupant une ressource, on a :

- séquence d’occupation d’une ressource par une circulation :



- Composition de l’espace de sécurité :

entre deux circulations c_1 et c_2 qui se succèdent immédiatement pour l’occupation de la ressource r , au moment de transition



On fait ici les hypothèses simplificatrices suivantes :

- les valeurs d’espace ε sont calculées pour la vitesse théorique de la circulation, (bien que les espacements dépendent des vitesses des circulations : de la vitesse sur la ressource suivante pour la durée de libération, sur la ressource précédente pour la durée de réservation)
- les vitesses sont supposées homogènes sur chaque arête. L’espace entre les deux circulations est donc respecté **si et seulement si** les espacements au départ et à l’arrivée de l’arête sont respectés.

Remarque : le principe de **durée d'occupation** recouvre indifféremment les notions de **durées d'arrêt** et de **durées de parcours**.

Illustration : Supposons que deux circulations de sens contraire doivent emprunter une portion de voie unique composée de cinq blocs successifs comme sur la figure 5.7.

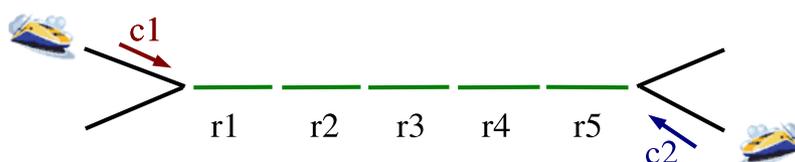


FIGURE 5.7 – Exemple : partage d'une voie unique (représentation naturelle)

Un exemple de mouvements possibles est représenté sur le graphique espace-temps 5.8 :

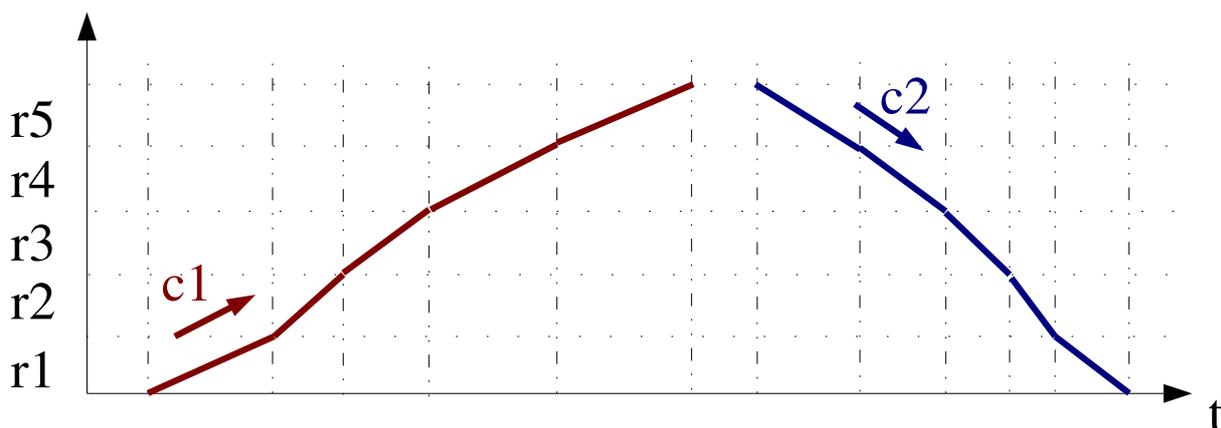


FIGURE 5.8 – GET : mouvements possibles pour le partage de la voie unique

Le graphique d'occupation des ressources correspondant est donné par la figure 5.9. Les cases représentent l'occupation de chaque ressource (qui comprend la durée de réservation, la durée de parcours de la circulation et la durée de libération).

5.2 Leviers d'action

A présent il convient d'identifier quels moyens sont à la disposition de la régulation pour agir sur le trafic ferroviaire. Les décisions de régulation sont de plusieurs ordres et ne sont pas toutes systématiquement envisagées. Ainsi, selon la gravité du problème rencontré, les régulateurs peuvent employer des moyens plus ou moins importants. On

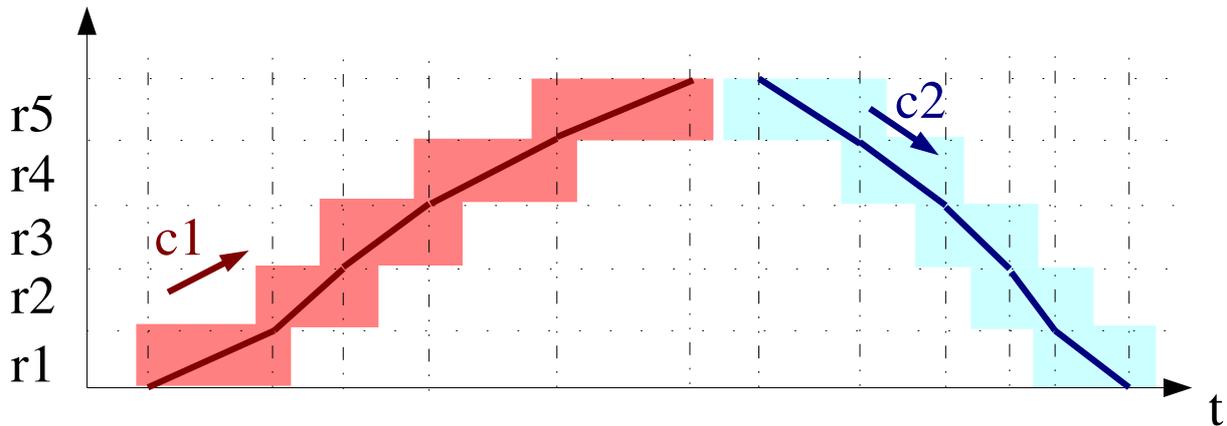


FIGURE 5.9 – Graphique d'Occupation des Ressources : partage de la voie unique

peut donc d'ores et déjà considérer que ces leviers d'action ne sont pas neutres et ne peuvent, pour certains, se justifier que par un gain appréciable en termes d'amélioration du trafic global. Autrement dit il conviendra d'évaluer un *contre coût* industriel lié à l'emploi des leviers les plus importants (cf. la section suivante sur les objectifs métiers).

Nous allons lister ces leviers d'action par ordre croissant d'importance (et donc par ordre décroissant de fréquence d'utilisation)

En premier lieu, la régulation du trafic ferroviaire consiste à **modifier les horaires**, c'est à dire les **vitesse des circulations**. En effet, c'est l'action la plus simple afin d'éviter l'apparition de conflits qui se traduit sur le terrain par (au moins) un avertissement par la signalisation ; donc un ralentissement ou un arrêt d'une des deux circulations. Or, du fait de l'inertie des circulations, la phase de réaccélération peut elle-même générer en cascade des retards supplémentaires.

Évidemment tout changement d'horaire important ne peut s'envisager sans **réordonnement des circulations**. Cela concerne directement les circulations qu'on désire faire se doubler, mais le réordonnement apparaît aussi à chaque fois qu'une ressource doit être utilisée séquentiellement comme un aiguillage ou une voie unique. Ce sont des décisions non négligeables sur le terrain ; car elles peuvent nécessiter le changement des programmations de séquences dans certains postes d'aiguillages. Ce sont aussi souvent les plus problématiques du point de vue de la résolution du problème mathématique, car elles représentent des choix disjonctifs (et ils sont nombreux).

Une autre catégorie de leviers d'action concerne le **re-routage** : soit sur des voies différentes sans changer de parcours global (par exemple : la gestion des voies ou des quais dans une gare), soit en modifiant le parcours initial en empruntant d'autres itinéraires prédéfinis (par exemple dans le cas de la coordination d'axes, ou de détournements). Nous avons décidé de regrouper ces deux aspects puisqu'il s'agit du même principe de choix des ressources d'infrastructure à emprunter, bien qu'il soit décliné à des niveaux différents.

Ensuite, la régulation peut consister à **modifier les missions-dessertes des circu-**

lations, c'est à dire des arrêts effectués par la circulation (ajouts ou retraits). On touche ici, directement aux engagements commerciaux contractualisés avec le voyageur. Ces décisions sont relativement courantes sur les dessertes de banlieue et dans les zones cadencées mais très rares pour le trafic grandes lignes. En effet dans le premier cas, on suppose que la prochaine mission (qu'on suppose avoir lieu peu de temps après) absorbera les flux qui n'ont pas été gérés. De plus, au niveau régional, l'intégration de la chaîne de décision facilite la démarche (par exemple sur le réseau Île de France, le GID local est représenté par le Centre Opération Transilien et l'opérateur est Transilien). Cela devient plus difficilement envisageable lorsque le GID et les transporteurs sont effectivement dissociés (comme sur les grandes lignes).

Enfin, le cas échéant, la régulation peut être amenée à **supprimer des circulations**, si il est impossible de les faire toutes circuler (par exemple, si un incident affectant l'infrastructure -ou une circulation longuement immobilisée- réduit la capacité de celle-ci à supporter le trafic initialement prévu). Aujourd'hui le processus qui mène à ces décisions est particulièrement lourd puisqu'en l'absence d'objectifs préétablis et contractualisés, un long processus (à l'échelle de notre problème) de négociation s'engage entre le gestionnaire d'infrastructure et les entreprises ferroviaires . En pratique, la simple réunion des différents acteurs en salle de crise peut prendre une heure.

Tous ces leviers d'action doivent donc être envisagés par la suite. Il existe des différences de traitement selon les catégories de trafic et les niveaux de décision, mais comme nous l'avons souligné dans les ambitions du projet, nous souhaitons développer un outil s'adaptant au cas général.

5.3 Objectifs métier

L'objectif général de la régulation pourrait se résumer ainsi : *maintenir le déroulement effectif du plan de transport le plus proche possible de sa définition théorique initiale*. Une autre formulation pourrait être : *minimiser les conséquences de l'incident*.

Mais comment mesurer les effets de l'incident et l'efficacité de nos actions ? Ensuite, quelles sont les options précises qui s'offrent alors aux acteurs de la gestion opérationnelle des circulations ?

5.3.1 Les outils de contrôle du système ferroviaire

Quel est l'outil de mesure (contrôle) du système ferroviaire ? Nous avons à notre disposition un outils de mesure unique³ : la date en quelques points du réseau de chaque circulation.

Comment l'utiliser ? Sur la base de cet outil de mesure, les objectifs métiers que nous détaillons par la suite peuvent s'appuyer sur :

- l'écart horaire (par comparaison avec les horaires théoriques) en chacun de ces lieux de chacune des circulations, (et, par déduction du précédent)
- la quantité de circulations retardées/détournées/supprimées,
- l'usage de l'infrastructure (taux d'occupation, capacité résiduelle, etc ...)

D'autres informations peuvent ensuite venir s'agréger à ces données de base pour enrichir le processus d'aide à la décision, par exemple :

- la quantité de voyageurs dans chaque circulation,
- la quantité de voyageurs en correspondance,
- la valeur commerciale de chaque circulation,
- les valeurs socio-économique des circulations.

5.3.2 Les principaux objectifs métiers de la régulation

La liste qui suit vise à définir les principales alternatives en termes d'objectifs métiers. Ces objectifs métiers traduisent des stratégies qui s'appuient sur les informations citées précédemment.

Viser le retour au plus tôt à la normale : il est naturel de vouloir revenir au plus tôt à un mode de fonctionnement nominal du système ferroviaire.

Limiter la propagation géographique : les conséquences de l'incident (et de sa gestion) peuvent impacter d'autres zones de régulation. De ce point de vue, une des actions à mener prioritairement peut consister à essayer de circonscrire les perturbations à la zone de régulation considérée.

3. En l'absence de géolocalisation permanente et généralisée, à ce jour.

Minimiser les changements à opérer : c'est une contrainte industrielle aisément compréhensible : tout changement possède un contre coût de mise en œuvre non négligeable (organisationnel, incertitudes sur les délais de mise en œuvre, risque d'erreurs).

Limiter le nombre de circulations impactées : pour des raisons tant opérationnelles (mise en œuvre des commandes) que pour limiter les risques d'apparition de sur-incident, il peut être préférable de minimiser le nombre de trains impactés.

Exemple : en l'absence de correspondance, il est parfois possible de limiter au seul train ayant subi l'incident les conséquences de ce retard. Exemple de stratégie simple : ce train cède la priorité à tout autre train avec lequel il rentre en conflit au cours du reste de son parcours. Une telle option est facile à mettre en œuvre, cependant elle peut mener à un énorme retard pour la circulation considérée -ce qui peut avoir de lourdes conséquences sur le reste du système si des correspondances sont prévues à son terminus-.

Limiter le nombre de passagers impactés : suivant le même principe on peut être amené à choisir la solution où le minimum de passagers sont retardés (mais il est probable que ceux qui le soient, le soient alors fortement).

Améliorer la ponctualité : on peut soit choisir de **minimiser le retard moyen** pondéré par le nombre de passagers, soit le retard maximum des circulations ou des passagers avec correspondance.

Gérer l'aspect économique : peut consister à minimiser la valeur économique du retard total (politiques de remboursements aux voyageurs) ou **minimiser les pénalités imputables au Gestionnaire d'Infrastructure** du fait des retards engendrés.

Le critère de robustesse : la solution proposée doit être celle qui a le plus de chance d'être réalisée; étant entendu qu'en cas de situation perturbée, d'autres événements peuvent survenir, de faibles variations horaires dans la mise en œuvre ne doivent pas remettre en cause la nouvelle solution de transport.

5.3.3 Options techniques concernant la mise en œuvre

Voici quelques exemples de questions à se poser concernant la comptabilisation des retards et les fonctions coûts correspondantes.

Où comptabiliser le retard ? En effet ceci n'est pas neutre en termes d'équité entre missions omnibus et directes, par exemple, si le décompte est effectué :

- au terminus de la mission (c'est la norme actuelle),
- à chaque arrêt commercial,
- à chaque point remarquable (points de mesure).

Comment comptabiliser le retard ? Ce facteur peut influencer sur la rapidité de convergence, la répartition des retards. On peut ainsi imaginer :

- les fonctions à seuils (comme celles qui correspondent à une démarche commerciale : garantie de service, ...)
- des fonctions coûts linéaires,
- des fonctions coûts quadratiques.

En termes métiers cela se traduit par des choix du type : faut-il préférer un train très en retard ou de multiples trains peu en retard ? (linéaire versus quadratique)

5.3.4 Analyse

La liste d'objectifs cités ci-dessus n'est pas exhaustive. On y retrouve pourtant les options les plus élémentaires qui sont déjà nombreuses. On notera que certaines sont difficiles à concilier voire antagonistes. Au final, les critères retenus devront probablement refléter une recherche de compromis.

Considérant qu'un consensus sur les objectifs métiers n'est pas encore atteint et dépend du problème considéré, voire de nouvelles stratégies (liées à de nouvelles offres commerciales ou à des changements de réglementation pour la gestion des circulations⁴), nous opterons par la suite pour le principe de la minimisation des écarts horaires (pondérés), laissant par la suite à l'expertise, le libre choix du paramétrage au travers de l'évaluation de ces poids.

Il est important de souligner que cette option permet par la suite de mettre en œuvre l'ensemble des objectifs métiers proposé ci-dessus.

5.4 Contraintes métiers

On peut distinguer trois catégories principales de contraintes métiers :

1. les contraintes d'exploitation sont assumées par le gestionnaire d'infrastructure. Ce sont des contraintes qui relèvent de la sécurité ou du fonctionnement même du réseau.
2. les contraintes opérationnelles sont très proches des contraintes d'exploitation, cependant nous avons fait le choix de les dissocier car elles sont assumées par les exploitants (EF) et non le gestionnaire d'infrastructure. Ce sont des contraintes qui relèvent des opérations indispensables au fonctionnement du système de production.
3. les contraintes commerciales regroupent les contraintes liées (uniquement) à la qualité du service fourni.

Nous allons d'abord détailler le contenu de ces trois catégories. Nous verrons ensuite que, du point de vue mathématique, ces contraintes appartiennent à un nombre assez restreint de classes différentes.

4. susceptible de subir de grandes évolutions avec l'arrivée de nouveaux acteurs.

5.4.1 Contraintes d'exploitation

Comme il s'agit de la sécurité et du fonctionnement même du réseau, aucune solution ne peut donc s'affranchir de les respecter.

D'une part, on retrouve les espacements de sécurité : ce type de contraintes regroupe l'ensemble des situations possibles sur le réseau : durées minimales de succession, durées de réoccupations des voies, durées de séparation sur voies banalisées, durées d'espacement entre itinéraires incompatibles . . .

D'autre part, les contraintes physiques concernent les limitations liées aux caractéristiques des infrastructures et des circulations. Les voies sont de nature hétérogène ; chacune possède donc une vitesse de circulation maximale. De même chaque catégorie de circulations possède ses propriétés dynamiques (qui sont susceptibles de changer au cours d'une mission : exemple train de banlieue plein, puis vide). Celles-ci couvrent également les situations de fonctionnement en mode dégradé telles que les travaux et perturbations affectant l'infrastructure (ralentissement temporaire aux alentours, réduction de capacité, . . .) et les incidents sur les mobiles (perte d'un bloc moteur, circuit de frein isolé, . . .).

Enfin on retrouve un ensemble de contraintes qui regroupe celles qui ne sont généralement pas exprimées car elles sont gérées de manière implicite par les opérateurs. Pourtant elles sont nécessaires pour modéliser de manière cohérente et exhaustive le fonctionnement du système ferroviaire. Il s'agit pour la plupart de contraintes liées à l'activation de leviers d'action (changement d'ordonnements, de voies, . . .).

Contraintes d'espacement : les contraintes d'espacement permettent d'assurer la sécurité dans le transport ferroviaire. On distingue deux types d'espacement. La première catégorie d'espacement correspond à la durée minimale séparant la réoccupation d'une ressource de capacité unitaire. La seconde catégorie correspond à l'utilisation simultanée d'une même ressource ; les espacements s'appliquent alors entre les entrées successives dans la ressource, tout comme entre les sorties successives de la ressource.

Les contraintes liées à une **réoccupation** sont définies pour deux circulations utilisant à tour de rôle une ressource. Il est donc nécessaire de garantir l'espacement entre la sortie de la première et l'entrée de la seconde. On retrouve ce type de situations pour les cantons ou voies en gare, la réoccupation de voies à quai, les couples d'itinéraires incompatibles entre circulations, ou encore les couples de circulations de sens contraires sur voie banalisée (ou IPCS).

Les contraintes liées à une **succession** sont définies pour deux circulations (de même sens) utilisant simultanément la même ressource (pouvant donc accueillir plusieurs circulations). Il suffit de garantir l'espacement minimal entre chaque entrée ainsi qu'entre chaque sortie pour garantir un espacement suffisant durant tout le parcours de la ressource (nous faisons l'hypothèse d'une vitesse homogène au sein de la ressource⁵). Ces situations se présentent pour des parcours de ressources agrégées (dans le cas de gestion capacitaire d'une gare ou d'une ligne), ou de zones avec des ressources flottantes (à ce

5. sinon il faut désagréger la ressource en parties de vitesses homogènes

jour uniquement les LGV avec signalisation de type ERTMS).

Contraintes horaires : une circulation, même retardée circule toujours dans une plage horaire bornée (une circulation qui aurait 18h de retard serait probablement annulée!).

Contraintes de durée d'occupation : dans le cas du simple **parcours** de la ressource, celle-ci correspond à la durée minimale nécessaire à une circulation pour parcourir la ressource. Elle est fonction des caractéristiques matérielles de la circulation et de la section d'infrastructure, mais aussi de la réglementation en place (système de signalisation).

En cas d'**arrêts**, la durée minimale d'occupation de la ressource doit intégrer la durée de stationnement mais aussi la durée des phases de décélération après l'entrée dans la ressource et d'accélération avant la sortie. Celle-ci est définie par les caractéristiques des matériels et de l'infrastructure.

Contraintes liées aux incidents : ce ne sont pas à proprement parler des contraintes mais uniquement des modifications (temporaires ou définitives) des propriétés des ressources en jeu.

Dans le cas des incidents **affectant les circulations**, il faut envisager : un arrêt (définitif ou le temps d'une réparation) ou une modification des caractéristiques des mobiles (irréversible) tels qu'une limitation de la vitesse maximale, des capacité d'accélération ou décélération, ...

Concernant les incidents **affectant l'infrastructure**, cela peut se traduire par une réduction de capacité de la ressource d'infrastructure ou une limitation de vitesse spécifique (donc un allongement du temps de parcours minimum).

Contraintes liées à l'ordonnement des circulations : dès que l'ordre des circulations n'est plus fixé, il faut assurer une cohérence dans la succession des différents évènements.

Les contraintes d'espacement s'appliquent différemment selon l'ordre des circulations. Le modèle doit être complété afin de tenir compte de l'éventuel **changement d'ordonnement** (choix disjonctif).

En cas de succession de deux circulations, la **contrainte d'interdiction de dépassement** se traduit ainsi : l'ordre des circulations doit être le même en entrée et en sortie de la ressource. Cette contrainte est complémentaire de la contrainte d'*espacement* en cas de succession.

En cas de **conflit** (pour l'occupation unitaire de la ressource), la contrainte d'espacement de type conflit exige que l'entrée de la seconde circulation ait lieu seulement après la sortie de la première.

Contraintes liées au choix d'itinéraire : dans la mesure où l'itinéraire n'est pas totalement prédéterminé (choix d'itinéraire), il faut s'assurer de la cohérence de l'option qui sera retenue.

- Les contraintes qui dépendent du choix d’itinéraire ; comme précédemment, le modèle doit être complété en conséquence : les contraintes horaire, de durée de parcours et d’espacement ne s’appliquent que si la (les) circulation(s) emprunt(ent) la ressource d’infrastructure considérée.
- Les circulations doivent être affectées de manière **unitaire** aux ressources ; c’est à dire sur (une et) une seule ressource à la fois.
- Le parcours emprunté doit être composé d’une **succession cohérente de ressources** (cette succession -agrégation- est un itinéraire -ou parcours-).

Contraintes liées à la suppression d’une circulation : si la suppression est un levier d’action possible il faut redéfinir les contraintes horaire, de durée de parcours, d’espacement et de correspondance de manière à ce qu’elles ne s’appliquent que si la circulation est bien effective.

Inertie et écarts de vitesse : si on autorise l’**arrêt imprévu** d’une circulation, on doit aussi prendre en compte dans ce cas l’allongement de la durée de parcours minimale qui modélise l’inertie de la circulation (i.e. phases d’accélération et de décélération). De plus, deux ressources adjacentes (surtout si elles sont courtes) ne peuvent être parcourues avec des **variations de vitesse** trop importantes par une même circulation à cause de l’inertie des circulations (cf. fret 1500 tonnes).

5.4.2 Contraintes opérationnelles

Puisque ces contraintes relèvent des opérations indispensables au fonctionnement du système de production, aucune solution ne peut donc s’affranchir de les respecter non plus.

Contraintes horaires : les personnels et matériels nécessaires aux opérations ne sont disponibles que dans la limite de certains créneaux horaires : ouverture des gares, disponibilité des conducteurs ou des matériels roulants.

Contraintes de durée d’occupation : les **durées minimales de stationnement** doivent être suffisantes pour faire face aux opérations techniques et aux besoins habituels du service : embarquement et débarquement de voyageurs, procédures de lancement... Elles sont déterminées pour chaque circulation, par le gérant de cette circulation, en accord avec les services locaux.

Ainsi, on doit prendre en compte les durées minimales d’arrêt pour les opérations d’accouplement, pour embarquement et débarquement, ou de retournement.

Contraintes de correspondance : une correspondance entre deux circulations nécessite une durée minimale (en deçà de laquelle l’opération d’embarquement ne pourra pas effectivement se terminer).

Les contraintes liées aux **roulements de personnel ou de matériel** se gèrent de la même manière qu’une correspondance.

Ainsi pour deux circulations associées, c'est à dire dépendant d'une navette, les (deux *numéros* de) circulations utilisent en réalité successivement le même matériel (exemple : *c'* prends la relève de *c*). Il faut donc lier les horaires d'arrivée et de départ (puisque c'est le même matériel qui change simplement d'identifiant commercial). De même si l'agent de conduite d'une circulation assure successivement le service d'une autre circulation, cela s'apparente aussi à une correspondance.

5.4.3 Contraintes commerciales

Étant liées à la qualité du service fourni, ces contraintes peuvent éventuellement être violées si il n'existe pas d'autres solutions. Cependant, comme cela correspond à un service dégradé, le choix de ces options doit être pénalisé.

Plusieurs options de modélisation mathématiques s'offrent alors : soit sous forme de contraintes classiques avec variables d'écart et pénalités dans la fonction objectif, soit via l'intégration directe dans les composants de la fonction objectif.

Ces contraintes commerciales peuvent recouvrir des contraintes de durées minimales de correspondance (confortables pour les voyageurs), d'entretiens (ravitaillements, nettoyage des rames), mais aussi des durées maximum de parcours ou d'attente liées à des engagements commerciaux.

Contraintes horaires : la circulation doit respecter ses horaires commerciaux avec une certaine marge (dans la mesure du possible, sans entraîner de pénalités).

Contraintes de durée d'occupation : la **durée maximale d'occupation** est fonction de contraintes commerciales. Exemple : sur un trajet TGV Marseille-Paris, on ne souhaite pas un arrêt supérieur à 30min en gare d'Avignon TGV.

Contraintes de correspondance : deux circulations sont en correspondance lorsque la durée écoulée entre l'arrivée de la première et le départ de la seconde est bornée entre deux valeurs fixées par le gérant de la circulation. Le gérant de deux circulations peut souhaiter définir une plage horaire pour la correspondance confortable de ses voyageurs.

La **relève** des voyageurs d'une circulation par une autre : La durée minimale entre l'arrivée de la première et le départ de la seconde doit être respectée.

L'**échange de voyageurs** entre deux circulations est une relève mutuelle.

5.4.4 Classes de contraintes mathématiques

Du point de vue mathématique ces contraintes métiers partagent pour la plupart des structures semblables. Finalement on retiendra les catégories de contraintes génériques suivantes :

1. liées aux créneaux horaires,
2. borne sur les durées d'occupation des ressources,
3. gestion des espacements de type succession et réoccupation,
4. contraintes liées implicitement aux autres leviers d'action.

Contraintes horaires : ce sont les contraintes liées au respect de l'horaire initial (en particulier une circulation ne doit pas partir d'un arrêt commercial avant l'horaire théorique).

Contraintes de durée d'occupation : elles regroupent les contraintes métiers associées aux durées d'occupation de ressources :

- durée minimale de parcours,
- durée minimale d'arrêt,
- durée maximale d'occupation,
- durées minimales d'arrêt pour les opérations d'accouplement,
- durées minimales d'embarquement et de débarquement,
- durées minimales d'accouplement, de retournement.

Contraintes d'espacement : elles recouvrent les contraintes métiers relatives aux espacements entre circulations (donc au respect des principes de sécurité et plus généralement des modes de fonctionnement des systèmes installés sur le terrain) :

- espacements de sécurité (succession),
- espacement sur conflit (réoccupation),
- respect des ordonnancements associés (interdiction de dépassement).

Contraintes implicites : elles regroupent les contraintes métiers qui ne sont pas explicitement exprimées mais qui sont néanmoins indispensables :

- modifications liées aux ordonnancements, choix de voie, et suppressions (arrêts),
- affectation unitaire,
- affectation des itinéraires (cohérence).

5.5 Données

A présent nous allons formaliser le contenu des flux de données nécessaires en amont et données résultats en aval.

5.5.1 Données requises

On distingue trois catégories de données indispensables qui concernent : l'infrastructure, les circulations et enfin celles qui permettent d'articuler le tout.

Pour la partie **infrastructure**, les données concernent d'abord la **topologie** (liste chaînée de ressources), puis les caractéristiques propres aux ressources : longueurs, vitesses maximales, pentes, type de systèmes au sol (sécurité, signalisation).

Ensuite pour la partie **circulations** les données nécessaires sont les caractéristiques propres telles que la longueur (utile pour l'affectation aux quais, le calcul des durées de libération), la masse et les capacités d'accélération.

Enfin, il y a un ensemble de données qu'on ne peut rattacher (uniquement) à une des catégories précédentes, ce sont : la grille théorique, les missions, les contraintes commerciales et les contraintes métiers s'appliquant aux circulations et ressources si elles ne sont pas directement déductibles des données ci-dessus.

5.5.2 Solution

La solution est un nouveau plan de transport : c'est à dire une grille horaire dont on peut déduire l'ensemble des commandes et en particulier les nouvelles consignes horaires et de vitesses ainsi que les ordonnancements (à reprogrammer sur certains systèmes) et les changements de mission éventuels (modification de l'information voyageurs).

FORMULATIONS

Dans ce chapitre nous commencerons par présenter les notations communes aux trois modèles mathématiques que nous décrirons par la suite.

6.1 Notations

Cette section va nous permettre de présenter les principales notations que nous utiliserons par la suite.

6.1.1 Ensembles

C : ensemble des circulations, notées c ,

R : ensemble des ressources d'infrastructure, notées r ,

$R(c)$: ensemble des ressources pouvant être empruntées par la circulation $c \in C$,

$C(r)$: ensemble des circulations pouvant utiliser la ressource $r \in R$

$Succ(c, r), Pred(c, r)$: respectivement ensemble des ressources succédant immédiatement à $r \in R$ et ensemble des ressources précédant immédiatement $r \in R$, utilisables par la circulation $c \in C$. Si ces ensembles ne possèdent qu'un seul élément, ceux-ci seront notés respectivement $succ(c, r)$ et $pred(c, r)$,

$It(c, r_1, r_2)$: ensemble des itinéraires possibles pour la circulation c depuis la ressource r_1 vers r_2 (ces itinéraires sont eux-mêmes considérés comme des ressources d'infrastructure agrégées $r_i \in R$),

Iic : ensemble des couples (r_i, r_j) d'itinéraires incompatibles,

$R^{corres}(c_1, c_2)$: ressources où il y a une contrainte métier de correspondance entre $c_1 \in C$ et $c_2 \in C$

$C^{corres}(r)$: ensemble des couples de circulations en correspondance dans la ressource $r \in R$,

$EchangeVoyageurs(r)$: ensemble des couples de circulations échangeant leurs voyageurs à la ressource $r \in R$.

6.1.2 Relations

$c_1 \prec c_2$: relation de succession entre c_1 et c_2 (c_1 est avant c_2),

$r_i \subset r$: relation d'inclusion où la ressource r_i est une composante de la ressource agrégée r telle que $r = \cup_i r_i$,

6.1.3 Données

Données horaires :

$a_{théo}(c, r)$: date d'entrée théorique de la circulation c dans la ressource r ,

$a_{réel}(c, r)$: date d'entrée réelle de la circulation c dans la ressource r .

(si la circulation est déjà entrée à la date de début de l'étude, elle remplace la valeur théorique)

$d_{théo}(c, r)$: date de sortie théorique de la circulation c de la ressource r ,

$d_{réel}(c, r)$: date de sortie réelle de la circulation c de la ressource r .

(si la circulation est déjà sortie, elle remplace la valeur théorique)

Données d'exploitation :

Capacite(r) : capacité d'accueil de la ressource $r \in R$,

$p_{(c,r)}^{théo}, p_{(c,r)}^{min}, p_{(c,r)}^{max}$: durées théoriques, minimales et maximales de parcours de la ressource $r \in R$ (comprenant les durées d'arrêt éventuel) pour la circulation $c \in C$,

$\varepsilon_{(c,r)}^l, \varepsilon_{(c,r)}^r$: durées de libération et de réservation de la ressource $r \in R$ par la circulation $c \in C$,

$\varepsilon_{(c_1, c_2, r)}$: espacement minimum entre deux circulations c_1 et c_2 , utilisant la ressource r , et telles que $c_1 \prec c_2$.

$$\varepsilon_{(c_1, c_2, r)} = \varepsilon_{(c_1, r)}^l + \varepsilon_{(c_2, r)}^r$$

Données commerciales :

$\varepsilon_{corres}^{min}(r, c_1, c_2), \varepsilon_{corres}^{max}(r, c_1, c_2)$: durées de correspondance minimale et maximale entre les circulations c_1 et c_2 associées à la ressource r .

6.1.4 Notations mathématiques

M : une valeur très grande déduite des données du problème.

Par la suite, pour chaque modèle, on définira des **variables horaires** qui permettront de calculer les horaires des circulations. Ces variables horaires seront associées à chaque arrivée ou départ d'une circulation sur chaque ressource. Les départs et les arrivées sont considérés sous le terme générique d'*événements* notés e et leurs variables associées seront notées $*$ (qui désignera alors indifféremment une variable de départ ou d'arrivée).

Nous définirons aussi des **variables d'ordonnement** entre les événements lorsque cela est nécessaire.

Enfin lorsque plusieurs ressources concurrentes peuvent être utilisées par une circulation, nous introduirons une variable de choix de ressource (ou **variable de choix d'itinéraire** i.e. de ressource agrégée).

6.2 Formulation mathématique en temps continu

6.2.1 Variables de décision

Horaires

Pour chaque circulation, pour chaque ressource parcourue, on définit :

$a_{c,r}$: date d'entrée de la circulation c dans la ressource $r \in R(c)$,

$d_{c,r}$: date de sortie de la circulation c de la ressource $r \in R(c)$, où

$$\forall c \in C, \forall r \in R(c), \quad d_{c,r} \in \mathbb{R} \text{ et } a_{c,r} \in \mathbb{R}.$$

Pour être conforme avec la convention de mesure des horaires définie précédemment, les départs et arrivées de chaque ressource successives sont liés de la manière suivante :
 $\forall c \in C, \forall r \in R(c)$

$$\begin{aligned} a_{c,r} &= d_{c,r_i} \quad , \quad \forall r_i \in \text{Pred}(c, r) \\ d_{c,r} &= a_{c,r_j} \quad , \quad \forall r_j \in \text{Succ}(c, r). \end{aligned} \tag{6.1}$$

▷ *L'horaire de sortie d'une circulation d'une ressource correspond à l'horaire d'entrée dans la ressource qui suit.*

Ordonnancement

Pour gérer l'ordre des évènements entre circulations sur chaque ressource, il faut introduire des variables booléennes. Ainsi, pour chaque ressource $r \in R$ on peut définir la variable d'ordre entre couples d'évènements pour deux circulations $(c_i, c_j) \in C(r)$, donc entre les variables horaires associées (e_i, e_j) où $e_i \in \{a_{c_i,r}, d_{c_i,r}\}$, et $e_j \in \{a_{c_j,r}, d_{c_j,r}\}$:

$$o_{(e_i \leq e_j)} = \begin{cases} 1 & , \text{ si l'évènement lié à } e_i \text{ se produit avant l'évènement lié à } e_j \\ 0 & , \text{ sinon} \end{cases}$$

(i.e. $o_{(e_i \leq e_j)} = 1 \Leftrightarrow e_i \leq e_j$)

avec les contraintes qui assurent la cohérence avec les variables horaires :

$$\begin{aligned} e_2 &\geq e_1 - M \cdot (1 - o_{(e_1 \leq e_2)}) \\ e_1 &\geq e_2 - M \cdot o_{(e_1 \leq e_2)} \end{aligned} \tag{6.2}$$

et :

$$o_{(e_1 \leq e_2)} + o_{(e_2 \leq e_1)} = 1$$

Exemple : la variable d'ordre entre l'évènement *arrivée de c_1* et l'évènement *départ de c_2* de la ressource r est définie telle que :

$$o_{(a_{c_1,r} \leq d_{c_2,r})} = \begin{cases} 1 & , \text{ si l'arrivée de } c_1 \text{ a lieu avant le départ de } c_2 \text{ de la ressource } r \\ 0 & , \text{ sinon} \end{cases}$$

(i.e. $o_{(a_{c_1,r} \leq d_{c_2,r})} = \Leftrightarrow a_{c_1,r} \leq d_{c_2,r}$)

avec :

$$\begin{aligned} d_{c_2,r} &\geq a_{c_1,r} - M.(1 - o_{(a_{c_1,r} \leq d_{c_2,r})}) \\ a_{c_1,r} &\geq d_{c_2,r} - M.o_{(a_{c_1,r} \leq d_{c_2,r})} \end{aligned} \quad (6.3)$$

Routage, choix des ressources

Pour gérer le routage des circulations entre plusieurs voies ou plusieurs itinéraires (ressources agrégées), on introduit des variables booléennes d'affectation de la circulation c pour chaque ressource r possible :

$$v_{c,r} = \begin{cases} 1 & , \text{ si la circulation } c \text{ emprunte la ressource } r \\ 0 & , \text{ sinon} \end{cases}$$

Contrainte associée au choix de ressource : il faut affecter une circulation sur (une et) une seule ressource. Ainsi, pour chaque circulation $c \in C$, et pour chaque choix d'un itinéraire r parmi $It(c, r_e, r_s)$, entre les ressource $r_e \in R(c)$ et $r_s \in R(c)$:

$$\sum_{r_i \in It(c, r_e, r_s)} v_{c,r_i} = 1 \quad (6.4)$$

▷ *cette contrainte impose un choix unique parmi plusieurs itinéraires possibles.*

Lien entre ressources agrégées : soit une ressource agrégée telle que $r = \cup_i r_i$, alors l'utilisation de la ressource agrégée équivaut à l'occupation de chacune des ressources qui la compose :

$$\forall c \in C, \forall r \in R(c), \forall r_i \subset r$$

$$v_{c,r} = v_{c,r_i} \quad (6.5)$$

6.2.2 Contraintes

L'expression mathématique de la plupart des contraintes dépend de la liberté de choix (métier) donnée pour chaque circulation ; en particulier de la possibilité d'affecter une circulation sur une autre ressource ou de changer l'ordre des circulations (donc de certains évènements). Par la suite, par souci de clarté, nous présenterons chaque contrainte en raffinant son expression de la formulation la plus simple (pas de choix de réordonnement ni de routage) vers la plus exhaustive possible.

Contraintes horaires

Horaire au plus tôt : les évènements ne peuvent avoir lieu avant leur horaire théorique (i.e. pas d'avance).

$\forall c \in C, \forall r \in R(c) :$

$$\begin{aligned} d_{c,r} &\geq d_{théo}(c,r) \\ a_{c,r} &\geq a_{théo}(c,r) \end{aligned} \quad (6.6)$$

En pratique, il s'agit de définir la valeur de la borne inférieure des variables.

Durée minimale d'occupation : c'est la durée minimale incompressible durant laquelle une ressource sera utilisée par une circulation. Elle est fonction des caractéristiques de la circulation et de l'infrastructure parcourue.

$\forall c \in C, \forall r \in R(c) :$

$$d_{c,r} - a_{c,r} \geq p_{(c,r)}^{min} \quad (6.7)$$

Cette contrainte lie les variables de départ et d'arrivée : elle garantit la cohérence des horaires vis à vis de la flèche du temps.

En cas de **choix de ressources ou d'itinéraires**, seule la contrainte associée à la ressource utilisée doit être active et les contraintes 6.7 associées à chaque choix peuvent s'écrire sous une forme agrégée. Par exemple pour une circulation $c \in C$, pouvant emprunter plusieurs itinéraires r_i parmi l'ensemble $It(c, r_e, r_s)$, l'ensemble des contraintes 6.7 associées à chaque r_i devient :

$$a_{c,r_s} - d_{c,r_e} \geq \sum_{r_i \in It(c,r_e,r_s)} v_{c,r_i} \cdot p_{(c,r_i)}^{min} \quad (6.8)$$

N.B. $\forall r_i \in It(c, r_e, r_s)$, $a_{c,r_s} = d_{c,r_i}$ et $d_{c,r_e} = a_{c,r_i}$, d'après 6.1.

Contraintes d'espacement

Du point de vue mathématique, on distinguera deux types d'espacements qui permettent de recouvrir l'ensemble des contraintes d'espacement définies dans le modèle métier :

La première catégorie correspond aux contraintes d'espacement entre circulations qui se **succèdent** lors de l'utilisation simultanée d'une même ressource ($Capacite(r) > 1$). Exemple : succession de plusieurs circulations de même sens sur une ressource partagée de type tronçon de LGV. Il s'agit de garantir le respect de l'ordre des circulations et leurs espacements en entrée et en sortie de la ressource considérée (on fait l'hypothèse de vitesses homogènes sur cette ressource).

L'autre catégorie correspond aux contraintes d'espacement liées à un **conflit** pour l'occupation exclusive d'une ressource ($Capacite(r) = 1$), consécutivement. Exemples : cantons, réoccupation d'un quai en gare ou occupation de zones (ressource agrégée) correspondant à des itinéraires incompatibles.

Successions : $\forall r \in R, \forall (c_1, c_2) \in C(r)$, les circulations c_1 et c_2 étant de même sens et $c_1 \prec c_2$:

Sans choix d'ordonnement ou d'itinéraire,

$$\begin{aligned} d_{c_2,r} &\geq d_{c_1,r} + \varepsilon_{(c_1,c_2,r)} \\ a_{c_2,r} &\geq a_{c_1,r} + \varepsilon_{(c_1,c_2,r)} \end{aligned} \quad (6.9)$$

▷ La circulation c_2 suit c_1 avec un espacement minimal imposé en entrée et en sortie de la ressource, on considère que les vitesses sont suffisamment homogènes au sein de cette ressource pour que l'espacement minimal soit respecté à tout moment (hypothèse liée à la modélisation de l'infrastructure).

Avec le **choix d'ordonnement**,

– d'une part (**contrainte de respect des espacements** selon l'ordre d'utilisation) :

$$\begin{aligned} d_{c_2,r} &\geq d_{c_1,r} + \varepsilon_{(c_1,c_2,r)} - M.(1 - o_{(d_{c_1,r} \leq d_{c_2,r})}) \\ a_{c_2,r} &\geq a_{c_1,r} + \varepsilon_{(c_1,c_2,r)} - M.(1 - o_{(a_{c_1,r} \leq a_{c_2,r})}) \\ d_{c_1,r} &\geq d_{c_2,r} + \varepsilon_{(c_2,c_1,r)} - M.o_{(d_{c_1,r} \leq d_{c_2,r})} \\ a_{c_1,r} &\geq a_{c_2,r} + \varepsilon_{(c_2,c_1,r)} - M.o_{(a_{c_1,r} \leq a_{c_2,r})} \end{aligned} \quad (6.10)$$

▷ Ces contraintes expriment les deux cas possibles selon l'ordre des circulations.

– et d'autre part (**contrainte d'interdiction de dépassement**) :

$$\forall r \in R, \forall c_1 \text{ et } c_2 \in C(r),$$

$$o_{(d_{c_1,r} \leq d_{c_2,r})} = o_{(a_{c_1,r} \leq a_{c_2,r})} \quad (6.11)$$

▷ L'ordre entre les circulations est identique à l'entrée et en sortie de la ressource.

Avec le **choix d'itinéraire**, les contraintes d'espacement entre deux circulations $c_1, c_2 \in C(r)$ sont nécessaires seulement si elles utilisent effectivement la même ressource r . Ainsi les contraintes 6.10 s'expriment alors :

$$\begin{aligned} d_{c_2,r} - d_{c_1,r} &\geq \varepsilon_{(c_1,c_2,r)} - M.(1 - o_{(d_{c_1,r} \leq d_{c_2,r})}) \\ &\quad - M.(2 - v_{c_1,r} - v_{c_2,r}) \\ a_{c_2,r} - a_{c_1,r} &\geq \varepsilon_{(c_1,c_2,r)} - M.(1 - o_{(a_{c_1,r} \leq a_{c_2,r})}) \\ &\quad - M.(2 - v_{c_1,r} - v_{c_2,r}) \\ d_{c_1,r} - d_{c_2,r} &\geq \varepsilon_{(c_2,c_1,r)} - M.o_{(d_{c_1,r} \leq d_{c_2,r})} \\ &\quad - M.(2 - v_{c_1,r} - v_{c_2,r}) \\ a_{c_1,r} - a_{c_2,r} &\geq \varepsilon_{(c_2,c_1,r)} - M.o_{(a_{c_1,r} \leq a_{c_2,r})} \\ &\quad - M.(2 - v_{c_1,r} - v_{c_2,r}) \end{aligned} \quad (6.12)$$

▷ On s'assure de l'ordre des circulations et que les deux circulations empruntent bien la même ressource.

Conflit : $\forall r \in R$ tel que $Capacite(r) = 1, \forall (c_1, c_2) \in C(r)$, les circulations c_1 et c_2 étant en conflit pour l'occupation de la ressource r :

D'une part (**contrainte de respect des espacements** selon l'ordre) :

$$\begin{aligned} a_{c_2,r} &\geq d_{c_1,r} + \varepsilon_{(c_1,c_2,r)} - M \cdot o_{(a_{c_2,r} \leq d_{c_1,r})} \\ a_{c_1,r} &\geq d_{c_2,r} + \varepsilon_{(c_2,c_1,r)} - M \cdot o_{(a_{c_1,r} \leq d_{c_2,r})} \end{aligned} \quad (6.13)$$

▷ *Les deux circulations ne peuvent utiliser la ressource en même temps ; l'entrée de la seconde s'effectue donc nécessairement après la sortie de la première.*

D'autre part (**contrainte du respect de l'ordre**) :

$$\forall r \in R / Capacite(r) = 1, \forall c_1 \text{ et } c_2 \in C(r),$$

$$\begin{aligned} 1 - o_{(a_{c_2,r} \leq d_{c_1,r})} &= o_{(a_{c_1,r} \leq d_{c_2,r})} \\ \text{soit} \\ o_{(d_{c_1,r} \leq a_{c_2,r})} &= o_{(a_{c_1,r} \leq d_{c_2,r})} \end{aligned} \quad (6.14)$$

▷ *Ces contraintes assurent le respect d'un ordre cohérent entre les évènements.*

En couplant ainsi les départs et arrivées des circulations respectives, on s'assure qu'il n'y ait qu'une circulation à la fois qui occupe la ressource.

Dans le cas du **choix d'itinéraire**, les contraintes d'espacement entre les deux circulations $c_1, c_2 \in C$ s'appliquent seulement si elles utilisent effectivement la même ressource $r \in R(c_1) \cap R(c_2)$. Ainsi, l'expression de 6.13 devient :

$$\begin{aligned} a_{c_2,r} - d_{c_1,r} &\geq \varepsilon_{(c_1,c_2,r)} - M \cdot o_{(a_{c_2,r} \leq d_{c_1,r})} \\ &\quad - M \cdot (2 - v_{c_1,r} - v_{c_2,r}) \\ a_{c_1,r} - d_{c_2,r} &\geq \varepsilon_{(c_2,c_1,r)} - M \cdot o_{(a_{c_1,r} \leq d_{c_2,r})} \\ &\quad - M \cdot (2 - v_{c_1,r} - v_{c_2,r}) \end{aligned} \quad (6.15)$$

Contraintes commerciales

Ces contraintes métiers sont susceptibles d'être transgressées si aucune solution satisfaisante n'est obtenue. En conséquence, se pose la question de la meilleure option de modélisation mathématique : ajout de variables d'écarts, définition comme des coupes auxiliaires ou prise en compte dans la fonction objectif ? Voici leurs expressions en tant que contraintes.

Durées maximales d'occupation :

$$\forall c \in C, \forall r \in R(c)$$

$$d_{c,r} - a_{c,r} \leq P_{(c,r)}^{max} \quad (6.16)$$

▷ *La circulation ne peut utiliser la ressource plus longtemps que la durée maximale définie.*

Contraintes de durée minimale de correspondance : ou relève des voyageurs d'une circulation par une autre. Deux circulations sont en correspondance lorsque la durée entre l'entrée de la première et la sortie de la seconde dans une ressource d'infrastructure est bornée par le gérant des circulations.

$$\forall r \in R^{corres}(c_1, c_2), \forall (c_1, c_2) \in C^{corres}(r), c_1 \prec c_2$$

$$\varepsilon_{corres}^{min}(c_1, c_2) \leq d_{c_2, r} - a_{c_1, r} \leq \varepsilon_{corres}^{max}(c_1, c_2) \quad (6.17)$$

▷ *La durée de présence commune des deux circulations ne peut être inférieure au seuil défini.*

6.2.3 Fonction objectif

La fonction objectif retenue consiste à minimiser la somme évaluée des retards à l'arrivée dans chaque ressource :

$$z = \min \left\{ \sum_{c \in C} \sum_{r \in R(c)} (p_{c,r}^a \cdot w(c, r)) \right\}$$

où :

- $w(c, r)$ correspond au coût du retard d'une unité de temps à l'arrivée dans la ressource r pour la circulation c . En phase d'étude tous les retards sont considérés égaux $w(c, r) = 1$ -au moins pour les dessertes commerciales-. Mais on pourrait étudier des fonctions objectifs plus sophistiquées (exemple : fonctions quadratiques, etc ...).
- $p_{c,r}^a$ est le retard par rapport à l'horaire théorique :

$$p_{c,r}^a = a_{c,r} - a_{théo}(c, r), \quad p_{c,r}^a \geq 0.$$

Coût du changement : afin d'éviter d'affecter une circulation sur une voie autre que la voie par défaut si cela n'est pas nécessaire (et donc changer inutilement les itinéraires correspondants), on introduit dans la fonction objectif une pénalité $\epsilon_{c,r}$ aux variables d'affectation ne correspondant pas aux ressources par défaut. On peut faire de même ($\epsilon(e_1, e_2)$) pour l'ordre par défaut. En effet, du point de vue industriel tout changement d'organisation doit être réfléchi. La fonction devient donc :

$$z = \min \left\{ \begin{aligned} & \sum_{c \in C} \sum_{r \in R(c)} p_{c,r}^a \cdot w(c, r) \\ & + \sum_{c \in C} \sum_{r \in R(c)} -\epsilon_{c,r} \cdot u_{c,r} \\ & + \sum_{r \in R} \sum_{(c_1, c_2) \in C(r)} -\epsilon_{(e_1, e_2)} \cdot O_{(e_1, e_2)} \end{aligned} \right\} \quad (6.18)$$

6.2.4 Synthèse du modèle

Nous allons ici résumer l'ensemble du modèle.

Fonction objectif :

$$z = \min \left\{ \begin{aligned} & \sum_{c \in C} \sum_{r \in R(c)} p_{c,r}^a \cdot w(c, r) \\ & + \sum_{c \in C} \sum_{r \in R(c)} -\epsilon_{c,r} \cdot v_{c,r} \\ & + \sum_{r \in R} \sum_{(c_1, c_2) \in C(r)} -\epsilon_{(c_1, c_2)} \cdot o_{(c_1, c_2)} \end{aligned} \right\} \quad (6.19)$$

sous les contraintes :

Durée minimale d'occupation : $\forall c \in C, \forall r \in R(c) :$

$$a_{c, r_s} - d_{c, r_e} \geq \sum_{r_i \in It(c, r_e, r_s)} v_{c, r_i} \cdot p^{min}(c, r_i) \quad (6.20)$$

Espacement et respect de l'ordre :

$$\forall r \in R, \forall c_1 \text{ et } c_2 \in C(r),$$

succession

$$\begin{aligned} d_{c_2, r} - d_{c_1, r} &\geq \varepsilon_{(c_1, c_2, r)} - M \cdot (1 - o_{(d_{c_1, r} \leq d_{c_2, r})}) \\ &\quad - M \cdot (2 - v_{c_1, r} - v_{c_2, r}) \\ a_{c_2, r} - a_{c_1, r} &\geq \varepsilon_{(c_1, c_2, r)} - M \cdot (1 - o_{(a_{c_1, r} \leq a_{c_2, r})}) \\ &\quad - M \cdot (2 - v_{c_1, r} - v_{c_2, r}) \\ d_{c_1, r} - d_{c_2, r} &\geq \varepsilon_{(c_2, c_1, r)} - M \cdot o_{(d_{c_1, r} \leq d_{c_2, r})} \\ &\quad - M \cdot (2 - v_{c_1, r} - v_{c_2, r}) \\ a_{c_1, r} - a_{c_2, r} &\geq \varepsilon_{(c_2, c_1, r)} - M \cdot o_{(a_{c_1, r} \leq a_{c_2, r})} \\ &\quad - M \cdot (2 - v_{c_1, r} - v_{c_2, r}) \end{aligned} \quad (6.21)$$

et

$$o_{(d_{c_1, r} \leq d_{c_2, r})} = o_{(a_{c_1, r} \leq a_{c_2, r})} \quad (6.22)$$

conflit

$$\begin{aligned} a_{c_2, r} - d_{c_1, r} &\geq \varepsilon_{(c_1, c_2, r)} - M \cdot o_{(a_{c_2, r} \leq d_{c_1, r})} \\ &\quad - M \cdot (2 - v_{c_1, r} - v_{c_2, r}) \\ a_{c_1, r} - d_{c_2, r} &\geq \varepsilon_{(c_2, c_1, r)} - M \cdot o_{(a_{c_1, r} \leq d_{c_2, r})} \\ &\quad - M \cdot (2 - v_{c_1, r} - v_{c_2, r}) \end{aligned} \quad (6.23)$$

et

$$o_{(d_{c_1, r} \leq a_{c_2, r})} = o_{(a_{c_1, r} \leq d_{c_2, r})} \quad (6.24)$$

Correspondances : $\forall r \in R, \forall (c_1, c_2) \in C^{corres}(r), c_1 \prec c_2$

$$\begin{aligned} d_{c_2,r} - a_{c,r} &\geq \varepsilon_{corres}^{min}(c, c_2) \\ d_{c_2,r} - a_{c,r} &\leq \varepsilon_{corres}^{max}(c, c_2) \end{aligned}$$

Variables : $\forall c \in C, \forall r \in R(c),$

$$\begin{aligned} a_{c,r} &\geq a_{théo}(c, r) \quad , \quad a_{c,r} \in \mathbb{R} \\ d_{c,r} &\geq d_{théo}(c, r) \quad , \quad d_{c,r} \in \mathbb{R} \end{aligned}$$

$$o_{*1 \leq *2} \in \{0, 1\}$$

$$v_{c,r_i} \in \{0, 1\} \quad , \quad \forall i \in It(c, r_e, r_s)$$

et telles que

$$\begin{aligned} a_{c,r} &= d_{c,r_i} \quad , \quad \forall r_i \in Pred(c, r) \\ d_{c,r} &= a_{c,r_j} \quad , \quad \forall r_j \in Succ(c, r) \end{aligned} \tag{6.25}$$

$$\begin{aligned} e_2 &\geq e_1 - M \cdot (1 - o_{e_1 \leq e_2}) \\ e_1 &\geq e_2 - M \cdot o_{(e_1 \leq e_2)} \end{aligned} \tag{6.26}$$

et

$$\sum_{r_i \in It(c, r_e, r_s)} v_{c,r_i} = 1 \tag{6.27}$$

6.2.5 Extensions du modèle

Espacements dynamiques en fonction des vitesses

Les valeurs d'espacements dépendent des vitesses des circulations. Or dans le modèle précédent ces valeurs d'espacement sont des constantes calculées pour la vitesse nominale des circulations considérées. Ces valeurs d'espacements deviennent trop approximatives si on souhaite représenter un fonctionnement précis de la réalité où les consignes de vitesse peuvent varier fortement (cf. la thématique de la fluidification ou une faible imprécision peut engendrer, sur le terrain, un arrêt préventif du train par la signalisation et donc, à cause de l'inertie d'un train, de fortes perturbations par rapport à la marche prévue). Nous souhaitons donc pouvoir introduire dans le modèle linéaire des valeurs d'espacement variables, fonctions de la vitesse de chacune des circulations considérées (valeurs qui sont elles mêmes solutions de notre problème).

Jusqu'à présent les valeurs des durées d'espacements étaient supposées connues et fixées telles que (avec $c_1 \prec c_2$) : $\varepsilon_{(c_1, c_2, r)} = \varepsilon_{(c_1, r)}^l + \varepsilon_{(c_2, r)}^r$

En effet, on peut déterminer par simulation ces durées pour une vitesse donnée. Par défaut on a donc coutume d'utiliser la valeur correspondant à la vitesse théoriquement prévue par le plan de transport initial.

Cependant la modulation des vitesses de chaque circulation est un levier important de la régulation. Mais en modifiant ainsi les vitesses sur les ressources qui précèdent et suivent la ressource considérée, on modifie aussi les valeurs (durées) d'espacement minimales. (c.f.[48])

Or, considérer la vitesse d'une circulation sur une distance fixée revient à considérer son temps de parcours.

Nous formulerons donc l'hypothèse simplificatrice suivante :

L'accroissement de la durée minimale de séparation entre deux circulations est linéairement proportionnel à l'augmentation du temps de parcours de chacune des circulations sur les ressources adjacentes.

N.B. : Cette hypothèse est valable pour des cantons fixes sur lesquels les espacements minimaux $\varepsilon_{\text{theo}}$ ont été calculés pour la vitesse maximale d'exploitation V_{max} (associée à la durée de parcours p^{min}). De plus, $\forall c \in C, \forall r \in R(c)$, où $V(c, r)$ représente la vitesse de c sur r :

$$\frac{V_{\text{max}}(c, r)}{V(c, r)} = \frac{d_{c,r} - a_{c,r}}{p_{(c,r)}^{\text{min}}}$$

Ainsi, la valeur de l'espacement dynamique $\varepsilon_{\text{dyn}}(c_1, c_2, r)$ entre les circulations c_1 et c_2 ($c_1 \prec c_2$) sur la ressource $r \in R(c_1) \cap R(c_2)$, avec $\text{prec}(c_2, r) = r_{-1}$ et $\text{succ}(c_1, r) = r_{+1}$, s'exprime :

$$\varepsilon_{\text{dyn}}(c_1, c_2, r) = \varepsilon_{\text{dyn}}^l(c_1, r) + \varepsilon_{\text{dyn}}^r(c_2, r) \quad (6.28)$$

avec

$$\varepsilon_{\text{dyn}}^l(c_1, r) = (d_{c_1, r_{+1}} - a_{c_1, r_{+1}}) \times \frac{\varepsilon_{\text{min}}^r(c_1, r)}{p_{(c_1, r_{+1})}^{\text{min}}} \quad (6.29)$$

$$\varepsilon_{\text{dyn}}^r(c_2, r) = (d_{c_2, r_{-1}} - a_{c_2, r_{-1}}) \times \frac{\varepsilon_{\text{min}}^l(c_2, r)}{p_{(c_2, r_{-1})}^{\text{min}}} \quad (6.30)$$

Ceci donne pour le problème linéaire décrit précédemment les modifications des contraintes d'espacement sur le modèle suivant (exemple avec la contrainte 6.10) :

$$\begin{aligned}
a_{c_2,r} - d_{c_1,r} &\geq \varepsilon_{\text{dyn}}(c_1, c_2, r) - M.(1 - o_{(a_{c_1,r} \leq a_{c_2,r})}) \\
&\text{i.e.} \\
a_{c_2,r} - d_{c_1,r} &\geq \varepsilon_{\text{dyn}}^l(c_1, r) + \varepsilon_{\text{dyn}}^r(c_2, r) - M.(1 - o_{(a_{c_1,r} \leq a_{c_2,r})}) \\
&\text{i.e.} \\
a_{c_2,r} - d_{c_1,r} &\geq k_1 \times (d_{c_1,r+1} - a_{c_1,r+1}) \\
&\quad + k_2 \times (d_{c_2,r-1} - a_{c_2,r-1}) \\
&\quad - M.(1 - o_{(a_{c_1,r} \leq a_{c_2,r})})
\end{aligned} \tag{6.31}$$

avec les données

$$k_1 = \frac{\varepsilon_{\min}^r(c_1, r)}{p_{(c_1,r+1)}^{\min}} \text{ et } k_2 = \frac{\varepsilon_{\min}^l(c_2, r)}{p_{(c_2,r-1)}^{\min}} \tag{6.32}$$

Remarque : on peut aussi prendre en compte la donnée $p_{(c,r)}^{\max}$, pour en déduire une relation affine.

Variations de vitesse

Problème : d'une part les écarts importants de vitesse entre deux ressources ne sont pas nécessairement atteignables en réalité, surtout si ces ressources sont relativement courtes, compte tenu des capacités de freinage ou d'accélération, et des caractéristiques de l'infrastructure.

D'autre part, avec les espacements dynamiques, le fait de rouler au maximum des vitesses permises dans les nœuds ferroviaires (dans un but de fluidification du trafic) implique que les accélérations et décélérations se font dans les ressources adjacentes. Si on doit décélérer en début de la ressource suivante ou accélérer à la fin de la ressource précédente de manière trop importante, les calculs des durées d'espacement peuvent être faussés (car ils sont basés sur les vitesses moyennes de parcours de ces ressources et non précisément sur les vitesses de sortie ou d'entrée).

Autrement dit il faudrait prendre en compte les accélérations, mais nous nous heurtons à un problème (physique) quadratique (faisant intervenir l'énergie cinétique) qui n'est pas modélisable *a priori* par un problème (mathématique) linéaire.

Option palliative : le *lissage* consiste à pénaliser fortement les variations brusque de vitesse -i.e. de durée de parcours- par rapport au plan de transport théorique réalisable.

Le problème principal réside alors dans le choix des pénalités adéquates. Le comportement (et les bons paramétrages) étant difficilement prévisible (grande sensibilité ou mauvaise interprétation?), il s'agit plus d'une démarche palliative expérimentale qu'un résultat clairement généralisable.

Solution alternative : [58] propose de linéariser l'effet de l'accélération et de la décélération (en fonction de la masse de la circulation et de l'allongement de sa durée de parcours). Le principe pourrait être utilisé en modulant les durées de réservation et libération sur le même principe.

6.3 Formulation mathématique en temps discret

A présent nous allons développer une formulation mathématique basée sur le principe de la discrétisation du temps.

6.3.1 Discrétisation du temps

Dans un premier temps, nous avons choisi de présenter une discrétisation du temps en périodes de pas fixe δ_p . Plusieurs autres options peuvent néanmoins être envisagées ultérieurement.

Dans la suite de ce document nous adopterons la convention suivante.

Convention

Toute période notée \bar{t} correspond à un intervalle $]t - \delta_p, t]$, où $t = n \times \delta_p$, $n \in \mathbf{N}$.

Les durées d'espacements (comme les durées de parcours ou de correspondance) deviennent des multiples de périodes de temps :

$$\varepsilon_{c,r} = \delta_p \times \left\lceil \frac{\varepsilon'_{c,r}}{\delta_p} \right\rceil$$

où $\varepsilon'_{c,r} \in \mathbb{R}$ est la durée exacte de l'espacement (telle que définie dans le modèle précédent, par exemple).

Ainsi, les durées de réservation et libération sont majorées (hypothèses restrictives liées à des contraintes de sécurité) et l'erreur maximale est donc égale dans le pire des cas à deux périodes de temps ($2 \times \delta_p$).

Données spécifiques à ce modèle

T : ensemble des périodes de temps notées \bar{t} .

$T_{c,r}$: ensemble des périodes de temps $\bar{t} \in T$ associées à la présence de $c \in C$ en $r \in R$.

$w_{c,r,\bar{t}}^a$, $w_{c,r,\bar{t}}^d$: valeurs dans la fonction objectif associées au passage de la circulation $c \in C$ dans la ressource $r \in R$ à la période $\bar{t} \in T$,

$\bar{\alpha}_{(c,r)}^*$, $\bar{\beta}_{(c,r)}^*$: périodes associées aux horaires d'un évènement respectivement au plus tôt et au plus tard possible pour la circulation $c \in C$ pour la ressource $r \in R$.

6.3.2 Variables de décision

Indicatrices d'état

Concernant l'arrivée et le départ : soient les variables binaires suivantes $\forall c \in C$, $\forall r \in R(c)$, $\forall \bar{t} \in T_{c,r}$:

$$x_{c,r,\bar{t}}^a = \begin{cases} 1 & \text{, si la circulation } c \text{ est arrivée à la ressource } r \in R \text{ à la fin de la période } \bar{t} \text{ ,} \\ 0 & \text{, sinon .} \end{cases}$$

et

$$x_{c,r,\bar{t}}^d = \begin{cases} 1 & , \text{si la circulation } c \text{ est partie de la ressource } r \in R \text{ à la fin de la période } \bar{t} , \\ 0 & , \text{sinon.} \end{cases}$$

où, (avec $\text{succ}(c, r) = r'$) :

$$x_{c,r,\bar{t}}^d = x_{c,r',\bar{t}}^a$$

ou, dans le cadre d'un choix d'itinéraire (avec $r \in \text{It}(c, r_e, r_s)$) :

$$\sum_{r_i \in \text{Prec}(c, r_s)} x_{c,r_i,\bar{t}}^d = x_{c,r_s,\bar{t}}^a, \forall \bar{t} \in T_{r_s, c} \quad (6.33)$$

et

$$\sum_{r_j \in \text{Succ}(c, r_e)} x_{c,r_j,\bar{t}}^a = x_{c,r_e,\bar{t}}^d, \forall \bar{t} \in T_{r_e, c} \quad (6.34)$$

La figure 6.1 illustre les valeurs prises par les variables horaires $x_{c,n,\bar{t}}^a$ pour les périodes 1 à 7, en supposant que l'évènement *arrivée de la circulation c dans la ressource n* se produise durant la cinquième période :

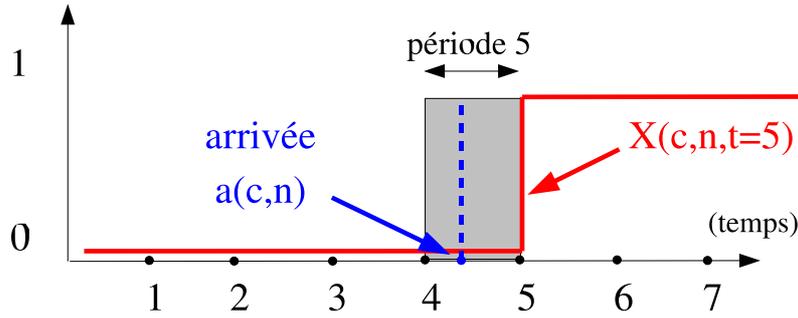


FIGURE 6.1 – Illustration de l'indicatrice d'état

Remarque : on peut définir les variables intermédiaires suivantes afin de faciliter la compréhension du modèle.

- Pour la ressource $r \in R$ l'indicatrice d'occupation $x_{c,r,\bar{t}}^o$ par la circulation $c \in C$ pour la période \bar{t} est définie par :

$$x_{c,r,\bar{t}}^o = (x_{c,r,(t+\varepsilon_{c,r}^r)}^a - x_{c,r,(t-\varepsilon_{c,r}^l)}^d)$$

La figure 6.2 illustre les valeurs prises par $x_{c,n,\bar{t}}^a$, $x_{c,n,\bar{t}}^d$ et $x_{c,r,\bar{t}}^o$ pour les périodes 1 à 7, en supposant que l'arrivée de la circulation c dans la ressource n se produise durant la quatrième période, le départ durant la cinquième et que les durées de réservation et de libérations valent chacune une période :

- Les indicatrices d'évènement $\forall c \in C, \forall r \in R(c)$ sont telles que :

$$\Delta_{c,r,\bar{t}}^a = \begin{cases} 1 & , \text{si l'évènement } \textit{arrivée} \text{ s'est produit durant la période } \bar{t} , \\ 0 & , \text{sinon.} \end{cases}$$

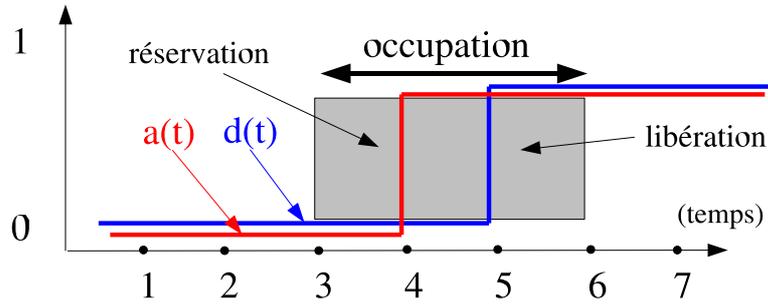


FIGURE 6.2 – Convention pour l'occupation d'une ressource

où :

$$\Delta_{c,r,\bar{t}}^a = (x_{c,r,\bar{t}}^a - x_{c,r,\bar{t}-\delta_p}^a) \quad \forall c \in C, \forall r \in R(c), \forall \bar{t} \in T_{c,r}$$

et

$$\Delta_{c,r,\bar{t}}^d = \begin{cases} 1 & , \text{si l'événement } \textit{départ} \text{ s'est produit durant la période } \bar{t} , \\ 0 & , \text{sinon.} \end{cases}$$

où :

$$\Delta_{c,r,\bar{t}}^d = (x_{c,r,\bar{t}}^d - x_{c,r,\bar{t}-\delta_p}^d) \quad \forall c \in C, \forall r \in R(c), \forall \bar{t} \in T_{c,r}$$

- Pour récupérer les horaires des évènements correspondants il suffit alors de considérer respectivement :

$$\sum_{\bar{t} \in T_{c,r}} t \cdot \Delta_{c,r,\bar{t}}^a, \text{ et } \sum_{\bar{t} \in T_{c,r}} t \cdot \Delta_{c,r,\bar{t}}^d, \quad \forall c \in C, \forall r \in R(c)$$

Variables d'ordre

Pour gérer l'ordre des évènements entre circulations, il faut introduire des variables booléennes. On peut envisager toutes les combinaisons possibles sur le modèle suivant :

$$o_{(e_1 \leq e_2)} = \begin{cases} 1 & , \text{si l'évènement } e_1 \text{ se produit avant l'évènement } e_2 \\ 0 & , \text{sinon} \end{cases}$$

Exemple : soient, $r \in R, c_1 \in C(r)$ et $c_2 \in C(r)$, alors $o_{(d_{c_1,r} \leq d_{c_2,r})}$ est la variable binaire d'ordonnement entre deux évènements (ici entre deux départs) telle que :

$$o_{(d_{c_1,r} \leq d_{c_2,r})} = \begin{cases} 1 & , \text{si le départ de } c_1 \text{ a lieu avant le départ de } c_2 \text{ de la ressource } r \\ 0 & , \text{sinon} \end{cases}$$

On fixe la variable avec les contraintes :

$$o(d_{c_1,r} \leq d_{c_2,r}) + o(d_{c_2,r} \leq d_{c_1,r}) = 1 \quad (6.35)$$

et (6.36) ou encore (6.37) :

$$1 + (x_{c_1,r,\bar{t}}^d - x_{c_2,r,\bar{t}}^d) \geq o(d_{c_1,r} \leq d_{c_2,r}) \geq (x_{c_1,r,\bar{t}}^d - x_{c_2,r,\bar{t}}^d), \forall \bar{t} \in T \quad (6.36)$$

ou :

$$\begin{aligned} o(d_{c_1,r} \leq d_{c_2,r}) &\geq (x_{c_1,r,\bar{t}}^d - x_{c_2,r,\bar{t}}^d), \forall \bar{t} \in T \\ o(d_{c_2,r} \leq d_{c_1,r}) &\geq (x_{c_2,r,\bar{t}}^d - x_{c_1,r,\bar{t}}^d), \forall \bar{t} \in T \end{aligned} \quad (6.37)$$

Remarques :

- les évènements à ordonner ne doivent pas se produire dans la même période $\bar{t} \in T$, sinon aucune contrainte (6.36,6.37) ne permet l'affectation de la variable (cela revient à dire que la période doit être inférieure à la valeur du plus petit espacement, sinon il faudrait introduire des contraintes supplémentaires -ce qui est assuré pour la définition des valeurs d'espacement 6.3.1 -).
- Lorsque les contraintes 6.35 sont intégrées au modèle, les contraintes 6.36 et 6.37 sont équivalentes :

$$\begin{aligned} (6.35) &\Leftrightarrow o(d_{c_1,r} \leq d_{c_2,r}) = 1 - o(d_{c_2,r} \leq d_{c_1,r}) \\ \text{et } (6.37) &\Rightarrow o(d_{c_1,r} \leq d_{c_2,r}) \leq 1 + (x_{c_1,r,\bar{t}}^d - x_{c_2,r,\bar{t}}^d) \end{aligned}$$

$$\text{donc } (6.37) \Rightarrow 6.36$$

$$\begin{aligned} \text{de plus } (6.35) \text{ et } (6.36) &\Rightarrow 1 + (x_{c_1,r,\bar{t}}^d - x_{c_2,r,\bar{t}}^d) \geq 1 - o(d_{c_2,r} \leq d_{c_1,r}) \\ \text{soit} &\quad (x_{c_2,r,\bar{t}}^d - x_{c_1,r,\bar{t}}^d) \leq o(d_{c_2,r} \leq d_{c_1,r}) \end{aligned}$$

$$\text{Conclusion : } (6.36) \Leftrightarrow (6.37).$$

- Dans le cas de conflits, les variables d'ordonnancement sont bien définies. En effet, les arrondis supérieur pour les durées de réservation et de libération impliquent qu'il existe au moins une période d'écart entre les deux évènements (arrivée/départ) à ordonner. Par contre, la contrainte 6.36 ne doit pas intégrer les valeurs d'espacement ε (liées aux réservations/libérations), sinon on pourrait être dans la situation où la libération et la réservation effectives se succèdent dans la même période et la variable ne serait alors pas définie!
- Dans le cas de successions il faut aussi mettre en place une contrainte d'espacement minimum.

6.3.3 Contraintes

Contraintes horaires

Cohérence entre indicatrices horaires :

$$\begin{aligned} \text{cohérence des horaires :} \quad & \forall c \in C, \forall r \in R(c), \forall \bar{t} \in T_{c,r} \\ x_{c,r,\bar{t}}^d & \geq x_{c,r,\overline{(t-\delta_p)}}^d \end{aligned} \quad (6.38)$$

$$x_{c,r,\bar{t}}^a \geq x_{c,r,\overline{(t-\delta_p)}}^a \quad (6.39)$$

▷ Une fois l'évènement produit, toutes les variables liées à des périodes ultérieures sont aussi fixées à la valeur 1 (conformément à la définition 6.3.2).

N.B. Les contraintes suivantes (durées minimales d'occupation) sont aussi absolument nécessaires pour assurer la cohérence du modèle.

Durées minimales d'occupation :

$$\begin{aligned} \text{Durée minimale d'occupation :} \quad & \forall c \in C, \forall r \in R(c), \forall \bar{t} \in T_{c,r} \\ x_{c,r,\bar{t}}^d & \leq x_{c,r,\overline{(t-p_{(c,r)}^{min})}}^a \end{aligned} \quad (6.40)$$

▷ Le départ ne peut avoir lieu que si l'arrivée s'est produite depuis une durée supérieure à $p_{(c,r)}^{min}$.

Contraintes d'espace

Contraintes de capacité :

$$\begin{aligned} \text{Contrainte de capacité :} \quad & \forall r \in R, \forall \bar{t} \in T \\ \sum_{c \in R(c)} (x_{c,r,\bar{t}}^o) & \leq \text{Capacite}(r) \\ \text{soit :} \\ \sum_{c \in R(c)} (x_{c,r,\overline{(t+\varepsilon_{c,r}^r)}}^a - x_{c,r,\overline{(t-\varepsilon_{c,r}^l)}}^d) & \leq \text{Capacite}(r) \end{aligned} \quad (6.41)$$

▷ Pour chaque période, l'ensemble des occupations ne doit pas excéder la capacité de la ressource considérée (NB : l'occupation telle que définie en 6.3.2 prend en compte les durées de réservation $\varepsilon_{c,r}^l$ et d'occupation $\varepsilon_{c,r}^r$).

Interdiction de dépassement : (succession)

Rappel : deux circulations peuvent occuper la ressource $r \in R$ en même temps ($\text{Capacite}(r) > 1$) si l'ordre est conservé et les espacements respectés aux extrémités de la ressource.

Soient deux circulations c_1 et c_2 parcourant r (dans le même sens), alors $\forall(c_1, c_2), \forall r \in R(c_1) \cap R(c_2)$:

$\forall(\bar{t}_1, \bar{t}_2)$ tel que ($t_2 \geq t_1$)

$$x_{c_1, r, \bar{t}_1}^a + (1 - x_{c_1, r, \bar{t}_2}^d) + (1 - x_{c_2, r, \bar{t}_1}^a) + x_{c_2, r, \bar{t}_2}^d \leq 3 \quad (6.42)$$

▷ *La deuxième circulation arrivée dans r ne peut en sortir avant la première.*
OU (avec les variables auxiliaires)

$$o_{(d_{c_1, r} \leq d_{c_2, r})} = o_{(a_{c_1, r} \leq a_{c_2, r})} \quad (6.43)$$

▷ *L'ordre des circulations au départ et à l'arrivée sont identiques.*

Remarques :

– Une variante (plus forte) de (6.42) :

$$x_{c_1, r, \bar{t}_1}^a + (1 - x_{c_1, r, \bar{t}_2}^d) + (1 - x_{c_2, r, \bar{t}_1 - \varepsilon_{(c_2, r)}^r}) + x_{c_2, r, \bar{t}_2 + \varepsilon_{(c_2, r)}^l}^d \leq 3$$

– (6.43) \Rightarrow (6.42) (-sans considérer les espacements-)

$$(6.37), \text{ et } (6.35) \Rightarrow (x_{c_1, r, \bar{t}}^d - x_{c_2, r, \bar{t}}^d) \leq o_{(d_{c_1, r} \leq d_{c_2, r})}, \forall \bar{t} \in T$$

$$\text{ et } o_{c_1^a, c_2^a, r'} \leq 1 - (x_{c_2, r, \bar{t}}^a - x_{c_1, r, \bar{t}}^a), \forall \bar{t} \in T$$

Donc

$$(6.43) \Rightarrow (x_{c_1, r, \bar{t}_1}^d - x_{c_2, r, \bar{t}_1}^d) \leq 1 - (x_{c_2, r, \bar{t}_2}^a - x_{c_1, r, \bar{t}_2}^a), \forall \bar{t}_1, \bar{t}_2$$

$$\text{ i.e. } x_{c_1, r, \bar{t}_1}^d + (1 - x_{c_1, r, \bar{t}_2}^a) + x_{c_2, r, \bar{t}_2}^a + (1 - x_{c_2, r, \bar{t}_1}^d) \leq 3, \forall \bar{t}_1, \bar{t}_2$$

Conclusion : (6.43) \Rightarrow (6.42)

– **Conséquence** : comme 6.43 \Rightarrow 6.42 et 6.42 nécessite $O(t^2)$ contraintes contre $O(t)$ pour 6.43, on préférera la contrainte 6.43.

Itinéraires incompatibles : (ou conflits considérés deux à deux)

Soient $c_1 \in C$ parcourant $r_1 \in R(c_1)$ et $c_2 \in C$ parcourant $r_2 \in R(c_2)$, telles que leurs parcours soient incompatibles $\{r_1, r_2\} \in Iic$, alors : $\forall \bar{t} \in T_{c_1, r} \cup T_{c_2, r}$

$$\left(x_{c_1, r_1, \bar{t} + \varepsilon_{(c_1, r_1)}^r}^a - x_{c_1, r_1, \bar{t} - \varepsilon_{(c_1, r_1)}^l}^d \right) + \left(x_{c_2, r_2, \bar{t} + \varepsilon_{(c_2, r_2)}^r}^a - x_{c_2, r_2, \bar{t} - \varepsilon_{(c_2, r_2)}^l}^d \right) \leq 1 \quad (6.44)$$

▷ Les deux circulations c_1 et c_2 ne peuvent occuper leurs itinéraires respectifs r_1 et r_2 en même temps (c'est un cas particulier de 6.41).

OU (avec les variables auxiliaires)

$$o_{(d_{c_1, r_1} \leq a_{c_2, r_2})} = o_{(a_{c_1, r_1} \leq d_{c_2, r_2})} \quad (6.45)$$

▷ Assure le respect d'un ordre cohérent entre les évènements (comme pour le modèle précédent).

Remarques :

- la contrainte d'itinéraires incompatibles peut être vue comme une contrainte de capacité (unitaire) telle que 6.41 de la ressource agrégée $r = r_1 \cup r_2$ pour chaque couple de circulations en conflit.
- implication (6.45 \Rightarrow 6.44) -sans considérer les espacements- :

$$\begin{aligned} 6.37 \text{ et } 6.35 &\Rightarrow (x_{c_1, r_1, \bar{t}}^a - x_{c_2, r_2, \bar{t}}^d) \leq o_{(d_{c_1, r_1} \leq a_{c_2, r_2})}, \forall \bar{t} \in T \\ \text{et } &o_{(d_{c_1, r_1} \leq a_{c_2, r_2})} \leq 1 - (x_{c_2, r_2, \bar{t}}^a - x_{c_1, r_1, \bar{t}}^d), \forall \bar{t} \in T \end{aligned}$$

Donc

$$\begin{aligned} 6.45 &\Rightarrow (x_{c_1, r_1, \bar{t}}^a - x_{c_2, r_2, \bar{t}}^d) \leq 1 - (x_{c_2, r_2, \bar{t}}^a - x_{c_1, r_1, \bar{t}}^d), \forall \bar{t} \in T \\ \text{i.e. } &x_{c_1, r_1, \bar{t}}^a - x_{c_1, r_1, \bar{t}}^d + x_{c_2, r_2, \bar{t}}^a - x_{c_2, r_2, \bar{t}}^d \leq 1, \forall \bar{t} \in T \end{aligned}$$

Conclusion : 6.45 \Rightarrow 6.44

- Conséquence : comme 6.44 exprime aussi la contrainte d'espacement, on utilisera 6.44 plutôt que 6.45
- Cas de la voie unique : on pourrait préférer une formulation compacte à l'énumération de nombreuses paires de conflits. Pour cela on peut introduire les variables intermédiaires $sens_1(r, \bar{t})$, et $sens_2(r, \bar{t}) \in \{0, 1\}$ et le système de contraintes suivant, afin d'autoriser l'occupation de r dans un seul sens à la fois (mais plusieurs circulations peuvent ainsi se succéder en même temps, dans le même sens) :

$$\begin{aligned} \forall \bar{t} \in T_{r,c} \\ \text{Capacite}(r) * sens_1(r, \bar{t}) &\geq \sum_{c \in C(r) \text{ de sens1}} \left(x_{c,r,t+\varepsilon_{c,r}^r}^a - x_{c,r,t-\varepsilon_{c,r}^l}^d \right) \\ \text{Capacite}(r) * sens_2(r, \bar{t}) &\geq \sum_{c \in C(r) \text{ de sens2}} \left(x_{c,r,t+\varepsilon_{c,r}^r}^a - x_{c,r,t-\varepsilon_{c,r}^l}^d \right) \\ \text{et} \\ sens_1(r, \bar{t}) + sens_2(r, \bar{t}) &\leq 1 \end{aligned}$$

Contraintes commerciales

Ici aussi, se pose la question de la meilleure option de modélisation : ajout de variables d'écart, définition comme des coupes auxiliaires ou prise en compte dans la fonction objectif. Voici néanmoins leurs expressions en tant que contraintes.

Relève des voyageurs : (en correspondances) de $c_1 \in C$ par $c_2 \in C$ dans la ressource $r \in Rc(c_1, c_2)$:

$$x_{c_2, r, \bar{t}}^d \leq x_{c_1, r, \overline{(t - \varepsilon_{\text{corres}}^{\min}(c_1, c_2, r))}}^a \quad \forall \bar{t} \in T_{r, c_1} \cup T_{r, c_2} \quad (6.46)$$

Durées maximales de parcours et d'arrêt :

$$\begin{aligned} \text{Durées maximales de parcours :} \quad & \forall c \in C, \forall r \in R(c), \forall \bar{t} \in T_{r, c} \\ x_{c, r, \bar{t}}^d & \geq x_{c, r, \overline{(t - p_{(c, r)}^{\max})}}^a \end{aligned} \quad (6.47)$$

6.3.4 Fonction objectif

La fonction objectif générique choisie consiste à minimiser la somme pondérée des retards (écarts entre les horaires théoriques et les horaires calculés dans la solution) :

$$\min \sum_{c \in C, r \in R(c), \bar{t} \in T_{r, c}} \left(w_{c, r, \bar{t}}^a \cdot x_{c, r, \bar{t}}^a + w_{c, r, \bar{t}}^d \cdot x_{c, r, \bar{t}}^d \right)$$

où $w_{c, r, \bar{t}}^a$ et $w_{c, r, \bar{t}}^d$ sont des poids à définir selon la stratégie qui sera retenue par l'expertise métier.

6.3.5 Synthèse du modèle

Nous allons ici résumer l'ensemble du modèle.

Fonction objectif :

$$\min \sum_{c \in C, r \in R, t} \left(w_{c,r,\bar{t}}^a \cdot x_{c,r,\bar{t}}^a + w_{c,r,\bar{t}}^d \cdot x_{c,r,\bar{t}}^d \right)$$

sous les contraintes :

Contraintes :

Cohérence entre variables horaires : $\forall c \in C, \forall r \in R(c), \forall \bar{t} \in T_{r,c}$

$$x_{c,r,\bar{t}}^* \geq x_{c,r,\bar{t}-\delta_p}^* \quad (6.48)$$

Durée minimale d'occupation : $\forall c \in C, \forall r \in R(c), \forall \bar{t} \in T_{r,c}$

$$x_{c,r,\bar{t}}^d \leq x_{c,r,\bar{t}-p_{(c,r)}^{min}}^a \quad (6.49)$$

Capacité et espacement : $\forall r \in R, \forall t \in \cup_{c \in C(r)} T_{r,c}$

$$\sum_{c \in C(r)} x_{c,r,\bar{t}}^d - x_{c,r,\bar{t}}^a \leq \text{Capacite}(r), \quad (6.50)$$

Respect de l'ordre : $\forall r \in R, \forall c_1 \text{ et } c_2 \in C(r)$

interdiction de dépassement (succession)

$$O_{(d_{c_1,r} \leq d_{c_2,r})} = O_{(a_{c_1,r} \leq a_{c_2,r})} \quad (6.51)$$

itinéraires incompatibles (conflit)

$$O_{(d_{c_1,r_1} \leq a_{c_2,r_2})} = O_{(a_{c_1,r_1} \leq d_{c_2,r_2})} \quad (6.52)$$

Relève des voyageurs : (en correspondance) de $c_1 \in C$ par $c_2 \in C$ dans la ressource $r \in Rc(c_1, c_2) : \forall \bar{t} \in T_{r,c_1} \cup T_{r,c_2}$

$$x_{c_2,r,\bar{t}}^d \leq x_{c_1,r,(\bar{t}-\varepsilon_{\text{corres}}^{min}(c_1,c_2,r))}^a \quad (6.53)$$

Variables :

$$\begin{aligned} x_{c,r,\bar{t}}^a, x_{c,r,\bar{t}}^d &\in \{0, 1\} \\ O_{*1,*2} &\in \{0, 1\} \end{aligned}$$

telles que :

Indicatrice d'état :

sans choix d'itinéraire

$$\begin{aligned} x_{c,r,\bar{t}}^a &= x_{c,r-1,\bar{t}}^d, \text{ prec}(c,r) = r-1 \\ x_{c,r,\bar{t}}^d &= x_{c,r+1,\bar{t}}^a, \text{ succ}(c,r) = r+1 \end{aligned} \quad (6.54)$$

sinon, avec $r \in It(c, r_e, r_s)$

$$\begin{aligned} \sum_{r_i \in \text{Prec}(c, r_s)} x_{c, r_i, \bar{t}}^d &= x_{c, r_s, \bar{t}}^a \\ \sum_{r_j \in \text{Succ}(c, r_e)} x_{c, r_j, \bar{t}}^a &= x_{c, r_e, \bar{t}}^d \end{aligned} \quad (6.55)$$

Variable d'ordre :

$$o_{*1 \leq *2} + o_{*2 \leq *1} = 1 \quad (6.56)$$

$$o_{*1 \leq *2} \geq (x_{*1, \bar{t}} - x_{*2, \bar{t}}), \forall \bar{t} \in T \quad (6.57)$$

$$o_{*2 \leq *1} \geq (x_{*2, \bar{t}} - x_{*1, \bar{t}}), \forall \bar{t} \in T \quad (6.58)$$

6.3.6 Mise en œuvre

En pratique, il est préférable de mettre en œuvre un prétraitement qui simplifie la gestion de ce grand nombre de variables et de contraintes.

Contraintes

Toutes les variables ne sont pas générées pour l'ensemble de la plage horaire de l'étude, mais seulement en fonction des $\alpha_{(c,r)}^*$ et $\beta_{(c,r)}^*$. Lors de la construction des contraintes, on peut donc vouloir manipuler des variables qui n'existent pas. Exemple : $x_{c,r,t-p_{(c,r)}^{min}}^*$, où $t - p_{(c,r)}^{min} \leq \alpha_{(c,r)}^*$

Dans le cas où la circulation n'a pas de choix d'itinéraire (ni d'arrêt définitif), on a :

$$\begin{aligned} x_{c,r,\bar{t}}^* &= 0, \text{ si } t \leq \alpha_{(c,r)}^* \\ x_{c,r,\bar{t}}^* &= 1, \text{ si } t > \beta_{(c,r)}^* \end{aligned}$$

Néanmoins, dans le cas général voici l'hypothèse que nous retiendrons :

$$\forall c \in C, \forall r \in R(c)$$

$$\begin{aligned} \alpha_{(c,r)}^a &\leq \alpha_{(c,r)}^d \quad \text{et} \quad \beta_{(c,r)}^a \leq \beta_{(c,r)}^d, \\ x_{c,r,\bar{t}}^* &= 0, \forall \bar{t} \in T, t \leq \alpha_{(c,r)}^*, \\ x_{c,r,\bar{t}}^* &= x_{c,r,\bar{\beta}_{(c,r)}^*}^*, \forall \bar{t} \in T, t > \beta_{(c,r)}^* \end{aligned}$$

Appliquée au modèle précédent voici l'écriture des contraintes, en pratique.

Cohérence des horaires :

$$\forall c \in C, \forall r \in R(c),$$

$$\begin{aligned} x_{c,r,\bar{t}}^d &\geq x_{c,r,\overline{(t-\delta_p)}}^d, & \alpha_{(c,r)}^d + \delta_p \leq t \leq \beta_{(c,r)}^d \\ x_{c,r,\bar{t}}^a &\geq x_{c,r,\overline{(t-\delta_p)}}^a, & \alpha_{(c,r)}^a + \delta_p \leq t \leq \beta_{(c,r)}^a \end{aligned}$$

Durée d'arrêt minimale :

$$\forall c \in C, \forall r \in R(c),$$

$$\begin{cases} x_{c,r,\bar{t}}^d = 0 & , \text{ si } \alpha_{(c,r)}^d \leq t \leq \alpha_{(c,r)}^a + p_{(c,r)}^{\min}, \\ x_{c,r,\bar{t}}^d \leq x_{c,r,\overline{(t-p_{(c,r)}^{\min})}}^a & , \text{ si } \alpha_{(c,r)}^a + p_{(c,r)}^{\min} \leq t \leq \min\{\beta_{(c,r)}^a + p_{(c,r)}^{\min}, \beta_{(c,r)}^d\}, \\ x_{c,r,\bar{t}}^d \leq x_{c,r,\overline{\beta_{(c,r)}^a}}^a & , \text{ si } \beta_{(c,r)}^a + p_{(c,r)}^{\min} \leq t \leq \beta_{(c,r)}^d. \end{cases}$$

Remarque : $\alpha_{(c,r)}^a + p_{(c,r)}^{\min} \leq \beta_{(c,r)}^d$, sinon le problème est infaisable.

Capacité :

$$\forall r \in R, \forall \bar{t} \in T, \quad \sum_{c \in C(r)} (x_{c,r,\bar{t}}^o) \leq \text{Capacite}(r)$$

où,

$$x_{c,r,\bar{t}}^o = \begin{cases} x_{c,r,\overline{(t+\varepsilon_{c,r}^r)}}^a & , \text{ si } \alpha_{(c,r)}^a - \varepsilon_{c,r}^r \leq t \leq \alpha_{(c,r)}^d + \varepsilon_{c,r}^1, \\ x_{c,r,\overline{(t+\varepsilon_{c,r}^r)}}^a - x_{c,r,\overline{(t-\varepsilon_{c,r}^1)}}^d & , \text{ si } \alpha_{(c,r)}^d + \varepsilon_{c,r}^1 \leq t \leq \beta_{(c,r)}^a - \varepsilon_{c,r}^r, \\ x_{c,r,\overline{\beta_{(c,r)}^a}}^a - x_{c,r,\overline{(t-\varepsilon_{c,r}^1)}}^d & , \text{ si } \beta_{(c,r)}^a - \varepsilon_{c,r}^r \leq t \leq \beta_{(c,r)}^d + \varepsilon_{c,r}^1, \\ 0 & , \text{ sinon .} \end{cases}$$

Interdiction de dépassement sur une ressource : deux circulations peuvent occuper la ressource $r \in R$ en même temps si l'ordre est conservé aux extrémités de la ressource et les espacements respectés

Soient deux circulations c_1 et c_2 parcourant r (dans le même sens), alors $\forall (c_1, c_2), \quad \forall r \in R(c_1) \cap R(c_2) :$

$$o_{(d_{c_1,r} \leq d_{c_2,r})} = o_{(a_{c_1,r} \leq a_{c_2,r})}$$

avec des contraintes d'espacements du type 6.42, 6.43.

Conflicts : (ou itinéraires incompatibles considérés deux à deux)

Soient $c_1 \in C$ parcourant $r_1 \in R(c_1)$ et $c_2 \in C$ parcourant $r_2 \in R(c_2)$ tels que leurs parcours soient incompatibles $\{r_1, r_2\} \in I$, alors :

$$x_{c_1,r_1,\bar{t}}^o + x_{c_2,r_2,\bar{t}}^o \leq 1, \quad \forall \bar{t} \in T,$$

avec

$$t_1 \leq t \leq t_2$$

où

$$t_1 = \max\{\alpha_{(c,r)}^d - \varepsilon_{c,r}^r, \beta_{(c,r)}^a + \varepsilon_{c,r}^l\}$$

$$t_2 = \min\{\beta_{(c,r)}^d - \varepsilon_{c,r}^r, \beta_{(c,r)}^a + \varepsilon_{c,r}^l\}$$

et

$$x_{c,r,\bar{t}}^o = \begin{cases} x_{c,r,(\bar{t}+\varepsilon_{c,r}^r)}^d & , \text{ si } \alpha_{(c,r)}^d - \varepsilon_{c,r}^r \leq t \leq \alpha_{(c,r)}^a + \varepsilon_{c,r}^l , \\ x_{c,r,\bar{\beta}_{(c,r)}^d}^d & , \text{ si } \beta_{(c,r)}^a + \varepsilon_{c,r}^l \leq t \leq \alpha_{(c,r)}^d - \varepsilon_{c,r}^r , \\ x_{c,r,(\bar{t}+\varepsilon_{c,r}^r)}^d - x_{c,r',(\bar{t}-\varepsilon_{c,r}^l)}^a & , \text{ si } t_1 \leq t \leq t_2 \\ x_{c,r,\bar{\beta}_{(c,r)}^d}^d - x_{c,r',(\bar{t}-\varepsilon_{c,r}^l)}^a & , \text{ si } \beta_{(c,r)}^d - \varepsilon_{c,r}^r \leq t \leq \beta_{(c,r)}^a + \varepsilon_{c,r}^l , \\ 0 & , \text{ sinon .} \end{cases}$$

Durée de correspondance : relève de $c_1 \in C$ par $c_2 \in C$ dans la ressource $r \in Rc(c_1, c_2)$:

$$\forall \bar{t} \in T_{r,c_1} \cup T_{r,c_2}$$

$$\begin{cases} x_{c_2,r,\bar{t}}^d = 0 & , \text{ si } \alpha_{(c_2,r)}^d \leq t \leq \alpha_{(c_1,r)}^a + \varepsilon_{\text{corres}}^{\min}(c_1, c_2, r) , \\ x_{c_2,r,\bar{t}}^d \leq x_{c_1,r,(\bar{t}-\varepsilon_{\text{corres}}^{\min}(c_1,c_2,r))}^a & , \text{ si } \alpha_{(c_2,r)}^a + \varepsilon_{\text{corres}}^{\min}(c_1, c_2, r) \leq t \leq \min\{\beta_{(c_1,r)}^a, \beta_{(c_2,r)}^d\} , \end{cases}$$

Durée d'occupation maximale : $\forall c \in C, \forall r \in R(c), :$

$$\begin{cases} x_{c,r,\bar{t}}^a = 0 & , \text{ si } \alpha_{(c,r)}^a + p_{(c,r)}^{\max} \leq t \leq \alpha_{(c,r)}^d , (*) \\ x_{c,r,\bar{t}}^d \geq x_{c,r,(\bar{t}-p_{(c,r)}^{\max})}^a & , \text{ si } \alpha_{(c,r)}^d \leq t \leq \beta_{(c,r)}^a + p_{(c,r)}^{\max} . \end{cases}$$

N.B. :

- la contrainte (*) doit pouvoir être transgressée si on autorise un stationnement définitif,
- 6.38 $\Rightarrow x_{c,r,\bar{t}}^d \geq x_{c,r,\bar{\beta}_{(c,r)}^a}^a$, si $\beta_{(c,r)}^a + p_{(c,r)}^{\max} \leq t \leq \beta_{(c,r)}^d$

Valeurs numériques

Toutes les données numériques doivent être arrondies en fonction de la taille des périodes choisies. $(\varepsilon_{(c,r)}^r, \varepsilon_{(c,r)}^l, p_{(c,r)}^{\min}, p_{(c,r)}^{\max})$

6.4 Formulation mixte

Les deux modèles précédents offrent une complémentarité qu'il serait dommage de ne pas exploiter. Pour cela, une première option consiste à générer parallèlement les deux modèles tels que définis précédemment et de les lier en générant une classe de contraintes supplémentaire intégrant les deux catégories de variables. Une deuxième option consiste à générer un modèle mixte ou subsiste pour chaque contrainte métier l'expression mathématique qui semble la plus favorable parmi les deux catégories de formulations. C'est cette dernière option qui est présentée ici (étant entendu que la première option se déduit facilement des deux chapitres qui précèdent et qu'on retrouvera ici aussi la contrainte liant les deux catégories de variables).

À présent les variables des deux modèles précédent vont coexister dans la même formulation. Par souci de lisibilité t désignera une date ($t \in \mathbb{R}$), et $\bar{t} \in T_{c,r}$ la période $]k - \delta_p, k]$ telle que $t \in \bar{t}$, où δ_p est la taille d'une période et $k = n \times \delta_p$, $n \in \mathbf{N}$.

6.4.1 Variables de décision

Horaires

$a_{c,r}$: date d'entrée de la circulation c dans la ressource $r \in R(c)$,

$d_{c,r}$: date de sortie de la circulation c dans la ressource $r \in R(c)$.

$$\forall c \in C, \forall r \in R, \quad d_{c,r} \in \mathbb{R} \text{ et } a_{c,r} \in \mathbb{R}$$

Indicatrices d'état

Soient les variables binaires suivantes $\forall c \in C, \forall r \in R(c), \forall \bar{t} \in T_{c,r}$:

$$x_{c,r,\bar{t}}^a = \begin{cases} 1 & , \text{ si la circulation } c \text{ est arrivée dans la ressource } r \in R \text{ à la fin de la période } \bar{t} , \\ 0 & , \text{ sinon .} \end{cases}$$

et

$$x_{c,r,\bar{t}}^d = \begin{cases} 1 & , \text{ si la circulation } c \text{ est partie de la ressource } r \in R \text{ à la fin de la période } \bar{t} , \\ 0 & , \text{ sinon.} \end{cases}$$

Routage

Pour gérer le routage des circulations entre plusieurs itinéraires (ressources), on a introduit des variables booléennes d'affectation de la circulation c sur la ressource r :

$$v_{c,r} = \begin{cases} 1 & , \text{ si la circulation } c \text{ emprunte la ressource } r \\ 0 & , \text{ sinon} \end{cases}$$

Ici, nous allons utiliser directement les variables précédentes. En effet pour $\bar{t} = \bar{\beta}_{(c,r)}^a$, dernière période possible pour l'arrivée de c dans r :

$$v_{c,r} = x_{c,r,\bar{\beta}_{(c,r)}^a}^a$$

Ordonnancement

Afin de gérer l'ordre des évènements entre circulations sur une ressource, il faut introduire des variables booléennes. Ainsi, pour chaque ressource $r \in R$ on peut définir la variable d'ordre entre couples d'évènements pour deux circulations $(c_i, c_j) \in C(r)$, donc entre les variables horaires associées (e_i, e_j) où $e_i \in \{a_{c_i,r}, d_{c_i,r}\}$, et $e_j \in \{a_{c_j,r}, d_{c_j,r}\}$:

$$o_{(e_1 \leq e_2)} = \begin{cases} 1 & \text{, si l'évènement lié à } e_1 \text{ se produit avant l'évènement lié à } e_2 \\ 0 & \text{, sinon} \end{cases}$$

avec les contraintes associées :

$$\begin{aligned} e_2 &\geq *_1 - M.(1 - o_{(e_1 \leq e_2)}) \\ e_1 &\geq *_2 - M.o_{(e_1 \leq e_2)} \end{aligned} \quad (6.59)$$

ou, alternativement :

$$o_{(e_1 \leq e_2)} + o_{(e_2 \leq e_1)} = 1 \quad (6.60)$$

$$\begin{aligned} &\text{et : } \forall \bar{t} \in T_{c,r} \\ 1 + (x_{c,r,\bar{t}}^1 - x_{c,r,\bar{t}}^2) &\geq o_{(e_1 \leq e_2)} \geq (x_{c,r,\bar{t}}^1 - x_{c,r,\bar{t}}^2) \end{aligned} \quad (6.61)$$

Exemple : la variable d'ordre entre l'arrivée de c_1 et le départ de c_2 de la ressource r est définie telle que :

$$o_{(a_{c_1,r} \leq d_{c_2,r})} = \begin{cases} 1 & \text{, si } c_1 \text{ arrive avant le départ de } c_2 \text{ dans la ressource } r, \\ 0 & \text{, sinon} \end{cases}$$

$$\text{i.e. } o_{(a_{c_1,r} \leq d_{c_2,r})} = 1 \Leftrightarrow a_{c_1,r} \prec d_{c_2,r}$$

avec :

$$\begin{aligned} d_{c_2,r} &\geq a_{c_1,r} - M.(1 - o_{(a_{c_1,r} \leq d_{c_2,r})}) \\ a_{c_1,r} &\geq d_{c_2,r} - M.o_{(a_{c_1,r} \leq d_{c_2,r})} \end{aligned}$$

ou :

$$\begin{aligned} o_{(a_{c_1,r} \leq d_{c_2,r})} + o_{(d_{c_2,r} \leq a_{c_1,r})} &= 1 \\ &\text{et : } \forall \bar{t} \in T_{c,r} \\ 1 + (x_{c_1,r,\bar{t}}^a - x_{c_2,r,\bar{t}}^d) &\geq o_{(a_{c_1,r} \leq d_{c_2,r})} \geq (x_{c_1,r,\bar{t}}^a - x_{c_2,r,\bar{t}}^d) \end{aligned}$$

6.4.2 Contraintes

Contraintes nécessaires à la cohérence du modèle :

$$\begin{aligned} \text{cohérence des horaires :} \quad & \forall c \in C, \forall r \in R(c), \forall \bar{t} \in T_{c,r} \\ & x_{c,r,\bar{t}}^d \geq x_{c,r,\overline{(t-\delta_p)}}^d \end{aligned} \quad (6.62)$$

$$x_{c,r,\bar{t}}^a \geq x_{c,r,\overline{(t-\delta_p)}}^a \quad (6.63)$$

$$\begin{aligned} \text{cohérence entre variables :} \quad & \forall c \in C, \forall r \in R(c) \\ d_{c,r} - \delta_p & \leq \alpha_{(c,r)}^d + (\delta_p) \cdot \sum_{\bar{t}} (1 - x_{c,r,\bar{t}}^d) \leq d_{c,r} \\ a_{c,r} - \delta_p & \leq \alpha_{(c,r)}^a + (\delta_p) \cdot \sum_{\bar{t}} (1 - x_{c,r,\bar{t}}^a) \leq a_{c,r} \end{aligned} \quad (6.64)$$

ou :

$$\begin{aligned} d_{c,r} - \delta_p & \leq x_{c,r,\bar{\alpha}_{(c,r)}^d}^d \cdot \alpha_{(c,r)}^d + \sum_{\bar{t}} t \cdot (x_{c,r,\bar{t}}^d - x_{c,r,\overline{(t-\delta_p)}}^d) \leq d_{c,r} \\ a_{c,r} - \delta_p & \leq x_{c,r,\bar{\alpha}_{(c,r)}^a}^a \cdot \alpha_{(c,r)}^a + \sum_{\bar{t}} t \cdot (x_{c,r,\bar{t}}^a - x_{c,r,\overline{(t-\delta_p)}}^a) \leq a_{c,r} \end{aligned} \quad (6.65)$$

où δ_p est la taille d'une période.

Remarques :

- Choix entre 6.64 et 6.65 : 6.64 équivaut successivement à :

$$\begin{aligned} a_{c,r} - \delta_p & \leq \beta_{(c,r)}^a - (\delta_p) \cdot \sum_t x_{c,r,\bar{t}}^a \leq a_{c,r} \\ \beta_{(c,r)}^a + \delta_p & \geq a_{c,r} + (\delta_p) \cdot \sum_t x_{c,r,\bar{t}}^a \geq \beta_{(c,r)}^a \end{aligned} \quad (6.66)$$

- Solutions fractionnaires : sous cet aspect il est probable que la formulation 6.65 soit la moins forte, les termes pouvant se compenser deux à deux si les valeurs sont très proches (ce serait différent s'il s'agissait d'une somme de termes positifs)
- Passage des variables du temps discret vers celles du temps continu :
 1. la fonction objectif tend à minimiser la valeur des variables,
 2. 6.66 $\Rightarrow a_{c,r} + (\delta_p) \cdot \sum_{\bar{t}} x_{c,r,\bar{t}}^a \geq \beta_{(c,r)}^a$.
- En cas d'arrêt définitif : si aucune variable discrète n'est affectée, alors la valeur de la variable continue ne correspond plus à la date de fin de la fenêtre d'étude mais à celle de la dernière période $\beta_{(c,r)}^*$. Cette convention posait d'ailleurs problème avec la formulation 6.65, où on avait supposé $x_{c,r,\bar{\beta}_{(c,r)}^a}^a = 1$ toujours vrai (ce qui aurait posé problème ensuite pour le choix d'itinéraire par exemple).

Durée minimale d'occupation :

On utilise les contraintes du modèle à temps continu pour la concision et la précision.

$$\begin{aligned} \text{durée minimale d'occupation :} \quad & \forall c \in C, \forall r \in R(c) \\ d_{c,r} - a_{c,r} & \geq p_{(c,r)}^{\min} \end{aligned} \quad (6.67)$$

$$\begin{aligned} \text{durée maximale d'occupation :} \quad & \forall c \in C, \forall r \in R(c) \\ d_{c,r} - a_{c,r} & \leq p_{(c,r)}^{\max} \end{aligned} \quad (6.68)$$

Contraintes d'espacement et de capacité

On utilise les contraintes du modèle à temps discret pour gérer les capacités :

$$\sum_c \left(x_{c,r,(t+\varepsilon_{c,r}^1)}^a - x_{c,r,(t-\varepsilon_{c,r}^1)}^d \right) \leq \text{Capacite}(r) \quad \forall r \in R, \forall \bar{t} \in T_{c,r} \quad (6.69)$$

et

$$\forall r \in R(c), \forall c_1, c_2 \in C(r)$$

conflit :

$$O(a_{c_1,r} \leq d_{c_2,r}) = O(d_{c_1,r} \leq a_{c_2,r}) \quad (6.70)$$

succession :

$$O(d_{c_1,r} \leq d_{c_2,r}) = O(a_{c_1,r} \leq a_{c_2,r}) \quad (6.71)$$

Contraintes d'allocation d'itinéraires

$$\begin{aligned} \text{continuité de l'itinéraire :} \quad & \forall c \in C, \\ & x_{c,r_e,\bar{t}}^d = \sum_{r_i \in \text{Succ}(r_e)} x_{c,r_i,\bar{t}}^a \\ & \sum_{r_j \in \text{Prec}(r_s)} x_{c,r_j,\bar{t}}^d = x_{c,r_s,\bar{t}}^a \end{aligned} \quad (6.72)$$

$$\begin{aligned} \text{durée de parcours :} \quad & \forall c \in C, \forall r \in R(c) \\ & d_{c,r} - a_{c,r} \geq \sum_i v_{c,r_i} \cdot p_{(c,r_i)}^{\min} \\ & (i.e.) \\ & d_{c,r} - a_{c,r} \geq \sum_i x_{(c,r_i),\bar{\beta}_{(c,r)}^a}^a \cdot p_{(c,r_i)}^{\min} \end{aligned} \quad (6.73)$$

Contraintes de correspondance

$$\begin{aligned} \text{relève des voyageurs :} \quad & \forall r \in R_{\text{corres}}(c_1, c_2), \forall (c_1, c_2) \in C_{\text{corres}}(r) \\ & \varepsilon_{\text{corres}}^{\min}(c_1, c_2) \leq d(c_2, r) - a(c_1, r) \leq \varepsilon_{\text{corres}}^{\max}(c_1, c_2) \end{aligned} \quad (6.74)$$

$$\begin{aligned} \text{échange des voyageurs :} \quad & \forall r \in R, \forall (c_1, c_2) \in \text{EchangeVoyageurs}(r) \\ & \varepsilon_{\text{corres}}^{\min}(c_1, c_2) \leq d(c_1, r) - a(c_2, r) \leq \varepsilon_{\text{corres}}^{\max}(c_1, c_2) \\ & \varepsilon_{\text{corres}}^{\min}(c_1, c_2) \leq d(c_2, r) - a(c_1, r) \leq \varepsilon_{\text{corres}}^{\max}(c_1, c_2) \end{aligned}$$

6.4.3 Fonction objectif

$$z = \min \left\{ \sum_{c \in C, r \in R(c)} (a_{c,r} \cdot w(c, r)) + \sum_{c \in C, r \in R(c), t} (w_{c,r,\bar{t}}^a \cdot x_{c,r,\bar{t}}^a + w_{c,r,\bar{t}}^d \cdot x_{c,r,\bar{t}}^d) \right\} \quad (6.75)$$

où $w_{c,r,\bar{t}}^a$ et $w_{c,r,\bar{t}}^d$ sont des poids à définir selon la stratégie qui sera retenue et $w(c, r)$ correspond au coût du retard d'une unité de temps à l'arrivée dans la ressource r pour la circulation c .

NB : on peut utiliser une pénalité linéaire pour les variables réelles et une pénalité non-linéaire (fonction à seuil ou quadratique par exemple) pour les variables discrètes. Ceci correspond bien aux critères d'évaluation actuels (d'une part somme des retards, d'autre part seuils déclenchant des pénalités).

6.4.4 Synthèse du modèle

Nous allons ici résumer l'ensemble du modèle.

Variables

Rappel : t désigne une date ($t \in \mathbb{R}$), et $\bar{t} \in T_{c,r}$ la période $]k - \delta_p, k]$ telle que $t \in \bar{t}$, où (δ_p) est la taille d'une période et $k = n \times \delta_p$, $n \in \mathbf{N}$.

Alors, $\forall c \in C, \forall r \in R(c)$,

$$d_{c,r} \in \mathbb{R}, a_{c,r} \in \mathbb{R}$$

et $\forall \bar{t} \in T_{c,r}$:

$$x_{c,r,\bar{t}}^a = \begin{cases} 1 & , \text{si la circulation } c \text{ est arrivée dans la ressource } r \in R \text{ à la fin de la période } \bar{t} , \\ 0 & , \text{sinon .} \end{cases}$$

et

$$x_{c,r,\bar{t}}^d = \begin{cases} 1 & , \text{si la circulation } c \text{ est partie de la ressource } r \in R \text{ à la fin de la période } \bar{t} , \\ 0 & , \text{sinon.} \end{cases}$$

$$o_{(e_1 \leq e_2)} = \begin{cases} 1 & , \text{si l'évènement lié à } e_1 \text{ se produit avant l'évènement lié à } e_2 \\ 0 & , \text{sinon} \end{cases}$$

Fonction objectif

$$z = \min \left\{ \sum_{c \in C, r \in R(c)} (a_{c,r} \cdot w(c, r)) + \sum_{c \in C, r \in R(c), t} (w_{c,r,\bar{t}}^a \cdot x_{c,r,\bar{t}}^a + w_{c,r,\bar{t}}^d \cdot x_{c,r,\bar{t}}^d) \right\} \quad (6.76)$$

sous les contraintes :

Cohérence des horaires :

$$\forall c \in C, \forall r \in R(c), \forall \bar{t} \in T_{c,r}$$

$$x_{c,r,\bar{t}}^d \geq x_{c,r,(\bar{t}-\delta_p)}^d \quad (6.77)$$

$$x_{c,r,\bar{t}}^a \geq x_{c,r,(\bar{t}-\delta_p)}^a \quad (6.78)$$

Durée minimale d'occupation :

$$\forall c \in C, \forall r \in R(c)$$

$$\underline{\text{durée minimale d'occupation :}} \quad \forall c \in C, \forall r \in R(c) \quad d_{c,r} - a_{c,r} \geq p_{(c,r)}^{\min} \quad (6.79)$$

Contraintes d'espace

$$\sum_c \left(\overbrace{x_{c,r,(t+\varepsilon_{c,r}^i)}^a}^{\text{Capacité :}} - \overbrace{x_{c,r,(t-\varepsilon_{c,r}^i)}^d} \right) \leq \text{Capacite}(r) \quad (6.80)$$

et

$$\forall r \in R(c), \forall c_1, c_2 \in C(r)$$

conflit :

$$o_{(a_{c_1,r} \leq d_{c_2,r})} = o_{(d_{c_1,r} \leq a_{c_2,r})} \quad (6.81)$$

succession :

$$o_{(d_{c_1,r} \leq d_{c_2,r})} = o_{(a_{c_1,r} \leq a_{c_2,r})} \quad (6.82)$$

Contraintes d'allocation d'itinéraires

$$\begin{aligned} \text{continuité de l'itinéraire :} \quad & \forall c \in C, \\ & x_{c,r_e,\bar{t}}^d = \sum_{r_i \in \text{Succ}(r_e)} x_{c,r_i,\bar{t}}^a \\ & \sum_{r_j \in \text{Prec}(r_s)} x_{c,r_j,\bar{t}}^d = x_{c,r_s,\bar{t}}^a \end{aligned} \quad (6.83)$$

$$\begin{aligned} \text{durée de parcours :} \quad & \forall c \in C, \forall r \in R(c) \\ & d_{c,r} - a_{c,r} \geq \sum_i v_{c,r_i} \cdot p_{(c,r_i)}^{\min} \\ & (i.e.) \\ & d_{c,r} - a_{c,r} \geq \sum_i x_{(c,r_i,\bar{\beta}_{(c,r)}^a)}^a \cdot p_{(c,r_i)}^{\min} \end{aligned} \quad (6.84)$$

Correspondance entre variables :

$$\begin{aligned} & \forall c \in C, \forall r \in R(c) \\ & \beta_{(c,r)}^d + \delta_p \geq d_{c,r} + (\delta_p) \cdot \sum_{\bar{t}} (1 - x_{c,r,\bar{t}}^d) \geq \beta_{(c,r)}^d \\ & \beta_{(c,r)}^a + \delta_p \geq a_{c,r} + (\delta_p) \cdot \sum_{\bar{t}} (1 - x_{c,r,\bar{t}}^a) \geq \beta_{(c,r)}^a \end{aligned} \quad (6.85)$$

6.4.5 Extensions du modèle**Description multiniveau**

Pour profiter de la complémentarité des deux formulations, nous avons présenté une combinaison des deux premières formulations mathématiques dans la formulation hybride

avec un même niveau de description pour chacune des parties. Cependant rien n'empêche d'adjoindre deux niveaux de modélisation distincts pour profiter à plein des concepts développés dans la partie modélisation.

En effet, la formulation hybride se prête bien à la superposition de deux niveaux de modélisation, schématiquement :

- une modélisation de type microscopique sur laquelle s'applique les contraintes de type temps continu (i.e. temps de parcours)
- une modélisation de type macroscopique sur laquelle s'applique les contraintes d'ordonnement. Dans cet autre niveau de description toutes les ressources où un changement d'ordonnement ne peut s'opérer sont agrégées.

Les contraintes d'espacement (expression de type temps continu) prennent alors pour référence les variables d'ordonnement définies sur la ressource macroscopique correspondante).

Cas de la fusion des deux formulations

Si on suit l'option d'une formulation mixte juxtaposant les deux formulations intégralement, on peut envisager une hypothèse moins restrictive où l'arrondi de toutes les valeurs serait à la partie inférieure dans la partie où le temps est discrétisé (contrairement aux formulations du modèle à temps discret seul ou du modèle mixte présenté ci-dessus).

C'est à dire que les nouvelles valeurs des durées d'espacement (comme les durées de parcours ou de correspondance) deviennent des multiples de périodes de temps :

$$\varepsilon'_{c,r} = \delta_p \times \left\lceil \frac{\varepsilon_{c,r}}{\delta_p} \right\rceil$$

où $\varepsilon_{c,r} \in \mathbb{R}$ est la durée exacte de l'espacement.

La sécurité est alors garantie par la partie de la formulation en temps continu qui utilise $\varepsilon_{c,r} \in \mathbb{R}$.

6.5 Mise en œuvre pratique des modèles mathématiques

Nous allons à présent compléter les modalités de mise en œuvre pratique de nos modèles mathématiques.

Modélisations et formulations

Les différents niveaux de modélisation peuvent être combinés. Deux raisons distinctes peuvent être à l'origine de cette combinaison.

Modélisation de comportements complexes : Certains modes de fonctionnement du système ferroviaire se modélisent aisément avec une combinaison de niveaux d'agrégation différents.

Par exemple : les enclenchements d'itinéraires (une fois une circulation engagée dans un itinéraire composé de plusieurs ressources - comme une série d'aiguillages en entrée de gare- la zone entière est verrouillée et dédiée à cet itinéraire uniquement) où :

- les contraintes d'espacement du type succession peuvent s'appliquer sur les ressources du niveau microscopique (si une circulation suit la première),
- les contraintes d'espacement du type conflit s'appliquent, elles, sur la ressource agrégée (la zone complète) entre familles de circulations aux itinéraires incompatibles.

Complexité du problème mathématique : Afin d'assurer le respect des principes de sécurité, on dispose de deux options : soit on considère des couples de conflits (par paires de trains et paires de ressources composant des itinéraires incompatibles), soit on considère une ressource (itinéraire) de capacité unitaire (chaque ressource n'accepte qu'une circulation à la fois). Cette option de modélisation (et de formulation) change aussi le fond du problème mathématique.

Ainsi, dans le premier cas, une zone complexe (en particulier une entrée en gare) peut nécessiter un très grand nombre de couples de conflits. Dans le second cas, on peut modéliser cette zone en détail avec une formulation du problème assez compacte. Le nombre des contraintes respecte les ordres de grandeurs suivants (par analogie avec les notations du chapitre précédent, c représente le nombre de circulations, r le nombre de ressources, t le nombre de périodes) :

- $O(c^2.r'^2)$ contraintes de type (6.13) et (6.14) associées à chaque couple de trains en conflits pour des couples de ressources (composant des itinéraires) incompatibles,
 - $O(r.t)$ contraintes de type (6.41) associées à chaque itinéraire (ressource agrégée),
- où $r \leq r'$ (r' étant le nombre de ressources élémentaires agrégées formant les itinéraires en nombre r).

6.5.1 Formulations et résolution

Une solution initiale

Principe : la simulation fournit (quasi instantanément) une solution initiale réalisable : c'est ce qui se produira sur le terrain si aucune action n'est prise. Par exemple : sans intervention de la part des opérateurs, le système de signalisation se comportera la plupart du temps comme un système FIFO. Attention cependant à certaines programmations qui attendront effectivement l'arrivée de tel train prévu avant d'engager toute autre action (MISTRAL).

Nous disposons d'autre part de la grille originelle qui correspond au déroulement prévu par le plan de transport théorique (qui, elle, n'est plus réalisable en l'état depuis l'apparition de l'incident). Notre objectif est de prendre des décisions visant à modifier cette solution initiale -qui s'exécuterait par défaut- afin qu'elle se rapproche au mieux de la solution optimale selon des critères prédéfinis (i.e. évaluation dans la fonction objectif). On est donc beaucoup plus proche d'un schéma de résolution de type recherche locale que dans le cas de la planification.

Remarque : dès le début de notre algorithme, nous disposons d'une solution réalisable et (donc) d'une borne maximale.

Extensions : des comportements simples (triviaux) de régulation peuvent être intégrés dans la simulation (exemple réaffectation automatique de quai si le quai est occupé par une circulation qui y est immobilisée suite à un incident). La solution initiale n'en sera que plus intéressante et proche de ce qui se déroulerait sans outils d'optimisation (action triviale de l'opérateur).

Éviter les symétries

Un facteur de performance important consiste à éviter autant que possible la symétrie. C'est à dire, pour la partie qui concerne les variables entières, éviter des choix de valeurs égales (dans la fonction objectif).

A cette fin, l'option qui correspond au choix du plan de transport théorique est donc systématiquement considérée comme préférable. Une pénalité (minime et croissante dans le cas de plusieurs choix) est appliquée pour les autres options : c'est le critère coût du changement dans la fonction objectif cf. 6.2.3, pour les choix d'ordonnancement (disjonctifs) ou choix d'itinéraires (quais, voies) alternatifs.

Par exemple : lors de l'affectation des voies en gare, supposons qu'il y ait trois quais disponibles. L'option choisissant le quai correspondant au plan de transport initial n'est pas pénalisée (supposons que ce soit le quai 1), le choix du quai 2 sera, lui associé à une pénalité ϵ , le choix du quai 3 sera associé à une pénalité $2.\epsilon$.

En effet, ces différentes options ne sont jamais réellement égales ; il y a bien souvent d'excellentes raisons métier de préférer le choix de l'une à une autre (dans l'exemple précédent : même quai pour la correspondance, dissociation des flux voyageurs, regroupement par zones de destination, etc).

Bien sûr, ces données ne sont pas modélisées ni même nécessairement accessibles aujourd'hui. Néanmoins, on peut faire le pari que le plan de transport d'origine est déjà le résultat de ces attentions. Il est inutile de déconstruire cette solution, si cela n'est pas motivé par un gain effectif en termes de régulation (qui surpasse le contre coût affecté au changement).

Ainsi on conjugue à la fois une meilleure efficacité dans les algorithmes de résolution et une meilleure qualité des réponses du point de vue métier.

MISE EN ŒUVRE INDUSTRIELLE ET INTÉGRATION SYSTÈME

En effet, quand on n'a pas le temps de poser au préalable de bons fondements, un très habile homme le peut faire dans la suite ; mais c'est avec une peine infinie pour l'architecte, et un péril éminent pour l'édifice.

Le Prince (1513)
Nicolas Machiavel

Les modèles décrits précédemment permettent de développer un logiciel d'optimisation. Pour devenir un outil d'aide à la décision efficace dans le cadre de la régulation du trafic en cas d'aléas il doit nécessairement s'insérer dans un système logiciel complexe.

Or, de l'architecture de ce système dépendent pour beaucoup les performances finales d'un outil opérationnel. C'est pourquoi une grande partie de la thèse s'est attachée à tester, définir et développer ce cadre.

Ainsi, nous aborderons l'aspect intégration système dans ce chapitre au travers de deux schémas fonctionnels : le premier décrit les modules et les projets qui nous ont permis dans la première phase de prototypage de donner jour à une vision cohérente du système logiciel global. Dans un deuxième temps, nous décrirons l'architecture que nous avons conçue pour le système cible, en opérationnel.

7.1 Intégration en phase d'étude

L'intégration de nos premiers travaux reprend les bases du projet LIPARI (cf. [21]). Le fonctionnement du système ferroviaire est reproduit au sein d'un simulateur ferroviaire (SISYFE). Des incidents sont injectés dans une grille horaire en suivant un plan d'expérience prédéterminée. L'outil doit détecter l'apparition de l'incident, calculer une nouvelle solution (optimisation) et donner des ordres de régulation en conséquence (exécutés par le simulateur). La partie plan d'expérience (analyse statistique des résultats de ces décisions, comme celle des catégories, fréquences et lieux des incidents à générer) a été confiée à l'outil *SARDAIGNE*.

7.1.1 Les modules en interaction

Les parties qui suivent s'attachent à décrire brièvement les parties gérées par chacun de ces modules afin de comprendre les interactions à l'œuvre et mieux appréhender la globalité du système logiciel.

SISYFE

(SIMulateur de SYstème FERroviaire)

Objectifs à l'origine du projet :

- évaluer par simulation si un système de contrôle-commande et de gestion des circulations standard, défini comme modèle de référence, peut acheminer un trafic extrêmement dense avec un niveau de régularité satisfaisant,
- définir l'architecture cible d'un nouveau système de gestion opérationnelle des circulations en s'appuyant sur le système de contrôle-commande ERTMS 2 et des modules semi-automatiques de régulation des circulations (moyens avancés de contrôle-commande et de supervision / pilotage des circulations),
- quantifier par simulation les gains en régularité procurés par l'un et l'autre composants du système cible et par le système cible complet par rapport au système de référence.

Méthodologie : les infrastructures du réseau, le système de contrôle-commande, les trains, la grille théorique de circulation et le mode de pilotage (programmé ou automatique) des systèmes sont modélisés et simulés via *SISYFE*.

La couche Gestion des Circulations dialogue avec *SISYFE* en simulation temps réel pour recevoir des données courantes sur les mouvements des circulations et transmettre des consignes pour l'ordonnancement, le routage et la fluidification du trafic.

SARDAIGNE

(Statistiques Appliquées à la Rationalisation des Décisions pour Aider à Gagner en Efficacité)

Le projet *SARDAIGNE* repose sur une méthode de gestion du plan d'expérience et d'analyse statistique pour comparer les systèmes ferroviaires. Il vise à développer et à mettre en œuvre une méthodologie d'évaluation de la robustesse des grilles horaires.

Dans un contexte de production proche de la saturation, avec des circulations de plus en plus denses et complexes, simuler les modifications des systèmes ferroviaires (infrastructures, grilles horaires, système de signalisation ...) est essentiel pour en tester la robustesse, c'est-à-dire la résistance à un aléa. Le moindre aléa risque en effet d'entraîner de lourdes conséquences. La simulation préalable permet d'éviter un test réel difficilement réalisable pour d'évidentes raisons de sécurité et de coûts. Mais comment s'assurer de la valeur de cette simulation ?

Savoir ce qu'il faut simuler : quel que soit le domaine d'activité, nombre d'entreprises sont aujourd'hui capables d'effectuer des simulations de moyens de production. La question n'est plus de savoir simuler, mais de savoir ce qu'il faut simuler. Dans le cas de la SNCF, les experts de l'infrastructure disposent de *SISYFE*. Cet outil de simulation permet de configurer le système ferroviaire et d'y introduire des incidents.

L'intérêt de *SARDAIGNE* est de proposer une démarche statistique pour choisir judicieusement les simulations. La comparaison des configurations incident par incident est en effet partielle : selon l'aléa retenu, différents systèmes ferroviaires peuvent paraître, chacun à leur tour, les meilleurs. Seule l'approche statistique prend en compte la diversité des cas.

SARDAIGNE est avant tout une méthodologie, en amont et en aval du simulateur. En amont, il aide l'expert à définir les configurations et les plans d'expériences à tester en se basant sur les données enregistrées du retour d'expérience qui concernent les catégories, durées et lieux d'incidents selon leur fréquence effective d'apparition sur le réseau étudié. En aval, il agrège et pondère les résultats et les présente sous forme d'indicateurs fiables avec, pour chacun d'eux, des intervalles de confiance.

7.1.2 La gestion des circulations

La partie régulation est elle-même subdivisée en trois modules.

Module de Détection Automatique d'Incidents

Sa mission consiste à détecter tout écart significatif entre le déroulement prévu et effectif de la grille de circulation (écart entre la grille théorique et grille réelle dans le simulateur). Il initie alors la phase de replanification.

Module de Régulation-Optimisation

C'est le module chargé de l'optimisation de la replanification des circulations. A partir des données métiers, le problème mathématique est généré, puis résolu. Le résultat (macroscopique) fait l'objet d'une interprétation dans le module suivant.

Module de Mise En Oeuvre

Ce dernier module traduit en commandes (sur une description d'infrastructure microscopique) pour le simulateur, les actions nécessaires au bon déroulement de la nouvelle grille horaire c'est à dire les consignes de vitesse, la programmation des ordonnancements et des itinéraires permettant de respecter les nouveaux horaires calculés.

7.1.3 Intégration

Le schéma [7.1.3](#) représente le système complet et l'intégration des différents éléments.

7.1.4 Conclusion

Si le prototype a apporté des résultats encourageants, il nous semble clair que nous avons encore des défis à relever avant d'aboutir à une architecture d'un système opérationnel efficace. Parmi ceux-là : le problème des données, un modèle qui facilite le passage de description microscopique à macroscopique et vice et versa, mais aussi le développement d'outils de prévision. Nous avons aussi développé la conviction que l'analyse détaillée du problème métier permettrait de mieux définir le problème mathématique et qu'il pourrait s'en suivre des gains appréciables tant en flexibilité face à des situations inenvisagées qu'en performances brutes. Enfin, cette première utilisation de la simulation et de l'optimisation nous a donné l'idée d'utiliser plus avant les synergies possibles, en opérationnel.

Ces réflexions ont mené à l'architecture qui suit et que nous avons commencé à mettre en place.

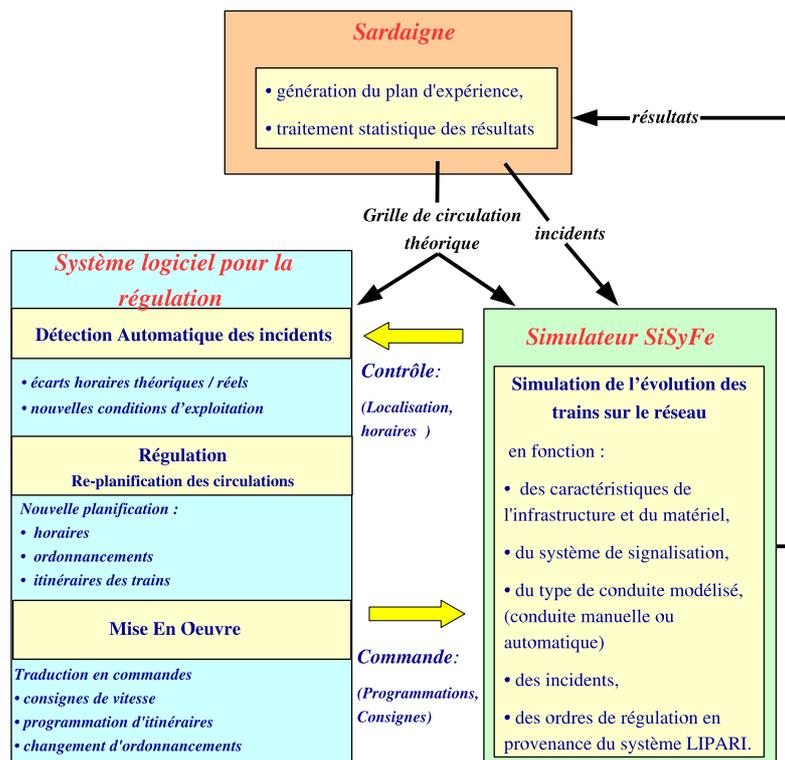


FIGURE 7.1 – Schéma d'intégration LIPARI en phase d'étude

7.2 Vers un outil opérationnel

Avant toute chose il est nécessaire de disposer (et fournir) en temps réel de toutes les données nécessaires à la mission de régulation. La première étape consiste donc à analyser les flux de données nécessaires à minima ainsi que ceux susceptibles d'améliorer encore l'efficacité de nos prises de décisions. L'étape suivante avant une quelconque mise en œuvre opérationnelle consiste à définir un modèle de données capable de décrire l'ensemble du système ferroviaire. Cette description sera par la suite commune à chacun des modules intervenant dans la régulation et la planification ferroviaire.

D'autre part, le module de replanification devra s'intégrer au sein des nouveaux outils de la gestion des circulations. Tous devront être étroitement reliés aux systèmes de contrôle-commande et en particulier aux systèmes d'information tant en amont qu'en aval du processus de décision. Or cette phase d'intégration peut bénéficier d'une opportunité exceptionnelle : le processus de rénovation en cours de la gestion du trafic ferroviaire en France (définition des projets Régulation du Futur, et particulièrement des Commandes Centralisées du Réseau). La description des interactions telles qu'elles seraient souhaitables sera l'objet de la deuxième partie.

Enfin, l'objectif étant de définir concrètement les bases d'un outil d'aide à la décision, la dernière partie détaillera le fonctionnement précis du module capable de proposer des solutions *opérationnelles* à l'opérateur humain. Ces propositions pourront ensuite être validées, refusées, ou enrichies par ce dernier.

7.2.1 Concernant les données : analyse des flux de données nécessaires

L'étape préliminaire à nos descriptions consiste à définir concrètement quels sont les flux de données nécessaires à nos outils. (voir [30])

Succinctement, voici les principales questions à se poser :

Quelles informations ?

- localisation spatio-temporelle des circulations (discrète ou continue),
- les destinations et les missions de chaque circulation,
- les caractéristiques (dont la capacité, la longueur et la vitesse limite) de chaque infrastructure,
- la description du type, la localisation et une estimation de la durée des incidents,

Qui doit en disposer ? Tous les acteurs de la GOC bien évidemment, mais plus particulièrement :

- le Gestionnaire d'Infrastructure,
- les Entreprises Ferroviaires,
- les équipes en charge directement de la régulation,
- les équipes d'intervention (internes : caténaires, rails et externes, samu, pompiers, ...)

- les clients (passagers et transporteurs).

Toutes les informations ne sont pas nécessaires à tous les acteurs. Cependant toutes ces informations précédentes doivent être accessibles par chacun des acteurs selon l'utilisation qu'il en aura. C'est la définition même d'un système d'information.

On trouvera plus d'informations sur les flux de données dans le rapport [66] par exemple.

Comment communiquer ? Les nouvelles technologies de l'information et de la communication inondent notre quotidien, pourtant la mise en place au niveau industriel de telles technologies à grande échelle semble prendre le pas. Quand un processus de déploiement est lancé il est peu probable qu'on fasse machine arrière ou même qu'on le remplace rapidement (au rythme de remplacement des produits grand public). On retrouve donc un décalage entre les solutions techniques possibles et ce qui est effectivement disponible. Pourtant ces solutions sont ou seront prochainement déployées.

- entre opérateurs : via le GSM-R (format GSM spécifique pour les communications ferroviaires), les réseaux IP,
- vers les utilisateurs : via les panneaux d'information (fixes : itinéraires bis/ dynamiques), les annonces (contrôleurs embarqués ou en gare), internet.

Les données sortantes : Les **équipes de régulation** doivent disposer de :

- un ensemble d'options de régulation,
- un critère clair d'évaluation de la performance,
- une estimation prédictive des flux de voyageurs,
- une adaptation des roulements (matériel et personnel),
- un ensemble de décisions de GOC (solution optimisée au regard des objectifs),
- l'état projeté du système ferroviaire correspondant,
- les consignes détaillées à mettre en œuvre.

L'information des clients passe par :

- la mise à jour des temps de transport,
- la mise à jour des correspondances.

Les **conducteurs** doivent connaître les nouvelles :

- consignes de vitesses,
- missions,
- information sur la situation (anticipation de la signalisation, consigne de comportement).

7.2.2 EIFEL et le modèle de données

(Echange d'Informations Ferroviaires Entre Logiciels)

Le projet EIFEL est la réponse aux premières questions soulevées en conclusion du prototype LIPARI. En effet, il est nécessaire de définir un modèle de données qui servira de plateforme d'échanges entre toutes les applications du système.

Le modèle de données d'EIFEL a pour ambition de couvrir l'ensemble du système ferroviaire. Nous nous contenterons de la description concernant la régulation ferroviaire : à savoir la description de l'infrastructure et des mobiles principalement.

Objectifs

Le modèle de données : en premier lieu, il s'agit de définir un modèle de données métiers unifié, générique (reposant sur les mécanismes de l'abstraction) et évolutif pour le développement des outils informatiques pour la conception, la supervision, la replanification et l'optimisation du système de production ferroviaire.

Ensuite, ce modèle doit être utilisable de façon statique (description de tous les objets métiers utiles) et dynamique (résultats d'optimisation et de simulations, acquisition de données en temps réel).

Compatibilité multiniveau : enfin, un objectif majeur est de permettre la définition d'un modèle adaptable à tout niveau de représentation et qui permette la définition d'une instance d'un plan de transport avec des niveaux hétérogènes d'agrégation. Par exemple, concernant l'infrastructure, certaines zones peuvent être décrites de façon plus détaillées que d'autres. Pour le matériels roulant, certains matériels peuvent être décrits précisément (caractéristiques de résistance à l'avancement, ...), d'autres de façon plus simple (valeurs forfaitaires de durées de parcours, ...). Enfin, on peut définir des fonctions d'agrégation des propriétés des « objets métiers » afin d'obtenir des propriétés « agrégées » cohérentes, comme pour le calcul des espacements sur une agrégation de canton.

EIFEL a donc été développé selon les spécification du Modèle pour les Outils de la Régulation Ferroviaire du chapitre 5 et permet ainsi les opérations d'agrégations-désagrégations qui y sont décrites.

Les services : le projet EIFEL offre aussi un ensemble de services métiers et techniques réutilisables, ainsi que des algorithmes de haut niveau sur ce modèle, tels que :

- la génération de données (signaux, consignes roses, ...)
- le parcours générique dans les données,
- la persistance des données,
- les tests de cohérence et d'intégrité des données,
- un calcul de marche,
- les moteurs d'optimisation et de simulation.

Acquisition et actualisation des données

Les types d'objets métiers (et sources de données) modélisés dans EIFEL sont les suivants :

- L'infrastructure et les travaux reposent sur le modèle RESEAU (+ ARMEN en 2009).
- Les sillons et circulations (horaires, missions, ...) sont générés par importation des données de HOUAT.
- Les matériels roulants suivent les caractéristiques techniques de type THOR.

– Les incidents sont issus de la base Bréhat pour le Rex en 2009.

Ainsi, des imports réguliers de RESEAU, HOUAT et THOR sont effectués (à terme, Armen, Bréhat, bases travaux et bases ressources matériels et humaines).

Néanmoins, l'outil en phase de développement devra collaborer avec différents modèles de données (métiers) existants en attendant une mise en place complète du modèle EIFEL en tant que socle commun aux données de toutes les futures applications.

Pour cela, en ce qui concerne ce document, un modèle (*design pattern* [26]) de type visiteur est mis en place. Il permet d'alimenter en données la partie spécifique à la régulation du modèle général EIFEL en particulier depuis les études de type Démiurge.

Architecture technique et liens avec le modèle mathématique : le modèle de données contient les contraintes métiers relatives à chaque élément (objet). Ce modèle métier sera au final transposé en problème mathématique (contraintes mathématiques) selon la formulation souhaitée.

La figure 7.2.2 illustre les différentes couches intermédiaires de modèles.

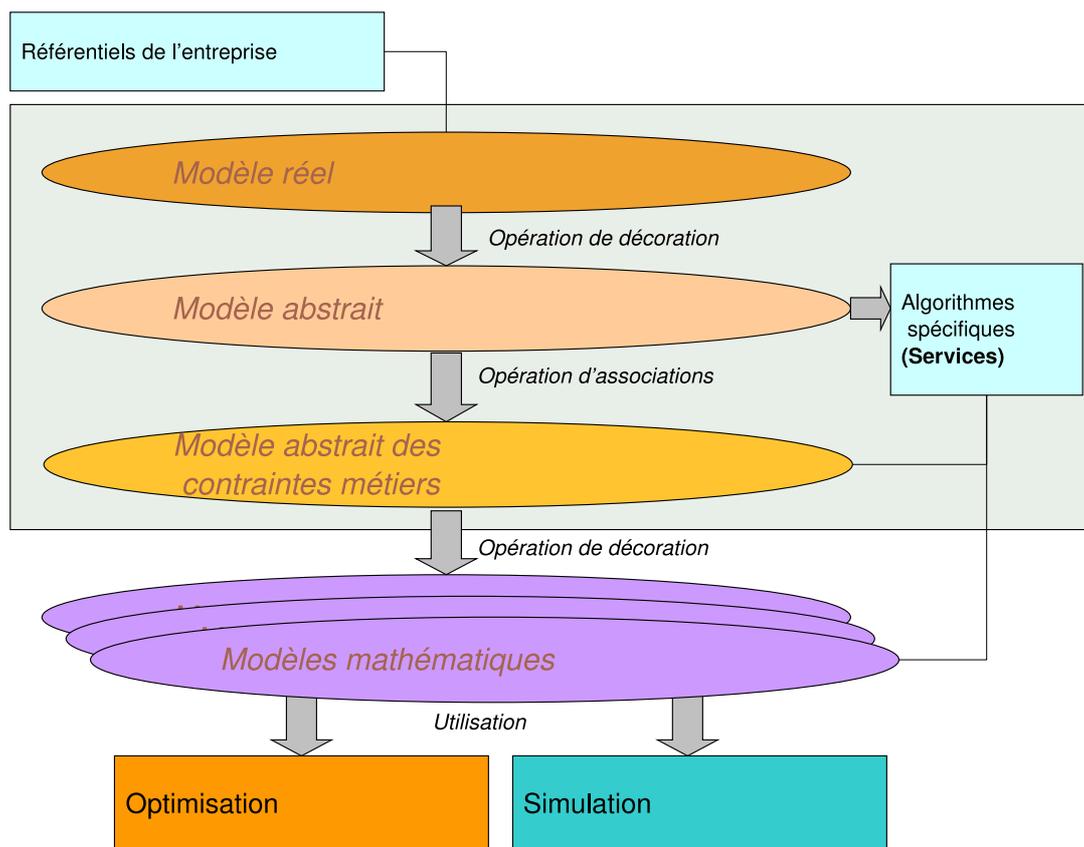


FIGURE 7.2 – Eiffel et les différentes couches de modèles

Afin d'aller plus loin dans la modélisation du monde métier d'EIFEL, on a défini

une **modélisation abstraite des contraintes** régissant le fonctionnement du système à partir des objets métiers abstraits du modèle. On peut les utiliser ensuite directement dans les services liés tels que le moteur d'optimisation, la simulation ou d'autres algorithmes spécifiques. En effet, celles-ci permettent d'associer : les objets métiers intervenant dans la contrainte, les variables de la contrainte et tous les paramètres de la contrainte. Ainsi, chaque (objet) contrainte du modèle abstrait peut être retranscrite en une contrainte (ou un lot de contraintes) du modèle mathématique correspondant.

7.2.3 Les autres principaux éléments du système logiciel

La simulation en temps réel du système ferroviaire

La prise de décision en gestion opérationnelle est actuellement réalisée par des experts ferroviaires qui s'appuient sur les systèmes d'information de l'entreprise et plusieurs outils facilitant essentiellement le suivi de la production ainsi que la représentation et la manipulation de l'information. Les décisions de modification du plan de transport demeurent dans une large mesure du ressort des experts, l'assistance de l'outil informatique pour faciliter la prise de décision étant restreinte, parfois inexistante.

Or les décisions prises par les experts impacteront le déroulement du plan de transport potentiellement de manière importante. Les régulateurs prennent des décisions dont ils ne peuvent pas toujours mesurer l'impact réel sur le terrain et qui reposent sur leur expertise et leur expérience tant les informations sur le problème sont parfois partielles. La construction mentale du scénario tel qu'il est susceptible de se dérouler est à la base même de toutes les décisions prises par les opérateurs (coordonnateur d'axe, coordonnateur régional, régulateur, chef de circulation, agent circulation ou opérateurs de futures CCR).

L'ambition de ce module est donc premièrement de permettre une visualisation du futur scénario qui va se dérouler. Nous désignerons ce processus par le terme de projection. Mais nous verrons plus loin que l'utilisation en temps réel de la simulation ferroviaire ouvre d'autres perspectives.

Objectifs techniques : Les progrès techniques permettent à présent d'automatiser la récupération des informations disponibles et de visualiser l'évolution probable du plan de transport en fonction de celles-ci. Un outil permettant entre autre de simuler l'avancement des trains (et donc de projeter l'évolution probable de la situation) sur un réseau conformément aux systèmes de contrôle a déjà été développé à la SNCF (cf. 7.1.1, [25]) démontrant la faisabilité du concept.

La simulation doit être capable de répondre à plusieurs objectifs, dont principalement :

1. **projeter** en temps réel la situation du système ferroviaire,
2. **valider** des choix de régulation (vérifier en détail le comportement du système),
3. simuler des grilles entières en mode conception (construction du plan de transport).

Interaction avec l'optimisation : la projection de l'état du système sans intervention particulière de la GOC (c'est à dire avec une régulation de type automate Mistral, ou règles de priorité) correspond toujours à une **première solution admissible**. C'est donc le point de départ naturel de l'optimisation.

Ce choix présente un double intérêt : d'une part la première solution est fournie instantanément -à l'échelle de la phase d'optimisation-, d'autre part la qualité de cette première solution correspond à une réalité industrielle. En effet, du point de vue métier, l'outil propose des actions supplémentaires pour améliorer une situation que l'expert maîtrise. Or, dans un environnement industriel il ne faut jamais oublier qu'une solution *optimale* n'a de sens que si elle est comprise et n'induit pas un bouleversement total du système de production (d'où l'intérêt de la réflexion sur la fonction objectif et la notion d'optimalité, en particulier avec la prise en compte d'une pénalité pour tout changement qu'implique la nouvelle solution).

Enfin, celle-ci sert de référence naturelle pour évaluer le gain apporté par les actions de régulation proposées.

Les outil de prévision des durées d'incident

Prévision de la durée d'incident : en opérationnel la durée de l'incident n'est pas connue avant la fin de celui-ci. À ce jour, les estimations de durée sont beaucoup trop vagues pour servir de base pertinente aux outils d'aide à la décision. Un outil dédié est donc nécessaire pour nous permettre de déterminer *a priori* la durée probable de l'incident. Cet outil a déjà fait l'objet d'études préliminaires. Il s'agit d'études statistiques sur la base des informations du retour d'expérience (REX) selon les catégories d'incidents, leurs localisations, les dates, etc. Cet outil est décrit dans [15].

Prévision de l'impact de l'incident : un outil similaire serait d'une grande utilité pour l'évaluation préalable de *l'impact de l'incident*, afin d'anticiper aussi la taille du problème à traiter. (i.e. à la fois l'étendue géographique et la plage horaire des perturbations consécutives à l'aléa, avant le retour à la normale)

En effet, à minima, l'outil précédent couplé à l'outil de simulation permettrait de déduire l'impact de l'incident sans intervention de la GOC. Cependant, cette projection est peu pertinente dès qu'une action de la GOC est nécessaire, justement (pour éviter une dérive infinie des circulations).

Cette première option garde donc un intérêt afin d'évaluer par exemple un seuil d'intervention (car si la grille horaire est suffisamment robuste, l'impact sera absorbé progressivement sans intervention). Dans tous les autres cas, une version plus évoluée est aussi nécessaire pour traiter à priori l'étendue de l'impact probable (de l'incident malgré les actions de régulation).

Nous proposons le même type d'outil statistique que précédemment, mais cette fois dédié à la prévision de l'impact de l'incident. Il est intéressant de noter que lorsque la base de données incidents sera suffisamment alimentée, l'outil devrait pouvoir anticiper l'effet (attendu) des actions de régulation (avant que celles-ci ne soient définies).

7.2.4 Fonctionnement du système logiciel

Nous proposons à présent une description chronologique de l'intervention de chacun de ces éléments. Cette description permettra de mettre en lumière les interactions et synergies que nous comptons utiliser dans le schéma opérationnel 7.2.4.

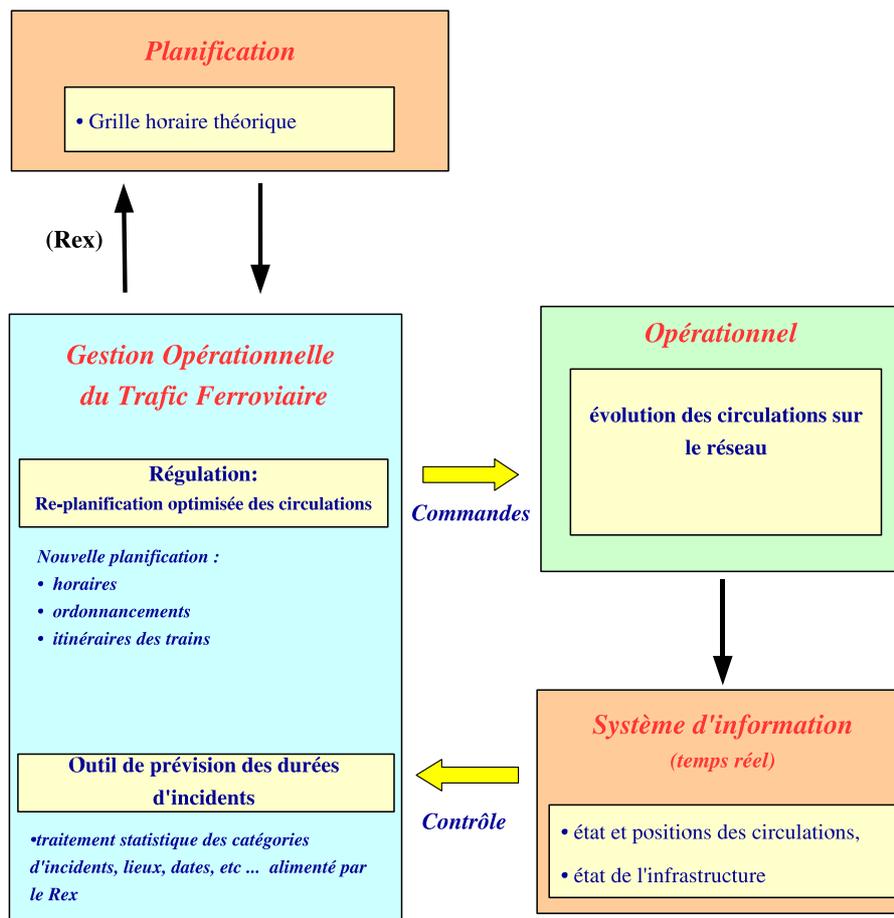


FIGURE 7.3 – Schéma d'intégration en opérationnel

Sont supposées acquises, les données concernant :

- l'état réel et théorique du système ferroviaire (Système d'Information),
- l'incident : identification, conséquences directes et estimation de la durée probable si non encore achevé,
- les conséquences de l'incident : conflits, retards,... (projection par simulation).

Détection d'incident

Le système traite en permanence le flux des données qui lui arrivent du terrain. Ces données sont traitées afin de générer continuellement des projection réactualisées. Comme

dans la pratique, le déroulement opérationnel diffère toujours du plan théorique, il apparaîtra toujours des différences entre la projection et la grille initiale. Cependant, au travers de ces projections les opérateurs seront en mesure d'observer la dérive des horaires et leurs conséquences prévisibles (bien avant leur apparition pour ces dernières). Au delà d'un certain seuil de déviation (par exemple il est détecté que la situation actuelle va déboucher sur un conflit, à terme), une alarme préventive est générée et le système entre dans la phase de processus suivante de recherche de solution.

Cependant le schéma général précédent ne couvre pas explicitement les incidents qui arrivent soudainement (par exemple une rupture de rail), nous allons donc détailler cette éventualité aussi. Dans ce cas, on doit attendre que le système d'information remonte au plus vite une information détaillée (identification du problème, durée prévisible), cependant le système permet de toute façon d'observer les conséquences dès qu'elles apparaissent au travers du monitoring (sans même que l'incident ne soit identifié). Au final, dans ce cas, l'étape suivante revêt seulement plus d'importance encore.

Analyse avant la phase d'optimisation

La projection précédente fournit d'importantes informations : la taille de la zone impactée, la nature des conséquences sur les autres circulations, etc. Un module spécifiquement dédié à l'analyse de ses informations peut améliorer immensément l'efficacité du processus complet.

Ce module de prétraitement des données doit s'appuyer sur une analyse statistique qui prend en compte la nature de l'incident, le lieu, la période du jour, etc, dans le but de fournir une **estimation de la durée prévisible de l'incident**. En effet, dans la plupart des cas la durée de l'incident n'est connue que lorsque celui-ci prend fin. À ce jour l'expertise seule guide l'opérateur (et malheureusement cette estimation est basée sur des informations très fragmentaires). Le principe de ce module est décrit dans [15].

Une autre information toute aussi cruciale concerne la **prévision de l'impact** de l'incident en cours. En effet, il est indispensable de déterminer le plus précisément possible la zone géographique concernée et la période durant laquelle les perturbations vont s'étendre. La solution la plus élémentaire consiste à lancer une nouvelle projection basée sur l'estimation de la durée d'incident. Cependant un autre module d'analyse statistique -basé sur le même principe que précédemment- pourrait apporter un surcroît de précision déterminant et une estimation de la fiabilité de la prédiction. Mais surtout, une fois le volume de données enregistrées suffisamment important, l'analyse statistique pourra aussi intégrer l'effet prévisible des opérations de régulation dans son estimation.

Bien sûr l'importance d'une intégration du système d'information est à nouveau mise en lumière ici, puisque les performances finales sont clairement dépendantes de la qualité et l'exhaustivité des données disponibles.

Au final les données brutes (de la projection principalement) doivent permettre de fournir :

- le domaine du sous-problème que nous allons optimiser (zone géographique et période temporelle de l'étude),
- un niveau (macroscopique) adéquat pour la description de l'infrastructure considérée,
- les valeurs précises des données nécessaires au problème (accès aux données stockées ou calcul par simulation -localisées-),
- la projection de l'évolution probable de la situation sans intervention ; ce sera la solution initiale de notre problème d'optimisation.

Recherche d'une meilleure solution

À présent que le problème métier est défini (i.e. le domaine d'étude, le niveau de description et les objectifs sont spécifiés), il est transcrit en problème mathématique, puis traité dans le but de fournir une solution (macroscopique) optimisée au regard des objectifs définis.

Cette étape prend fin dès que la limite de temps imparti au calcul est atteinte ou que la solution optimale est trouvée. Cependant, prouver l'optimalité de la meilleure solution disponible est un problème mathématique difficile. C'est donc une étape qui n'est pas prioritaire en opérationnel dans la mesure où elle est susceptible de consommer beaucoup trop de temps de calcul. Par conséquent, les méthodes de résolution ont intérêt à être paramétrées de manière à privilégier la recherche de nouvelles solutions plutôt que la preuve d'optimalité (de solutions déjà trouvées).

Dans le pire des cas (si la phase d'optimisation ne parvient pas à trouver une meilleure solution), la solution qui sera proposée par la suite à l'opérateur est donc similaire à la projection (qui est bien une solution réalisable du problème).

Analyse après la phase d'optimisation

La solution macroscopique fournie à la fin de la phase d'optimisation consiste en une grille horaire macroscopique (horaires uniquement disponibles pour les éléments d'infrastructures considérés -le plus souvent agrégés-), l'ensemble des ordonnancements entre circulations, les choix d'itinéraires (macroscopiques). Cette solution doit être réinterprétée (extrapolée) au niveau microscopique (par exemple en calculant précisément les consignes de vitesse pour chaque élément d'infrastructure).

Évidemment au cours de cette étape, il existe une multitude de variantes pour transformer la grille horaire macroscopique en horaires microscopiques. Par exemple, un train qui devrait rejoindre deux gares en 15 min. précisément peut le faire en maintenant une vitesse constante tout au long de son parcours, ou accélérer d'abord puis se laisser aller sur l'ère (sans traction moteur) ensuite. Dans les deux cas il est parfaitement possible que les contraintes opérationnelles soient respectées.

Durant cette phase, il peut donc être intéressant d'utiliser au maximum les informations collectées au cours de la phase d'optimisation afin d'améliorer encore la qualité de notre solution. Un post-traitement peut donc être effectué par exemple en s'appuyant sur la valeur des variables d'écart et ainsi améliorer un objectif secondaire (sans *backtracking*). Pour reprendre l'exemple précédent, on peut alors envisager de choisir l'option minimisant la consommation globale d'énergie ou privilégier la robustesse de la grille horaire solution.

Vérification détaillée via la simulation

La simulation permet alors de vérifier la validité d'une solution au niveau microscopique. Concrètement, la solution est pérenne tant que toute déviation constatée dans la simulation reste en deçà d'un seuil prédéfini (par exemple pas de conflit entre circulations, les ordonnancements calculés restent valides, ...).

Si on constate un écart trop important (qui ne dépend pas de la phase de post-traitement, mais qui révèle une infaisabilité effective de la solution calculée), le processus revient à la phase de prétraitement enrichie de ces nouvelles informations (c'est à dire avec des valeurs actualisées, de nouvelles contraintes métiers, ou avec une description plus détaillée des zones où apparaissent les problèmes).

Si les déviations constatées sont acceptables, la solution finale est fournie à l'expert.

Validation par l'opérateur

Si l'opérateur souhaite effectuer des modifications (notamment réorienter l'optimisation en modifiant les paramètres, comme les valeurs de la fonction objectif), une nouvelle simulation doit permettre de vérifier que cela n'a pas de conséquences inattendues (et permet de visualiser concrètement le nouveau déroulement souhaité des opérations). Si les modifications concernent des éléments importants, le processus revient au stade du prétraitement enrichi de ces nouvelles informations et d'une nouvelle solution initiale.

Quand la solution est finalisée et validée par l'expert, la grille horaire de référence est mise à jour et les commandes (consignes de vitesse, séquences, ...) sont envoyées automatiquement aux autres acteurs concernés via le système d'information.

Synthèse

On suppose disponibles :

- les caractéristiques statiques du système ferroviaire : infrastructure et mobiles (SI base de données statiques),
- l'état réel et théorique du système ferroviaire (SI dynamique),
- les données concernant l'incident (SI dynamique).

Le déroulement chronologique général suit les principales étapes suivantes :

1. acquisition continue des données du terrain¹ (contrôle),
2. projection en temps réel de l'état du réseau sans action de la GOC (simulation),
3. détection automatique de l'apparition d'un conflit dans la projection (DAI),
4. définition d'un périmètre géographique et d'une durée d'étude (simulation + outils de prévisions),
5. passage au niveau macroscopique de l'étude (modèle multiniveau),
6. traduction du problème abstrait métier en problème mathématique,
7. résolution du problème mathématique,
8. interprétation de la solution à l'échelle microscopique (MEO, modèle multiniveau),
9. vérification dans le simulateur de la validité,
10. validation par l'opérateur,
11. passage d'ordre (commande).

Le système propose :

- un ensemble de décisions de GOC (solution optimisée au regard des objectifs),
- l'état projeté du système ferroviaire correspondant,
- les consignes détaillées à mettre en œuvre.

1. (dont l'incident !)

RÉSULTATS ET ANALYSES

The game of science is, in principle, without end. He who decides one day that scientific statements do not call for any further test, and that they can be regarded as finally verified, retires from the game.

On the Problem of a Theory of Scientific Method
Karl Popper

Chronologiquement, nous avons d’abord étudié un outil de planification, puis la modélisation et les techniques de résolution retenues ont évolué vers la conception d’un prototype de re-planification dont les enseignements sont à la base du schéma que nous proposons. C’est pourquoi nous présentons ce chapitre en suivant cette même chronologie.

8.1 Enseignements de la planification

DEMIURGE est un outil de planification que nous avons fini de développer (pour la partie optimisation) au tout début de nos travaux. Il a été notre premier terrain d’expérimentation et a permis de fixer les principales options (de modélisation) pour le modèle mathématique à temps continu. DEMIURGE a par la suite continué d’évoluer -indépendamment de nos travaux- du point de vue fonctionnel afin de toujours mieux répondre aux attentes des utilisateurs sur le terrain. Aujourd’hui, DEMIURGE est un outil opérationnel en usage notamment sur le plateau commun pour la construction de grilles horaires.

Caractéristiques : DEMIURGE est un outil de gestion de capacité qui utilise un modèle simplifié nœuds-tronçons avec un niveau de description de type macroscopique. Il dispose de fonctionnalités étendues (gestion des itinéraires en gare, ...).

Méthodologie : nous avons testé le comportement en situation réelle d’utilisation. C’est à dire avec un temps de calcul limité à 5 min. En effet le logiciel est installé directement sur les postes des experts qui modélisent une infrastructure puis rentrent les données de l’étude (selon le cahier des charges retenu).

Ils sont ensuite amenés à modifier les données de l’étude selon les résultats obtenus au regard de leur expertise du problème. Cela les mène à tester de nombreux scénarios différents. C’est pourquoi il est nécessaire d’obtenir rapidement une solution pour valider ou infirmer la nouvelle hypothèse introduite dans l’étude (leur souhait étant idéalement une attente inférieure à 5 min, et surtout pas supérieure à 10 min).

Intérêt : les résultats (mathématiques) de cet outil sont intéressants car ils nous ont permis d’explorer les options de modélisation du problème, de formulation en temps continu, de paramétrage et de performances.

8.1.1 Paramétrage

Le logiciel utilisé *CPLEX* permet de nombreux réglages avancés [33, 34]. Nous avons mesuré l’influence de ces réglages sur les problèmes.

Réglages généraux

Afin d’optimiser le temps de calcul, nous avons donc défini une nouvelle liste de réglages par défaut. Ces réglages suivent les recommandations énoncées dans [37] et [35]. Ils ont été validés par de nombreux tests sur des études de références. Ces réglages ont d’ailleurs

Réglage	Originaux							Nouveaux						
	1 ^{ere} admissible			meilleure			Nb	1 ^{ere} admissible			meilleure			Nb
Solution	t	n	%	t	n	%		t	n	%	t	n	%	
Ordre 0	150	4669	8.14	230	6179	4.12	4	25	293	8.32	100	700	2.18	11
Ordre 1	-	-	-	-	-	-	0	90	998	7.20	260	2300	2.61	13
Ordre 2	240	5595	5.19	265	5887	2.34	5	50	930	5.28	200	2200	2.31	10

TABLE 8.1 – Comparaison de résultats obtenus avant et après paramétrage

été repris pour le logiciel LIPARI puis la partie optimisation d'EIFEL. Ils consistent à adapter la configuration du logiciel à la spécificité des problèmes et aux exigences des utilisateurs. Les paramètres importants concernent les options suivantes :

- on concentre la puissance de calcul sur la détection de solutions admissibles et non sur la preuve d'optimalité.
- on favorise les "sondages" lors de l'énumération pour choisir une bonne direction d'énumération.
- on permet à *CPLEX* d'utiliser des heuristiques (avec une fréquence automatiquement déterminée par le logiciel).
- on stocke l'arbre d'énumération sur le disque lorsqu'il dépasse une certaine quantité pour libérer de la mémoire vive.

Les coupes

On peut définir pour chaque type de coupes plusieurs stratégies de génération (voir [36]) parmi l'ensemble suivant : {aucune, automatique, modérée, agressive, très agressive}.

Or la suppression comme l'intensification des degrés de génération de coupes n'ont pas apporté de gains systématiques. Peut-être dans certains cas particuliers qui restent à caractériser nous pourrions en tirer parti (certains problèmes et coupes de Gomory par exemple). Néanmoins, nous avons conclu que les réglages par défaut (choix du niveau d'intensité de génération des coupes librement décidé par le logiciel) était les plus efficaces.

Résultats

Les résultats présentés dans le tableau 8.1 sont obtenus lors de calculs limités à 300 secondes (5 minutes) sur un même problème (ligne D) mais avec des données triées dans un ordre différent (un objectif métier consistait aussi à rendre la résolution moins sensible à l'ordre d'insertion des données par les utilisateurs) :

Où,

- *Ordre 0* correspond à l'ordre "naturel",
- *Ordre 1* au trie dans l'ordre alphabétique,
- *Ordre 2* idem en sens inverse (ante alphabétique).

Et,

- t est une approximation du temps de calcul nécessaire pour arriver à la solution considérée (il n'y a pas d'information précise concernant le temps pour trouver une

Étude	trains	nœuds	var réelles	var entières	contraintes
Tlse Nord-est	32	39	3553	4467	14828
LigneC cad moyen	24	45	1148	2720	12376
Lyon-Turin 1	15	24	433	1202	8094
Lyon-Turin 5	15	24	457	1347	32316
Saturation Nantes-Angers	104	15	1965	6677	35148
Tlse JLK hyp 3.2	36	39	6493	2958	9648
Base-Demande	24	45	1306	1363	1329
test saturation	118	15	1965	6141	31963
Tlse Nord-ouest	98	44	10041	5458	20219
RER D	210	41	4281	19755	84363

TABLE 8.2 – Taille des études métiers

- solution particulière avec *CPLEX*),
- n est le nombre de nœuds déjà traités,
 - % représente le saut de dualité (*gap*),
 - *Nb* indique le nombre de solutions admissibles trouvées au cours de la résolution.

On peut constater que les réglages apportés permettent des résultats à la fois plus performants et moins sensibles à l'ordre des données. Par la suite, pour chaque nouveau projet il a été impossible de trouver des réglages qui surpassent ceux résultant de cette première étude.

8.1.2 Analyse sur la recherche de solutions

Nous venons de voir que nous avons choisi un paramètre préférant la recherche de solution à la recherche de la preuve d'optimalité. Les résultats suivants justifient ce choix (au delà de l'aspect industriel du problème).

Exemple de tailles d'études traitées

Le tableau 8.2 présente quelques données concernant la tailles des études traitées en opérationnel qui ont servi de références pour nos tests.

Remarque complémentaire : la difficulté du problème ne peut être jugée sur la simple base du nombre de trains et de nœuds. Ce tableau a seulement un caractère indicatif. D'autres facteurs rentrent en jeu comme la topologie du réseau.

Qualité des solutions en fonction des durées de calcul

Voici résumés dans le tableau 8.3 les résultats obtenus en fonction de plusieurs durées de calcul (le symbole "*" signifie que la solution est prouvée optimale à la fin de la résolution).

(Les calculs sont effectués sur une station de travail de type Pentium 4 cadencé à 3,4 Ghz et équipée de 1 Go de ram, CPLEX 9.0, environnement Windows.)

Étude	1min	5min	15min	60min	borne
Tlse Nord-est	3439	3324	3304	3286	2654
LigneC cad. moyen	37646	37424	37424	37424	37092
Lyon-Turin 1	-396932	-396932	-396932	-396932	-396964
Lyon-Turin 1 Q	-397729	*	*	*	-
Lyon-Turin 2	-297296	*	*	*	-
Lyon-Turin 3	-496795	*	*	*	-
Lyon-Turin 5	-698516	-798340	-798351	-798370	-798555
Lyon-Turin 6	-396766	*	*	*	-
Lyon-Turin 7	-396533	*	*	*	-
Lyon-Turin 8	2107	*	*	*	-
Lyon-Turin 9	781	*	*	*	-
Sat. Nantes-Angers	-129996173	*	*	*	-
Tlse JLK hyp 3.2	27167	16776	16534	12151	2504
L3C routage	6542	*	*	*	-
Base-Demande (base)	1363	1329	1324	1318	1260
Base-Demande (demande)	-39998655	-39998674	-39998674	-39998683	$-3.9999e^7$
idem (base) + intervalle commercial	1348	1336	1336	1336	1260
idem (demande) + intervalle commercial	-39998531	-39998664	-39998669	-39998669	$-3.9999e^7$
test mémoire saturation	-79996560	*	*	*	-
routage Nantes-Angers	129996185	*	*	*	-
Violations ciblées					
Tlse Nord-ouest	113876	74672	74672	65627	967
RER D ordre expert	6772	6731	6729	6725	6579
RER D ordre 1	6798	6729	6729	6727	6578
RER D ordre 2	6727	6725	6725	6725	6579
Stéphane B. 1	6893	6759	6754	6722	6576
Stéphane B. 2	6498	6449	6449	6449	6345
Stéphane B. 3	-	7278	7250	7240	6934

TABLE 8.3 – Évolution de la recherche de solutions en fonction de la durée de calcul

Au bout de cinq minutes la qualité de la solution ne progresse plus. On constate aussi qu'on est souvent proche d'une solution optimale du problème.

Conclusions

Un premier constat est que toutes les études (sauf la dernière) ont une solution au bout d'une minute. Ensuite celles-ci sont résolues pour beaucoup à l'optimum en moins de cinq minutes.

D'autre part, on remarque que les solutions sont quasiment identiques au delà de 15 min. et qu'aucune nouvelle preuve d'optimalité n'apparaît (ce qui est peut être lié aux réglages CPLEX).

Enfin il faut noter que ces délais de résolution correspondent à une utilisation réaliste des outils d'aide à la décision (ils répondent parfaitement au cahier des charges des utilisateurs).

8.2 Prototype de replanification

Nous avons ensuite développé la partie régulation du prototype LIPARI.

Caractéristiques : le prototype LIPARI s'appuie sur un modèle avec deux niveaux de représentation : un détaillé (microscopique), l'autre simplifié (macroscopique). L'architecture logicielle est basée sur la collaboration entre ces deux niveaux et l'interaction avec un simulateur (qui utilise la description détaillée). Le plan d'expérience a été supervisé par l'outil Sardaigne. La formulation mathématique utilisée ici prend en compte la totalité de la formulation du temps continu décrite au chapitre 6. Enfin, ce prototype a aussi servi de test pour la construction du modèle en temps discret, en particulier pour tester l'implémentation des méthodes pour la gestion des listes (*containers*) de variables lors de l'écriture des contraintes.

Intérêt : cet outil nous a beaucoup appris. Il nous a permis d'appréhender pleinement des problèmes de régulation avec la formulation en temps continu et de la faire évoluer. La dernière partie de cette section résumera d'ailleurs l'ensemble des aspects positifs et négatifs qu'il a permis de mettre en valeur.

8.2.1 Explosion des retards induits

Le tableau 8.4 présente une série de résultats (somme des retards exprimée en secondes induits après régulation -uniquement des vitesses-) obtenus en fonction de plusieurs durées d'incident. L'étude porte sur la LN1 : le TGV 6627 et retenu en gare du Creusot (solutions optimales calculées en moins d'une seconde).

On constate l'effet de seuil (ici entre 2 et 5 minutes) au delà duquel le retard n'affecte plus seulement la circulation initiale (qui parvient d'ailleurs à compenser son retard dans cette première phase) mais impacte aussi les autres circulations. L'explosion des retards induits est alors très rapide.

retard initial	1min	2min	5min	10min	15min
LN1	12	12	2261	14707	33985

TABLE 8.4 – Évolution du retard induit en fonction de la durée d'incident

Étude	circulations	(nœuds, tronçons)	var binaires	(colonnes, lignes)
Tours-Bdx	90	(43,42)	97412	(110013, 109323)
LN1	40	(44,45)	45373	(51673,52593)
Eole	542	(120, 139)	1789711	(18587330,3824570)

TABLE 8.5 – Taille des problèmes étudiés

N.B. : il ne s'agit pas de retards en fin de parcours mais bien du cumul de tous les retards -exprimé en secondes- à chaque point de mesure ; c'est pourquoi les valeurs deviennent vite spectaculaires (ceci est vrai pour toutes nos études).

8.2.2 Exemples de taille d'études traitées

Le prototype LIPARI a servi pour plusieurs projets prospectifs au cours desquels il a permis d'étudier des problèmes terrains divers. Le tableau 8.5 montre quelques exemples de taille des études traitées :

Ces tailles sont les tailles maximales traitées (modèle complet, avec tous les leviers d'action possibles). On ne peut bien sûr traiter les cas les plus compliqués que grâce à l'apport de la solution initiale. Néanmoins, plus la perturbation à gérer est importante (i.e. plus le nombre de valeurs de variables entières changent dans la solution optimale), plus la recherche de la solution sera lente. On sort alors du cadre d'une utilisation opérationnelle dans les cas extrêmes comme sur le réseau Eole (on peut d'ailleurs saturer la mémoire de la machine). Cela étant, il est peu probable qu'un problème sur le terrain fasse réellement intervenir l'ensemble des 542 circulations de l'étude (l'étude consistait à étudier le projet de liaison souterraine entre le réseau Paris St Lazare et Paris Nord -qui sont déjà, chacun pris indépendamment, parmi les réseaux les plus complexes du système ferroviaire français-).

8.2.3 Résultats du prototype

Aspects positifs

Tout d'abord, au niveau de la résolution du problème mathématique, ce projet a démontré que le modèle mathématique à temps continu associé à une résolution via CPLEX permet déjà une mise en pratique concrète de l'optimisation pour la régulation ferroviaire.

Il a d'ailleurs permis de raffiner le modèle jusqu'à prendre en compte les espacements dynamiques. La qualité métier des résultats a été validée par les experts au cours du projet d'étude de faisabilité d'un système de fluidification (l'étude portait sur la zone de Baudrecourt).

D'autre part, l'expérience montre la pertinence du concept d'architecture. En particulier, la rapidité de la simulation a ouvert des perspectives de collaboration simulation-optimisation très intéressantes.

Aspects négatifs

Pourtant ce prototype nous a aussi permis de tirer quelques leçons pour le projet opérationnel.

La principale difficulté provient de la disponibilité des données : c'est un écueil majeur, car nous avons pu constater à quel point le problème est sensible ; le problème d'optimisation ne peut se contenter de valeurs estimées. Généralement, l'estimation est volontairement pessimiste (principe de sécurité : les espacements sont surévalués par exemple) ; il est alors difficile de rattraper le retard (voire certaines estimations trop imprécises peuvent même ajouter inutilement du retard). Inversement, toute estimation trop optimiste risque d'invalider la solution au niveau microscopique (la simulation montre que la grille ne se déroule pas comme prévue, les risques sont grands de voir apparaître des surincidents générés par de mauvaises commandes de régulation). La disponibilité des données (au moins statiques) ne doit pas interférer dans le problème de régulation, cela ne relève pas du problème de recherche. Pourtant cette expérimentation souligne encore une fois l'urgence de lancer une initiative de collecte et de mise à disposition de ces données sous forme numérique.

Le deuxième enseignement concerne le passage problématique d'un modèle macroscopique à un modèle microscopique (i.e. interpréter les résultats de l'optimisation au niveau détaillé de la simulation). L'automatisation en l'état s'est révélée quasi impossible. Elle a monopolisé beaucoup de nos ressources.

En effet, les modèles de données sur lesquels reposent ces deux outils sont distincts ; l'outil opérationnel doit reposer sur un modèle de données unique (comme pour tous les outils de la chaîne horaire : de la phase de planification à la replanification en opérationnel).

D'autre part, le passage d'un niveau de description macroscopique à un niveau microscopique n'est pas unique ; une étude métier serait nécessaire pour définir une stratégie générale voire adaptative (celle-ci devrait probablement être liée à la réflexion sur les objectifs de la régulation) et ses conclusions pourraient être implémentées dans les règles d'agrégation-désagrégation.

Ce que le prototype n'a pas montré

Ici, les paramètres de l'étude sont totalement prédéfinis : la zone géographique de l'étude et la durée de l'incident étaient fixées *a priori* (la plage horaire de l'étude aussi). Le prototype ne répond donc pas à la question -cruciale en pratique- : comment déterminer les valeurs de ces paramètres de manière pertinente ?

De la même façon, le passage du modèle microscopique au modèle macroscopique était simplifié : les deux niveaux de représentation étaient fixés (la description par les experts a d'ailleurs demandé plusieurs semaines de travail). Dans la réalité la question du choix des niveaux de représentation se posera comme celle de la taille de l'étude : il

faut d'une part savoir définir un périmètre pertinent et d'autre part générer le modèle correspondant dynamiquement, sans compter qu'il faudra estimer la durée de l'incident et non se contenter de constater sa fin effective avant d'engager nos calculs !

Enfin, le protocole présente aussi une simplification vis à vis de l'aspect temporel : la simulation s'interrompt pendant la phase de calcul. Or, en réalité le temps va continuer à s'écouler et l'impact de l'incident se propager pendant la phase de calcul. L'état initial pris en compte pour l'optimisation ne doit donc pas être celui du temps du **début** de l'optimisation, mais bien celui de la **fin du processus**. Autrement dit une projection de l'état du système à l'échéance prévue de la fin du calcul est nécessaire (par exemple à +5 min.).

8.3 Une première version d'EIFEL

L'état d'avancement de la base de données EIFEL n'a pas permis à la fin de ce projet de tester ses caractéristiques dans le cadre de la régulation. En effet, seuls des jeux de données issus d'extrapolation des modèles précédents ont pu être générés ; des cas simples permettent de tester le fonctionnement élémentaire des trois modèles mathématiques, mais pas d'étudier leur efficacité réelle. En effet, une seule étude conséquente a pu être malgré tout exploitée mais la qualité des données ne permet pas de s'assurer de la cohérence du problème (au niveau métier et par conséquence de la cohérence des valeurs et même des types de contraintes générées). Par exemple, l'extraction est incomplète en matière d'horaires théoriques et génère des incohérences pour les contraintes de connexité de la formulation à temps discret (problème qu'on dépasse en permettant une grande plage horaire pour toutes les variables, mais demande en contrepartie de générer beaucoup plus de variables que nécessaire).

8.3.1 Taille des formulations

Cet exemple permet néanmoins de se faire une meilleure idée des tailles de problèmes associées à chaque formulation. Ainsi le modèle à temps discret et le modèle mixte sont implémentés comme précisé au chapitre 6, mais la valeur du pas de discrétisation est fixée à 30 secondes et la plage horaire de validité des variables discrètes est de 1800 secondes (une demi-heure). Cela génère déjà 600 variables discrètes par variable métier (le nombre des contraintes croît aussi en conséquence avec par exemple 599 contraintes de cohérence par variable métier).

Le tableau 8.6 illustre les données métiers (communes) de l'étude Lyon-Turin (impact sur le réseau Lyon Pardieu) :

circulations	ressources	contraintes métiers	variables métiers
9	69	1898	844

TABLE 8.6 – Taille de l'étude Lyon Turin

La taille des formulations mathématiques associées est illustrée dans le tableau 8.7 :

Catégorie de contraintes	temps continu	temps discret	modèle mixte
Durée de parcours min.	422	25320	25742
Cohérences	-	49796	49796
Connexités	412	24720	25132
Espacements	2404	82271	82271
Équivalences	-	-	844
Total	3238	182107	183785

TABLE 8.7 – Indicateurs sur le nombre de contraintes de la formulation mathématique

8.3.2 Tests préliminaires

Parmi les problèmes de données, il semblait clair que les valeurs de capacité de chaque ressource n'étaient pas bonnes (ce qui induisait une mauvaise catégorie d'espacements). Nous avons donc tenté de ne générer que des espacements de type ressource multicapacité (contraintes écrites par paires de trains) et de laisser tourner une optimisation sans incident pour corriger les problèmes et prendre cette solution comme référence (en soustrayant ses premiers résultats aux valeurs de retard qui correspondent aux instances avec les incidents). Malgré cela, il n'a pas été possible de faire fonctionner d'instance pour le problème mixte, et les résultats pour le temps continu et discret ne sont pas interprétables avec certitude.

En effet, les structures des fonctions objectifs diffèrent et ne les rendent pas directement comparables, car pour le modèle à temps discret, le domaine de validité d'une variable horaire est borné par le nombre limité de variables discrètes créées qui correspondent à autant de pas de temps associés. Si on donnait les mêmes valeurs au retard dans les deux fonctions objectifs des deux modèles, le coût de l'arrêt d'un train dans la fonction objectif serait très différent. Dans le cas du modèle à temps continu, celui-ci serait égal à la somme des retards en chaque ressource suivant l'arrêt, soit la différence entre l'horaire théorique et la borne maximale de l'étude (48 heures, ici). Dans le cas du modèle à temps discret, il serait seulement égal à la taille du domaine de validité de la variable horaire, c'est à dire la durée recouverte par l'ensemble des pas de temps pour lesquels les variables entières sont nulles, c'est à dire 3600 secondes ici. Pour éviter que la solution d'arrêter un train ne soit préférée, on surévalue dans la fonction objectif la valeur du dernier pas de temps associé à chaque variable horaire (cette valeur est équivalente à la taille de l'étude soit 48 heures ou 176800 secondes).

Nous présentons néanmoins les valeurs que nous avons obtenues pour une étude où l'incident consistait à retarder un train de fret dans une zone critique de la grille horaire de 82 puis 300 secondes. Le tableau 8.8 donne les valeurs des fonctions objectifs, en fonction des meilleures solutions primales fournies par les approches heuristiques génériques de CPLEX au bout de 300 secondes de calcul. La solution sans incident est ensuite fournie comme solution initiale aux calculs suivants.

Ces résultats semblent indiquer que les performances des méthodes primales de CPLEX

incident	temps continu	temps discret
0	347777	954900
retard de 82s	352167	2034030
retard de 300s	374473	2046480

TABLE 8.8 – Retards au bout de 300 secondes de calcul

seraient meilleures pour le modèle continu que pour le modèle discret. Ceci serait toutefois prévisible aux vues des différences de taille et de flexibilité des deux modèles. En effet, de nombreuses études de la littérature rapportent ce type de conclusions : quand un modèle comprend un très grand nombre de variables, les heuristiques d'arrondi de la solution linéaire (ou de plongeon dans l'arbre de *branch and bound*) ont du mal à sélectionner les variables pertinentes à arrondir parmi un plus grand nombre de choix possibles. Si par ailleurs, la fixation d'une variable restreint plus les solutions comme c'est le cas pour le modèle discret moins flexible, les arrondis aboutissent plus souvent à de mauvaises solutions. Cela n'indiquerait donc pas que le modèle discret est moins apte à induire de bonnes solutions primales que le modèle continu, mais seulement que les stratégies par défaut de CPLEX ne sont pas adaptées. Il conviendrait donc sûrement de développer des heuristiques primales *ad hoc*. On pourra notamment projeter la solution du modèle discret dans l'espace des variables de la formulation continue pour pratiquer l'arrondi dans ce dernier espace. Enfin, il faut ajouter que les réglages CPLEX n'ont pas été adaptés au modèle à temps discret (qui reprend intégralement les réglages calibrés pour le temps continu).

De plus, on a constaté pour le modèle à temps discret qu'au bout d'une heure de calcul la borne primale n'est plus que de 645570 pour l'incident de 300 secondes. On peut remarquer que cette borne primale a été grandement améliorée au bout d'une heure, en comparaison avec celle obtenue au bout de 5 minutes. Ce qui permet bien d'espérer pouvoir obtenir de bonnes bornes primales à partir de la formulation en temps discret, soit en améliorant la formulation, soit en développant des heuristiques plus adaptées.

Pour motiver d'avantage les espoirs que porte le modèle discret quant à son aptitude à induire de bonnes solutions primales réalisables, on aurait aussi souhaité comparer les bornes duales obtenues à la racines pour ces trois modèles. Le tableau 8.9 présente ces bornes duales.

incident	temps continu	temps discret
0	90797	357097
retard de 82s	94487	361238
retard de 300s	104297	373087

TABLE 8.9 – Bornes duales

Les bornes duales du modèle discret sont plus grandes que les bornes primales du

modèle continu car les retards ne sont pas pris en compte de la même façon. Néanmoins, le fait que la borne duale de modèle discret soit du même ordre de grandeur que la borne primale du modèle continu, laisse supposer que la borne duale du modèle discret est de bonne qualité et permet donc une bonne appréciation de la qualité d'une solution. De plus, cela montre que la borne primale du modèle continu est elle aussi de bonne qualité.

Enfin, le tableau 8.10 donne la valeur de la somme des retards en fin de parcours, qui correspond aux meilleures solutions primales précédentes.

incident	temps continu	temps discret
0	1155	4989
retard de 82s	1401	-
retard de 300s	1619	-

TABLE 8.10 – Valeurs du retard correspondant

Malgré les précautions concernant la valeur du dernier pas de temps, les deux solutions des problèmes avec incident de la formulation en temps discret (au bout de 300 secondes) font s'arrêter un train, ce qui explique d'ailleurs les valeurs élevées des fonctions objectif associées (à cause des pénalités).

8.3.3 Conclusion

Nous ne pouvons pas établir de conclusion définitive suite à cette étude sur les qualités comparatives de chacun de ses modèles.

Il est néanmoins possible que la modélisation en variables discrètes permette une meilleure estimation de la solution entière optimale. Par conséquent des solutions d'arrondis obtenues à partir de cette meilleure estimation duale pourraient donner de meilleures solutions primales. À la suite de cette pré-étude nous recommandons donc de développer des heuristiques primales spécifiques pour le modèle discret et le modèle mixte dès qu'il sera possible d'accéder à des jeux de données complets et fiables (issus directement d'EIFEL et non d'extractions).

Par ailleurs, l'étude du modèle à temps discret (et mixte) doit s'accompagner des autres modules de prétraitement afin de le rendre réellement viable; en particulier, en ce qui concerne la définition d'une durée d'étude et une taille du domaine de définition des variables horaires adaptée. Sinon, on doit générer des quantités inutilement importantes de variables (et de contraintes).

Enfin, ce type de formulation peut se révéler supérieur pour modéliser certaines parties d'infrastructure particulières (comme expliqué à la section 6.5). C'est dans ce cadre aussi qu'il serait nécessaire de pouvoir les tester et choisir le niveau de description en même temps que la zone où s'applique chaque modèle.

PERSPECTIVES ET CONCLUSIONS

9.1 Perspectives

9.1.1 Définition des objectifs

Il faudra que l'ensemble des acteurs (EF, GI, GID) arrêtent des objectifs pérennes et efficaces pour la régulation. Cette approche rationnelle pourrait d'ailleurs s'appuyer sur l'expérimentation via la simulation afin de mieux percevoir les conséquences des différentes options possibles.

Durant cette thèse il n'a pas été possible de définir de tels objectifs, c'est pourquoi nous avons retenu l'option la plus généraliste : minimiser une somme pondérée de retards. Pourtant, comme nous l'avons vu, la palette des options potentielles est large entre le système de règles aujourd'hui théoriquement en œuvre et les systèmes d'enchères qui sont parfois évoqués.

Par ailleurs, il est probable que cette réflexion ne puisse faire l'impasse de la remise à plat complète de la chaîne horaire. La phase initiale d'achat du sillon n'est pas complètement décorrélée de la manière dont on souhaite le voir géré en cas d'aléas. De plus, les besoins des clients ne sont pas homogènes : un sillon pour le fret et un sillon TER ne partagent pas les mêmes besoins. Par exemple, le premier peut généralement s'accommoder d'une certaine flexibilité des horaires mais requiert une garantie absolue d'acheminement, alors que dans le second cas, la régularité et une présence *suffisante* aux heures de pointes sont des critères primordiaux ; ainsi dans le cas d'une séquence de TER (surtout dans le cadre du cadencement), l'un d'entre eux peut être éventuellement supprimé.

Tous ces aspects peuvent être contractualisés différemment et de manière opposable par chacun des acteurs en apportant souplesse et clarté dans la gestion des circulations. Par contre ces attentes paraissent difficilement gérables à terme avec le système actuel surtout avec l'arrivée de nouveaux acteurs (d'autres entreprises ferroviaires).

L'efficacité finale du système ferroviaire dépendra donc grandement des suites de cette piste de réflexion.

9.1.2 De la gestion efficace des interactions entre les CCR

L'outil envisagé permet de gérer des périmètres plus vastes que ceux définis par le rayon d'action des futures Commandes Centralisées du Réseau (CCR). Mais le mode de résolution choisi écartait par choix un schéma du type *résolution collaborative* entre les différents CCR. Une telle approche mérite pourtant un important travail de recherche.

En effet, la question se pose du mode d'interaction le plus efficace. Si certaines pistes sont évoquées ou envisagées par ailleurs [51] le système ne pourra pas se contenter d'un système à base de règles pour être efficace.

Par ailleurs, si la logique aurait voulu que le découpage des zones CCR soit guidé par un travail de recherche et d'optimisation (afin de définir les zones les plus indépendantes possibles), il n'en est rien dans la réalité et cet aspect risque d'amputer l'efficacité de la coordination de la régulation.

La gestion des interactions entre CCR est donc un problème de premier ordre tant sur le plan théorique que sur le plan pratique si on considère le système tel que défini à ce jour.

9.1.3 Couplage avec les autres problématiques ferroviaires

Dans un premier temps on peut imaginer un outil se contentant de prendre en compte (en temps *qu'oracle*) le reste de la production ferroviaire. Ceci concerne principalement les ressources humaines et matérielles, puisqu'en opérationnel, il est fondamental que les roulements soient effectivement réalisables après une replanification des circulations.

Dans un deuxième temps il serait intéressant de coupler explicitement les problèmes mathématiques des roulements (de personnel et de matériel) à la problématique de la replanification car si l'organisation actuelle tend à dissocier ces deux aspects (le premier aspect étant du ressort de l'entreprise ferroviaire, le second relevant du gestionnaire d'infrastructure) personne ne peut se contenter de solutions de régulation irréalisables pour cause de manque de matériel ou de personnel.

9.1.4 Architecture technique et méthodes de résolution

L'amélioration des performances de l'optimisation sont clairement possibles. Pour commencer, ces améliorations pourraient s'appuyer sur une architecture spécifique d'optimisation **parallèle** à laquelle le problème pourrait bien se prêter.

Bien sûr l'autre piste concerne l'emploi d'autres **techniques de résolution**, sur la base de méthodes exactes comme heuristiques. Or comme nous l'avons vu, le panel de techniques déjà utilisées dans le milieu académique est vaste. Cependant, comme cela a aussi déjà été mis en lumière, il manque encore à ce jour une base de comparaison. Nul doute que la diffusion de standards (RailML) ouvrira la voie prochainement à des bibliothèques de jeux de tests de type Miplib ferroviaire. Ces collections de problèmes ferroviaires réa-

listes et industriellement pertinents permettront alors une réelle évaluation comparative de chaque nouvelle méthode.

D'autre part, la gestion opérationnelle de processus industriels est une entreprise qui n'est pas déterministe. En ce sens il serait intéressant d'envisager **l'optimisation stochastique**. L'apport de telles techniques serait probablement intéressant surtout durant la phase amont de conception de grille, avec des préoccupations de robustesses en tête. En effet cela pourrait être un moyen d'étudier et diminuer la sensibilité des grilles horaires face à l'aléa. Toutefois il faut se garder de vouloir l'utiliser à contre sens : parer au manque de données (par ailleurs accessible par la technique) via cette approche ne serait pas raisonnable ; des données incomplètes ou trop peu précises ne permettront pas une optimisation efficace avec des résultats précis (et nous avons souligné combien la précision est importante pour cette application industrielle) à partir de données lacunaires.

9.1.5 Le Système d'Information

Pour finir, le développement d'un système d'information intégré peut paraître hors du champ des problèmes de recherche. Pourtant du point de vue de la mise en œuvre industrielle cela représente probablement le défi le plus long et le plus dur à réaliser.

Or la clef de voûte de l'aide à la décision (de la décision tout court) est la disponibilité de données pertinentes. Aujourd'hui les conditions ne sont pas réunies pour une mise en œuvre rapide. Il réside donc dans cet aspect une source de gain d'efficacité important bien que mal exploitée en production.

Il est certain que la recherche académique peut ici aussi aider à définir ou identifier de nouvelles architectures pour des systèmes d'information efficaces.

9.2 Conclusions

Dans la première partie de la thèse nous nous sommes attachés à mieux définir les besoins inhérents à notre sujet et notamment faire évoluer notre démarche au regard de la bibliographie existante sur le sujet spécifique de la replanification mais aussi, plus généralement, par rapport aux techniques et modèles utilisés également en conception de grille horaire. En effet, ces deux problèmes possèdent beaucoup de similitudes du point de vue strictement mathématique. Pourtant le contexte opérationnel rend les techniques utilisées pour la première, inadaptées en l'état pour la seconde.

L'étude de la bibliographie nous a révélé que de nombreuses techniques ont été appliquées aux problèmes ferroviaires. Malheureusement il manquait jusqu'à présent des modélisations formelles exhaustives du système ferroviaire, des bases de comparaison et des études réalistes. C'est avec beaucoup d'intérêt que nous avons participé à l'émergence de projets de recherche en Europe pour faire avancer ces thématiques. Nul doute que le proche avenir sera riche en avancées déterminantes dans ces directions.

Fort de ces réflexions, nous avons entrepris de bâtir un nouveau modèle abstrait de représentation du système ferroviaire. Ce modèle comprend en particulier une description de l'infrastructure qui permet plusieurs niveaux de représentation qui incluent ses différents modes de fonctionnement (contraintes d'exploitation). Ce travail de formalisation comprend aussi une analyse des autres contraintes métiers (opérationnelles et commerciales) ainsi que des objectifs possibles de la régulation. Ce modèle pourra servir dans l'avenir de socle commun à l'ensemble des applications liées à la production ferroviaire : depuis la phase de planification, jusqu'à la régulation en temps réel, en passant la réaffectation d'itinéraires, la fluidification, la simulation, . . .

Sur la base de cette représentation, trois modèles mathématiques du problème ont été définis. Le premier, basé sur une représentation en temps continu, présente une innovation intéressante qui est la prise en compte dynamique des vitesses pour les espacements. La deuxième formulation repose sur une discrétisation du temps et une version hybride de ces deux formulations peut aussi s'accorder avec une description multiniveau décrite précédemment.

Ensuite, nous avons décrit l'architecture système qui paraît nécessaire pour la mise en œuvre opérationnelle de ce projet. Dans cette partie, une place importante est dédiée à l'explication des synergies attendues de la coopération entre un outil d'optimisation et un outil de simulation. D'autre part, des outils importants en pratique pour permettre la résolution du problème sont aussi esquissés. Surtout, nous nous sommes attachés à décrire un schéma de système logiciel et les interactions entre ses éléments, au fil d'une description chronologique illustrant une phase de régulation. Ce schéma général correspond à l'environnement logiciel tel qu'il devrait être mis en place en opérationnel afin d'aider efficacement la régulation du système ferroviaire en temps réel.

Puis, nous avons présenté les résultats que nous avons obtenus au cours des différentes expérimentations que nous avons menées sur plusieurs prototypes en milieu industriel. Celles-ci prouvent qu'une mise en œuvre pratique est souhaitable dès que les données nécessaires seront disponibles en opérationnel.

Enfin, en conclusion, nous proposons un ensemble de perspectives de recherches qui font appel à tout un panel de domaines et qui semblent indispensables pour l'accomplissement d'outils opérationnels efficaces.

ANNEXE A

ANNEXES

GLOSSAIRE

Acronymes ferroviaires :

CCR : Commandes Centralisées du Réseau (postes de régulation).

ERTMS : European Traffic Management System (système de signalisation).

LGV : Lignes à Grande Vitesse : infrastructure dédiée aux TGV.

EF : Entreprise Ferroviaire, exploitant de l'infrastructure.

GI : Gestionnaire d'Infrastructure.

Définitions générales :

Système : Assemblage d'éléments fonctionnant de manière unitaire et en interactions permanentes.

(Système) Temps Réel : un système est dit temps réel lorsque l'information après acquisition et traitement est encore pertinente.

Système de production : le système de production décrit l'ensemble du processus grâce auquel l'entreprise produit un bien ou un service apte à satisfaire une demande à l'aide de facteurs de production acquis sur le marché.

Système d'information (SI) : un système ou sous-système d'équipements, d'informatique et de télécommunications, interconnectés dans le but de l'acquisition, du stockage, de la structuration, de la gestion, du déplacement, du contrôle, de l'affichage, de l'échange (transmission ou réception) de données.

Modèle : représentation structurée des données, des traitements et des flux d'informations entre entités.

Modélisation : étape permettant de construire le modèle, elle est nécessaire à la mise en place du système d'information.

Définitions pour la modélisation :

Système ferroviaire : ensemble des ressources du monde ferroviaire et des interactions qui les régissent (la modélisation est une transposition de ce système).

Ressource : le système ferroviaire est composé d'un ensemble de ressources dont les états peuvent changer. Exemples : ressources matérielles, ressources humaines, ressources d'infrastructure. ... Ces entités du monde réel sont transposées dans la modélisation.

État : chaque ressource se trouve dans un unique état, à un instant donné. Exemple : signal fermé/ouvert/HS...

Événement : notification du changement d'état d'une ressource (dans la réalité ou dans le simulateur).

Incident, aléas : perturbation initiale dans le fonctionnement normal du système ferroviaire (en référence au déroulement théorique du plan de transport). À distinguer de ses propres conséquences et du bruit du système.

Conséquences, impact (d'un incident) : ensemble des perturbations impliquées par l'incident. La conséquence initiale est la première perturbation qui en résulte (il peut y avoir plusieurs conséquences initiales en parallèle : exemple le déraillement d'un train implique l'indisponibilité de la voie utilisée et la réduction des vitesses sur la voie de sens opposée). Les conséquences de rangs suivant sont les répercussions de cette (ces) toute(s) première(s) perturbation(s). Exemple : retards en cascade, effet boule de neige.

Bruit, dispersion de conduite : écarts du système contrôle-commande dus à l'inertie des matériels et l'hétérogénéité des comportements humains, normalement absorbée par les marges de régularité.

Robustesse : le concept est encore mal défini. Il semble qu'il couvre deux aspects d'une part l'optimisation des marges de régularité afin de palier efficacement au bruit du système, d'autre part structurer les grilles de manière à rendre le plus faible possible les conséquences d'un éventuel incident (cette deuxième partie semble assez peu considérée).

Instabilité : se mesure à l'aune des conséquences d'un incident sur une grille.

Marges de régularité : la marge de régularité est une fraction de durée supplémentaire associée à l'occupation d'une ressource. Elle est destinée à faire face à l'imprécision du système de contrôle-commande.

Règles de sécurité : le bon fonctionnement du système ferroviaire est assuré par le respect de dispositions de sécurité : par exemple le respect des espacements de sécurité (sur le terrain, au moyen du cantonnement, de l'ERTMS,...). Ces règles de sécurité constituent des contraintes intégrées dans la modélisation et auxquelles on ne peut déroger ni pour la simulation, ni pour l'optimisation.

Définitions ferroviaires de base :

Plan de transport : ensemble des dispositions destinées à organiser les ressources (sillons, matériels, agents, ...) mises en place pour transporter des biens ou des personnes par voie ferrée.

Infrastructure : ressources fixes (plate-forme, ouvrages en terre, ouvrages d'art, voie, signalisation ...).

Circulation : ressource associant un matériel roulant (train, circulation, manœuvre) et une mission.

Matériel roulant : ensemble de ressources matérielles en mouvement (les locomotives, les automoteurs et autorails, les wagons, les voitures, les engins de travaux).

Mission, desserte : succession des points d'arrêt commerciaux ou techniques (arrêts de service, ...) utiles à l'entreprise ferroviaire pour la réalisation de son plan de transport.

Sillon : ressource d'infrastructure ferroviaire requise par une circulation pendant une période donnée. Ce terme correspond à la notion de « slot » pour le trafic aérien. ([RFF :] "Il existe des sillons-catalogue, des sillons sur mesure, des sillons spot. ")

Interopérabilité : capacité des matériels à opérer entre eux, et particulièrement, d'une circulation internationale à fonctionner avec des règles techniques reconnues par chaque réseau ferroviaire national sur son trajet.

Blanc travaux : ressource d'infrastructure ferroviaire requise exclusivement pour la maintenance préventive, pendant une période donnée (i.e. interdite aux circulations).

Exploitation ferroviaire : toutes les activités directement liées au service de transport : la gestion des circulations, des manœuvres, des activités intermodales, des activités dans les gares et les terminus, etc ...

Graphique des Circulations ,Graphique Espace-Temps (GET) : représentation graphique bidimensionnelle (espace-temps) de l'ensemble des ressources sillons (et blanc travaux) en fonction des ressources infrastructures et du temps. Permet de visualiser l'espacement entre les circulations.

Graphique d'Occupation des Voies (GOV) : représentation graphique bidimensionnelle de l'ensemble des sillons en fonction des ressources voies et du temps. Permet de visualiser précisément l'utilisation des voies, principalement en gare.

Régulation : action de la GOC comprenant la redéfinition de l'ordre des circulations pour respecter le plan de transport et la sécurité. Elle s'appuie sur la planification de la circulation, les écarts horaires, la disponibilité du réseau géré, les règles de gestion formalisées (comme les priorités et les correspondances).

Conflit : demande d'utilisation concurrente d'une ressource nécessitant un arbitrage (ou engageant un système automatique).

Gestion Opérationnelle des Circulations (GOC) : ensemble des opérations visant à gérer les circulations en phase de production.

Règles de priorités ou Règles de la GOC : Au sens IN 2670 : procédures, moyens régissant la priorité à accorder entre circulations en conflits. A priori, les règles de la GOC devraient illustrer le *moyen* d'atteindre l'*objectif* de la GOC.

Accès au réseau : le gestionnaire de l'infrastructure met son infrastructure à la disposition de plusieurs entreprises ferroviaire, contre rémunération (cf. prix du sillon).

Prix du sillon, Péage, Redevance d'utilisation : prix qu'une entreprise ferroviaire (EF) doit payer au gestionnaire de l'infrastructure pour pouvoir utiliser un sillon.

Itinéraire, parcours : succession de ressources d'infrastructure à emprunter pour rejoindre deux ressources éloignées (ensemble ordonné de ressources connexes). Chemin dans le graphe de connexité. Généralement un parcours fait référence à une description plus macroscopique, un itinéraire à une description plus microscopique (exemple : parcours Paris-Nice, itinéraire en gare.)

Navette : une circulation qui utilise le même matériel qu'une autre circulation. Autrement dit, un même matériel est réutilisé à la fin de sa mission pour une nouvelle mission avec un nouveau numéro de circulation.

Point remarquable (PR) : Point géographique auquel peut être rattaché un horaire. Historiquement, les points remarquables comprennent des points aisément identifiables (visuellement) par un agent de conduite (ex : passerelle du kilomètre 4,0 , à la sortie de Paris - Gare de Lyon).

Marche type : vitesse de référence d'une circulation sur un parcours donné.

Détente : vitesse inférieure à la marche type d'une circulation.

Désheurage : l'horaire initial d'une circulation est modifié.

Fluidification : opération du contrôle local de circulation ferroviaire, consistant en l'utilisation des consignes de vitesse pour prévenir les situations conflictuelles. Compte tenu des inerties, anticiper un ralentissement vaut bien mieux que subir un arrêt obligatoire (puis relancer la circulation).

Capacité (d'une ressource d'infrastructure) : nombre maximum de circulations pouvant utiliser cette ressource par unité de temps.

Saturation : Etat d'une infrastructure ayant atteint sa pleine capacité (i.e. lorsqu'il n'est plus possible d'insérer une circulation dans le plan de transport.)

Canton : section élémentaire d'infrastructure encadrée par deux signaux (BAL, BAPR, BM). Zone élémentaires pour le système de sécurité.

Définitions pour la simulation :

Simulation : Méthode utilisée pour étudier des résultats sans effectuer l'expérimentation dans le monde réel. La simulation numérique n'est qu'une représentation du réel assise sur le modèle théorique sous-jacent.

Simulateur : outils permettant d'effectuer des simulations. Dans le monde ferroviaire il permet d'étudier l'évolution des circulations sur le réseau.

Plan d'incidents : description des paramètres incidents nécessaires à la génération d'un unique incident pour chacune des simulations.

Définitions pour l'optimisation :

Aide à la décision (AD) : Les outils d'aide à la décision permettent d'apporter des réponses pertinentes à des problématiques mettant en œuvre plusieurs choix possibles , d'aider au diagnostic et, plus généralement, de faciliter la prise de décision stratégique ou opérationnelle. Elle fait appel pour cela à l'optimisation.

Optimisation : processus visant à améliorer un résultat selon un (des) objectif(s) prédéfini(s). En général cela consiste à trouver un ensemble d'état des variables satisfaisant certaines contraintes et pour lequel la fonction objectif atteint une valeur optimale.

Leviers d'actions, choix, variables de décisions (GOC) Il s'agit des variables sur lesquelles on peut intervenir. Dans le cadre de la gestion des circulations, les leviers d'actions possibles sont les suivants : modification des vitesses, changement d'ordonnement, changements de voies, suppression/ajout d'arrêts, modification de mission/itinéraires, suppression de circulations. Ces actions doivent permettre d'atteindre les objectifs de la GOC.

Contrainte : (métier) réduit la liberté d'action, limite le domaine des possibles. (mathématiques) Relation limitant les valeurs possibles des variables. L'ensemble des contraintes délimite donc l'ensemble des solutions. Exemples : *les règles de sécurité*, les lois de la physique (accélérations, vitesses) sont des contraintes.

Objectifs : cible, but que l'on cherche à atteindre, le point où l'on se propose d'arriver, ce que l'on vise. Les objectifs de GOC doivent être définis clairement et permettre une évaluation quantifiable du résultat.

Fonction objectif : c'est la fonction qui permet d'évaluer (quantifier) la qualité nos décisions au regard de nos objectifs. Elle exprime donc pour chacune des variables et selon leur état une valeur correspondante. Ainsi la fonction objectif permet de déterminer parmi les solutions possibles (cf. contrainte) la (les) meilleure(s). On l'appelle aussi "fonction économique" ou "fonction coût" car un critère économique est souvent choisi.

TABLE DES FIGURES

3.1	Graphique Espace Temps	23
3.2	Graphique d’Occupation des Voies	24
3.3	Principe du cantonnement	25
3.4	Aiguillages en entrée de gare	25
3.5	Itinéraires incompatibles	26
3.6	Principe du cadencement	28
3.7	Détail d’une grille horaire cadencée (RER ligne C)	29
3.8	Principe de la fluidification	35
3.9	Exemple : LGV Sud-Est : détournement par Dijon	36
5.1	Représentation métier de la gare TGV Lyon St Exupéry	62
5.2	Différents niveaux de représentations pour le trajet du TGV 9835	65
5.3	Axes de circulation en gare de Bordeaux - Saint Jean	66
5.4	Illustration du graphe de connexité	67
5.5	Agrégation en série	68
5.6	Agrégation en parallèle	69
5.7	Exemple : partage d’une voie unique (représentation naturelle)	70
5.8	GET : mouvements possibles pour le partage de la voie unique	70
5.9	Graphique d’Occupation des Ressources : partage de la voie unique	71
6.1	Illustration de l’indicatrice d’état	96
6.2	Convention pour l’occupation d’une ressource	97
7.1	Schéma d’intégration LIPARI en phase d’étude	123
7.2	Eiffel et les différentes couches de modèles	127
7.3	Schéma d’intégration en opérationnel	130
A.1	Diagramme UML du modèle de données (abstrait) EIFEL	152

INDEX

- bibliographie, [157](#)
- cadencement, [26](#)
- contraintes commerciales, [77](#)
- contraintes d'exploitation, [74](#)
- contraintes métiers, [73](#)
- contraintes opérationnelles, [76](#)
- DEMIURGE, [134](#)
- description du sujet, [12](#)
- EIFEL, [123](#)
- gestion des ressources humaines, [16](#)
- gestion du matériel, [16](#)
- gestion opérationnelle des circulations, [16](#)
- identification de la demande, [16](#)
- LIPARI, [118](#)
- module de détection automatique des incidents, [120](#)
- module de mise en œuvre, [120](#)
- module de régulation, [120](#)
- notations mathématiques, [81](#)
- planification, [16](#)
- platforming, voir routage
- production ferroviaire, [15](#)
- replanification, voir gestion opérationnelle des circulations
- représentation macroscopique, [61](#)
- représentation microscopique, [61](#)
- routage, [16](#)
- SARDAIGNE, [119](#)
- signalisation, [22](#)
- SISYFE, [118](#)
- sommaire, [3](#)
- structure du document, [12](#)

RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] ARRIVAL, Algorithms for Robust online Railway optimization : Improving the Validity and reliability of Large scale systems, <http://arrival.cti.gr/>.
- [2] ASSAD, A. A. (1980), Models for rail transportation, *Transportation Research Part A : General*, 14(3) : p. 205–220.
- [3] BENHIZIA, F., DE ALMEIDA, D., DAUZÈRE-PÉRÈS, S. et DESSAGNE, G. (2010), Optimisation du plan de transport par planification intégrée des ressources, dans *ROADEF 2010*.
- [4] BJØRNER, D. (2007), *The Domain Book : A Compilation of Reports and Papers on Domain Models*, chapitre Railways, p. 157–185, Technical University of Denmark.
- [5] BOUF, D., CROZET, Y., LÉVÈQUE, J. et ROY, W. (2005), Etude comparée des systèmes de régulation ferroviaire : Grande-Bretagne, France et Suède. Analyse des règles du jeu et de leur mise en œuvre, enseignements pour la France, Rapport technique, Laboratoire d'économie des transports pour la direction de la stratégie de la SNCF.
- [6] BOUVAREL, P. (2006), Cours de chemin de fer. exploitation, Signalisation, Installations de traction électrique (notions).
- [7] BRANNLUND, U., LINDBERG, P. O., NOU, A. et NILSSON, J. E. (1998), Railway Timetabling Using Lagrangian Relaxation, *Transportation Science*, 32(4) : p. 358–369.
- [8] BUSSIECK, M. R., WINTER, T. et ZIMMERMANN, U. T. (1997), Discrete optimization in public rail transport, *Mathematical Programming*, 79(1-3) : p. 415–444.
- [9] CACCHIANI, V., CAPRARA, A. et TOTH, P. (2007), A column generation approach to train timetabling on a corridor, *4OR : A Quarterly Journal of Operations Research*.
- [10] CAIMI, G. (2009), *Algorithmic decision support for train scheduling in a large railway network*, Thèse de doctorat, ETH Zürich.
- [11] CAIMI, G., FUCHSBERGER, M., BURKOLTER, D., HERRMANN, T., WÜST, R. et ROOS, S. (2009), Conflict-free train scheduling in a compensation zone exploiting the speed profile, dans *ISROR RailZurich2009*.

- [12] CAPRARA, A., FISCHETTI, M. et TOTH, P. (2002), Modeling and solving the train timetabling problem, *Operations Research*, 50 : p. 851–861.
- [13] CAPRARA, A., KROON, L., MONACI, M., PEETERS, M. et TOTH, P. (2006), *Passenger Railway Optimization*, tome Transportation, p. 129–187, Elsevier.
- [14] CAPRARA, A., MONACI, M., TOTH, P. et GUIDA, P. L. (2006), A Lagrangian heuristic algorithm for a real-world train timetabling problem, *Discrete Appl. Math.*, 154(5) : p. 738–753.
- [15] CHANDESIRIS, M. (2008), Dynamic and real-time prediction of duration of incident, dans *World Congress on Railway Research 2008, Seoul*.
- [16] CORDEAU, J.-F., TOTH, P. et VIGO, D. (1998), A survey of Optimization Models for Train Routing and Scheduling, *Transportation Science*, 32(4) : p. 380–404.
- [17] CORMAN, F. (to appear), *Real-Time railway traffic management : dispatching in complex, large and busy railway networks*, Thèse de doctorat, Delft University of Technology.
- [18] CORMAN, F., D’ARIANO, A., PRANZO, M. et HANSEN, I.A. (2009), Effectiveness of dynamic reordering and rerouting of trains in a complicated and densely occupied station area, dans *ISRO RailZurich2009*.
- [19] COUR DES COMPTES (2008), Le réseau ferroviaire : une réforme inachevée, une stratégie incertaine. rapport public thématique., Rapport technique, Cour Des Comptes.
- [20] D’ARIANO, A. (2008), *Improving Real-Time Train Dispatching : Models, Algorithms and Applications*, Thèse de doctorat, Delft University of Technology.
- [21] DESSAGNE, G. et LERIN, C. (2005), LIPARI : optimisation de la gestion des circulations en trafic dense, dans *ROADEF 2005*.
- [22] DJELLAB, H., REBAÏ, R. et DE VULPILLERES, V. (2004), *Projet DEMIURGE, Modèle mathématique*, SNCF and Eurodecision, SNCF Direction de la Recherche RP, 45 rue de Londres, 75379, Paris Cedex 08, France.
- [23] DUPUIS, M. (2004), La construction du graphique horaire, *Revue générale des chemins de fer*, p. 7–23.
- [24] ERTMS, European Railway Traffic Management System, <http://www.ertms.com/>.
- [25] FONTAINE, M. et GAUYACQ, D. (2001), SISYFE : a toolbox to simulate the railway network functioning for many purposes. Some cases of application, dans *World Congress on Railway Research 2001, Köln*.
- [26] GAMMA, E., HELM, R., JOHNSON, R. et VLISSIDES, J. (1995), *Design patterns : Elements of reusable object-oriented software*, Addison-Wesley Longman Publishing Co., Inc.
- [27] GÉLY, L. et DESSAGNE, G. (2007), Rescheduling @SNCF, dans *Robust planning and Rescheduling in Railways*.
- [28] GIBSON, S. (2003), Allocation of capacity in the rail industry, *Utilities Policy*, 11(1) : p. 39–42.
- [29] HAENEL, H. (2008), Une organisation ferroviaire à la hauteur de nos ambitions, Rapport technique, Ministère de l’écologie, du développement et de l’aménagement durables.

- [30] HANSEN, I.O. (2007), Feedback and feedforward rescheduling, TRAIL masterclass.
- [31] HUISMAN, D. (2004), *Integrated and Dynamic Vehicle and Crew Scheduling*, Thèse de doctorat, Tinbergen Institute, Erasmus University Rotterdam.
- [32] IAROR, International Association of Railway Operations Research, <http://www.iaror.org/>.
- [33] ILOG (2003), *CPLEX 9.0 Parameters*.
- [34] ILOG (2003), *CPLEX 9.0 User's Manual*.
- [35] ILOG, CPLEX Documentation, *CPLEX Performance Tuning for Mixed Integer Programs*.
- [36] ILOG Online Documentation, *Cuts, Parameters Affecting Cuts*, dans (Discrete Optimisation/Solving MIP/Using the Mixed Integer Optimizer).
- [37] ILOG, Online Documentation, *Performance Features of the Mixed Integer Optimizer*, dans (Discrete Optimisation/Solving MIP/Using the Mixed Integer Optimizer).
- [38] KREUGER, P., CARLSSON, M., OLSSON, J., SJÖLAND, T. et ÅÅSTRÖM, E. (1997), The TUFF train scheduler : Two duration trip scheduling on single track networks, Proceedings of the workshop on Tools and Environments for (Constraint) Logic Programming at the International Logic Programming Symposium ILPS'97, New York, USA.
- [39] KROON, L. (1996), Routing Trains through a Railway Station, *Transportation Science*, 30(3) : p. 181–194.
- [40] KROON, L., HUISMAN, D. et MARÓTI, G. (2008), *Railway timetable and Traffic. Analysis, Modelling, Simulation.*, chapitre Optimisation Models for Railway Timetabling, EurailPress.
- [41] LARROCHE, Y., MOULIN, R. et GAUYACQ, D. (1996), SEPIA : A real-time expert system that automates train route management, *Control Engineering Practice*, 4(1) : p. 27–34.
- [42] LIEBCHEN, C. (2006), *Periodic Timetable Optimization in Public Transport*, Thèse de doctorat, Technische Universität Berlin Institut für Mathematik.
- [43] LINDNER, T. et ZIMMERMANN, U.T. (2003), *Mathematics-Key Technology for the Future : Joint Projects Between Universities and Industry*, chapitre Train Schedule Optimization in Public Rail Transport, p. 703–716, Springer, Berlin.
- [44] MASCIS, A. et PACCIARELLI, D. (2002), Job-shop scheduling with blocking and no-wait constraints, *European Journal of Operational Research*, 143(3) : p. 498–517.
- [45] MERCIER, A. et SOUMIS, F. (2007), An integrated aircraft routing, crew scheduling and flight retiming model, *Computers and Operations Research*, 34(8) : p. 2251–2265.
- [46] MIPLIB, <http://miplib.zib.de>.
- [47] OLIVEIRA, E. et SMITH, B. (2000), A job-shop scheduling model for the single-track railway scheduling problem, Research Report University of Leeds.
- [48] PAÑERO HUERGA, J. A. (2004), *Conception d'un outil d'aide à la construction des sillons ferroviaires*, Thèse de maître, École nationale des ponts et chaussées.

- [49] PACHL, J. (2008), *Railway timetable and Traffic. Analysis, Modelling, Simulation.*, chapitre Timetable Design Principle, EurailPress.
- [50] PACHL, J. (2008), *Railway timetable and Traffic. Analysis, Modelling, Simulation.*, EurailPress.
- [51] PACIARELLI, D. (2007), Schedule coordination among dispatchers, TRAIL master-class.
- [52] PEETERS, L. (2003), *Cyclic Railway Timetable Optimization*, Thèse de doctorat, Erasmus University Rotterdam.
- [53] PETITJEAN, G. (2004), *DEMIURGE, conception de grille horaire, étude de capacité*, SNCF and Axlog, SNCF Direction de la Recherche RP, 45 rue de Londres, 75379, Paris Cedex 08, France.
- [54] QUERO, D. (2002), Le système MISTRAL de commande-contrôle des postes d'aiguillage, *Revue générale des chemins de fer*, p. 55–61.
- [55] RADTKE, A. (2008), *Railway timetable and Traffic. Analysis, Modelling, Simulation.*, chapitre Infrastructure Modelling, EurailPress.
- [56] RAILML, <http://www.railml.org/>.
- [57] RÉTIVEAU, R. (1987), *La signalisation ferroviaire*, Paris : Presses de l'École nationale des ponts et chaussées.
- [58] RODRIGUEZ, J. (2005), A constraint programming model for real-time train scheduling at junctions, *Transportation Research, Part B : Methodological*, 41 : p. 231–245.
- [59] SAHIN, G., CUNHA, C. B. et AHUJA, R. K. (2005), New approaches for the train dispatching problem, *Transportation Science*.
- [60] SVATEK, C. (2004), Priority rules in operations, Rapport technique, Rail Net Europe : working group Quality and Operations.
- [61] TÖRNQUIST, J. (2006), Computer-based decision support for railway traffic scheduling and dispatching : A review of models and algorithms, dans L. Kroon et R. Möhring (réds.), *5th Workshop on Algorithmic Methods and Models for Optimization of Railways*, Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss Dagstuhl, Germany.
- [62] TÖRNQUIST, J. (2006), *Railway traffic disturbance management : an experimental analysis of disturbance complexity, management objectives and limitations in planning horizon*, Thèse de doctorat, Blekinge Institute of Technology, School of Engineering.
- [63] TÖRNQUIST, J. et PERSSON, J. A. (2005), Train Traffic Deviation Handling Using Tabu Search and Simulated Annealing.
- [64] TÖRNQUIST, J. et PERSSON, J. A. (2007), N-tracked railway traffic re-scheduling during disturbances, *Transportation Research Part B : Methodological*, 41(3) : p. 342–362.
- [65] WEGELE, S., CORMAN, F. et D'ARIANO, A. (2008), Comparing the effectiveness of two real time train re-scheduling systems in case of perturbed traffic conditions, dans *Computer in Railways XI*, p. 534–544.

- [66] ZHANG, Y. (2006), *Les nouveaux outils de gestion des circulations*, Thèse de maître, École nationale des ponts et chaussées.