

UNIVERSITÉ DE MONTRÉAL

MÉTHODES D'OPTIMISATION POUR UN PROBLÈME STOCHASTIQUE
D'HORAIRE D'AUDIENCES JUDICIAIRES

VINCENT MARTIN-GAUTHIER
DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(MATHÉMATIQUES APPLIQUÉES)
AVRIL 2013

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

MÉTHODES D'OPTIMISATION POUR UN PROBLÈME STOCHASTIQUE
D'HORAIRE D'AUDIENCES JUDICIAIRES

présenté par : MARTIN-GAUTHIER Vincent

en vue de l'obtention du diplôme de : Maîtrise ès Sciences Appliquées

a été dûment accepté par le jury d'examen constitué de :

M. ANJOS Miguel F., Ph.D., président

M. DESAULNIERS Guy, Ph.D., membre et directeur de recherche

M. HERTZ Alain, Doct. ès Sc., membre et codirecteur de recherche

M. LABIB Richard, Ph.D., membre et codirecteur de recherche

M. GENDREAU Michel, Ph.D., membre

À Félicia Toda, dont les espérances au sujet de mon avenir académique n'auront finalement pas été trop déçues.

REMERCIEMENTS

Je puis, dans ces lignes, remercier celles et ceux qui m'ont aidé à faire de cette maîtrise une réalité. Je commencerai par remercier Julien Major, bien souvent malgré lui placé au cœur des péripéties de mes recherches. Dans la même veine, je salue Buck, pionnier du second cycle, qui fut d'une aide précieuse. Pour leur soutien moral et le délassément qu'ils ont toujours su m'apporter, je remercie également Jean-Sébastien, Phil et l'ensemble des sauces.

Je me dois de témoigner ma gratitude à mes directeurs, les professeurs Desaulniers et Hertz, qui ont cru en mes capacités et qui ont fait le pari de ce projet. Je mentionne à part le professeur Labib, pour bien prendre le temps de souligner d'une part sa rigueur, mais également le fait qu'il n'a pas craint de monter dans un train en marche lorsque ses lumières furent requises. À eux trois je dois de pouvoir aspirer à ce grade de maître.

Malgré mon horaire erratique, j'ai pu bénéficier de la camaraderie de mes semblables. Je chéris mes échanges avec Hocine, Charles et Maikel, les passionnantes discussions politiques avec Guillaume ou les fameuses parties de cartes avec Mélisende et tous les autres. Leur dynamisme et l'éclat de leur potentiel forment un terreau fertile pour le dépassement de soi.

Ce projet de recherche a pris naissance lors d'un stage que j'ai eu la chance d'accomplir au Palais de justice de Montréal. Pour le lancer, j'ai reçu le soutien de M^e Bourque, la collaboration de M^e Richard-Cossette et le privilège de la confiance de M^e Beaulieu. Je voudrais rappeler également combien Isabelle fut pour moi un soutien éprouvé et, je l'espère, une proche. Ces personnes et les autres du soutien à la justice, œuvrent de façon remarquable et font honneur, par leur ouverture et leur souci du service, à la vocation de fonctionnaire.

Même accompagné, ma capacité à mener ce projet fut tributaire de l'appui matériel, financier et technique consenti par le Conseil de recherche en sciences naturelles et génie, par le Groupe d'études et de recherche en analyse des décisions, par le palais de justice de Montréal et enfin par les talents de programmeur de François Lessard.

Le proverbe dit qu'on trouve de grandes femmes derrière tout grand homme. Sans aspirer moi-même à quelque grandeur, je ne déroge pas à la règle. C'est pourquoi, Anne Marie, Bianca, Anne, Jeannine et Diane, je vous dis merci. Enfin, comme pour tout ce que j'entreprends, j'ai su à mes côtés les deux femmes qui illuminent mon existence et à qui, plus que des remerciements, je témoigne ici de mon amour, Chantal et Madeleine.

Encore de l'espace? Je rendrai donc spécialement hommage au professeur Yves Nobert, qui m'a mené vers la recherche opérationnelle et à Jean-Philippe Meloche, qui m'a si souvent montré à quel point elle pouvait aiguïser l'esprit.

RÉSUMÉ

Ce mémoire s'intéresse à la problématique de l'élaboration des rôles dans un palais de justice de grande taille. Ce contexte est défini sous la forme du problème stochastique d'horaire d'audience judiciaires, qui se rattache à la famille des problèmes d'horaire. Cette formulation se distingue par la présence d'évènements (les procès) présentant une incertitude sur leur durée, ce qui permet des stratégies de surplanification pour favoriser la minimisation des délais juridiques et une bonne utilisation des capacités.

Les objectifs de la recherche rapportée sont de présenter un modèle mathématique en lien avec la littérature scientifique et les contraintes particulières du domaine judiciaire. La résolution proposée doit réussir à incorporer des éléments stochastiques pour rendre compte de l'incertitude sur la durée des procès. Cette dernière est expérimentée sur des jeux de données élaborés en collaboration avec le service des rôles du palais de justice de Montréal.

Inspirée des problèmes d'horaire, deux formulations sont proposées en vue de permettre la minimisation des délais juridiques tout en imposant une tolérance maximale sur la probabilité d'encombrement du rôle. Un seuil limite le nombre de causes pouvant avoir une probabilité non nulle de procéder dans une période et un type de salle donné. Les solutions entières obtenues par énumération implicite avec ce seuil subissent une évaluation stochastique pour déterminer la probabilité que le nombre de procès ayant lieu dépasse le nombre de salles disponibles. Des coupes sont ajoutées en cours de résolution lorsqu'une solution présente des périodes pour lesquelles la probabilité d'encombrement dépasse la tolérance de risque désirée.

Les résultats obtenus démontrent que l'utilisation de différentes méthodes d'évaluation des probabilités selon la taille des ensembles de procès permet de résoudre plus rapidement le problème d'optimisation posé que le recours à une seule méthode. Il est également établi qu'une résolution alliant un calcul exact des probabilités, une simulation de Monte-Carlo et une approximation par la loi normale parvient à produire des horaires de meilleure qualité et comportant moins de risque qu'une assignation séquentielle des causes proche de la pratique actuelle de faire dans le domaine judiciaire.

Le problème stochastique d'horaire d'audiences judiciaires qui est proposé est l'un des rares problèmes d'horaire considérant l'incertitude sur la durée des évènements. Sa résolution présente également l'intérêt de linéariser des contraintes aléatoires par l'intégration dynamique de plans coupants amenant l'obtention de solutions réalisables. Les résultats obtenus suggèrent que l'administration québécoise de la justice pourrait bénéficier de l'élaboration mathématique des rôles pour optimiser l'utilisation de ses ressources et réaliser des gains de temps pour les justiciables.

ABSTRACT

This master's thesis tackles the issue of roll making in a high-sized courthouse. This context is formalised as the judiciary hearing stochastic timetabling problem, subtype of the general timetabling problem class. This problem features events of uncertain duration, paving way to overbooking strategies seeking to minimise justice lead time and an appropriate use of ressources.

The main objective of the research shown is to present a mathematical model taking into account context-specific constraints while keeping in line with the state-of-the-art in timetabling. The solving proccess reported aims to use stochastic means to achieve a proper consideration of the uncertainty on trial length. Experimentation has been conducted using data obtained from the management of rolls of the Montréal courthouse.

Based on timetabling problems, two models are presented for optimising justice lead time while enforcing a maximum risk tolerance on the overcrowded roll probability. First, for any given period and courtroom type, the number of trial taking place with an actual probability is capped. Integer solutions founded by an branch-and-bound algorithm subject to this cap are tested by a stochastic method asserting the probability of having a greater number of trials than corresponding courtrooms for any given period. Cuts are added during the solving process when a integer solution exceeds the tolerated risk probability on one period (or more) for any courtroom type.

Testing results indicates that multiple probability evaluation methods in the same solving process can lead to lower computation times than a singe-method solving. Furthermore, while using exact probability computation, Monte-Carlo simulation and approximation by the normal law, solving the judiciary hearing stochastic timetabling problem produces better solutions implying lower risk and justice lead times than sequential assignment methods close to the actual roll management procedure.

The proposed judiciary hearing stochastic timetabling problem is one of the few timetabling problem dealing with events of uncertain duration. The solving process presented achieve linearisation of chance constraints by dynamically adding cutting planes leading to feasible solutions. Results also support the idea that Québec justice administration would benefit implementing a mathematical roll-making process by less lead times and a more consistent use of ressources.

TABLE DES MATIÈRES

DÉDICACE	iii
REMERCIEMENTS	iv
RÉSUMÉ	v
ABSTRACT	vi
TABLE DES MATIÈRES	vii
LISTE DES TABLEAUX	x
LISTE DES FIGURES	xi
LISTE DES SIGLES ET NOTATION	xii
CHAPITRE 1 INTRODUCTION	1
1.1 Problématique	1
1.1.1 Le soutien administratif dans un palais de justice	2
1.1.2 Le cheminement d’une cause	2
1.1.3 La mise au rôle	3
1.2 Pertinence académique	4
1.2.1 Cadre d’optimisation	4
1.2.2 Intérêt	5
1.3 Objectifs de recherche	6
1.4 Plan du mémoire	7
CHAPITRE 2 REVUE DE LITTÉRATURE	8
2.1 Typologie des problèmes d’horaire	8
2.1.1 Horaires dans l’éducation	9
2.1.2 Horaires dans le sport	13
2.1.3 Horaires dans le transport	14
2.1.4 Horaires de personnel	16
2.2 Composer avec l’incertitude	18
2.2.1 Programmation stochastique	19

2.2.2	Programmation par contraintes	19
2.2.3	Programmation probabiliste	20
2.2.4	Programmation floue	20
2.2.5	Robustesse et récupération	21
2.3	Résolution et proximité théorique	22
2.3.1	Problème d'horaire de classe	22
2.3.2	Les salles dans la résolution des problèmes d'éducation	23
2.3.3	Résolution en demande incertaine	24
2.4	Apport du mémoire	26
2.4.1	Durée incertaine des évènements	26
2.4.2	Énumération implicite et plans coupants	27
CHAPITRE 3	FORMULATIONS ET ALGORITHME DE RÉOLUTION	28
3.1	Formulation du problème stochastique d'horaire d'audiences judiciaires	28
3.1.1	Généralités	28
3.1.2	Modèle avec durée fixe et surcapacité	30
3.1.3	Modèle avec probabilités	31
3.2	Méthodes d'évaluation des probabilités d'encombrement du rôle	32
3.2.1	Méthode exacte	33
3.2.2	Méthode approchée	35
3.2.3	Méthode simulée	36
3.2.4	Synthèse	37
3.3	Méthode de résolution	38
3.3.1	Énumération implicite	38
3.3.2	Insertion des plans coupants	38
3.3.3	Algorithme	40
CHAPITRE 4	EXPÉRIMENTATION ET RÉSULTATS	42
4.1	Données utilisées	42
4.1.1	Portrait des données transmises par le palais de justice de Montréal	42
4.1.2	Construction des jeux de données pseudo-réalistes	46
4.2	Outils logiciels	49
4.2.1	Les programmes de soutien	51
4.2.2	<i>CourtCausesScheduler</i>	53
4.3	Résultats expérimentaux	53
4.3.1	Impact des méthodes d'évaluation des probabilités de débordement	53
4.3.2	Évaluation de la performance de <i>CourtCausesScheduler</i>	61

4.4	Discussion	68
CHAPITRE 5 CONCLUSION		71
5.1	Bilan de recherche	71
5.2	Retombées académiques	72
5.3	Retombées pratiques	72
5.4	Avenues de poursuites	73
RÉFÉRENCES		75

LISTE DES TABLEAUX

Tableau 2.1	Méthodes de résolution pour le problème d'horaire de classe	22
Tableau 2.2	Méthodes de résolution du problème d'horaire de cours d'université . .	23
Tableau 2.3	Méthodes de résolution du problème d'horaire d'examen	24
Tableau 2.4	Méthodes de résolution des problèmes de transport	26
Tableau 2.5	Méthodes de résolution des problèmes de personnel	26
Tableau 3.1	Comparatif des méthodes d'évaluation de la probabilité de débordement des salles d'audience	37
Tableau 4.1	Nombre de jours-salles utilisées par les causes selon la durée prévue au rôle	47
Tableau 4.2	Probabilités de génération des besoins de salles	51
Tableau 4.3	Catégorie d'exemplaires utilisés pour les tests d'évaluation de la probabilité d'encombrement du rôle	54
Tableau 4.4	Résultats détaillés de la résolution à une seule méthode	56
Tableau 4.5	Temps moyen de résolution (s) selon le choix d'une méthode d'évaluation	58
Tableau 4.6	Temps moyen d'évaluation des probabilités (ms) par solution entière . .	59
Tableau 4.7	Temps moyen de résolution selon la variation du paramètre σ_1	60
Tableau 4.8	Temps moyen de résolution selon la variation du paramètre σ_2	61
Tableau 4.9	Résultats détaillés de la résolution hybride	62
Tableau 4.10	Tableau comparatif de la méthode hybride	63
Tableau 4.11	Caractéristiques des exemplaires pseudo-réalistes	64
Tableau 4.12	Résultats détaillés en utilisant <i>CourtCausesScheduler</i>	65
Tableau 4.13	Valeur de la solution du constructeur séquentiel selon le taux de surcharge	66
Tableau 4.14	Nombre de périodes-types de salle dont le risque d'encombrement excède 5 %	68

LISTE DES FIGURES

Figure 2.1	Classification des problèmes d’horaire	9
Figure 2.2	Relations entre les problèmes d’horaires du domaine ferroviaire	15
Figure 4.1	Comparaison du volume judiciaire	43
Figure 4.2	Origine des salles d’audience utilisées	44
Figure 4.3	Typologie des causes	44
Figure 4.4	Exactitude de la prévision de la durée	45
Figure 4.5	Proportion comparée des issues des causes	46
Figure 4.6	Proportion comparée des causes d’une durée réelle nulle	48
Figure 4.7	Proportion et taux proposé des causes ayant dépassé leur durée prévue	49
Figure 4.8	Comparatif des sommes des causes de l’échantillon de durée non-nulle sans excès de la durée prévue et de la distribution proposée	50
Figure 4.9	Comparaison entre la répartition des causes selon leur durée prévue et une distribution Gamma	50
Figure 4.10	Risque d’encombrement pour solution de l’approche séquentielle pour $r-350$	66
Figure 4.11	Risque d’encombrement pour solution de l’approche séquentielle pour $r-400$	67
Figure 4.12	Risque d’encombrement pour solution de l’approche séquentielle pour $r-450$	68

LISTE DES SIGLES ET NOTATION

- C : Ensemble des causes, préassignées ou non.
- C^A : Sous-ensemble des causes préassignées.
- C^R : Sous-ensemble des causes régulières, soit celles devant être inscrites au rôle.
- P : Ensemble des périodes appartenant à l'horizon de planification du rôle.
- S : Ensemble des types de salles d'audience disponibles pour l'élaboration du rôle.
- C_{ps}^A : Ensemble des causes préassignées se déroulant dans une salle de type s telles que celles-ci risquent de couvrir la période p avec une probabilité non-nulle.
- K : Ensemble des paires de causes en conflit.
- P_c : Ensemble des périodes de début valides pour la cause c .
- P_{cp} : Ensemble des périodes de début de la cause c telles que celles-ci risquent de couvrir la période p avec une probabilité non nulle.
- S_c : Ensemble des types de salles valides pour la cause c .
- $1 - \alpha$: Seuil de probabilité d'encombrement du rôle.
- β : Taux de capacité permis par la surplanification.
- γ : Nombre de tirages effectués dans la simulation de Monte-Carlo lors de l'évaluation d'une probabilité d'encombrement du rôle.
- σ_1 : Nombre de causes impliquées dans une évaluation de la probabilité d'encombrement du rôle à partir duquel la méthode simulée sera préférée à la méthode exacte.
- σ_2 : Nombre de causes impliquées dans une évaluation de la probabilité d'encombrement du rôle à partir duquel la méthode approchée sera préférée à la méthode simulée.
- M_c : Pénalité associée à l'incapacité de mettre au rôle la cause c .
- x_{cps} : Variable de décision prenant la valeur 1 si la cause c débute à la période p dans une salle de type s et la valeur 0 sinon.
- e_c : Variable d'écart prenant la valeur 1 si la cause c n'a pas pu être mise au rôle et la valeur 0 sinon.
- Y_{cps} : Variable aléatoire prenant la valeur 1 si la cause c procède à la période p et la valeur 0 sinon.
- π_{cp} : Probabilité que la cause c soit en cours à la période p .
- π_{cd}^R : Probabilité que la cause régulière c procède sur d périodes ou plus.
- N_s : Nombre de salles de type s disponibles pour une période durant l'horizon.
- $\hat{x}$: Solution entière au problème d'horaire de salles d'audience.
- G_{ps} : Ensemble de couples (c, q) tel que le procès de la cause c , s'il commence à la période q dans une salle de type s , aura une probabilité non nulle d'être en cours à la période p .

$\overline{G}_{ps}$: Ensemble de couples (c, q) tel que la probabilité cumulée que les procès des causes c , s'ils débutent à leur période q respective, occasionne un encombrement du rôle à la période p pour la salle de type s .

Γ_{ps} : Ensemble des ensembles $\overline{G}_{ps}$ induisant une violation des contraintes sur l'encombrement du rôle.

$\hat{G}_{ps}$: Sous-ensemble de couples (c, q) de G_{ps} reprenant tous les couples comprenant des c appartenant à C^A ainsi que les c appartenant à C^R tel que, dans une solution entière $\hat{x}$ du problème d'horaire de salles d'audience, la variable x_{cqs} associée prend la valeur 1.

$\hat{G}_{ps}^R$: Sous-ensemble de couples (c, q) de $\hat{G}_{ps}$ reprenant les couples pour lesquels c appartient à C^R .

$\hat{G}'_{ps}$: Sous-ensemble de couples (c, q) de $\hat{G}_{ps}^R$ de cardinalité minimale appartenant à Γ_{ps} .

CHAPITRE 1

INTRODUCTION

Affirmer que la recherche opérationnelle est un domaine méconnu du grand public relève presque de l’euphémisme. Au gré des vulgarisations, elle devient une branche des mathématiques dédiées à l’optimisation des manières de faire, le pendant stratégique de la gestion des opérations, ou encore la jonction entre les mathématiques et la gestion. Une métaphore plus cryptique consiste à déclarer qu’elle est en quelque sorte l’antithèse de la macroéconomie néoclassique. Non sans humour, on peut en effet considérer qu’en recherche opérationnelle, des modèles mathématiques sont élaborés et raffinés de façon à correspondre le mieux possible à la réalité, alors que les pontes de la théorie walrassienne¹ semblent plutôt tentés de forcer la réalité à correspondre au modèle mathématique de l’économétrie !

En alliant la puissance des mathématiques aux capacités de l’informatique contemporaine, la recherche opérationnelle décuple nos capacités à gérer de façon optimale des situations logistiques complexes. Avec elle, le génie humain fracasse les frontières de ce qui est familièrement connu comme le « gros bon sens » pour non seulement ordonner ce qui serait resté chaotique, mais en orienter l’organisation dans la direction souhaitée. Les développements en matière d’optimisation permettent de réaliser des économies humaines, matérielles et financières qui auraient tenu du miracle il y a à peine un demi-siècle. Si ses bénéfices se font sentir dans une variété de domaines d’application, de larges pans de l’activité sont encore à conquérir. Avec humilité, nous espérons pouvoir agir comme tête de pont dans l’un de ces futurs champs de déploiement de la recherche opérationnelle.

1.1 Problématique

Le présent mémoire se propose d’offrir une compréhension mathématique d’une problématique du domaine de la gestion publique, plus précisément du domaine de l’administration de la justice. Les dernières décennies ont vu se diffuser au sein des fonctions publiques canadiennes les préceptes de la nouvelle gestion publique (Pollitt et Bouckaert, 2011, p. 247-255). Ce courant de pensée fait la part belle aux notions de performance, entre autres la gestion par résultats (Emery et Giaque, 2005, p. 88). Néanmoins, il subsiste des cas de figure où l’aide à la décision pourrait se faire à grand renfort d’outils issus du savoir-faire en recherche

1. de Léon Walras (1834-1910). Fondateur de l’école de pensée dite de Lausanne, cet ingénieur de formation apporta des contributions déterminantes à la pensée économique, notamment la théorie de l’équilibre général et le concept d’utilité marginale. Son héritage est encore aujourd’hui très présent dans le discours économique.

opérationnelle.

1.1.1 Le soutien administratif dans un palais de justice

En vertu de l'organisation des pouvoirs hérités de la coutume britannique et décrite dans les documents constitutifs du Canada et du Québec, la séparation des pouvoirs prévaut entre l'exécutif et le judiciaire (1982 c. 11 (U.K.)). Le premier est chargé de la gouverne et de l'administration alors que le second doit veiller à ce que la justice soit prononcée conformément à l'État de droit. L'administration de la justice est l'un des rares domaines d'interaction entre ces deux pouvoirs fondamentaux. Au Québec, la gestion des différents palais de justice de la province et la gamme de services qui agissent en soutien de la bonne marche du système judiciaire sont autant de tâches qui reviennent au ministère québécois de la Justice (Québec (Province). Ministère de la justice (2011)), chapeauté par un ministre qui agit également en tant que procureur général (L.R.Q., c. M-19).

Le territoire québécois était divisé en 2012 en trente-six districts judiciaires différents (L.R.Q., c. D-11), chacun desservi par au moins un palais de justice. La charge administrative propre à un palais varie énormément d'un endroit à l'autre, de Campbell's Bay à Montréal en passant par Saguenay. Cependant, elle tourne essentiellement autour des mêmes sphères générales : le soutien à la magistrature locale, les services judiciaires civils et criminels, la tenue du greffe et des registres (archives et dossiers judiciaires). L'accès à la justice dans des délais raisonnables dépend de la célérité avec laquelle le ministère de la Justice du Québec s'acquitte de sa responsabilité à l'égard du judiciaire, un état de fait notamment souligné par la jurisprudence (Villaggi, 2005, p. 137).

1.1.2 Le cheminement d'une cause

Considérons l'exemple d'une poursuite civile entre deux ou plusieurs individus ou organisations. L'appellation « civile » qualifiant l'affaire devant être jugée comme ne relevant pas du droit criminel, dans lequel il incombe à l'État de traduire en justice les accusés. Du point de vue du système judiciaire, la poursuite est intentée dès lors qu'une partie, qualifiée de demanderesse, procède à l'ouverture d'un dossier au greffe civil du district judiciaire dans lequel elle souhaite instruire l'affaire. Dans un esprit d'équité, il est tenu qu'un huissier de justice signifie à ce moment aux autres parties en cause, qualifiées de défenderesses, qu'elle font l'objet d'une poursuite. Par la suite, le dossier passera au plumeitif, où les différentes parties y verseront les éléments sur lesquels elles entendent baser leurs plaidoiries respectives.

En parallèle des dépôts au dossier, les parties peuvent déposer des procédures diverses. Ces dernières peuvent prendre la forme d'une injonction, de demande face au processus judiciaire

ou encore de contestation de la recevabilité de pièces portées au dossier. Les procédures doivent obligatoirement concerner la pratique, ou la « forme », et non le litige faisant l'objet de la poursuite en lui-même. Par exemple, considérons un divorce dans lequel les parents se disputent le droit de garde des enfants. Une procédure recevable serait une demande d'injonction pour obtenir du juge une décision provisoire sur le droit de garde pour le temps des Fêtes, en attendant que la procédure statue sur le régime de droit de garde définitif. Ces procédures sont mises au rôle pratique, qui constitue l'ordre du jour des audiences sur la pratique (par opposition au mérite, qui constitue le « fond » d'une cause).

Après dépôt au dossier d'éléments par les différentes parties, un greffier attestera du statut complet du dossier. Si cette attestation est délivrée, les parties vont statuer sur les temps de plaidoirie qu'elles estiment nécessaires. Selon la durée totale prévue pour le procès, le dossier sera révisé par un greffier spécial qui ajustera, au besoin, la durée prévue selon l'expérience pratique. Une fois qu'une durée est déterminée, le maître des rôles proposera des dates d'audiences aux parties qui effectueront une conciliation de leurs disponibilités respectives. Il découlera de cette étape le choix d'une date pour l'ouverture du procès et la mise au rôle de mérite de la cause. Ce rôle servira de base à l'horaire des salles d'audience.

Entre la détermination de la date et l'ouverture du procès, des conférences de règlements à l'amiable ont lieu, sous la supervision d'un juge. Ces séances visent à permettre d'obtenir un accord entre les parties sans devoir se soumettre à l'arbitrage d'un magistrat. Véronique Hivon, ancienne ministre et critique en matière de justice, considère par ailleurs comme un « principe fondament[al ...] le règlement des différends par des modes alternatifs. » (Éditeur officiel du Québec, 2012, p. 6357). Les règlements à l'amiable font partie de ces modes alternatifs permettant de limiter l'engorgement du système judiciaire en aval.

À l'approche de la date du procès et en l'absence de la notification d'un règlement à l'amiable ou d'un abandon des procédures, les services judiciaires du palais de justice vont, conformément au rôle, préparer l'une des salles d'audience. Durant toute la durée du procès, il est possible que celui-ci soit écourté par un désistement ou encore un règlement survenu entre les deux parties. Si tel n'est pas le cas, la ou le juge prendra la cause en délibéré avant de prononcer un jugement qui sera enregistré au greffe et porté au dossier. Selon le cas, ce dernier peut être contesté par un appel de la décision. Dans ce cas, le dossier sera transféré à une cours de juridiction supérieure pour révision, ce qui relancera le processus.

1.1.3 La mise au rôle

Du processus décrit à la section 1.1.2, il est possible d'isoler l'élaboration du rôle pratique, qu'on peut qualifier de problème stochastique d'horaire d'audiences judiciaires. Ce problème consiste ainsi à planifier le moment de début d'audience pour des procès dans un palais

de justice. Un palais de justice contient un certain nombre de salles comportant un certain nombre d'équipements particuliers. Les procès sont mis au rôle et doivent se dérouler dans une seule salle. Comme cela est exposé dans la section 1.1.2, les procès ont une durée prévue donnée mais une durée effective variable, voire nulle. L'enjeu est donc, pour chacune des causes devant passer en procès, de limiter au maximum les délais dûs à la mise au rôle, tout en minimisant le recours à des salles supplémentaires en raison de l'encombrement de l'horaire.

1.2 Pertinence académique

Maintenant que le problème stochastique d'horaire d'audiences judiciaires a été délimité au niveau de la problématique de gestion qu'il suscite, il convient désormais de démontrer l'intérêt mathématique d'une résolution de ce problème. En étayant la présence d'un cadre d'optimisation, il sera possible d'évoquer les apports qu'une telle recherche apporte à la fois au domaine de la recherche opérationnelle ainsi qu'à celui de la gestion des services judiciaires.

1.2.1 Cadre d'optimisation

Un palais de justice peut contenir un nombre variable de salles. Ces salles disposent de différents équipements pour répondre aux besoins spécifiques de certaines causes. Par exemple, un banc des accusés pour les causes criminelles, de l'espace de travail pour un grand nombre d'avocats ou pour accueillir davantage de public, etc. Généralement, la présence d'un équipement n'entraînera pas nécessairement la présence ou l'absence d'autres équipements. Les équipements dont disposent les salles sont considérés comme fixes et comme des paramètres du problème.

Un procès consiste en une ou plusieurs audiences devant un juge pour arbitrer un litige entre des parties (l'État peut être partie dans une cause criminelle). La cause entendue lors de ce procès peut être caractérisée par le besoin d'un ou plusieurs équipements particuliers, qui devront se retrouver dans la salle retenue. Ces caractéristiques seront considérées comme fixes et comme des paramètres du problème. La durée totale des audiences requises pour un procès est estimée au moment de la planification. Cependant, la durée véritable n'est pas connue à ce moment et représente un élément stochastique. Il est possible qu'un règlement hors-cours avant l'ouverture du procès fasse en sorte que cette durée soit nulle dans les faits. De même, différents facteurs peuvent amener la durée à être inférieure ou supérieure au temps prévu à l'origine. Les probabilités pour chacune des durées possibles de la cause seront considérées comme connues et comme des paramètres du problème. Les procès sont considérés comme se déroulant sur des périodes consécutives, peu importe la durée.

Les parties représentées dans un procès ont un certain nombre de disponibilités. Que ce soit en raison de l'emploi du temps d'un témoin expert, de celui des avocats ou pour un autre motif, il existe des périodes pour lesquelles l'une des parties ne peut plaider ou être présente lors de la plaidoirie de l'autre partie. Pour qu'une audience puisse avoir lieu durant une ou des périodes consécutives données, il faut que l'ensemble des parties soient disponibles. L'intersection de ces disponibilités représente les périodes possibles pour le procès. Pour chacun des procès à planifier, l'ensemble des périodes où le procès peut être planifié est considéré comme fixe et comme un paramètre du problème.

L'ensemble de ces paramètres forment le cadre à l'intérieur duquel doit être formé un rôle, soit un horaire dans lequel figurera une partie ou la totalité des causes pouvant y être assignées dans le respect des aspects déterministes et stochastiques du problème. Les causes laissées de côté dans ce rôle sont reportées dans le prochain exercice de planification. Ainsi, il est aisé de remarquer que pour être valide, cet horaire devra respecter les disponibilités des différentes causes et la capacité des salles du palais de justice. Ces éléments, tout comme la présence de conflits entre les différentes causes, constituent autant de contraintes s'ajoutant au cadre issu des paramètres et permettant de déterminer si un rôle correspond ou non à une solution réalisable. De là, il est possible de formuler comme un objectif mathématique la réduction des délais induits par la forme de l'horaire comme le moyen d'évaluer la qualité des rôles assemblés et, par conséquent, de procéder à l'optimisation de ces derniers.

1.2.2 Intérêt

Sur la base de témoignages recueillis au palais de justice de Montréal, l'un des plus grands en Amérique du Nord en termes de volume annuel de causes, il est établi qu'une approche mathématique comme celle qui découlera d'une optimisation faisant appel à la recherche opérationnelle n'est pas présentement utilisée pour l'élaboration des rôles de mérite. Une telle approche ne semble pas non plus avoir eu cours par le passé. Si tel a déjà été le cas, l'usage s'en est perdu dans la pratique régulière, ce qui maintient l'intérêt de la présente démarche. De plus, ces travaux sont en mesure de déterminer si une diminution substantielle des délais judiciaires peut être réalisée sans augmenter de façon appréciable les risques d'encombrement du rôle. Advenant cet état de fait, les administrateurs publics pourront déterminer, sur la base de cette démonstration, la pertinence de réformer concrètement la formation du rôle dans le réseau québécois de la justice.

Bien entendu, l'originalité du champ d'application n'a qu'un effet marginal sur la pertinence du modèle mathématique par rapport à la communauté académique. Mais la formulation retenue comporte un ajout dynamique de contraintes probabilistes qui ne sont pas strictement linéaires dans leur formulation large. Cet aspect, plus longuement décrit aux sec-

tions 3.1.3 et 3.2, n'est pas inédit, mais constitue une manière intéressante de s'attaquer à l'incertitude de la durée réelle des causes.

1.3 Objectifs de recherche

Les sections 1.1 et 1.2 ont permis de mettre en lumière l'existence et les contours du problème stochastique d'horaire d'audiences judiciaires, ainsi que d'en déterminer qu'il peut constituer un problème mathématique d'intérêt au point de vue académique. Les travaux de recherche menés par la suite ont eu comme objectif de développer un modèle mathématique décrivant le problème stochastique d'horaire d'audiences judiciaires tel qu'il se présente dans un palais de justice de grande taille. Il est important pour garantir un potentiel de retombées concrètes que ce modèle soit conçu de telle sorte qu'il puisse servir de base à un outil de gestion qui permette d'améliorer la qualité des horaires de tribunaux. Cette amélioration est déterminée par des délais moindres pour les parties en cause et par une utilisation plus efficiente des salles d'audiences disponibles. Pour satisfaire cette orientation générale, il importe de s'attarder sur ces différentes facettes une par une.

- **Caractériser les contraintes particulières d'un palais de justice de grande taille :** Si on aspire à régler une problématique concrète en élaborant un modèle mathématique, ce dernier doit être à même de représenter la réalité avec assez de justesse pour permettre qu'il soit à la base d'une implantation éventuelle ;
- **Développer un modèle mathématique basé sur l'état de la littérature en matière de problèmes d'horaire et sur les contraintes identifiées ci-haut :** Autant l'aspect particulier de la problématique doit être considéré, autant il ne serait pas proprement scientifique de ne pas s'appuyer sur les réalisations passées de la recherche opérationnelle et de s'inscrire dans la littérature spécialisée. Le domaine des problèmes d'horaires, qui constitue la grande famille du problème stochastique d'horaire d'audiences judiciaires, représente une source indéniable d'informations pertinentes quant à l'élaboration d'un modèle mathématique apte ;
- **Déterminer les éléments d'optimisation stochastique pouvant servir à rendre la réalité fluctuante du processus judiciaire :** La durée des différentes causes n'est pas connue avec précision, ce qui induit le risque d'en arriver à un encombrement du rôle dû à un nombre de causes devant procéder supérieur aux capacités en termes de salles ;
- **Valider et améliorer le modèle en l'utilisant sur des données réalistes et en le comparant avec les approches empiriques utilisées dans le milieu :** Plutôt que de produire un échafaudage théorique d'un niveau de raffinement indéniable, l'intérêt

de la démarche effectuée est de vérifier si des gains peuvent être faits sur la pratique actuelle.

1.4 Plan du mémoire

Ce mémoire présentera au sein des prochains chapitres le fruit des travaux et des efforts qui ont permis de mener à bien les objectifs décrits à la section 1.3. Le second chapitre porte sur la revue de la littérature scientifique associée aux problèmes d'optimisation d'horaires. Les typologies de ces problèmes seront présentées dans un premier temps. Puis, les axes par lequel l'incertitude est prise en compte seront couverts, ainsi que les méthodes de résolution des problèmes similaires au problème stochastique d'horaire d'audiences judiciaires. Enfin, ce chapitre se conclura sur une description de l'apport théorique de ce mémoire sur la littérature actuelle.

Le troisième chapitre est consacré au modèle mathématique formulé pour résoudre le problème stochastique d'horaire d'audiences judiciaires. La formulation en tant que telle est abordée en premier lieu, suivi des variantes présentant un intérêt pour la résolution. Le reste du chapitre expose les méthodes élaborées pour évaluer la probabilité d'un débordement des salles disponibles. Après qu'elles aient été tour à tour exposées, la dernière section de ce chapitre permettra de voir comment ces méthodes viennent bonifier le processus de résolution.

Le quatrième chapitre présente le volet expérimental des travaux de maîtrise, ainsi que les résultats qui y ont été obtenus. Des exemplaires du problème stochastique d'horaire d'audiences judiciaires sont présentés en ouverture du chapitre, de même que les moyens techniques à disposition. Ces exemplaires sont par la suite utilisés dans deux séries de tests expérimentaux. La première vise à valider l'impact des méthodes d'évaluation des contraintes probabilistes sur le comportement de la résolution. La seconde, quant à elle, a pour objet de démontrer que la résolution mathématique du problème stochastique d'horaire d'audiences judiciaires permet d'obtenir des solutions de meilleure qualité que la pratique actuelle.

Enfin, un dernier chapitre conclura cet ouvrage en effectuant un retour sur les objectifs de recherche définis en introduction et en soulignant les contributions autant académiques que pratiques des travaux réalisés au cours de cette maîtrise.

CHAPITRE 2

REVUE DE LITTÉRATURE

L'étude d'un problème d'optimisation ne pourrait être une démarche scientifique probante si elle se déroulait en vase clos. Un vecteur crucial de l'avancement du savoir humain est la diffusion des résultats antérieurs et leur prise en compte dans le cadre des travaux d'aujourd'hui et de demain. L'étude du problème stochastique d'horaire d'audiences judiciaires s'inscrit donc dans un courant plus grand, soit l'optimisation des problèmes d'horaire. Une typologie de cette famille sera présentée. Le même procédé permettra de revenir sur les grands courants qui caractérisent la gestion de l'incertitude dans l'optimisation. Une fois le contexte général dépeint, il est possible de se concentrer sur les formulations et les méthodes de résolution qui se rapprochent du problème stochastique d'horaire d'audiences judiciaires. Le chapitre se terminera par un retour sur les éléments originaux de l'étude en cours.

2.1 Typologie des problèmes d'horaire

La présente section reprend largement la typologie générale qui fut définie par Burke *et al.* (2004). Cette classification est basée d'abord sur le champ d'application auquel le problème d'horaire se rattache. Les grandes sphères identifiées sont ainsi l'éducation, le sport, le transport et les horaires de personnel. Ce découpage (reproduit à la figure 2.1) ne peut prétendre être exhaustif, mais il demeure largement valable au vu de la littérature scientifique contemporaine. Les contextes propres à ces différents domaines sont à la naissance de formulations variables d'un champ à l'autre. À l'intérieur d'un champ donné, une série de contributions peuvent faire émerger un sous-type défini, avec des contraintes reconnues. Par exemple, la multiplicité des problèmes liées à l'optimisation du transport ferroviaire et aux problèmes de tournées de véhicules font en sorte qu'ils y sont représentés comme un sous-type marqué.

Toujours dans (Burke *et al.* (2004)), il est établi également les caractéristiques de ce qui peut être considéré comme un problème d'horaire. Tous les problèmes d'horaire comportent ainsi sous une forme ou une autre trois grands ensembles, soit des événements, des ressources et un horizon temporel. Ces trois ensembles sont finis et sont unis par la contrainte fondamentale des problème d'horaire : deux événements ne peuvent utiliser les mêmes ressources au même moment dans le temps. Ainsi, une même ressource peut être affectée à de multiples événements pourvu que l'horizon temporel de ces derniers ne se recoupe pas. Autrement dit, il est possible d'inscrire au même moment plusieurs événements si et seulement si chacun

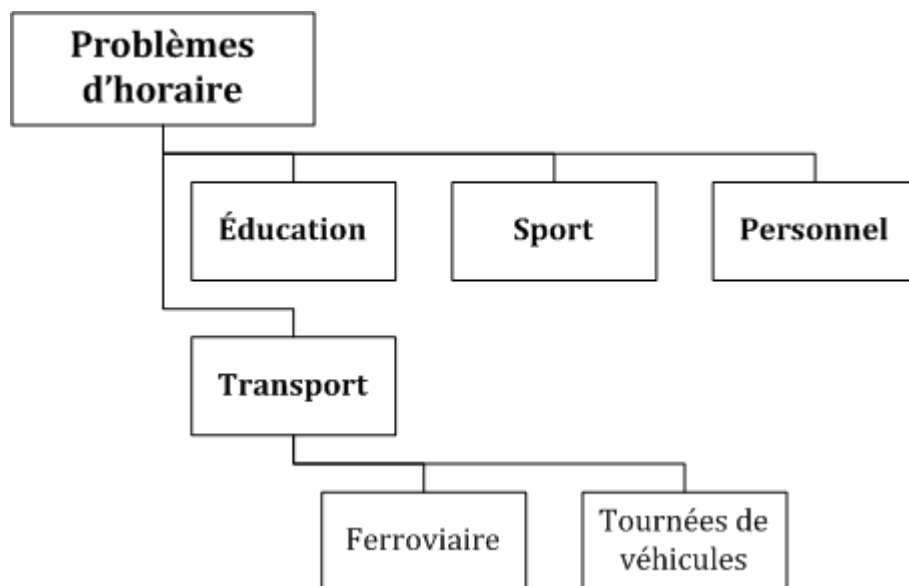


Figure 2.1 Classification des problèmes d'horaire selon Burke *et al.* (2004)

d'entre eux font appel à des sous-ensembles de ressources mutuellement exclusifs. Les autres contraintes, en quantités finies, vont baliser la faisabilité et les préférences en matière d'allocation de périodes et de ressources aux événements, ainsi que les interactions entre ces trois ensembles.

Le plus souvent, la formulation mathématique est spécifique à un champ donné et peut prendre des formes très diverses autrement. Même s'il forme un espace distinct, le temps peut être modélisé sous forme de variables ou de contraintes d'espace-temps, ce qui l'associe à une partie des ressources (voir Cordeau *et al.* (1998), Kliewer *et al.* (2006), Ranson et Ahmadi (2006), Cicerone *et al.* (2009a), Cicerone *et al.* (2009b) et Nguyen *et al.* (2010)). Dans le même ordre d'idée, parfois, la résolution s'attaque au problème d'optimisation de façon globale, mais, selon De Causmaecker et Vanden Berghe (2010), les problèmes d'horaire de personnel vont fréquemment dissocier l'assemblage des horaires de l'affectation de ces derniers au personnel.

2.1.1 Horaires dans l'éducation

Sans surprises, les problèmes d'horaire concernant l'éducation ont su susciter l'intérêt des chercheurs universitaires. En effet, depuis Gotlieb (1963), l'optimisation des horaires de cours, dans les universités et ailleurs, a joui d'une attention constante du milieu de la recherche opérationnelle. Il faut souligner que ces problèmes présentent un certain niveau de difficulté ; Socha *et al.* (2002) comparent avantageusement le problème d'horaire de cours d'université à

celui de la coloration de graphe. En plus de ce niveau de difficulté, Pearl (1984) expose que la taille des exemplaires ayant un intérêt pratique permet d’atteindre un niveau de complexité suffisant pour justifier le recours à des métaheuristiques.

Les sections suivantes présentent trois grandes formulations de problèmes d’horaires touchant l’enseignement. Le problème d’horaire de classe est la forme la plus épurée, alors que le problème d’horaire d’universités est souvent le problème de référence pour toute la famille. Le problème d’horaire d’examens, quant à lui, s’intéresse à un contexte légèrement différent, ce qui induit une variation suffisante de sa formulation pour qu’il soit pertinent de le traiter comme un sous-type propre.

Le problème d’horaire de classe

Le problème d’horaire de classe est dépeint par Schaerf (1999) comme se concentrant sur l’assignation des charges de cours ou d’enseignement à l’intérieur d’un horaire. De façon plus large, autant Schaerf (1999) que Burke *et al.* (2004) conçoivent ce problème comme visant à construire un horaire dans lequel il n’y a aucun conflit au niveau des cours, du personnel enseignant ou des salles, les solutions présentant des conflits de ce type n’étant pas considérées réalisables. Schaerf (1999) repertorie par ailleurs d’autres contraintes, comme certains cours requérant la présence de plus d’un professeur, nécessitant qu’un cours puisse être donné à plusieurs groupes au même endroit ou encore l’assignation au préalable d’une certaine quantité de cours. Il est à noter que dans ce dernier cas, le problème d’horaire de classe est démontré $\mathcal{NP}$ -Complet. Par ailleurs, cette $\mathcal{NP}$ -Complétude s’étend également aux cas de disponibilités imparfaites des professeurs ou des cours, ainsi qu’aux situations où un cours doit se tenir dans des salles particulières, car Burke *et al.* (2004) ont démontré par le biais d’activités fictives que ces situations étaient équivalentes à une préassignation.

Schaerf (1999) propose deux transformations en temps polynomial pour passer du problème d’horaire de classe à, respectivement, un problème de multigraphe biparti et un problème de coloration d’arêtes, cette dernière transformation est également repertoriée par Csimá (1965). Dans le cadre de cette dernière transformation, Burke *et al.* (2004) utilisent une couleur pour chacune des périodes possibles dans l’horaire. Si le même cours contient plusieurs séances avec le même groupe, des arêtes parallèles sont utilisées. Dans le cas restreint où ces séances doivent être consécutives dans l’horaire, il est possible toujours pour Burke *et al.* (2004) de passer de cette forme à un problème de remplissage de sacs (*bin packing*).

Le problème d'horaire d'université

Alors que le problème d'horaire de classe utilise le groupe d'étudiant comme base pour construire les horaires de ces derniers, le problème d'horaire d'université considère les étudiants sur une base individuelle dans la composition des horaires. Ces derniers sont néanmoins tenus de s'inscrire à un sous-ensemble de cours prescrits par leur cursus. Plusieurs groupes, ou sections, peuvent être constitués pour le même cours si le nombre d'étudiants susceptibles de s'y inscrire est suffisant. Dans ce cadre, repris entre autre par Schaerf (1999), Burke *et al.* (2004), Najdpour et Feizi-Derakhshi (2010), Nguyen *et al.* (2010), Burke *et al.* (2010) et Shiau (2011), un professeur doit être prévu pour chacune des sections de chacun des cours, ainsi que des locaux et une période à l'horaire. Pour Socha *et al.* (2002), ce problème est essentiellement une reformulation du problème d'horaire de classe où les étudiants sont considérés un à un. Enfin, Nguyen *et al.* (2010) et Abdullah *et al.* (2010) précisent que le problème peut être abordé soit d'une perspective étudiante, soit du point de vue du corps professoral, selon que les contraintes retenues et la formulation visent davantage à produire les horaires de cours des différents étudiants ou plutôt l'horaire de travail des professeurs.

Parmi les autres observations générales, Burke *et al.* (2004) soulignent la proximité théorique entre la notion de temps et celle de l'espace, qui prennent respectivement la forme de période et de salle de classe. Burke *et al.* (2010) amènent également le concept de multi-salle. En présence d'un certain nombre de salles présentant les mêmes caractéristiques, une réduction de la symétrie lors de la résolution peut survenir en considérant ces salles identiques non pas comme plusieurs salles distinctes, mais comme une grande salle théorique. Celle-ci peut alors accommoder simultanément un nombre donné de cours correspondant au nombre de salles distinctes qui sont remplacées. Pour respecter la capacité des salles, le nombre d'inscrits pour chacun des cours assignés à une multi-salle ne peut dépasser la taille réelle des salles qui sont ainsi abstraites. Werra (1985) propose pour ce problème une transformation en coloration de graphe pour laquelle Burke *et al.* (2004) utilisent comme objectif de réduire le poids des arêtes reliant deux sommets de la même couleur.

Les contraintes liées à la disponibilité imparfaite des cours, salles et professeurs qui ont été mentionnées au problème précédent sont également valables pour le problème d'horaire de cours d'université. Il en va de même pour les interactions qui les caractérisent. D'autres contraintes de faisabilité peuvent être ajoutées au problème, comme l'interdiction rapportée par Shiau (2011) que certaines périodes servent de pause-repas aux étudiants. Ahmed et Zhoujun (2010) présentent une formulation dans laquelle les cours peuvent avoir des durées excédant une période, ce qui est pris en charge selon Hertz (1991), Nguyen *et al.* (2010) et Shiau (2011) par la construction de patrons pour la tenue des sections de ce cours.

La littérature recèle également nombre de contraintes souples, dont le respect ou la vio-

lation minimale ajoute du relief à l'espace des solutions et permet de dénoter la qualité des horaires réalisables produits. De façon non exhaustive, on retrouve :

Contraintes basées sur les étudiants

- limiter le nombre de cours par jour (Ahmed et Zhoujun (2010)) ;
- éviter les journées ne comportant qu'un seul cours pour un étudiant (Socha *et al.* (2002)) ;
- favoriser les pauses entre des cours de la même journée (Abdullah *et al.* (2010)) mais sans qu'elles soient trop longues (Najdpour et Feizi-Derakhshi (2010)) ;
- regrouper les cours d'un même cursus dans la semaine de classe (Burke *et al.* (2010)) ;
- regrouper géographiquement les cours d'un même étudiant (Yang *et al.* (2010)).

Contraintes basées sur les professeurs

- limiter le nombre de cours par jour pour un professeur donné (Ahmed et Zhoujun (2010) et Najdpour et Feizi-Derakhshi (2010)) ;
- favoriser la dispersion dans la semaine de classe des charges de cours (Burke *et al.* (2010) et Yang *et al.* (2010)) ;
- valoriser les horaires comportant des périodes de repos entre deux charges de cours d'un même professeur (Najdpour et Feizi-Derakhshi (2010)).

Autres contraintes

- pénaliser un trop grand nombre de cours par semaine, autant pour les étudiants que les professeurs (Nguyen *et al.* (2010)) ;
- éviter d'assigner des cours après certains autres, comme les séances de sport (Yang *et al.* (2010)) ;
- organiser la journée de façon à ce que certains cours cruciaux soient prioritairement donnés dans des périodes de haute fréquentation (Yang *et al.* (2010)) ;
- limiter le recours à la dernière période de la journée de classe (Socha *et al.* (2002)).

Une multitude de ces contraintes pourraient encore être citées, puisqu'autant Shiau (2011) que Nguyen *et al.* (2010) ou Yang *et al.* (2010) relèvent l'intérêt pour une formulation de prendre en compte les différentes préférences formulées autant par les étudiants que par les professeurs.

Le problème d'horaire d'examen

Le problème d'horaire d'examen vise à organiser la période d'examen complétant une session universitaire. Pour Schaerf (1999), il s'agit dès lors de créer un horaire dans lequel les examens de cours appartenant au même cursus ne doivent pas se chevaucher. Pour Burke *et al.* (2004), la formulation se rapproche du problème d'horaire de classe en cela qu'il s'agit d'apparier locaux et périodes aux différents examens. Elle diffère cependant du fait qu'il est possible de scinder un groupe en plusieurs salles ou encore d'affecter plusieurs examens à la même salle, si sa capacité le permet. Pour peu que les durées d'examens sont semblables, Ranson et Ahmadi (2006) suggèrent de formuler les salles et les périodes dans un même jeu de ressources.

Dans ce contexte, l'assignation préalable et la disponibilité imparfaite peuvent être considérées comme des contraintes de faisabilité, au même titre que d'autres mesures mentionnées par Schaerf (1999), telles un nombre maximal d'examen pour chaque étudiant ou un intervalle au sein duquel l'ensemble des examens doivent se situer. Ranson et Ahmadi (2006) évoquent d'éviter qu'un étudiant ait plus d'un examen par jour alors que McCollum (2006) va plus loin en amenant l'idée de favoriser une dispersion maximale des examens d'un même étudiant à l'intérieur des temps prévus. Les mêmes auteurs insistent cependant sur la difficulté d'établir des contraintes souples qui peuvent être généralisées à l'extérieur du contexte d'une application pratique en particulier. Ces dernières sont également, selon McCollum (2006) d'une complexité généralement supérieure aux problèmes typiques de la diffusion contemporaine.

2.1.2 Horaires dans le sport

Le sport professionnel offre probablement l'un des cadres de déploiement de l'optimisation d'horaire les plus connus du grand public, à l'insu de ce dernier. Pour Burke *et al.* (2004) et Arjen van Weert et Jan A.M. Schreuder (1997), le problème d'horaire sportif consiste à déterminer le calendrier des rencontres entre les différentes équipes d'une ligue ou d'un tournoi sportif. La notion de domicile, soit le lieu où se déroule une rencontre, est déterminante. Rasmussen et Trick (2008) lui lient différentes notions qui évoquent les qualités d'un horaire. Tout d'abord, un horaire sera équilibré si, pour toute les équipes, la quantité de parties disputées à domicile est équivalente au nombre de rencontres pour lesquelles l'équipe se produisait à l'extérieur. La notion d'équité prévoit que lors des rencontres opposant les deux mêmes équipes, chacune aie aussi souvent « l'avantage du domicile ». Le plus souvent, cette équité est atteinte par l'usage d'un horaire renversé ; un horaire est établi pour la moitié des rencontres requises, puis le même horaire est utilisé de façon inverse pour combler le calendrier sportif, en permutant les équipes hôtes et celle en visite dans le cours de l'exercice. Arjen van

Weert et Jan A.M. Schreuder (1997) apportent la notion d'horaire de ronde minimale qui vise à composer des fragments d'horaire (ronde) en répartissant les équipes dans différents groupes et exigeant que les rondes amènent chacune des équipes à rencontrer une vis-à-vis provenant de chacun des groupes établis de la sorte.

Parmi les contraintes fréquemment rencontrées, on retrouve selon Rasmussen et Trick (2008) l'incapacité des sites à recevoir les matchs pour certaines périodes, l'interdiction pour les équipes d'avoir un calendrier ne prévoyant aucune rencontre au second et à l'avant-dernier jour du calendrier. Le système de ligues parallèles fondées sur le calibre va également occasionner des contraintes particulières pour les équipes ayant récemment changé de division ou étant en passe d'être reclassées. Dans la variante du problème de calendrier sportif, Burke *et al.* (2004) relatent comme objectif la minimisation des déplacements des équipes, une réalité importante pour le sport professionnel en Amérique du Nord ou pour les circuits amateurs. Rasmussen et Trick (2008) considèrent qu'il importe dès lors d'ajouter des contraintes pour espacer les rencontres lors d'un déplacement, et amènent la capacité de transformer le problème de calendrier sportif en un problème de voyageur de commerce.

2.1.3 Horaires dans le transport

Par la force des choses, le domaine du transport est fondamentalement marqué par la dynamique de l'espace, mais sa prévalence comme champ d'application de la recherche opérationnelle a amené plusieurs de ses problématiques à s'inscrire dans le giron des problèmes d'horaire. La complexité de ses opérations font du transport ferroviaire un champ d'application fertile, pour lequel un échantillon de problèmes permettra de saisir les diverses problématiques d'horaires qui se combinent et s'influencent l'une l'autre (comme l'illustre la figure 2.2), en plus de la question centrale de l'incertitude (dont le traitement est étayé à la section 2.2). Bien qu'une analyse du problème de tournées de véhicules à dépôts multiples amène à le considérer comme pouvant appartenir à la classe des problèmes d'horaire, ce type de problème bénéficie déjà d'un courant entier de la littérature qui se consacre à son optimisation, distinct des problèmes d'horaire participant au tour d'horizon actuel.

Le domaine ferroviaire

Cordeau *et al.* (1998) recensent plusieurs problèmes qui, quoique distincts, participent tous à l'horaire général d'un réseau ferroviaire. Le premier est le problème d'organisation des trains, qui, comme son nom le suggère, consiste à déterminer, depuis les gares de triages, comment les wagons seront assemblés en train et le moment où ils quitteront la gare. L'organisation interne d'une gare de triage est à elle seule une question d'optimisation à part entière. Déterminer la

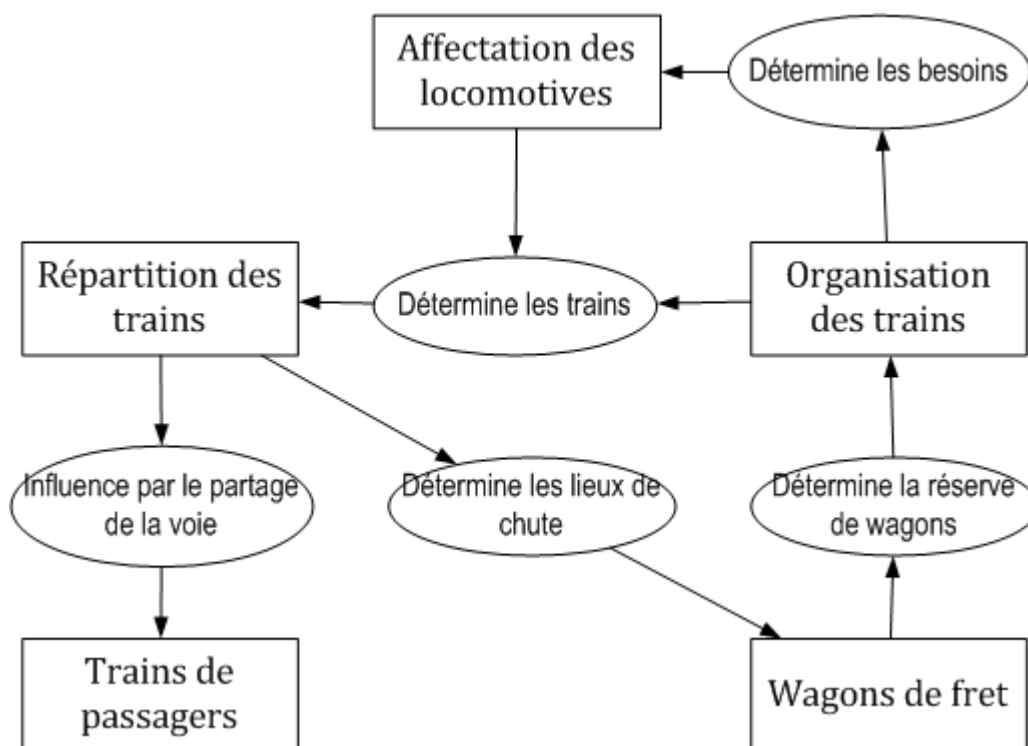


Figure 2.2 Relations entre les problèmes d'horaires du domaine ferroviaire

composition des trains et le moment de leur départ pour minimiser les délais de livraison ou les retards est déjà un problème d'une complexité suffisante. En fait, la résolution peut se faire de façon successive, les trains étant assemblés dans un premier temps et leur moment de départ choisi par la suite. Or, les wagons ne font pas que des voyages avec du chargement. Les points de livraison et d'expédition des marchandises ne se recoupant pas parfaitement, le problème des wagons de fret consiste à planifier la relocalisation des wagons de fret en cherchant à réduire au minimum les occasions où les wagons parcourent le réseau ferroviaire sans charge utile. Cordeau *et al.* (1998) exposent le problème en mettant l'accent sur le caractère incertain de la demande pour des wagons de fret, ce qui induit le besoin de constituer des réserves à différents endroits sur le réseau. De la même façon qu'il faut procéder au redéploiement des wagons, le repositionnement des locomotives, lorsqu'on lui allie la nécessité de couvrir les départs de convois avec une locomotive d'une puissance suffisante, constitue un autre problème d'horaire ferroviaire, l'affectation des locomotives. Son objectif est de minimiser les déplacements « inertes » d'une locomotive (ex. remorquée par une autre, ou sans wagon) tout en réduisant le nombre de locomotives nécessaires pour couvrir les besoins des différents triages. D'une certaine façon, les trois problèmes présentés jusqu'à maintenant constituent les appuis logistiques du problème de répartition des trains, qui, quant à lui, s'intéresse aux

trajets effectués par les trains entre les dépôts et gares. Son contexte d’optimisation est de minimiser les retards sur l’horaire de livraison tout en limitant les temps d’attente des trains en dehors des gares, par exemple sur des voies de dépassement.

Au-delà du transport de marchandises, Cicerone *et al.* (2009a) soulèvent le problème d’horaire de train de passagers qui cherche à minimiser les temps de déplacement des passagers. Fischetti *et al.* (2009) considèrent que la qualité d’une solution est plutôt sa capacité à recouper les horaires de desserte souhaités pour les différentes gares. Ces deux articles citent cependant l’incertitude sur les temps de trajets comme un facteur à considérer. Cicerone *et al.* (2009b) ajoutent que cette incertitude peut entraîner l’échec d’une correspondance pour les passagers, ce qui doit être pris en compte comme impact négatif d’un retard. Le problème d’horaire de ligne de chemin de fer est vu par Cicerone *et al.* (2009b) comme découlant du problème général puisqu’il a comme finalité d’optimiser la desserte d’un tronçon du réseau par différents niveaux de lignes ferroviaires, permettant un service local, régional et express entre grands centres. En cherchant à produire des horaires robustes (une notion exposée à la section 2.2.5), Cicerone *et al.* (2009a) introduit le problème des temps d’attente, pour optimiser la durée des arrêts en gare permettant à l’horaire de service « d’encaisser » les contretemps et imprévus influençant la durée des trajets. Ce dernier n’est pas un problème d’horaire au sens strict, mais Cicerone *et al.* (2009b) a établi qu’il est $\mathcal{NP}$ -dur, ce qui ne manque pas d’avoir un impact sur le problème auquel il se rattache.

2.1.4 Horaires de personnel

Alors que les problèmes précédents ont été regroupés selon la nature des opérations à effectuer, les problèmes d’horaire de personnel doivent leur particularité à la sphère de gestion qui leur est propre. Ernst *et al.* (2004) évoquent comme dénominateur commun de comporter des heures-personnes comme ressource et de faire de leur allocation l’enjeu de l’optimisation. La typologie utilisée par cet article est reprise dans cette section pour présenter les traits distinctifs des problèmes généraux appartenant à cette branche des problèmes d’horaire. Les paramètres dictant les besoins en personnel sont fréquemment déterministes, mais peuvent contenir une dimension incertaine ou provenir de normes plus ou moins arbitraires. Ernst *et al.* (2004) segmentent les étapes fondamentales de construction d’un horaire comme suit : prévision de la demande, planification des jours de repos, planification des quarts, composition des horaires, assignations de tâches particulières, assignation des horaires et tâches.

Le personnel du transport

Si les problèmes évoqués à la section 2.1.3 comportait une dimension spatiale importante, les problèmes de personnel liés ont une emphase similaire. Ernst *et al.* (2004) rappellent que dans le transport aérien, l'espace est un enjeu de l'optimisation. Ce problème est souvent apparenté au problème de transport équivalent, voire lié dynamiquement. Les itinéraires à couvrir peuvent être fragmentés en tournées partielles qui sont analogues aux quarts de travail qu'on retrouverait dans un problème d'horaire conventionnel. Des coûts associés aux déplacements entre les tournées peuvent être incorporés en tant que voyages à vide pour les employés.

Le personnel infirmier

La principale particularité soulevée par De Causmaecker et Vanden Berghe (2010) au sujet des problèmes d'horaire du personnel infirmier est la notion de qualification. Au-delà d'assurer une couverture des quarts, l'enjeu de ce type de problèmes est d'allouer des quarts à du personnel de qualification variable de façon à ce que certaines tâches puissent être réalisées à l'intérieur de l'horizon de planification. Ces tâches peuvent requérir du personnel disposant de certaines qualifications précises, ou encore plus d'une personne ainsi qualifiée sur le même quart. Cette facette se rapproche des contraintes pouvant être rencontrées dans les problèmes de planification industrielle. Bien entendu, d'autres contraintes sont plus proches des autres problèmes de personnel, comme la présence de disponibilités imparfaites au niveau des employés, ainsi que des restrictions quant à la succession de quarts pouvant être effectuée. En somme, l'objectif poursuivi est de minimiser les violations sur la couverture des quarts, sur les tâches à assurer et sur les restrictions de quarts en lien avec le personnel.

Le personnel de la sécurité publique

Le principal indice de performance des services de sécurité publique (corps policiers, services de pompiers, ambulances, etc.) est le temps de réponse. Ernst *et al.* (2004) y voient une caractéristique de l'évaluation de la demande dans les problèmes d'horaire de personnel rattaché. En raison des conséquences funestes de faillir de remplir certains critères de temps de réponse maximaux, l'espace des solutions devient très contraint, une rigidité dans la résolution qui peut être amplifiée par une gestion inadéquate de l'incertitude. Cette organisation du problème fait en sorte que la formation des quarts est plus facile que pour la plupart des problèmes d'horaires, alors que l'effort d'optimisation se reporte davantage sur l'élaboration des horaires et leur assignation aux divers employés.

Le personnel des centres d'appel

Comme les autres problèmes de personnel présentés jusqu'à maintenant, la confection des horaires pour un centre d'appel comporte une particularité au niveau de sa modélisation mathématique. Selon Henderson et Mason (1998), le cadre de ce problème est la planification de l'allocation du personnel pour couvrir différentes périodes pour lesquelles la demande est variable. Ernst *et al.* (2004) estiment que l'incertitude au niveau de la demande est un élément central de cette famille de problèmes, essentiellement car la demande varie plus rapidement que les capacités en personnel ne peuvent raisonnablement le faire. Henderson et Mason (1998) considèrent un découpage du temps en périodes suffisamment courtes pour rendre compte de la variation de la demande et de l'accumulation de cette dernière dans une file d'attente. Ainsi, lorsque la demande excède les capacités du personnel en poste, elle s'accumule pour la période suivante. Des paramètres permettent alors de déterminer la part des appels entrants qui demeure en place et celle qui se termine par un abandon de l'utilisateur. Cette demande résiduelle s'ajoute à celle de la nouvelle période en propre. Avant d'assembler les horaires, il importe donc de former des quarts qui permettent de maintenir à un niveau acceptable le taux d'abandon, soit d'insatisfaction de la demande, ainsi que de réduire les délais d'attente.

2.2 Composer avec l'incertitude

En gestion, l'incertitude est omniprésente. Parvenir à modéliser ce qui est inconnu représente un défi. Pourtant, laisser de côté la part incertaine d'une réalité limite la force d'une formulation. Plusieurs problèmes d'horaire dépeints précédemment (surtout aux sections 2.1.3 et 2.1.4) comportent une part d'incertitude intrinsèque, mais d'autres paramètres des problèmes peuvent être modélisés avec une valeur fluctuante ou imprécise. Dans l'optique de la production de solutions ayant une valeur appliquée, il devient ainsi pertinent de parvenir à produire une formulation ou un mode de résolution permettant d'inclure cet état de fait. À l'intérieur de cette section, la typologie s'inspire de Cicerone *et al.* (2009b) en présentant d'abord deux approches déterministes, la programmation stochastique et la programmation par contraintes, puis d'autres techniques comme la programmation probabiliste et floue, pour ensuite exposer les notions de robustesse et de récupération. D'autres manières de faire font partie de la littérature. À titre d'exemple, citons Cordeau *et al.* (1998) qui mentionnent l'utilisation directe d'une distribution statistique sans traitement plus sophistiqué, ou encore Henderson et Mason (1998) pour qui le recours à une simulation peut permettre d'identifier des bornes sur un paramètre aléatoire.

2.2.1 Programmation stochastique

Très usité dans le domaine des tournées de véhicules selon Verderame *et al.* (2010), la programmation stochastique par recours permet la gestion de l'incertitude sous forme de scénarios. Ravi et Sinha (2005) et Fischetti *et al.* (2009) font état d'une formulation dans laquelle le nombre de scénarios est fini et où chacun d'entre eux peut survenir selon une probabilité qui lui est propre. Selon la construction des variables de recours, Fischetti *et al.* (2009) soulignent que même si le nombre de scénarios induits peut être très grand, des techniques d'échantillonnage peuvent guider la résolution. Toute la gestion de l'incertitude réside dans ces scénarios, qui sont représentés à l'intérieur du second stade, avec les variables de recours. En effet, pour Sahinidis (2004), un programme stochastique cherchera à déterminer une configuration des variables de décision (premier stade) qui ne sera pas optimale du simple point de vue de ces variables, mais qui intégrera une minimisation des coûts additionnels liés aux variables de recours (second stade), qui décrivent l'impact sur la solution des différentes incertitudes pouvant survenir. Pour Birge et Louveaux (1997), l'enchaînement des stades est liée à la levée de l'incertitude. Les décisions prises dans le cadre du premier stade doivent composer avec l'incertitude. À l'opposé, les variables de recours vont agir une fois qu'il n'y a plus d'incertitude, mais elles seront liées les choix posés au premier stade. Ravi et Sinha (2005) expliquent que les deux stades permettent de séparer la partie déterministe du problème du contexte incertain, qui lui est pris en charge dans le second stade. C'est dans celui-ci, pour citer Dimitrakopoulos et Ramazan (2008), que les recours sont déterminés et inclus dans l'optimisation générale, sous la forme de scénarios. En somme, la programmation stochastique par recours est limitée par le besoin de connaître au préalable les probabilités associées à l'incertitude modélisée et par la taille de l'éventail de scénarios possibles. Par contre, la présence d'une résolution en deux stades permet d'incorporer au sous-problème de recours d'autres méthodes de gestion de l'incertitude.

La capacité d'encapsulation permise par la programmation stochastique lui permet de se contenir elle-même. Sahinidis (2004) qualifie de programmation stochastique dynamique le fait de résoudre en succession plusieurs problèmes de programmation stochastique pour représenter les facteurs d'interférence stochastique. Il convient toutefois de rappeler qu'il découle de cette pratique une certaine lourdeur, que Sahinidis (2004) recommande d'atténuer en utilisation des algorithmes d'approximation là où un gain de temps peut être effectué.

2.2.2 Programmation par contraintes

Pour résoudre des programmes en nombres entiers, Lombardi *et al.* (2010) capitalisent sur la décomposition de Benders, afin d'inclure des éléments de résolution provenant de la pro-

grammation par contraintes. Bien que cette forme de programmation partage une proximité avec les programmes en nombres entiers, il est nécessaire de procéder à une transformation pour bénéficier des outils de résolution typique des programmes en nombres entiers. Ainsi, les contraintes sont d’abord représentées sous la forme d’un graphe de tâches conditionnelles, qui, lui, peut être décrit par un programme en nombres entiers. Dans ce dernier, des variables d’activation représentent l’actualisation d’un scénario qui est lui-même issu de sous-graphes. En utilisant des branchements probabilistes pour produire ces derniers, on parvient à créer une résolution où le problème maître va rechercher l’optimalité pendant que les sous-problèmes vont statuer sur la faisabilité. Lombardi *et al.* (2010) mettent cependant en garde contre l’utilisation de cette technique sur de trop petits exemplaires, car le manque de relief dans l’espace des solutions peut rendre fastidieux la recherche d’une solution optimale.

2.2.3 Programmation probabiliste

Alors que les techniques déterministes exposées ci-haut vont utiliser des variables pour représenter l’incertitude, la programmation probabiliste a plutôt recours aux contraintes pour incorporer cette réalité. Sahinidis (2004) relate que c’est avec une classe de contraintes dite aléatoires que le caractère incertain des événements peut être pris en compte. Verderame *et al.* (2010) ajoutent que si elles sont limitées à des incertitudes dont la distribution est connue, les contraintes aléatoires peuvent en revanche représenter différents types d’incertitudes à la fois. Pour Golamnejad *et al.* (2006), la programmation probabiliste adopte en ce sens une optique de « faire au mieux », plutôt que de chercher « la » meilleure solution. Souvent, les contraintes aléatoires vont prendre la forme d’un seuil de tolérance par rapport à une incertitude. Si le niveau de risque est jugé acceptable, la contrainte est dès lors satisfaite. La non-linéarité de la plupart des contraintes aléatoires représente cependant un inconvénient majeur. Sahinidis (2004) rapporte que des contraintes linéaires peuvent être utilisées en substitution, sous réserve d’en démontrer l’équivalence. Verderame *et al.* (2010) précisent que certaines conversions peuvent se faire en utilisant la structure de la distribution de l’incertitude sur laquelle est basée la contrainte aléatoire.

2.2.4 Programmation floue

Œuvrant autant avec les contraintes que les variables, cette forme de programmation mathématique permet de s’affranchir de la dépendance aux distributions des éléments incertains qu’on cherche à modéliser. En effet, Verderame *et al.* (2010) assurent que la programmation floue peut donner des résultats probants même en l’absence de toute distribution. Sahinidis (2004) expose brièvement que la théorie des ensembles flous donne une assise à la program-

mation floue en utilisant des contraintes floues qui ont une influence sur des variables (elles aussi) floues. Sahinidis (2004) rapporte comment une fonction d'appartenance linéaire peut être utilisée pour prendre en compte les parties floues d'un programme, permettant une résolution qui utilise le programme flou comme recours. Verderame *et al.* (2010) tirent de cette linéarisation une possibilité de coupler la programmation floue avec d'autres méthodes de gestion de l'incertitude. Les auteurs soulignent également que des facteurs complexes, comme ceux tirés de comportements sociologiques, peuvent être agrégés en un ou plusieurs ensembles flous. Dans un contexte d'optimisation, Sahinidis (2004) présente l'objectif comme la minimisation de la violation des contraintes floues, souvent selon des bornes définies au préalable.

2.2.5 Robustesse et récupération

Issue de la programmation stochastique avec recours selon Sahinidis (2004), la notion de robustesse vise à déterminer une solution pouvant satisfaire, lors de l'actualisation de l'incertitude, un seuil de tolérance. Cicerone *et al.* (2009a) et Henderson et Mason (1998) préfèrent parler de la qualité que possède une solution non optimale. Confrontée à une légère variation sur des paramètres incertains, une solution robuste subira une détérioration de sa valeur jugée moindre que celle que subirait en de pareilles circonstances la valeur de la solution optimale selon une résolution déterministe. Verderame *et al.* (2010) reconnaît à l'idée de robustesse la capacité de couvrir différents types d'incertitude alors que Sahinidis (2004) estime que les solutions robustes risquent d'être trop conservatrices pour être utiles dans un contexte d'application. Surtout, les auteurs constatent que certaines formulations vont compenser ce travers en minimisant sciemment les risques de recours.

Pour conserver des solutions intéressantes, Cicerone *et al.* (2009b) proposent plutôt le concept de robustesse légère, qui vise à qualifier des solutions qui conservent leur faisabilité dans un sous-ensemble de scénarios couvrant une relaxation acceptée comme étant raisonnable. Cette notion vise surtout à concurrencer la robustesse stricte, qui elle cherche à garantir la faisabilité dans l'ensemble des scénarios. Avec la robustesse légère, on introduit un relâchement dans la prudence, ne cherchant à se prémunir que contre une combinaison raisonnable de facteurs, ou encore en considérant qu'il est plus plausible que les variations dues à l'incertitude s'annulent partiellement entre elles plutôt que de toutes repousser la solution dans le même sens.

Après la phase de planification, Cicerone *et al.* (2009a) et Cicerone *et al.* (2009b) décrivent la récupération comme un ensemble de méthodes heuristiques qui procèdent à des ajustements sur une solution après actualisation de l'incertitude. Une bonne récupération permet de rendre une solution à nouveau réalisable. La qualité de la récupération peut se mesurer par la valeur

qui a pu être recouvrée par rapport à la valeur attribuée à la solution lors de la phase de planification. Cicerone *et al.* (2009b) combinent d’ailleurs les stratégies de récupération avec l’idée de robustesse légère pour démontrer qu’il est possible d’amener des solutions robustes à être concurrentielle dans un contexte d’applications.

2.3 Résolution et proximité théorique

Il a été fait brièvement état à la section 1 que le domaine d’application du problème stochastique d’horaire d’audiences judiciaires n’est pas l’un des champs généraux en matière de problèmes d’horaire. Néanmoins, des similitudes sont présentes. Les prochaines sections vont permettre d’une part d’identifier les éléments communs du problème stochastique d’horaire d’audiences judiciaires et des problèmes définis à la section 2.1, d’autre part de relever comment ces éléments sont pris en compte dans les autres travaux scientifiques.

2.3.1 Problème d’horaire de classe

Forme la plus fondamentale du problème d’horaire d’université, le problème d’horaire de classe est également proche du problème stochastique d’horaire d’audiences judiciaires de par sa structure. On y retrouve le personnel enseignant comme faisant partie des ressources à assigner, ce qui permet de représenter les conflits entre les causes du problème de salles d’audience comme issues de l’utilisation du même avocat (ou d’une même partie). Il suffit que ces conflits deviennent des cours ayant recours au même enseignant, les salles d’audience deviennent des salles de classe et le problème de salles d’audience devient une forme du problème d’horaire de classe dont les cours ont une durée variable, généralement supérieure à une période. L’incertitude sur la durée des causes n’est pas très compatible avec le contexte général des horaires de classes, mais sinon, l’articulation de cette équivalence présente un certain intérêt.

Tableau 2.1 Méthodes de résolution pour le problème d’horaire de classe selon Hertz (1991), Schaerf (1999) et Burke *et al.* (2004)

Exact	Métaheuristique	Autre
algorithmes de flot	algorithmes génétiques recuit simulé recherche tabou	permutation prog. par contraintes

Schaerf (1999) parle du problème d’horaire de classe comme d’un problème dur mais d’une complexité relativement peu élevée. On peut voir une synthèse des méthodes de résolu-

tion couramment utilisées dans le tableau 2.1. Les heuristiques de permutation fréquemment utilisées durant les années soixante et soixante-dix ont cédé la place à des méthodes métaheuristiques diverses. Burke *et al.* (2004) signalent qu’une succession de problèmes de flots à coût minimum permettent de prendre en compte un grand nombre de contraintes souples. Schaerf (1999) expose que les algorithmes génétiques peuvent évoluer en marge d’un espace de solutions trop étroit pour tenter de retrouver une faisabilité intéressante. Hertz (1991) propose de réduire les conflits d’affectation des enseignants avec une méthode tabou dont le déplacement permute les cours ou les enseignants en charge. Enfin, Schaerf (1999) évoque l’utilisation du recuit simulé ou la programmation par contraintes en utilisant comme recherche locale la descente en gradient (*hill climbing*).

2.3.2 Les salles dans la résolution des problèmes d’éducation

La proximité avec le problème d’horaire de classe présenté à la section 2.3.1 permet des rapprochements avec les autres problèmes de la même famille. La ressource que représente les salles (tantôt de cours de justice, tantôt de cours académique) procède de la même logique dans les deux cas. Il apparaît donc fondé de penser que les modes de formulation ayant cours dans les problèmes d’horaire de l’éducation pour prendre en compte les enjeux liés aux salles trouveront une résonance dans la formulation mathématique du problème stochastique d’horaire d’audiences judiciaires.

Le tableau 2.2 résume les méthodes de résolution du problème d’horaire de cours d’université. Toujours selon la logique d’une résolution successive de problèmes de flot à coût minimum, Schaerf (1999) considère l’appariement des cours aux salles comme l’un des éléments pouvant faire l’objet d’un graphe de flot. Avec le recours à la programmation par contraintes, Yang *et al.* (2010) énoncent une méthode en deux stades qui construit un horaire en relaxant les contraintes de salles tout en conservant un plafond pour éviter les débordements. Un sous-groupe de contraintes permet par la suite d’amener cette solution à être

Tableau 2.2 Méthodes de résolution du problème d’horaire de cours d’université selon Hertz (1991), Schaerf (1999), McCollum (2006), Burke *et al.* (2010), Najdpour et Feizi-Derakhshi (2010) Yang *et al.* (2010), Lutuksin et Pongcharoen (2010) et Shiau (2011)

Exact	Métaheuristique	Autre
relaxation lagrangienne algorithme récursif algorithme de flot	recherche par voisinage essaimage algorithme génétique algo. colonie de fourmis	prog. nombres entiers décomposée prog. par contraintes

réalisable. L'avantage de cette façon de faire est qu'elle permet l'interception d'une solution et sa modification selon un arbitrage humain. Salwani Abdullah *et al.* (2010) reprennent l'approche par étape à l'intérieur d'une recherche à voisinage multiples dont un voisinage consiste à déplacer des cours d'une période à l'autre dans la même salle alors qu'un autre propose de déplacer un cours conflictuel dans une autre salle et ailleurs dans l'horaire. Les résultats obtenus permettent de démontrer une bonne performance pour ces deux voisinages, particulièrement sur les petits exemplaires énoncés par Socha *et al.* (2002). Pour résoudre les conflits liés entre autres aux salles, Shiau (2011) ont recours à une métaheuristique par essaimage dans lequel un vecteur représente la solution. Un déplacement aléatoire des cours provoquant un conflit permet de réparer les solutions. Cette méthode incorpore les préférences des usagers tout en conservant une performance comparable aux algorithmes génétiques contemporains, comme celui de Najdpour et Feizi-Derakhshi (2010), fondé sur la permutation des cours entre les parents jusqu'à obtenir un enfant réalisable. Cet algorithme génétique, qui supprime les procédés traditionnels, utilise comme procédure de mutation un déplacement aléatoire d'un cours ailleurs dans l'horaire.

En ce qui concerne le problème d'horaire d'examens, Schaerf (1999) mentionne que l'effort particulier pour assurer une répartition des cours dans les salles est souvent cantonné à des heuristiques de construction de la solution initiale. Or, pour reprendre McCollum (2006), il importe de remettre la recherche de l'optimalité au coeur de la démarche et de ne pas viser la simple réalisabilité. Le tableau 2.3 souligne que les méthodes de résolution pour ce problème ne sont que partiellement similaires à celles qui prévalent pour le problème d'horaire de cours d'université.

2.3.3 Résolution en demande incertaine

Si les problèmes d'horaire de l'éducation présentés à la section 2.1.1 ont une structure à laquelle s'apparente le problème stochastique d'horaire d'audiences judiciaires, une facette importante de ce problème n'est pas exprimée dans ces derniers. La principale problématique

Tableau 2.3 Méthodes de résolution du problème d'horaire d'examen selon Schaerf (1999), Burke *et al.* (2004) et McCollum (2006)

Exact	Métaheuristique	Autre
relaxation lagrangienne	recuit simulé algorithmes génétiques recherche taboue	algorithmes de réseaux programmation floue prog. de Pareto

d'optimisation du rôle judiciaire demeure l'incertitude quant à la durée des causes. C'est pourquoi il est permis de revenir également sur les problèmes pour lesquels l'incertitude de la demande fait partie des éléments déterminants. Dans le domaine du transport ferroviaire et dans celui des horaires de personnel, la demande peut être d'un volume incertain. Généralement, cette incertitude se retrouve dans le nombre des événements à inclure à l'horaire, mais l'organisation de la résolution comporte néanmoins un intérêt.

Qualifiées d'approches dynamiques par Cordeau *et al.* (1998), quelques méthodes de résolution du problème de répartition des trains prennent en compte l'incertitude. Dans un programme en nombres entiers, les liaisons réalisables pour les différents trains potentiels prennent la forme de variables binaires. La résolution par énumération implicite est viable en termes de temps de calcul pour les petits exemplaires. Si la faisabilité de portions de trajets est maintenue, un recuit simulé peut agir comme récupération. Enfin, un algorithme génétique peut constituer une population à partir des solutions potentielles et observer leur survie dans un environnement où des incertitudes sont introduites. Pour le problème des wagons de fret, les mêmes auteurs suggèrent une programmation stochastique où les différentes demandes font office de variables de recours. L'énumération avec plans coupants est une autre approche aux résultats satisfaisants pour introduire les éléments incertains. Sans être strictement une méthode exacte, l'optimisation robuste offre également une bonne performance dans un cadre où la vitesse des convois peut être considérée comme fixe. Les méthodes de résolution repertoriées pour les différents problèmes de transport évoqués à la section 2.1.3 sont regroupées dans le tableau 2.4.

Du côté des problèmes d'horaire de personnel, Henderson et Mason (1998) soulèvent l'intégration à un programme en nombres entiers des fruits d'une simulation permettant d'anticiper les niveaux de demande pour un centre d'appel. Des coupes sont utilisées pour borner la résolution à des patrons respectant un seuil de couverture acceptable. Côté *et al.* (2009) exposent un programme en nombres entiers reposant sur un problème de flot. Ce dernier est construit en utilisant des contraintes aléatoires sur la demande. La performance de cette approche face à d'autres programmes complètement linéaires n'est pas démontrée par les auteurs, mais cette formulation permet une réduction notable de la complexité. Dans le domaine du transport aérien, Ernst *et al.* (2004) mentionnent le recours à la programmation floue couplée à une recherche par voisinages. D'autres formulations linéaires présentent selon les mêmes auteurs des faiblesses si la linéarisation de l'incertitude n'est pas bien démontrée.

Tableau 2.4 Méthodes de résolution des problèmes de transport selon Cordeau *et al.* (1998) et Kliewer *et al.* (2006)

Exact	Métaheuristique	Autre
algorithmes de flot génération de colonnes relaxation lagrangienne énumération implicite prog. stochastique énum. avec plans coupants	algorithmes génétiques heuristique de court chemin	prog. dynamique réseau neuronal transformation en inventaire trans. en tournée de véhicule trans. en ordonnancement

2.4 Apport du mémoire

Au terme de cette revue de la littérature scientifique contemporaine, il est possible de supputer les aspects du problème stochastique d'horaire d'audiences judiciaires, de sa formation et de la résolution proposée qui dénotent un intérêt mathématique. Les prochaines sections aborderont la particularité de l'incertitude à laquelle le problème est confrontée au niveau des causes, puis l'hybridation entre la programmation en nombres entiers et la programmation probabiliste qui apporte une méthode de résolution singulière pour un problème d'horaire.

2.4.1 Durée incertaine des évènements

Il a été évoqué à la section 2.3.3 que le problème stochastique d'horaire d'audiences judiciaires connaissait une incertitude sur sa demande en termes de durée des évènements la composant, soit les procès. Dans les cas répertoriés dans le présent chapitre, l'incertitude tient davantage à une quantité d'éléments qu'à la durée de ceux-ci. Pour le traitement, la plupart des méthodes de formulation énoncées à la section 2.2 demeurent valides, mais il n'en demeure pas moins qu'une partie de la structure du problème sera différente. De plus, la possibilité de procès à durée nulle constitue un élément inusité dans les articles scientifiques

Tableau 2.5 Méthodes de résolution des problèmes de personnel selon Henderson et Mason (1998), Ernst *et al.* (2004) et Côté *et al.* (2009)

Exact	Métaheuristique	Autre
énum. avec plans coupants	algorithmes génétiques	prog. probabiliste programmation floue prog. par contraintes

et ouvrages couverts. Quelques formulations de problèmes d'horaire vont s'attarder sur la variation de la durée des événements. On peut signaler les problèmes issus de la gestion des blocs opératoires, où la durée des interventions chirurgicales peut varier. Par contre, la prise en considération d'événements potentiellement sans durée semble apparaître comme un élément novateur. La section 3.1 présente d'ailleurs des variantes du problème stochastique d'horaire d'audiences judiciaires qui permettent cette prise en compte.

2.4.2 Énumération implicite et plans coupants

Comme il est possible de le constater dans les tableaux 2.1 à 2.4, les différentes familles de problèmes d'horaire vont mettre l'accent sur différentes techniques de résolution. Dans le domaine de l'éducation, les métaheuristiques sont très prisées en raison de la taille de l'espace des solutions, alors que dans le domaine du transport, la programmation en nombres entiers a un ascendant indéniable. Tel que relaté à la section 2.2.3, il est possible de linéariser les contraintes aléatoires, ce qui permet de les inclure dans un programme en nombres entiers. Chez Côté *et al.* (2009), les deux types de programmation sont associées pour résoudre un problème d'horaire de personnel alors que les contraintes aléatoires sont représentées sous la forme d'un graphe. Ce dernier est alors transposé en contraintes linéaires au sein du programme en nombres entiers. Dans la formulation présentée à la section 3.3, la jonction entre programme en nombres entiers et programmes probabilistes ressemble davantage au procédé présenté par Henderson et Mason (1998), toujours dans le contexte d'un problème d'horaire de personnel. Plutôt que d'utiliser l'insertion de coupes pour intégrer graduellement des bornes sur l'incertitude, les coupes vont servir à intégrer au modèle une linéarisation des contraintes aléatoires sans qu'il soit nécessaire de recourir à une transformation supplémentaire.

CHAPITRE 3

FORMULATIONS ET ALGORITHME DE RÉOLUTION

Afin de poursuivre l'étude du problème stochastique d'horaire d'audiences judiciaires, il convient d'en présenter des formulations mathématiques et de mettre de l'avant un algorithme de résolution. Le chapitre s'ouvre sur la description des deux formulations retenues pour modéliser mathématiquement le problème stochastique d'horaire d'audiences judiciaires. Dans le contexte de la résolution qui sera présentée en fin de chapitre, des méthodes d'évaluation de probabilités sont par la suite abordées.

3.1 Formulation du problème stochastique d'horaire d'audiences judiciaires

Dans cette section seront présentés successivement deux modèles de programmation mathématique représentant le problème stochastique d'horaire d'audiences judiciaires. Les notions communes aux deux modèles seront exposées en introduction. L'algorithme de résolution présenté à la section 3.3 met à profit ces deux modèles pour dégager une solution optimale.

3.1.1 Généralités

La section 1.2.1 a brossé un premier portrait du problème stochastique d'horaire d'audiences judiciaires. En guise de rappel, ce problème comporte un ensemble C^R fini de causes régulières qu'on tente d'inscrire dans un horaire composé d'un nombre fini de périodes P dans un palais de justice contenant un nombre fini de types S de salles. Les causes $c \in C^A$, les causes assignées, ont été mises au rôle précédemment et ne peuvent être déplacées. Chacune des causes engendre, une fois à l'horaire, un procès dont la durée peut varier selon une distribution connue. Un nombre donné et connu de causes entrent en conflit deux à deux en raison d'un usage concurrent de ressources (avocats, témoins, juges, etc.), leurs procès respectifs ne peuvent donc se dérouler sur les mêmes périodes. Si trois ou plusieurs causes sont en conflit, il est quand même possible de les modéliser deux à deux sans perdre de généralité. Enfin, les causes peuvent nécessiter des équipements ou des installations seulement présents dans certains types de salles d'audience, le nombre de salles disponibles pour chacun de ces types étant lui aussi fini.

Le problème d'optimisation posé consiste donc à fixer dans un horaire des audiences pour les procès tout en réduisant les délais. Les formulations exposées aux sections 3.1.2 et 3.1.3 utilisent le concept de « multi-salles » proposé par Burke *et al.* (2010). Pour l'expliquer

brèvement, il faut considérer que puisque, par définition, les salles d'un même type disposent des mêmes équipements, il est possible de les remplacer par une salle pouvant accueillir pour une période donnée autant de procès qu'il y a de salles d'audience de ce type dans le palais de justice. Cet artifice théorique permet de simplifier la résolution en diminuant la symétrie engendrée par les salles.

Considérant le recours aux multi-salles, voici les contraintes proposées, qui sont exposées plus en détail au paragraphe suivant :

Contraintes fermes

- CF1 : Les causes en conflit ne doivent pas procéder durant la même période ;
- CF2 : Le procès d'une cause doit débiter à une période valide ;
- CF3 : Chaque cause à l'horaire reçoit un type de salle d'audience disposant des équipements appropriés ;
- CF4 : Pour chaque période, la probabilité que le nombre de procès en cours dans un type de salle soit supérieur au nombre de salles de ce type est inférieure à $1 - \alpha$, un seuil de tolérance défini comme paramètre ;
- CF5 : Les causes préassignées ne peuvent être déplacées.

Contraintes souples

- CS1 : Toutes les causes doivent être inscrites au rôle et ainsi recevoir une période de début.

Objectif

Minimiser la somme des temps d'attente avant l'ouverture de chacun des procès. L'ensemble K permet d'identifier les paires de causes pouvant, en entrant en conflit, induire une violation de la contrainte (CF1). À l'inverse, l'ensemble P_c ne retient, pour chaque cause, que les périodes de début assurant le respect de la contrainte (CF2). La construction de l'ensemble S_c permet, selon le même raisonnement, le respect de la contrainte (CF3). La notion de multi-salles est traduite par le nombre N_s qui, pour chaque type s , indique la quantité de salles qui y sont consolidées. La souplesse de la contrainte (CS1) est assurée par l'ajout de variables d'écart e_c dans les contraintes d'affectation des causes au rôle, variables qui sont fortement pénalisées dans la fonction-objectif. Dans sa formulation actuelle, l'objectif implique que les causes n'ont pas de priorité entre elles dues à la nature du litige, à sa médiatisation ou à la durée prévue des audiences. Cette conception repose sur l'équité des justiciables aux yeux de l'État.

3.1.2 Modèle avec durée fixe et surcapacité

Le premier modèle présenté s'inspire le plus directement de la pratique empirique prévalant dans l'administration de la justice. Cette dernière établit, selon un taux de surcharge, idée ici reprise, une capacité dans laquelle est fixée de façon séquentielle la place des différentes causes. La formulation utilisée ci-après est caractérisée par la notion de durée fixe puisqu'elle considère que le procès de chaque cause aura automatiquement la plus longue durée prévue. Comme la durée réalisée des causes ne peut qu'être inférieure (jusqu'à en être nulle) ou égale à cette durée déterminée, les solutions vont présenter des vides par rapport à la planification initiale. Le recours à la surplanification permet de diminuer les temps d'attente avant le début des différents procès. À l'aide d'un paramètre, cette surplanification peut être ajustée de manière à réaliser un arbitrage entre la réduction globale des temps d'attente et le risque de produire un horaire occasionnant des scénarios d'encombrement de rôle lorsque l'incertitude est levée sur la durée des causes. Voici la formulation qui incorpore ces éléments :

Variables

x_{cps} : Variable binaire de décision sur la période p de début de la cause c et sur le type de salle s retenu.

e_c : Variable binaire d'écart sur la non-mise au rôle de la cause c .

Programme

$$\min_{x,e} \quad \sum_{c \in C^R} \sum_{p \in P_c} \sum_{s \in S_c} p \cdot x_{cps} + \sum_{c \in C^R} M_c e_c \quad (3.1)$$

$$\text{s.c.} \quad \sum_{p \in P_c} \sum_{s \in S_c} x_{cps} + e_c = 1 \quad \forall c \in C^R \quad (3.2)$$

$$\sum_{c \in C^R} \sum_{q \in P_{cp}} x_{cqs} + |C_{ps}^A| \leq \beta N_s \quad \forall s \in S, p \in P \quad (3.3)$$

$$\sum_{s \in S_i} \sum_{q \in P_{ip}} x_{iqs} + \sum_{s \in S_j} \sum_{q \in P_{jp}} x_{jq s} \leq 1 \quad \forall (i, j) \in K, p \in P \quad (3.4)$$

$$x_{cps}, e_c \in \{0, 1\} \quad \forall c \in C^R, p \in P_c, s \in S_c \quad (3.5)$$

La fonction-objectif (3.1) minimise deux éléments distincts. D'abord, elle somme les délais d'attente, représentés par l'indice p des variables de décision x_{cps} ayant pris la valeur 1. Puis, la fonction-objectif somme les pénalités M_c associées aux causes c n'ayant pas été inscrites au rôle. Ces pénalités viennent du non-respect de la contrainte (CS1). Pour une cause c

donnée, la valeur du paramètre M_c doit être suffisamment grande pour qu'il soit en tout temps préférable d'inscrire la cause c au rôle, *ceteris paribus*.

Les contraintes d'affectation (3.2) forcent le choix d'une période de début et d'un type de salle pour chacune des causes, dans le respect des contraintes (CF2) et (CF3). Si la contrainte (CS1) n'est pas respectée, la variable d'écart e_c correspondante va provoquer l'ajout de la pénalité dans la fonction-objectif. Les contraintes de surplanification (3.3) limitent le nombre de causes pouvant se dérouler lors d'une période p par rapport au nombre de salles du type s . La sommation effectuée ne tient compte que des périodes de début valide pour chaque cause c . Il faut cependant lui ajouter l'ensemble C_{ps}^A des causes préassignées procédant dans une salle de type s et dont la durée maximale couvre la période p . Le paramètre β détermine la surcapacité maximale autorisée. Si β prend la valeur 1, il n'y aura pas de surplanification, mais la contrainte (CF4) sera assurément respectée et les solutions seront très conservatrices. Si β prend toute valeur supérieure à 1, cette garantie de respect est perdue. Les contraintes (3.4) empêchent deux causes étant en conflit de procéder à la même période, selon la contrainte (CF1). Enfin, les contraintes de binarité (3.5) découlent de la définition des variables qui représentent pour une cause c le choix, ou l'absence de choix d'une période p et d'un type de salle s comme le début du procès associé.

3.1.3 Modèle avec probabilités

La section précédente a présenté un modèle à durée fixe linéaire mais qui ne peut garantir le respect de la contrainte (CF4). Pour obtenir cette garantie, le présent modèle délaisse l'utilisation d'une durée fixe pour les causes afin d'introduire des variables aléatoires représentant l'incertitude sur la durée des causes. La valeur de ces variables va indiquer, avec la levée de l'incertitude, si le procès d'une cause c donnée se concrétise durant une période p dans une salle de type s . Suivant une loi de Bernoulli, la valeur de ces variables obéit aux probabilités suivantes :

$$Y_{cps} = \begin{cases} 1 & \text{avec une probabilité de } \pi_{cp} & \text{si } c \in C^A \\ & \text{avec une probabilité de } \sum_{q \in P_{cp}} x_{cqs} \left(\pi_{c(p-q+1)}^R \right) & \text{si } c \in C^R \\ 0 & \text{avec une probabilité de } (1 - \pi_{cp}) & \text{si } c \in C^A \\ & \text{avec une probabilité de } \sum_{q \in P_{cp}} x_{cqs} \left(1 - \pi_{c(p-q+1)}^R \right) & \text{si } c \in C^R. \end{cases} \quad (3.6)$$

Il est possible de formuler des probabilités π_{cp} directement avec les causes préassignées car si leur durée demeure incertaine, leur période de début et le type de salle alloué est déjà arrêté par définition. Pour les causes régulières, la probabilité dépend de la valeur des variables de

décision x_{cps} , puisque le procès d'une cause régulière c a une probabilité π_{cd} de s'étendre sur au moins d périodes. On peut ainsi considérer l'exemple d'une cause pouvant durer jusqu'à quatre périodes. Si deux moments de départs valides se retrouvent à l'intérieur d'un horizon de quatre périodes, la valeur de Y_{cps} tiendra compte des probabilités associées à ces deux périodes de départ. Comme les variables aléatoires expriment l'incertitude sur la durée des causes, elles sont utilisées pour formuler les contraintes probabilistes suivantes :

$$P \left[\sum_{c \in C^A \cup C^R} Y_{cps} > N_s \right] \leq 1 - \alpha \quad \forall s \in S, p \in P. \quad (3.7)$$

Ces nouvelles contraintes doivent être respectées pour chaque type de salle et pour toutes les périodes de l'horizon. Le membre de droite représente le niveau de risque maximal autorisé pour qu'une solution soit jugée acceptable. La formulation reprend donc les contraintes du modèle avec durée fixe et surcapacité (3.1)-(3.5), en substituant les contraintes (3.7) aux contraintes (3.3). Sous cette forme, les contraintes (3.7) ne peuvent être intégrées dans une formulation linéaire. Pour une période p et un type de salle s donnés, on peut identifier un procès pouvant survenir avec une probabilité non-nulle par un couplet composé de la cause procédant et de sa période de départ. L'ensemble $\overline{G}_{ps}$ regroupe ces couplets à la condition que la probabilité que tous ces procès aie lieu simultanément dépasse le seuil $1 - \alpha$ retenu dans les contraintes (3.7). En supposant que la composition de $\overline{G}_{ps}$ est connue, on peut formuler les contraintes linéaires suivantes :

$$\sum_{(c,q) \in \overline{G}_{ps}} x_{cqs} \leq |\overline{G}_{ps}| - 1 \quad \forall \overline{G}_{ps} \in \Gamma_{ps}, p \in P, s \in S. \quad (3.8)$$

L'ensemble Γ_{ps} regroupe les divers ensembles $\overline{G}_{ps}$. La taille de Γ_{ps} peut être considérable selon le nombre de causes et leurs distributions. Il en découle un fort grand nombre potentiel de contraintes (3.8), ce pourquoi elles seront ajoutées au besoin en cours de résolution.

3.2 Méthodes d'évaluation des probabilités d'encombrement du rôle

En cours de résolution, les contraintes probabilistes (3.7) violées sont identifiées et remplacées dans les modèles par les contraintes (3.8) correspondantes. Or, l'ajout de ces dernières suppose qu'un ensemble $\overline{G}_{ps}$ est connu. Mais cet ensemble est défini par la violation qu'il provoque des contraintes (3.7). Pour permettre une résolution satisfaisante, il est nécessaire d'être en mesure de tester si une solution entière du problème stochastique d'horaire d'au-

diances judiciaires satisfait, pour chaque période et chaque type de salles, ces contraintes probabilistes. Pour la suite de la section, le cas d'une période p et d'un type de salle s est retenu. Trois méthodes différentes sont proposées et comparées pour tenter de permettre une évaluation versatile de cette probabilité.

L'ensemble $\overline{G}_{ps}$ a été décrit à la section 3.1.3 comme regroupant les causes et leur période de début entraînant un encombrement du rôle. De façon plus générale, pour inclure dans le calcul de probabilité les causes préassignées et l'ensemble des causes régulières pertinentes, l'ensemble G_{ps} est dénoté formellement comme suit :

$$\overline{G}_{ps} \subseteq G_{ps} = \{(c, q) \mid c \in C, q \in P_c \text{ et } P[Y_{cps} = 1] > 0\}. \quad (3.9)$$

De par la construction de G_{ps} , il est possible de restreindre l'évaluation de la probabilité qu'il y ait débordement des salles aux seules causes faisant partie de l'ensemble, et d'écrire la probabilité devant être évaluée comme :

$$P \left[\sum_{c \in C} Y_{cps} > N_s \right] = P \left[\sum_{c: (c, q) \in G_{ps}} Y_{cps} > N_s \right]. \quad (3.10)$$

L'obtention d'une solution entière $\hat{x}$ permet de déterminer un sous-ensemble $\hat{G}_{ps}$ de G_{ps} pouvant être défini comme suit :

$$\hat{G}_{ps} = \{(c, q) \mid c \in C^A \text{ et } q \in P_{cp}\} \cup \{(c, q) \mid c \in C^R, \hat{x}_{cqs} = 1 \text{ et } q \in P_{cp}\}. \quad (3.11)$$

En se fondant sur l'égalité (3.10), on peut trouver la relation suivante :

$$P \left[\sum_{c: (c, q) \in \hat{G}_{ps}} Y_{cps} > N_s \right] > 1 - \alpha \Rightarrow \hat{G}_{ps} \in \Gamma_{ps}. \quad (3.12)$$

3.2.1 Méthode exacte

La première méthode consiste à déterminer la valeur de la probabilité de la relation (3.12) par un calcul exact. En s'appuyant sur chacun des ensembles $\hat{G}_{ps}$, la probabilité se décompose en une sommation comme suit :

$$\mathbb{P} \left[\sum_{c: (c,q) \in \hat{G}_{ps}} Y_{cps} > N_s \right] = \begin{cases} \sum_{i=N_s+1}^{|\hat{G}_{ps}|} \mathbb{P} \left[\sum_{c: (c,q) \in \hat{G}_{ps}} Y_{cps} = i \right] & \text{si } |\hat{G}_{ps}| > N_s \\ 0 & \text{sinon.} \end{cases} \quad (3.13)$$

Le reste de la présente section va porter sur le cas de figure où $|\hat{G}_{ps}| > N_s$, puisque la probabilité d'encombrement du rôle est triviale dans l'autre cas. Comme l'illustre la sommation (3.13), une seconde probabilité doit être évaluée pour toutes les tailles de sous-ensembles suffisamment grandes pour excéder le nombre de salles. Prenons en considération un ensemble $\hat{G}_{ps}$ contenant cinq couplets de causes-périodes. Si $N_s = 3$, il faut donc évaluer successivement la probabilité qu'il y ait quatre, puis cinq causes simultanément. Pour chacune de ces tailles, un sous-ensemble U des couples de $\hat{G}_{ps}$ représenteront exclusivement les couples comprenant des causes en cours avec une probabilité équivalant au produit des probabilités individuelles des causes liées aux couples appartenant à U par l'inverse des produits des autres causes issues des couples de $\hat{G}_{ps}$, soit, formellement :

$$\mathbb{P} \left[\sum_{c: (c,q) \in \hat{G}_{ps}} Y_{cps} = i \right] = \sum_{U \subseteq \hat{G}_{ps} : |U|=i} \left(\prod_{c: (c,q) \in U} \pi_{cp} \right) \left(\prod_{c: (c,q) \in \hat{G}_{ps} \setminus U} (1 - \pi_{cp}) \right). \quad (3.14)$$

D'où, pour le cas de figure $|\hat{G}_{ps}| > N_s$, de façon générale :

$$\mathbb{P} \left[\sum_{c: (c,q) \in \hat{G}_{ps}} Y_{cps} > N_s \right] = \sum_{i=N_s+1}^{|\hat{G}_{ps}|} \sum_{U \subseteq \hat{G}_{ps} : |U|=i} \left(\prod_{c: (c,q) \in U} \pi_{cp} \right) \left(\prod_{c: (c,q) \in \hat{G}_{ps} \setminus U} (1 - \pi_{cp}) \right) \quad (3.15)$$

En d'autres mots, si on se réfère à notre exemple d'un ensemble de cinq couplets représentant des procès potentiels dans trois salles, l'un des cas de figure pour qu'il y ait débordement serait que quatre de ces procès aient lieu simultanément. Dans ce scénario, cela pourrait être tous les procès de l'ensemble sauf celui associé au troisième couplet de l'ensemble. La probabilité que cet évènement survienne est le produit de la probabilité que chacun des quatre procès individuellement aient lieu. Il faut toutefois ajouter un produit supplémentaire, par l'inverse de la probabilité que le procès du troisième couplet aie lieu, car le scénario évalué affirme que seuls les quatre autres causes procèdent. Puisque cette évaluation est faite sur des solutions entières, les probabilités π_{cp} propres à chacun des procès sont connues, même pour les causes régulières, puisqu'un seul $\hat{x}_{cps}$ pertinent peut avoir la valeur 1 à la fois. On

peut affirmer :

$$\pi_{cp} = \pi_{c(p-q+1)}^R \text{ si } \exists \hat{x}_{cqs} = 1,$$

ce qui permet une simplification d'une écriture déjà complexe. L'essentiel au sujet de l'équation (3.15) est d'en comprendre que le calcul de la probabilité d'encombrement du rôle peut se faire par une sommation de produits des probabilités individuelles des différentes causes impliquées. Par la méthode exacte, on peut obtenir une évaluation de cette probabilité qui sera aussi valable que les probabilités π_{cp} sur lesquelles elle se fondent, ce qui n'est pas négligeable.

Or, il faut parvenir à effectuer cette somme de produits de probabilités. Si, par convention, n est utilisé pour désigner la valeur de $(N_s + 1)$ et m pour désigner la taille de l'ensemble $\hat{G}_{ps}$, on peut évaluer les temps de calcul nécessaires. En soi, la première sommation représente un temps de $O(m - n)$. Selon la taille des sous-ensembles U par rapport à $|\hat{G}_{ps}|$, le nombre de ces sous-ensembles est variable, mais sera toujours de $O\left(\frac{m!}{(m-n)!n!}\right)$. Enfin, le nombre de multiplications au sein de la sommation du membre droit de l'équation (3.14) est constant et de $O(m - 1)$. Par ailleurs, comme ne sont considérés que les cas où $|\hat{G}_{ps}| > N_s$, les expressions $(m - n)$ et $(m - 1)$ sont toutes deux de $O(m)$, d'où, pour l'évaluation de la probabilité d'encombrement du rôle avec la méthode exacte, des temps de calculs de $O\left(\frac{m^2 m!}{(m-n)!n!}\right)$. La présence de factorielle et d'un terme au carré tend à occasionner une explosion combinatoire sur les temps de calculs associés à cette méthode, ce qui n'est pas, non plus, négligeable.

3.2.2 Méthode approchée

Comme mentionné à la section 3.2.1, la méthode exacte implique un grand nombre de calculs dès que le nombre de causes liées à la probabilité se met à augmenter. D'où la pertinence de procéder non pas à l'évaluation exacte de cette probabilité, mais plutôt à son approximation. Procéder de la sorte peut permettre une économie de calcul s'il est possible de rapprocher le comportement de la probabilité (3.10) d'une distribution connue, pour autant, bien sûr, que cette dernière possède des propriétés appropriées. À cette fin, les variables Y_{cps} , permettent de décrire un scénario potentiel une fois l'incertitude levée. En effet, la cause c sera en cours dans une salle de type s durant la période p si $Y_{cps} = 1$. La probabilité que cet événement survienne est calée sur la distribution de la durée potentielle de la cause c et sur la période de début p qui provient d'une solution donnée. Ainsi décrites, les variables Y_{cps} forment un ensemble de variables indépendantes suivant chacune une distribution de Bernoulli. Le théorème central limite généralisé sous la condition de Lyapunov, qui est, entre autre, rapporté par Billingsley (2012), stipule que la distribution d'une somme telle celle comprise dans la probabilité (3.10) va tendre vers la loi normale si la valeur de $|\hat{G}_{ps}|$ est substantielle et ce, même si les variables Y_{cps} n'ont pas les mêmes probabilités. La distribution

résultante étant une gaussienne de la forme suivante :

$$\mathcal{N} \left(\sum_{c: (c,q) \in \hat{G}_{ps}} \pi_{cp}, \sum_{c: (c,q) \in \hat{G}_{ps}} \pi_{cp}(1 - \pi_{cp}) \right). \quad (3.16)$$

La principale force de la méthode approchée est la rapidité des calculs qu'elle implique, l'aire sous une gaussienne pouvant se calculer ou s'interpoler à partir de la moyenne et de l'écart-type. Pour reprendre l'évaluation faite à la section 3.2.1, cette méthode implique des calculs de $O(m)$. Elle est donc à l'abri de l'explosion combinatoire affligeant la méthode exacte.

Cependant, les gains en temps de calcul sont réalisés au prix de la perte de l'exactitude. La méthode approchée n'est qu'une heuristique dont l'imprécision diminue avec l'augmentation de la cardinalité de $\hat{G}_{ps}$. Lorsque celle-ci est faible, l'approximation par la loi normale retourne une probabilité qui peut être passablement différente de la valeur exacte.

3.2.3 Méthode simulée

La méthode simulée qui est proposée cherche à obtenir une approximation dont la qualité est indépendante de l'exemplaire sur lequel la probabilité doit être calculée. En fait, plutôt que de tenter de calculer ou d'approcher la probabilité (3.7), cette méthode propose d'effectuer une simulation de Monte Carlo afin d'obtenir des statistiques pour estimer les probabilités d'encombrement du rôle. En se fondant sur la loi des grands nombres, les caractéristiques se dégageant des tirages effectués sont considérées comme tendant vers les caractéristiques de la distribution simulée si le nombre de tirages est suffisant. Comme l'explique Mooney (1997), ce type de simulation peut se baser sur des variables de Bernoulli. En utilisant la probabilité π_{cp} associée à aux variables Y_{cps} , une valeur pseudo-aléatoire est assignée à celles-ci, ce qui permet de déterminer si la somme de leurs valeurs est supérieure ou non à N_s . Cette procédure est répétée pour un nombre γ de tirages de manière à dégager une proportion de cas pouvant servir d'estimé de la probabilité de débordement des salles. Plus la valeur du paramètre γ est grande, plus l'estimation produite par la simulation de Monte Carlo tend à s'approcher de la probabilité réelle qui est approximée, sous réserve de l'absence de biais dans la génération des valeurs pseudo-aléatoires.

La méthode simulée permet ainsi de produire une approximation dont le degré de validité n'est pas lié au nombre de causes impliquées dans la probabilité à déterminer. Par conséquent, cette méthode est plus versatile que celle décrite à la section 3.2.2. Au niveau des temps de calculs, ils sont avec cette méthode de $O(\gamma m)$.

La simulation de Monte Carlo permet certes de connaître le comportement d'une proba-

bilité complexe, mais elle ne peut que l'estimer. En comparaison avec la méthode d'approximation proposée à la section 3.2.2, l'adaptabilité de cette simulation est acquise au prix d'un temps de calcul plus long.

3.2.4 Synthèse

Le choix de la méthode de calcul des probabilités dépend essentiellement de deux critères. Dans un premier temps, la méthode retenue doit être précise. Il s'agit de déterminer si une solution viole une contrainte probabiliste et, le cas échéant, un ensemble $\overline{G}_{ps}$. Une méthode imprécise pourrait retenir comme réalisable une solution ne respectant pas les contraintes, ou encore rejeter une solution les satisfaisant. Puis, comme ce test peut être effectué jusqu'à $|P||S|$ fois sur chaque solution entière retournée par l'énumération implicite, il ne doit pas recourir à un trop grand nombre de calculs, ce qui plomberait la performance générale de la résolution. Le tableau 3.1 permet de comparer en un clin d'oeil les caractéristiques des différentes méthodes en lien avec ces critères.

En supposant un nombre de causes suffisant, la méthode approchée se démarque, car ses temps de calculs sont les plus faibles. Sous la réserve de cette hypothèse, elle offre une approximation convenable, ce qui est également le cas de la méthode simulée, mais cette dernière est plus lente. Cependant, si le nombre de causes est plutôt faible, un autre raisonnement prévaut. Pour un petit nombre de causes, l'explosion combinatoire liée à la résolution de la méthode exacte est plus modeste, ce qui rend moins prohibitif de se livrer à un calcul exact. En comparaison, la méthode approchée peine à fournir une approximation convenable. Par contre, la méthode simulée reste concurrentielle, sa précision étant indépendante du nombre de causes. Même avec un grand nombre de tirages, son temps de calcul peut déclasser celui de la méthode exacte. Le critère d'efficacité peut cependant demeurer pertinent, car une seule méthode est réellement exacte. Cette garantie échappe à la méthode simulée, même si avec un nombre faramineux de tirages, elle peut en venir à une précision d'une finesse très élevée.

Tableau 3.1 Comparatif des méthodes d'évaluation de la probabilité de débordement des salles d'audience

Méthode	Exacte	Approchée	Simulée
Exactitude	oui	non	non
Ordre de calcul	$O\left(\frac{ \hat{G}_{ps} ^2 \hat{G}_{ps} !}{(\hat{G}_{ps} - N_s - 1)!(N_s + 1)!}\right)$	$O(\hat{G}_{ps})$	$O(\gamma \hat{G}_{ps})$

3.3 Méthode de résolution

3.3.1 Énumération implicite

Pour résoudre le problème stochastique d'horaire d'audiences judiciaires, la méthode proposée incorpore des éléments de l'ensemble de ce chapitre. Il s'agit d'une énumération implicite avec plans coupants, soit les coupes (3.8). Ces dernières étant fort nombreuses, la résolution démarre l'énumération implicite en utilisant les contraintes (3.3) assorties d'une grande valeur pour β . Une faible valeur du paramètre β produira des solutions dont la surplanification ne violera pas les contraintes probabilistes, une assurance qui se paie par une minimisation trop conservatrice. Inversement, les contraintes (3.3) comportant une valeur de β excessivement grande peuvent être respectées par des solutions pourtant très loin de satisfaire les contraintes (3.7). Ainsi, la résolution tire profit de la formulation avec durée fixe et surcapacité comme de la formulation probabiliste.

De façon plus concrète, la résolution du modèle (3.1)-(3.5) par énumération implicite est effectuée par un solveur commercial, en l'occurrence *CPLEX* utilisant la gamme de paramètres par défaut. Après la résolution de la relaxation linéaire associée à un sommet de l'arbre de branchement, la méthode évalue si la solution obtenue respecte les contraintes de binarité. Si la solution est fractionnaire et que la borne inférieure au sommet ne permet d'élaguer cette branche, deux nouveaux sommets sont créés par branchement sur la variable binaire dont la valeur dans la relaxation linéaire est la plus proche de 0.5. Une valeur respective de 0 et de 1 est imposée à la variable dans chacun des sommets-fils ainsi formés. À chaque solution entière rencontrée en cours ou au terme de la résolution de la relaxation linéaire, l'une des méthodes d'évaluation de la probabilité d'encombrement du rôle décrite à la section 3.2 est employée pour vérifier le respect des contraintes probabilistes. Ce choix se fonde, selon la période p et le type de salle s , sur le nombre de causes composant le sous-ensemble $\hat{G}_{ps}$ correspondant. Si l'évaluation conclut à la violation d'une contrainte (3.7), la coupe (3.20) correspondante est ajoutée à l'ensemble de l'arbre de branchement, mais seulement s'il s'agit de la solution optimale de la relaxation linéaire du sommet courant. Autrement, la solution entière est déclarée non réalisable et la résolution de la relaxation linéaire se poursuit sans heurts.

3.3.2 Insertion des plans coupants

La section 3.2 a passé en revue les trois méthodes d'évaluation du membre de gauche des contraintes (3.7). Elles sont toutes utilisées dans la résolution. La méthode exacte fournit, comme son nom l'indique, l'évaluation la plus précise de la probabilité d'encombrement du rôle, mais sur un ensemble $\hat{G}_{cps}$ de grande taille, elle est sujette à l'explosion combinatoire de ses temps de calcul. C'est alors que la méthode approchée, forte de son temps de calcul

réduits, peut entrer en scène et jouer un rôle complémentaire, alors sa précision laisse à désirer si la taille de $\hat{G}_{ps}$ n'est pas substantielle. Le recours à la méthode simulée permet de faire le pont pour une valeur de $|\hat{G}_{ps}|$ hors de la zone de confort des deux méthodes précédentes. Deux seuils paramétriques (σ_1 et σ_2) sont utilisés pour déterminer la méthode qui sera utilisée pour procéder à l'évaluation. L'évaluation de la probabilité se fait par la méthode exacte si $|\hat{G}_{ps}| \leq \sigma_1$ et par la méthode approchée si $|\hat{G}_{ps}| \geq \sigma_2$. Pour les valeurs telles que $\sigma_1 \leq |\hat{G}_{ps}| \leq \sigma_2$, la méthode simulée est privilégiée. Comme sa précision est indépendante de $|\hat{G}_{ps}|$ mais qu'elle repose sur un temps de calcul supérieur à celui de la méthode approchée, la méthode simulée permet de faire le relai entre les deux seuils.

Peu importe la méthode utilisée, lorsque la probabilité d'encombrement du rôle excède $1 - \alpha$, pour une salle de type s et une période p , la solution entière $\hat{x}$ n'est pas réalisable au sens des contraintes (3.7). Le rôle des coupes sera de séparer ces solutions non réalisables. De l'ensemble $\hat{G}_{ps}$, on peut ne retenir que les couples impliquant des causes régulières, qu'on exprime

$$\hat{G}_{ps}^R = \left\{ (c, q) \in \hat{G}_{ps} \mid c \in C^R \right\}, \quad (3.17)$$

duquel on pourra dégager la coupe suivante pour séparer $\hat{x}$:

$$\sum_{c: (c,q) \in \hat{G}_{ps}^R} x_{cqs} \leq |\hat{G}_{ps}^R| - 1. \quad (3.18)$$

Toutefois, la coupe obtenue peut être renforcée s'il existe un ensemble $\overline{G}_{ps}$ tel que :

$$|\overline{G}_{ps} \cap \hat{G}_{ps}^R| < |\hat{G}_{ps}^R|.$$

C'est pourquoi la coupe sera resserrée sur un sous-ensemble $\hat{G}'_{ps} \subseteq \hat{G}_{ps}^R$ de cardinalité minimum. Ce sous-ensemble appartient lui aussi à Γ_{ps} et se retrouve en classant les couples (c, q) selon la valeur de $P[Y_{cps} = 1]$ pour la variable de Bernoulli appropriée. La coupe reprendra le plus petit sous-ensemble de $\hat{G}_{ps}^R$ tel que la solution entière sera séparée et qu'aucun autre sous-ensemble de même taille ne permet cette séparation. Ainsi, l'ajout successif de coupes redondantes entre elles sera évité.

Par définition, les variables Y_{cps} prennent la valeur 1 si la cause c procède à la période p dans une salle de type s . Les procès s'ouvrent et se poursuivent jusqu'à terme de façon continue sur un certain nombre de périodes. L'absence d'interruption dans les procès permet d'effectuer le lien suivant entre les variables Y_{cps} :

$$P[Y_{cps} = 1] > 0 \implies P[Y_{cps} = 1] \geq P[Y_{c(p+1)s} = 1] \quad (3.19)$$

La relation (3.19) s'explique par le fait que si la cause c procède à la période $p + 1$ sans avoir débuté à cette même période, le procès était forcément en salle à la période p . L'inverse n'est cependant pas vrai, car une cause peut procéder à la période p , s'y terminer et ne pas utiliser de locaux à la période suivante. Dans le cadre des coupes, cette relation présente un intérêt. Pour une solution entière, une seule variable x_{cqs} aura influencé $P[Y_{cps} = 1]$ pour $c \in C$. Selon la relation (3.19), si le choix de la période q amène la cause c à contribuer à l'encombrement du rôle, toutes les autres périodes de début valides de la cause c situées entre q et p aurait provoqué, *ceteris paribus*, un encombrement du rôle. On peut donc les inclure de proposer la coupe suivante :

$$\sum_{c: (c,q) \in \hat{G}'_{ps}} \sum_{t \in P_{cp} \cap \{q, q+1, \dots, p\}} x_{cts} \leq |\hat{G}'_{ps}| - 1 \quad (3.20)$$

3.3.3 Algorithme

L'algorithme 1 présente de façon plus concise le fonctionnement de la méthode de résolution proposée. Les étapes initiales n'ont pas de portée spécifique et décrivent la logique de résolution d'une énumération implicite effectuée selon les paramètres par défaut de *CPLEX*. L'intérêt de présenter cet algorithme réside davantage dans les étapes 8 et suivantes, qui concernent la procédure d'évaluation de la probabilité d'encombrement du rôle et de la génération des coupes. C'est par ces étapes qu'on peut passer d'une résolution à surcapacité à une solution respectant les contraintes probabilistes. De fait, comme on peut le voir à l'étape 20, le classement d'une solution comme (meilleure) solution réalisable ne se fait que sur la relaxation linéaire d'un sommet où le seuil de tolérance $1 - \alpha$ est respecté. Inversement, une solution entière réalisable selon la formulation stricte du modèle avec durée fixe et surcapacité peut être écartée.

Comme mentionné à la section 3.3.1, la résolution des relaxations linéaires et la gestion de l'arbre de branchement est effectuée par *CPLEX* et s'effectue selon les paramètres par défaut. Lorsqu'une violation d'une contrainte (3.7) est constatée, l'insertion d'une coupe (3.20) est effectuée pour l'ensemble de l'arbre de branchement, ce qui implique que la coupe est retenue pour la résolution de l'ensemble des relaxations linéaires subséquentes.

Algorithme 1 Méthode de résolution

- 1: Transmission à *CPLEX* de la formulation avec durée fixe et surcapacité du problème stochastique d'horaire d'audiences judiciaires.
 - 2: Inscription du sommet initial de l'arbre d'énumération dans la liste des sommets à résoudre.
 - 3: **tant que** les sommets à résoudre ont une borne inférieure sur le sommet parent permettant d'améliorer la meilleure solution réalisable **faire**
 - 4: Considérer le prochain sommet dans la liste des sommets à résoudre
 - 5: Résolution par *CPLEX* de la relaxation linéaire.
 - 6: **si** la solution est entière **alors**
 - 7: **pour tout** $p \in P, s \in S$ **faire**
 - 8: Évaluer la probabilité d'encombrement du rôle.
 - 9: **si** $|\hat{G}_{ps}| < \sigma_1$ **alors**
 - 10: Calcul des produits de probabilités selon l'équation (3.15).
 - 11: **sinon**
 - 12: **si** $|\hat{G}_{ps}| < \sigma_2$ **alors**
 - 13: Pour γ tirages, attribuer des valeurs pseudo-aléatoires aux variables Y_{cps} pertinente.
 - 14: Approcher la probabilité d'encombrement du rôle par les statistiques issues de la méthode simulée utilisant γ tirages.
 - 15: **sinon**
 - 16: Approximation de la probabilité d'encombrement du rôle par la gaussienne (3.16).
 - 17: **si** $P \left[\sum_{c: (c,q) \in G_{ps}} Y_{cps} > N_s \right] > (1 - \alpha)$ et la solution courante est issue de la fin de la relaxation linéaire **alors**
 - 18: Déterminer une coupe (3.20) de cardinalité minimale et l'appliquer au modèle courant.
 - 19: **sinon**
 - 20: **si** la valeur de la solution courante est inférieure à celle de la meilleure solution réalisable **alors**
 - 21: Classer la solution obtenue comme meilleure solution réalisable
 - 22: Classer la solution obtenue comme solution réalisable.
 - 23: **ARRET.** Terminer la résolution et retourner la meilleure solution réalisable comme solution optimale.
-

CHAPITRE 4

EXPÉRIMENTATION ET RÉSULTATS

Le problème stochastique d'horaire d'audiences judiciaires a été décrit et apparenté aux autres problèmes d'horaire. En s'inspirant des efforts et des avancées de la communauté scientifique en optimisation, une méthode de résolution a été élaborée et dans le présent chapitre, sa rigueur sera éprouvée. En s'appuyant sur un transfert de connaissance en provenance des institutions judiciaires de Montréal, des jeux de données s'apparentant aux conditions réelles de planification des procès ont pu être élaborés. De même, afin de servir d'étalon comparatif, le processus séquentiel qui a traditionnellement cours a été modélisé. Les résultats de cette expérimentation sont présentés et discutés au terme de ce chapitre, précédés de résultats intermédiaires.

4.1 Données utilisées

D'un point de vue strictement mathématique, la pertinence d'une résolution probabiliste du problème stochastique d'horaire d'audiences judiciaires n'est pas liée au type d'exemplaires utilisé. Or, si une perspective d'application pratique est employée, il est aisé de constater que des résultats positifs sur la méthode exposée à la section 3.3 seront d'autant plus probants si les exemplaires utilisés lors des tests se rapprochent des conditions réelles de planification judiciaire. La collaboration avec le service du maître des rôles du palais de justice de Montréal a permis d'obtenir des renseignements sur les causes civiles relevant la Cour supérieure du Québec dans le district de Montréal. Les sections subséquentes présentent l'analyse des données ainsi obtenues et présente comment elles sont utilisées pour générer des jeux de données qui partagent certaines caractéristiques.

4.1.1 Portrait des données transmises par le palais de justice de Montréal

L'analyse a porté sur le rôle civil de quatre mois d'une même année. Ces mois, qui seront désignés comme A , B , C , et D ont été choisis en raison de leur caractère représentatif des variations du niveau d'activité de la Cour supérieure tout au long du calendrier judiciaire. La figure 4.1 permet de constater que cette fluctuation est présente autant en termes de nombre de causes individuelles prévues au rôle qu'en terme de jours-causes en procès. La plupart des causes civiles ne requiert pas d'équipement particulier, ce qui entraîne une faible différenciation entre les différentes salles d'audience. Cependant, les salles sont partagées entre

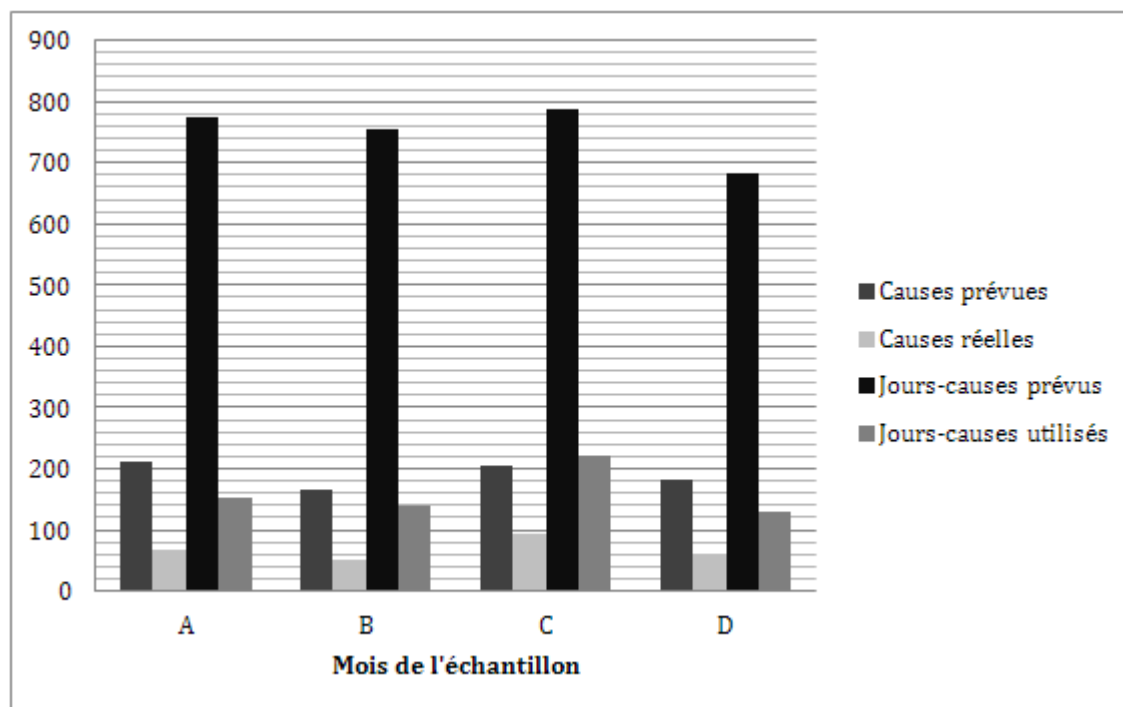


Figure 4.1 Comparaison du volume judiciaire

un bassin attribué en premier lieu à la Cour supérieure et un autre relevant au premier chef de la Cour du Québec. Toutefois, des prêts de salles sont pratique courante pour absorber les aléas de surplanification, comme l'illustre la figure 4.2. Les vingt-quatre salles du bassin de la Cour supérieure furent mises a profit pour la majorité des besoins de la Cour, mais pour l'ensemble des mois de l'échantillon, le recours aux salles de la Cour du Québec fut nécessaire couvrir les besoins. La principale matière du litige est relativement variée dans l'échantillon reçu, comme la figure 4.3 permet de le constater. Il convient de souligner que le type, en quelque sorte, de la cause, n'a généralement aucun impact sur la priorité du procès en découlant au rôle. Les poursuites liées à la réclamation de dommages ainsi qu'aux contrats représentent cumulativement un peu plus de la moitié des causes de l'échantillon. Les procès pour divorce et concernant les comptes sont également bien représentés, avec environ un dixième de l'échantillon. Hormis les injonctions, toutes les autres matières comptent pour moins de cinq pour cent de l'échantillonnage reçu.

Comme cela a été mentionné à la section 1.2, un élément capital de l'intérêt mathématique du problème stochastique d'horaire d'audiences judiciaires est le caractère incertain de la durée des causes. En pratique, au moment de l'inscription au rôle, les causes comportent une durée prévue (voir section 1.1.2). La figure 4.4 permet de constater que la proportion des causes qui ont requis une salle d'audience pour la durée prévue est somme toute faible. L'axe

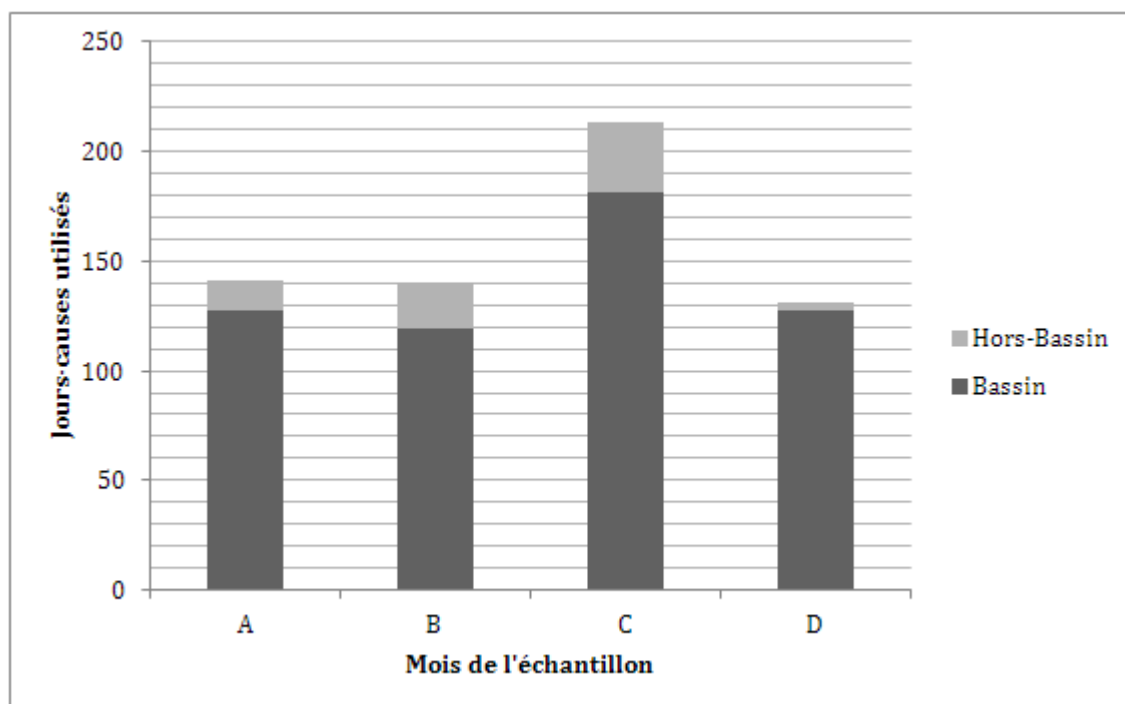


Figure 4.2 Origine des salles d'audience utilisées

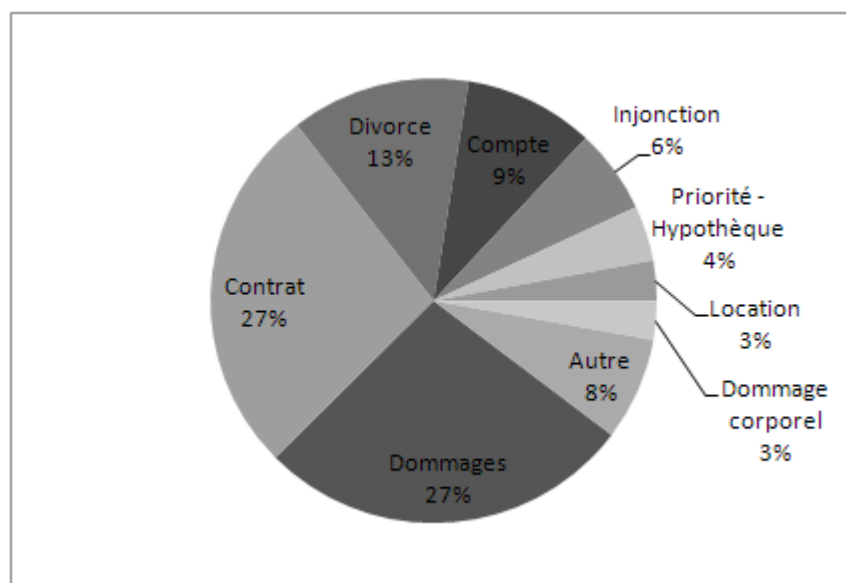


Figure 4.3 Typologie des causes

horizontal y regroupe les causes selon le nombre de jours prévus pour l'audition sur le rôle. La moitié des causes ayant une durée prévue d'une seule journée ont requis une salle, mais la proportion chute à 15 % pour les causes devant durer deux jours, puis fléchit légèrement de deux points de pourcentage pour les causes devant durer trois jours, et enfin fond de façon continue sur les autres intervalles pour atteindre un maigre 3 % pour les causes ayant une durée prévue supérieure à sept jours. Le cas de figure de loin le plus fréquent est celui des causes n'ayant finalement pas eu besoin d'une salle d'audience. Ces causes à durée nulle représentent au moins la moitié du nombre total de causes dans chacune des classes de durées proposées. En fait, cette proportion dépasse même les trois quarts dans le cas des causes ayant une durée prévue de six à sept jours. Cela permet de constater que d'inscrire une cause sur le rôle n'implique pas l'utilisation d'une salle à tous les coups.

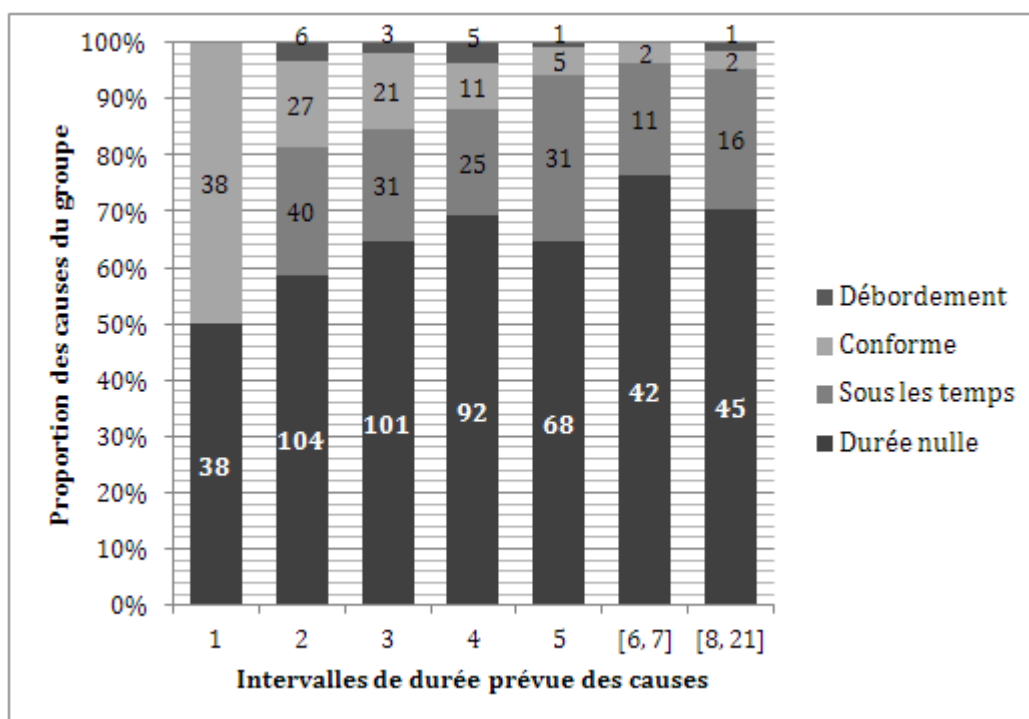


Figure 4.4 Exactitude de la prévision de la durée

Le rôle permet de s'informer sur l'issue des différentes causes. Une synthèse de ces informations est fournie de façon visuelle dans la figure 4.5. La catégorie la plus importante, « accord », regroupe les situations pour lesquelles les parties en causes sont parvenues à un accord avant ou pendant la durée du procès. Au contraire, l'étiquette « procès » caractérise les causes pour lesquelles une décision du juge a été rendue au terme du procès. La troisième catégorie en importance est celle des remises, qui comprend les causes pour lesquelles une nouvelle audience sera fixée ultérieurement, pour diverses raisons. Les retraits peuvent

se distinguer des rejets en ce qu'ils constituent une rétractation de l'une des parties, alors que le rejet a également pour effet de faire cesser les procédures, il émane d'une décision judiciaire, comme par exemple une requête en rejet d'action. Le grand nombre d'accords et de remises fournit une piste d'explication pour les causes n'ayant pas finalement nécessité de salle d'audience.

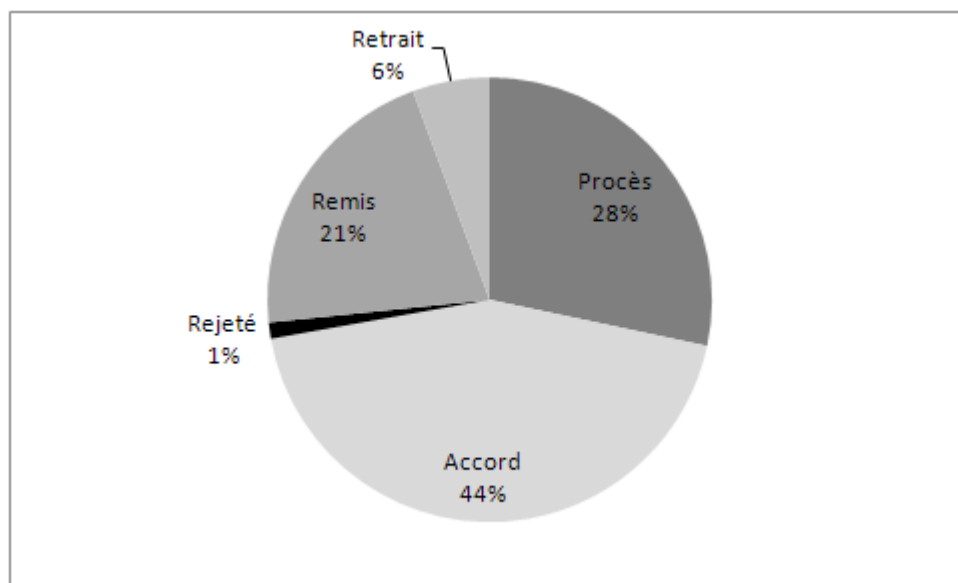


Figure 4.5 Proportion comparée des issues des causes

4.1.2 Construction des jeux de données pseudo-réalistes

Le principal intérêt de détenir un échantillon de causes ayant figuré sur un véritable rôle est de pouvoir en dégager de l'information permettant de formuler des distributions. Avec ces dernières, il est possible de constituer à volonté des jeux de données réalistes au sens du partage des caractéristiques de l'échantillon fourni. Le nombre de causes reçues est trop faible pour permettre de raffiner le traitement selon les différents types d'éléments sur lesquels portent les causes. Or, une analyse peut être faite au niveau global entre la durée prévue des différentes causes et la durée réelle qui fut constatée au rôle. La relation entre la durée prévue des causes et le nombre de jours où une salle d'audience fut réellement occupée est compilée dans le tableau 4.1. En conformité avec les proportions illustrées par la figure 4.4, les causes à durée nulle représente la plus grande partie des 766 causes de l'échantillon. Il est également possible de constater une borne supérieure relative sur la durée des causes. En effet, si certaines causes ont requis une salle d'audience pour un nombre de jours excédant leur durée prévue, en aucun cas cette utilisation supplémentaire n'a excédé un jour.

Tableau 4.1 Nombre de jours-salles utilisées par les causes selon la durée prévue au rôle

Durée Prévue	Réelle												Total
	0	1	2	3	4	5	6	7	10	11	12	20	
1	38	38											76
2	104	40	27	6									177
3	101	15	16	21	3								156
4	92	15	6	4	11	5							133
5	68	10	7	8	6	5	1						105
6	24	3		2	3	1	2						35
7	18	1			1								20
8	18	1			1	1		1					13
9	6	1				3							10
10	8	1	1					1	2				13
11	2										1		3
12	1												1
13	1									1			2
14	1			1		1							3
15	2												2
17	1												1
18	2												2
20	2						2						3
21	1											1	2
Total	490	125	57	42	25	16	4	2	2	1	1	1	766

Les deux caractéristiques évoquées au paragraphe précédent permettent de construire la distribution des durées possibles d'une cause selon sa durée prévue. Dans un premier temps, on peut diviser la distribution en durée nulle et durée non nulle. La figure 4.6 compare, selon la durée initialement prévue, le nombre de causes ayant eu une durée nulle à celles ayant procédé. Les causes d'une durée prévue d'un seul jour se démarquent des autres par une parité au niveau de la durée réelle nulle et non nulle. De leur côté, pour les autres durées prévues, le ratio s'élève environ à 2 : 1 en faveur des causes n'ayant finalement pas procédé. On peut remarquer que, pour les durées prévues supérieures à six jours, les proportions sont plus variables, largement en raison du bassin plus restreints. Dans la figure 4.7, la proportion des causes ayant excédée leur durée prévue est adossée à un taux uniforme de 2 %. Ce taux rend compte avec une précision acceptable de la situation là où il y a bel et bien lieu dépassement. La principale divergence provient des trois causes ayant une durée prévue de onze jours, dont une donna lieu à un procès de douze jours.

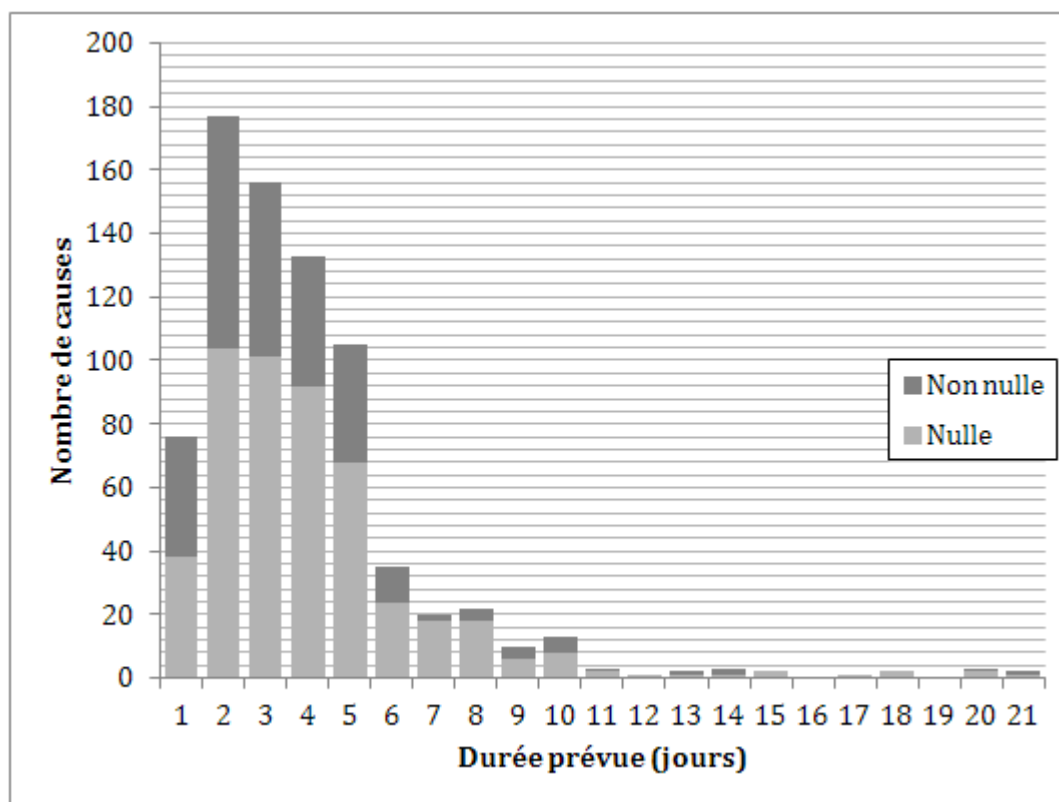


Figure 4.6 Proportion comparée des causes d'une durée réelle nulle

Les durées possibles sont ainsi caractérisées en nulle, non nulle et non nulle excédant la durée prévue. En raison du faible nombre de causes pour plusieurs durées prévues, il n'a pas été possible d'identifier une distribution cohérente. Par ailleurs, la figure 4.8 illustre

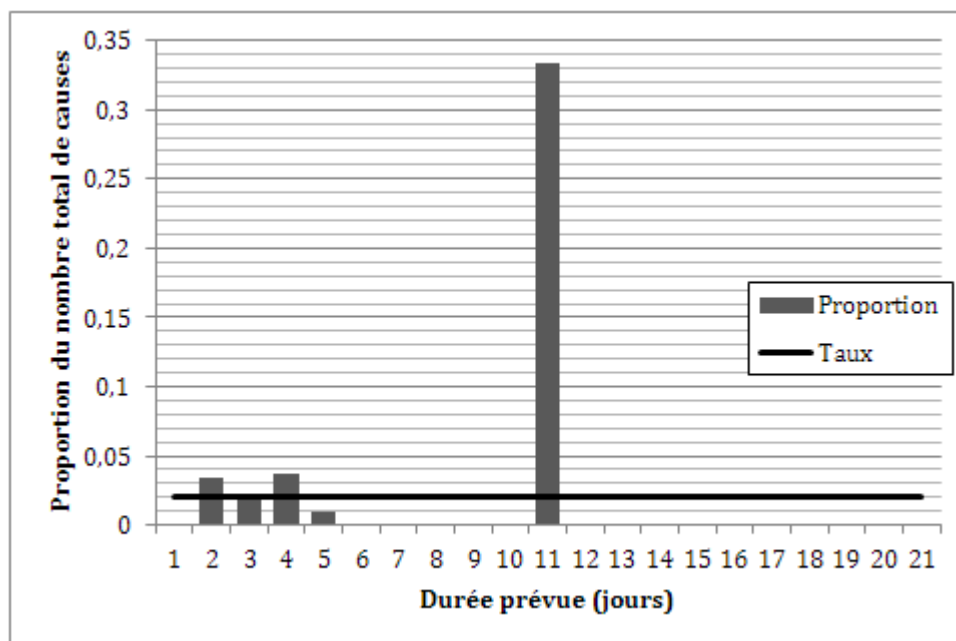


Figure 4.7 Proportion et taux proposé des causes ayant dépassé leur durée prévue

que la dispersion des causes dans l'échantillon est très proche d'une distribution aléatoire sur les durées possibles pour chaque durée prévue. Dans cette figure, la courbe illustrant la distribution aléatoire utilise des valeurs continues plutôt que discrètes par souci d'arriver à un nombre total de causes équivalent à celui de l'échantillon.

Après avoir établi comment, par une analyse de l'échantillon de données, construire une distribution de durées possibles à partir de la durée prévue d'une cause, il faut encore déterminer comment allouer à une cause une durée prévue réaliste. Comme en témoigne la figure 4.9, il est possible d'approximer la répartition des causes selon leur durée au sein de l'échantillon en utilisant la loi Gamma. Un exercice d'ajustement a permis d'identifier un paramètre de forme $\alpha = 2,0038$ et un paramètre d'intensité $\beta = 1,6429$, qui furent utilisés par la distribution illustrée ainsi que dans celle retenue pour construire les jeux de données pseudo-réalistes.

4.2 Outils logiciels

Sans les capacités de l'informatique moderne, la mise en œuvre des principes mathématiques développés en optimisation serait tout bonnement impossible. C'est pourquoi cette section présentera les différents logiciels et l'infrastructure impliqués dans le processus de mise à l'épreuve du modèle du problème stochastique d'horaire d'audiences judiciaires et de la résolution proposée à la section 3.3. Deux programmes de soutien ont été élaborés pour

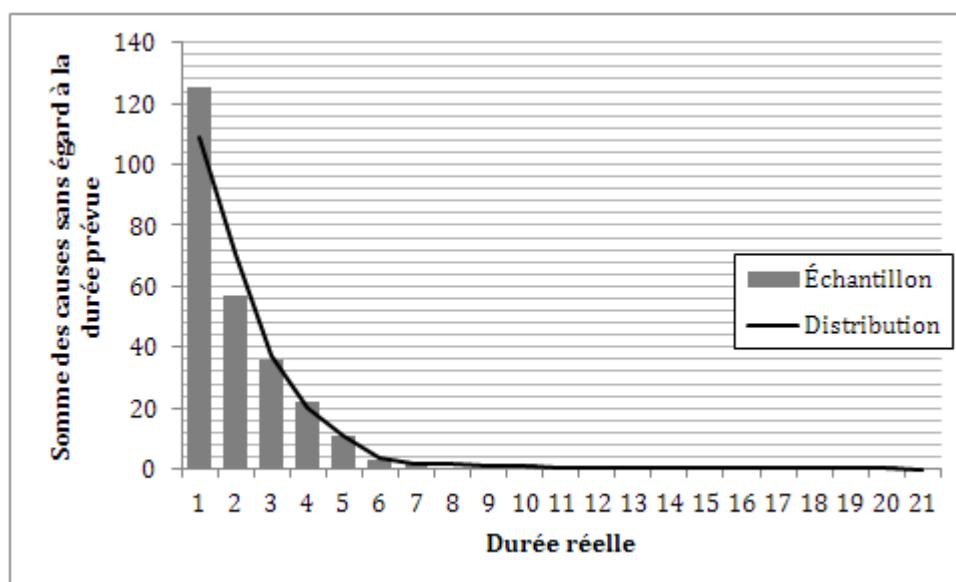


Figure 4.8 Comparatif des sommes des causes de l'échantillon de durée non-nulle sans excès de la durée prévue et de la distribution proposée

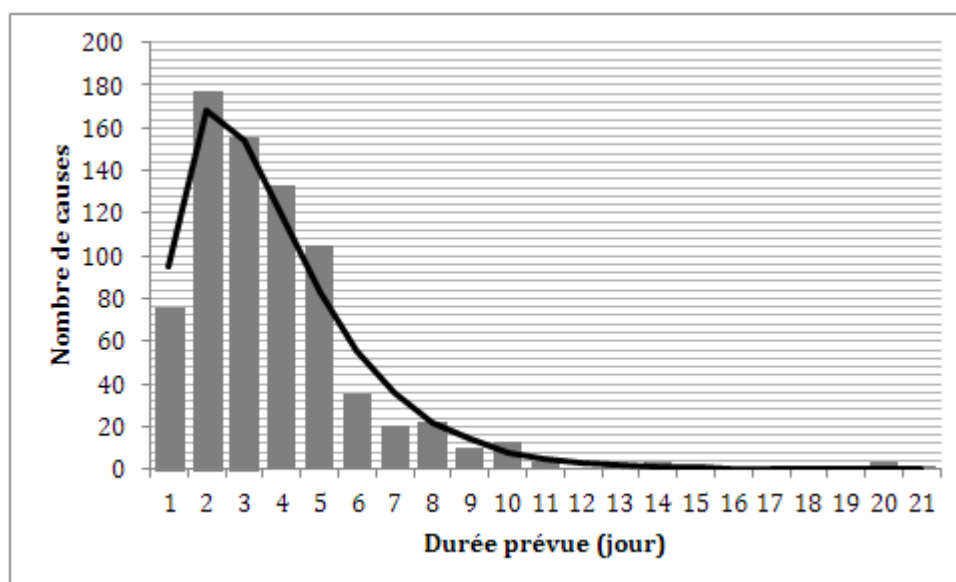


Figure 4.9 Comparaison entre la répartition des causes selon leur durée prévue et une distribution Gamma

permettre, respectivement, l'élaboration de jeux de données et offrir une base de comparaison tentant de rester fidèles aux us constatés dans le milieu judiciaire. Par contre, pour appliquer la méthode de résolution mathématique proposée, un logiciel a été développé et calibré pour répondre aux besoins spécifiques suscités.

4.2.1 Les programmes de soutien

Dans le but d'éviter les biais induits par l'édification d'une méthode de résolution fondée sur quelques exemplaires seulement, un générateur de données a été mis au point pour permettre de produire différents exemplaires respectant les caractéristiques générales présentées à la section 4.1.2. Ce générateur de données permet de doter de façon pseudo-aléatoire les causes de disponibilités, de conflits et d'un modèle de durée. Quant aux besoins en type de salle, la répartition est faite d'une façon pondérée par la proportion de salles des différents types, comme on peut le voir dans le tableau 4.2. Différents patrons de besoins sont présentés en autant de lignes. La dernière colonne identifie la proportion des causes qui sont générées avec le patron de cette ligne. La dernière ligne donne le nombre de salles de chaque type. Les trois types de salles d'audiences modélisés comportent deux types principaux, ainsi qu'un sous-type.

Tableau 4.2 Probabilités de génération des besoins de salles

Types	I	II	III	Proportion
	✓			0,22
	✓	✓	✓	0,30
	✓		✓	0,06
		✓		0,42
N_s	12	24	3	

Afin de fournir une base de comparaison permettant de juger de la performance de *Court-CausesScheduler*, un constructeur séquentiel d'horaires a été établi pour tenter d'émuler le processus traditionnel-empirique qui prévaut dans le milieu judiciaire. Les différentes causes sont énumérées et le constructeur tente de placer chacune d'entre elles à la date de départ la plus avantageuse. Une fois une assignation réalisée, le constructeur passe à la cause suivante et ne peut revenir en arrière. Le choix d'une période de départ est limité par un taux de surcharge. Si plusieurs types de salles sont permis, l'assignation séquentielle va privilégier le type de salle comportant le plus grand nombre de places restantes selon le taux de surcharge pour la durée complète de la cause. Si la limite de surcharge était enfreinte pour l'ensemble

des combinaisons de périodes et de type de salle permis, le constructeur indique que la cause n'est pas inscrite au rôle et lui assigne la même pénalité que *CourtCausesScheduler*. L'algorithme 2 reprend de façon plus systématique ce fonctionnement. Bien que cet algorithme reprenne la notion de durée fixe, les variables Y_{cps} y sont utilisées pour représenter si la cause c est en cours à la période p dans une salle de type s pour des raisons de simplicité. De plus, il faut noter que les valeurs de β utilisées dans les tests ne représentent pas nécessairement les taux de surcharge utilisés par le maître des rôles du palais de justice de Montréal lors de sa planification.

Algorithme 2 Assignment séquentielle

```

1: pour tout  $p \in P, s \in S, c \in C^A$  faire
2:   si  $c \in C_{ps}^A$  alors
3:      $Y_{cps} = 1$ 
4:   sinon
5:      $Y_{cps} = 0$ 
6: Conserver en mémoire la composition initiale des  $P_c$  et  $S_c$ .
7: pour tout  $c \in C^R, p \in P, s \in S$  faire
8:    $Y_{cps} = 0$ 
9: pour tout  $c \in C^R$  faire
10:  tant que  $\sum_{p \in P_c} \sum_{s \in S_c} x_{cps} + e_c = 0$  faire
11:    Considérer  $p$ , la période la plus tôt de  $P_c$ .
12:    Considérer  $s$ , le type de salle de  $S_c$  minimisant  $\beta N_s - \sum_{d \in C} Y_{dps}$ 
13:    Déclarer  $x_{cps} = 1$  et considérer la faisabilité.
14:    pour tout  $q \in P_{cp}$  faire
15:       $Y_{cqs} = 1$ 
16:      si  $\sum_{d \in C} Y_{dqs} > \beta N_s$  alors
17:        Retirer  $s$  de  $S_c$ .
18:         $x_{cps} = 0$ .
19:      si  $x_{cps} = 0$  alors
20:        pour tout  $q \in P_{cp}$  faire
21:           $Y_{cqs} = 0$ 
22:        si  $S_c = \emptyset$  alors
23:          Restaurer la composition initiale de  $S_c$ .
24:          Retirer  $p$  de  $P_c$ .
25:        si  $P_c = \emptyset$  alors
26:          Restaurer la composition initiale de  $P_c$ .
27:         $e_c = 1$ 
28: ARRÊT Terminer la résolution et retourner la solution courante.

```

Enfin, une simulation de Monte Carlo a été programmée pour évaluer la qualité des solutions produites lors des différents tests. Le nombre de tirages effectués peut être modifié

en tant que paramètre de la simulation.

4.2.2 *CourtCausesScheduler*

Le logiciel *CourtCausesScheduler* a été créé en 2011 afin de permettre la gestion de la résolution mathématique du problème stochastique d'horaire d'audiences judiciaires, particulièrement en ce qui a trait à l'insertion des plans coupants. La résolution brute est effectuée par un solveur commercial mais les trois méthodes d'évaluation de la probabilité de débordement des salles exposées à la section 3.2, la génération des coupes ainsi que la collecte de certaines statistiques sont effectuées par le logiciel. La résolution d'exemplaires par *CourtCausesScheduler* est effectuée par un ordinateur du Groupe d'étude et de recherche en analyse des décisions (GERAD) en utilisant un processeur Intel i7-2600 d'une puissance de 3,40 GHz et 15,56 Go de mémoire vive, ainsi que le solveur IBM ILOG CPLEX version 12.5.0.0.

4.3 Résultats expérimentaux

Deux séries de tests ont été réalisées avec *CourtCausesScheduler*. La première série a servi à observer l'impact du choix des méthodes d'évaluation des probabilités de débordement sur le comportement de la résolution, sa durée et la qualité de la solution obtenue. La seconde série, quant à elle, a démontré l'intérêt de l'approche mathématique sous-tendue par *CourtCausesScheduler* par rapport à la pratique traditionnelle. Les sections suivantes vont ainsi présenter les deux séries de tests, leurs résultats respectifs et la discussion générale qui s'en suit.

Pour tous les tests présentés, le seuil de tolérance au risque d'encombrement fut fixé à $1 - \alpha = 0,05$. Avec le calendrier judiciaire en vigueur dans le district de Montréal ce niveau de risque se traduira vraisemblablement par des débordements à quelques reprises dans l'année, mais la possibilité d'éviter l'encombrement du rôle par un prêt de salles issues du bassin de la Cour du Québec (voir section 4.1.1) permet de tolérer cette éventualité. Toutes les résolutions effectuées avec *CourtCausesScheduler* en utilisant la méthode simulée utilisent le même nombre de tirages. Fixée à $\gamma = 10000$, cette valeur a été choisie arbitrairement pour obtenir des statistiques au millième près. Les tests de calibrage ont révélé que cette valeur de γ permet une approximation très proche des valeurs obtenues avec la méthode exacte, sans pour autant provoquer des calculs prohibitifs.

4.3.1 Impact des méthodes d'évaluation des probabilités de débordement

Les premiers résultats présentés dans ce chapitre concerne l'impact des méthodes d'évaluation présentée à la section 3.2. Plus précisément, la section 3.2.4 présente des hypothèses

sur le comportement de ces méthodes, largement tributaire de l'impact de la taille des ensembles $\hat{G}_{ps}$ manipulés. Ceux-ci sont eux-même influencés par la valeur des différents N_s et par le taux de surcharge β utilisé durant la résolution.

Afin de permettre de vérifier les relations entre la taille des ensembles $\hat{G}_{ps}$ et le comportement des méthodes d'évaluation, des exemplaires ont été générés en variant la valeur des N_s . Au sein de chaque exemplaire, les trois types de salles ont reçu la même valeur de N_s . La génération pseudo-réaliste a été ajustée afin que les besoins en type de salles, dont découle la composition des ensembles S_c , reflète une proportion équitable résultant de cette décision. Le tableau 4.3 présente les sept valeurs de N_s retenues et les catégories d'exemplaires en découlant. La troisième colonne présente le nombre de causes régulières. Ce dernier provient d'essais de résolution visant à permettre une saturation du rôle sans rendre excessifs les temps de calculs. Les exemplaires de calibrage n'ont pas été utilisés dans les présents tests. Les causes préassignées ont été réduites au minimum, soit une seule par exemplaire, par volonté de laisser le plus de place possible à l'effet de la variation de N_s d'une catégorie à l'autre. La dernière colonne présente la plus grande valeur possible de $\hat{G}_{ps}$, selon un taux de surcharge $\beta = 2,2$. Cette valeur, utilisée pour les tests, fut fixée par la volonté d'avoir une faible surcharge, mais une qui soit suffisante pour permettre l'introduction de coupes même sur la catégorie $s3$.

Tableau 4.3 Catégorie d'exemplaires utilisés pour les tests d'évaluation de la probabilité d'encombrement du rôle

Nom	N_s	$ C^R $	$\lfloor \beta N_s \rfloor$
s3	3	50	6
s5	5	250	11
s10	10	575	22
s15	15	900	33
s20	20	1400	44
s25	25	1450	55
s30	30	2200	66

Le second objectif de cette série de tests est de déterminer une valeur pour les paramètres σ_1 et σ_2 décrits à la section 3.3.2. L'utilisation de ces paramètres permet de combiner les trois méthodes d'évaluation des probabilités à l'intérieur d'une même méthode générale de résolution. Cette hybridation va permettre, à la lumière des résultats obtenus, d'évaluer les probabilités d'encombrement du rôle à l'aide d'une méthode appropriée à la situation, accélérant la résolution générale.

Résultats comparés

Vingt-et-un exemplaires, soit trois pour chacune des catégories décrites en termes généraux à la section précédente, ont été résolues par *CourtCausesScheduler*. Tout d’abord, une seule des trois méthodes d’évaluation fut utilisée à la fois. Le détail des résultats ainsi obtenus est exposé dans le tableau 4.4. Ces résultats concernent ainsi trois résolutions pour chacun des trois exemplaires d’une même catégorie. Pour chaque ligne, le tableau présente le nombre de coupes ajoutées par le solveur *CPLEX*, la valeur de la meilleure solution obtenue et le saut d’intégrité associé à cette dernière, relativement à la valeur de la borne obtenue au sommet racine. Les colonnes suivantes indiquent le nombre de sommets énumérés au sein de l’arbre de branchement, puis le temps système écoulé de façon globale puis spécifiquement consacré à l’évaluation de la probabilité d’encombrement. Les trois dernières données présentées sont le nombre de solutions entières retournées par l’énumération implicite, le nombre d’entre elles satisfaisant les contraintes probabilistes et le nombre de coupes générées pour s’en assurer. Une seule exemplaire, *s5-2*, s’est révélé trop coriace pour l’ensemble des méthodes d’évaluation.

On peut également observer le comportement des différentes méthodes. Ainsi, la méthode exacte vient très rapidement à bout des quatre premiers exemplaires, mais sa rapidité fléchit dans les cas comportant cinq ou dix salles de chaque type. Sur cette dernière catégorie d’exemplaires, les deux autres méthodes prennent clairement le dessus. À partir des exemplaires *s15*, les temps d’évaluation des solutions entières par énumération des probabilités deviennent tels que la progression de l’énumération implicite n’arrive plus à se faire convenablement. C’est d’ailleurs pour cette raison que les résultats des tentatives de résolution ne sont pas inclus pour les exemplaires suivants dans le tableau 4.4.

Les résultats sur les premières catégories d’exemplaires montrent que la méthode simulée (avec $\gamma = 10000$) vise juste lorsqu’on compare la valeur des solutions qu’elle retourne à celles obtenues par la méthode exacte. L’unique divergence dans les valeurs des solutions survient sur l’exemplaire *s5-2*, qui n’a pas été résolu à optimalité. Tel que mentionné ci-haut, les temps de calcul démarquent les deux méthodes. La simulation de Monte-Carlo est laissée sur place par l’énumération sur de petits ensembles, mais, exempte d’explosion combinatoire, elle s’impose à partir des exemplaires à dix salles par type.

Enfin, quelques mots pour décrire la performance de la méthode approchée. La qualité de son approximation, si on la compare avec celle de la méthode simulée, s’améliore grandement sur les exemplaires plus grands. Cependant, elle conserve une volatilité dans la valeur comparée de sa solution, soit sur les exemplaires *s15-2*, *s20-2*, *s30-1* et *s30-2*. On peut également souligner que la méthode approchée s’enlise et ne parvient pas à résoudre cinq des petits exemplaires même si, à l’exception de *s5-1* et *s5-2*, les temps de résolution des deux autres

Tableau 4.4 Résultats détaillés de la résolution à une seule méthode

Exemplaire	Méthode ^a	Coupes <i>CPLEX</i>	Valeur solution	Saut d'intégrité (10^{-3})	Sommets énumérés	Durée (s)	Temps coupes (s)	Solutions entières	Solutions retenues	Coupes générées
s3-0	E	2	442	2,27	4	0,06	0,01	3	2	1
s3-0	S	2	442	2,27	4	0,09	0,06	3	2	1
s3-0 ^b	A	2	451	22,68	43703156	18000,00				10033
s3-1	E	0	464	10,89	364	0,17	0,10	193	3	1
s3-1	S	0	464	10,89	364	2,70	2,61	190	3	1
s3-1	A	0	468	19,61	2340	1,04	0,33	1436	2	4
s3-2	E	0	496	0,00	4442	1,95	0,71	2613	4	4
s3-2	S	0	496	0,00	4643	7,25	5,76	2731	4	4
s3-2 ^b	A	0	504	16,13	14244198	18000,00				28829
s5-0	E	6	3653	0,79	380	2,44	0,19	23	8	3
s5-0	S	6	3653	0,79	380	4,51	2,27	23	8	3
s5-0 ^b	A	6	3660	2,70	17839556	18000,00				25668
s5-1	E	9	3648	1,42	261095	799,94	414,65	178954	3	224
s5-1	S	9	3648	1,42	329837	4863,15	4380,35	223691	3	206
s5-1 ^b	A	9	3651	2,24	29325351	18000,00				5897
s5-2 ^b	E	7	3844	1,66	3384300	18000,00				306
s5-2 ^b	S	9	3845	2,03	135549	18000,00				324
s5-2 ^b	A	9	3858	5,31	7687777	18000,00				9577
s10-0	E	1	32570	0,11	1924	1818,21	1802,39	315	10	51
s10-0	S	1	32568	0,05	981	23,98	12,1	207	5	45
s10-0 ^b	A	1	32779	6,53	546916	18000,00				9734
s10-1	E	1	18603	0,07	135	767,23	764,51	73	3	3
s10-1	S	1	18603	0,07	180	14,82	11,95	102	3	5
s10-1	A	1	18603	0,07	242	4,14	0,40	132	3	12
s10-2	E	5	62909	0,11	173	357,15	350,84	9	8	1
s10-2	S	5	62909	0,11	173	9,97	3,61	9	8	1

Tableau 4.4 Résultats détaillés de la résolution à une seule méthode (suite et fin)

Exemplaire	Méthode ^a	Coupes <i>CPLEX</i>	Valeur solution	Saut d'intégrité (10^{-3})	Sommets énumérés	Durée (s)	Temps coupes (s)	Solutions entières	Solutions retenues	Coupes générées
s10-2	A	5	62908	0,09	1088	10,37	0,80	351	7	37
s15-0 ^c	E									
s15-0	S	0	88714	0,04	120	9,61	3,28	5	0	0
s15-0	A	0	88714	0,04	120	6,42	0,07	5	5	0
s15-1	S	2	131103	0,04	65	7,13	2,78	4	4	0
s15-1	A	2	131103	0,04	65	4,37	0,05	4	4	0
s15-2	S	1	104647	1,93	1792	64,41	43,96	139	9	23
s15-2	A	1	104814	3,53	86937	435,23	95,24	16462	9	110
s20-0	S	0	302783	0,06	3	6,41	3,41	2	2	0
s20-0	A	0	302783	0,06	3	2,23	0,04	2	2	0
s20-1	S	0	295778	0,08	15	11,57	5,62	6	6	0
s20-1	A	0	295778	0,08	15	5,98	0,12	6	6	0
s20-2	S	0	405523	0,02	3	6,44	4,42	4	4	0
s20-2	A	0	405530	0,04	95	11,03	0,27	21	4	6
s25-0	S	0	90486	0,07	85	16,47	5,34	5	5	0
s25-0	A	0	90486	0,07	85	11,08	0,09	5	5	0
s25-1	S	0	82277	0,01	15	9,7	4,6	4	4	0
s25-1	A	0	82277	0,01	15	5,16	0,09	4	4	0
s25-2	S	0	109051	0,06	3	6,68	4,08	3	3	0
s25-2	A	0	109051	0,06	3	2,7	0,09	3	3	0
s30-0	S	0	532183	0,01	3	7,41	4,78	3	3	0
s30-0	A	0	532183	0,01	1448	61,88	9,91	757	4	128
s30-1	S	0	499413	0,03	3	12,43	6,58	4	4	0
s30-1	A	0	499429	0,07	805	35,92	1,58	138	4	40
s30-2	S	0	561884	0,03	3	8,58	4,95	3	3	0
s30-2	A	0	561914	0,08	539	44,5	1,65	151	4	31

^aLes méthodes sont identifiées ainsi : *A* pour la méthode approchée, *E* pour la méthode exacte et *S* pour la méthode simulée.

^bRésolution incomplète pour cet exemplaire.

^cL'explosion des temps de calcul de la méthode exacte rend prohibitive toute tentative de résolution des grandes exemplaires.

méthodes sont faibles.

Lorsqu'on regroupe les résultats pour une même catégorie d'exemplaires, on peut mettre ces derniers en relation avec la taille des ensembles sur lesquels s'effectue l'évaluation des probabilités. Ainsi, le tableau 4.5 présente, pour chaque méthode d'évaluation, le temps de résolution moyen, en secondes, pour les sept tailles de type de salles considérées. Pour les cas où une méthode n'est pas parvenue à résoudre un exemplaire, un temps de 18 000 secondes a été retenu, comme dans la présentation du tableau 4.4. Ainsi, le gonflement dans les temps de résolution pour $N_s = 5$ est dû à l'exemplaire *s5-2*, sur laquelle les trois méthodes ont buté.

Tableau 4.5 Temps moyen de résolution (s) selon le choix d'une méthode d'évaluation

N_s	Exacte	Simulée	Approchée
3	0,73	3,35	12000,35
5	6267,29	7622,55	18000,00
10	980,86	16,27	6004,84
15	18000,00	27,05	148,67
20	18000,00	8,14	6,41
25	18000,00	10,95	6,31
30	18000,00	9,47	47,43

En observant particulièrement les lignes $N_s = 3$ et $N_s = 10$, on peut voir la croissance rapide des temps de calcul de la méthode exacte. Alors que ceux de la méthode simulée ont grossièrement quadruplé, de 3,35 à 16,27, la méthode exacte explose d'un temps moyen de 0,73 pour $N_s = 3$ à 980,86 pour $N_s = 10$. Pour les autres catégories d'exemplaires, la méthode exacte n'arrive plus à fournir des solutions. L'évolution des temps de calcul de la méthode approchée est tout autre. Les temps moyens sont très élevés sur les petites valeurs de N_s en raison des exemplaires incomplets où un grand nombre de solutions entières sont rejetées. Toutefois, après un gonflement dû à la résolution ardue de *s15-2*, le temps moyen avoisine 6,5 pour ensuite se remettre à augmenter avec une valeur moyenne de 47,43 pour $N_s = 30$. Enfin, en ce qui concerne la méthode simulée, les temps sont changeants. Les plus fortes valeurs moyennes surviennent pour les tailles d'ensembles médianes et il semble y avoir une légère tendance à la hausse puisque les temps triplent en moyenne de $N_s = 3$ à $N_s = 30$.

Pour obtenir une comparaison plus fidèle des temps d'évaluation liés au choix d'une méthode plutôt qu'une autre, le tableau 4.6 présente, avec la même facture que le tableau 4.5, le rapport entre le temps passé à évaluer les probabilités d'encombrement et le nombre de solutions entières rencontrées lors de la résolution. Le résultat est un estimé, en millisecondes, du temps passé sur chacune des solutions entières à effectuer l'évaluation de la probabilité

d'encombrement. Les valeurs impliquées dans ce rapport ne sont présentées que pour les exemplaires résolus, d'où l'absence de distorsion au niveau de $N_s = 5$ et de la méthode approchée.

Tableau 4.6 Temps moyen d'évaluation des probabilités (ms) par solution entière

N_s	Exacte	Simulée	Approchée
3	0,292	2,88	0,230 ^a
5	2,32 ^a	19,6 ^a	s,o ^b
10	7349,5	87,0	2,48 ^a
15	s,o ^b	338,0	5,79
20	s,o ^b	1120,3	14,8
25	s,o ^b	1168,3	22,5
30	s,o ^b	1631,0	12,7

^aValeur pour les exemplaires résolus seulement.

^bAucun exemplaire de cette catégorie ne fut résolu par cette méthode.

Les constats pouvant être faits à partir du tableau 4.5 au sujet de la croissance rapide des temps de la méthode exacte et de celle, plus lente, mais néanmoins notable, de la méthode approchée sont également visibles avec le temps moyen par solution entière. On peut remarquer une divergence au niveau de $N_s = 30$ pour la méthode approchée : la résolution globale a été plus longue, mais le temps par solution entière est près de deux fois plus faible par rapport à $N_s = 25$. Du côté de la méthode simulée, la croissance du temps de calcul moyen avec celle de N_s est sans équivoque. Par ailleurs, en aucun cas la méthode simulée ne parvient à s'imposer avec les meilleurs temps par solution entière.

Choix des seuils de transition

L'algorithme 1 présenté dans la section 3.3.3 propose une formulation hybride qui combine les différentes méthodes d'évaluation selon des seuils paramétriques sur la taille des ensembles $\hat{G}_{ps}$. Sous un premier seuil σ_1 , l'évaluation se fait avec la méthode exacte. Au-delà d'un second seuil σ_2 , c'est la méthode approchée qui est employée. Entre ces seuils, l'algorithme a recours à la méthode simulée. Si on se réfère au tableau 4.5, on peut constater la présence d'une tendance dans les temps les plus faibles, qu'on retrouve d'abord avec la méthode exacte, puis chez la méthode simulée et la méthode approchée.

Trois essais ont été menés pour identifier une valeur pour le premier seuil. Puisque la méthode exacte ne parvient pas à résoudre les exemplaires de catégorie $N_s = 15$, la valeur du paramètre σ_2 a été fixé à l'aide de la taille maximale d'ensembles tolérés rapportée dans

le tableau 4.3, d'où $\sigma_2 = 28$, la valeur médiane entre les tailles maximales pour $N_s = 10$ et $N_s = 15$. Le tableau 4.7, reprenant les formats précédents, permet de comparer les temps moyens de résolution selon la valeur donnée au paramètre σ_1 . Les moyennes de ce tableau excluent l'exemplaire *s5-2*, dont la résolution n'a pu être terminée. Le trait double entre les lignes $N_s = 5$ et $N_s = 10$ dénote le seuil à partir duquel les variations proposées du paramètre σ_1 ont une influence sur les temps obtenus. Il est aisé de constater que $\sigma_1 = 25$ laisse une place trop grande à la méthode exacte qui provoque, en moyenne, un allongement considérable de la résolution des exemplaires $N_s = 10$ et $N_s = 15$. Pour la dernière catégorie d'exemplaires, les moyennes sont comparables d'un seuil à l'autre. Toutefois, pour $N_s = 10$, la valeur $\sigma_1 = 20$ joue un rôle charnière entre des temps d'environ 16 secondes pour $\sigma_1 = 16$ et $\sigma_1 = 18$ et le double de cette durée pour $\sigma_1 = 22$, 34,69 pour être plus précis. Ce gain peut sembler mince, mais il est beaucoup plus tangible que les autres variations constatées. De plus, il survient dans une catégorie d'exemplaires directement visée par la variation, puisque la méthode exacte s'applique intégralement sur celle-ci si $\sigma_1 \geq 22$.

Tableau 4.7 Temps moyen de résolution selon la variation du paramètre σ_1

N_s	$\sigma_1 = 16$	$\sigma_1 = 18$	$\sigma_1 = 20$	$\sigma_1 = 22$	$\sigma_1 = 25$
3	0,71	0,70	0,70	0,69	0,74
5	397,92	395,76	391,98	371,96	393,17
10	16,09	16,01	21,61	34,69	986,78
15	199,70	199,05	197,35	190,99	976,63
20	6,57	6,44	6,43	6,15	6,57
25	6,47	6,39	6,29	6,11	6,46
30	47,82	47,58	47,42	46,37	48,04

Les essais précédents ont permis de retenir une valeur pour le paramètre σ_1 . En conservant cette valeur constante à $\sigma_1 = 18$, cinq valeurs ont été données au paramètre σ_2 , comme en témoigne le tableau 4.8. À la manière du tableau précédent, l'exemplaire *s5-2* n'est pas considéré. De même, le trait double marque le seuil à partir duquel les changements proposés dans la valeur du paramètre σ_2 peut avoir une influence concrète dans les résultats. Pour $\sigma_2 = 28$ et $\sigma_2 = 30$, le temps moyen de résolution de la catégorie $N_s = 15$ est nettement supérieur à ceux obtenus par les autres variations, qui se retrouvent autour de 28 secondes. Pour la catégorie $N_s = 20$, c'est la valeur $\sigma_2 = 44$ qui remporte la palme avec seulement 4,24, là où les autres seuils retournent entre 6,44 et 8,16. Cette valeur est par ailleurs la seule à se démarquer pour une catégorie d'exemplaires mais reste compétitive sur toutes les catégories, n'enregistrant la pire performance que pour $N_s = 30$, mais avec un temps très proche de celui

de $\sigma_2 = 30$.

Tableau 4.8 Temps moyen de résolution selon la variation du paramètre σ_2

N_s	$\sigma_2 = 28$	$\sigma_2 = 30$	$\sigma_2 = 36$	$\sigma_2 = 44$	$\sigma_2 = 55$
3	0,70	0,92	0,78	0,73	0,78
5	395,76	432,17	438,90	432,96	438,43
10	16,01	17,55	16,40	16,41	16,63
15	199,05	229,9	27,68	28,27	27,72
20	6,44	8,16	7,02	4,24	7,51
25	6,39	7,67	7,05	6,96	7,04
30	47,58	51,18	51,03	51,27	50,75

Les essais exposés ci-haut amène donc à proposer de retenir 20 et 44 comme valeurs respectives de σ_1 et σ_2 . Le tableau 4.9 présente les résultats détaillés de la résolution des exemplaires utilisés jusqu'à maintenant. La colonne variation donne la différence entre la valeur de la solution obtenue avec la méthode hybride et celle retournée en n'utilisant que la méthode simulée. Seulement trois de cinq exemplaires présentant une divergence ont une valeur notablement différentes. La section 4.3.1 montre que cette dernière méthode était elle-même une approximation acceptable de la solution issue d'une méthode exacte.

Des résultats plus raffinés figurent dans le tableau 4.10, qui combine les temps moyen de résolution globale, en secondes, et le temps moyen d'évaluation des probabilités par solution entière, en millisecondes. Pour les temps globaux, la colonne « Variation » permet de comparer le temps pris par la méthode hybride avec celui pris par la méthode qui, parmi les trois présentées à la section 3.2, a *le mieux performé*. Dans deux cas, la méthode hybride parvient à produire une légère amélioration des temps de calcul, autrement, elle affiche une performance légèrement au-dessus de la meilleure méthode. Une seule exception à ce constat, la catégorie d'exemplaires $N_s = 30$ où la méthode simulée avait mieux fait que la méthode approchée, sur laquelle la méthode hybride s'est davantage alignée. Les temps d'évaluation moyens par solution entière laisse penser, lorsqu'on les compare aux valeurs correspondantes du tableau 4.6, qu'hormis sur les exemplaires avec une faible valeur de N_s , que la méthode hybride se fonde rarement sur une seule méthode dans sa résolution.

4.3.2 Évaluation de la performance de *CourtCausesScheduler*

Sur la base des résultats comparant les méthodes d'évaluation des probabilités de débordement et en utilisant la méthode hybride proposée à la section 4.3.1, trois exemplaires ont

Tableau 4.9 Résultats détaillés de la résolution hybride

Exemplaire	Coupes <i>CPLEX</i>	Valeur solution	Variation	Saut d'intégrité (10^{-3})	Sommets énumérés	Durée (s)	Temps coupes (s)	Solutions entières	Solutions retenues	Coupes générées
s3-0	2	442	-	2,27	4	0,05	0,01	3	3	1
s3-1	0	464	-	10,89	364	0,17	0,05	193	3	1
s3-2	0	496	-	0,00	4442	1,97	0,65	2613	4	4
s5-0	6	3653	-	0,79	380	2,50	0,13	23	8	3
s5-1	9	3648	-	1,42	261095	863,41	435,01	178954	3	224
s5-2	7	3844	-1	1,66	3405430	17,9 ks	12,6 ks	2,06 M	20	306
s10-0	1	32568	-2	0,05	981	24,56	12,10	207	5	45
s10-1	1	18603	-	0,07	180	14,59	11,66	102	3	5
s10-2	5	62909	-	0,11	173	10,07	3,48	9	8	1
s15-0	0	88714	-	0,04	120	10,04	3,42	5	5	0
s15-1	1	131103	-	0,04	65	8,58	3,63	4	4	0
s15-2	1	104647	+167	3,53	1792	66,19	44,67	139	9	23
s20-0	0	302783	-	0,06	3	3,31	0,45	2	2	0
s20-1	0	295778	-	0,08	15	6,78	0,58	6	6	0
s20-2	0	405523	-	0,04	3	2,64	0,50	4	4	0
s25-0	0	90486	-	0,07	85	11,9	0,24	5	5	0
s25-1	0	82277	-	0,01	15	5,85	0,27	4	4	0
s25-2	0	109051	-	0,06	3	3,12	0,18	3	3	0
s30-0	0	532183	-	0,01	1448	67,36	10,71	757	4	128
s30-1	0	499429	+16	0,07	805	38,91	1,63	138	4	40
s30-2	0	561914	+30	0,08	539	47,55	1,83	151	4	31

Tableau 4.10 Tableau comparatif de la méthode hybride

N_s	Temps moyen de résolution (s)		Temps par solutions (ms)
	Hybride	Variation	
3	0,74	+0,01	0,73
5	6266,0	-1,29	4,46 ^a
10	21,84	+5,57	101,2
15	27,41	+0,36	342,1
20	3,84	-2,57	111,7
25	6,61	+0,30	57,5
30	49,04	+39,57	12,8

^aCette valeur ne tient pas compte de l'exemplaire *s5-2*.

été générés en utilisant des données pseudo-réalistes afin d'offrir une base de compétition entre *CourtCausesScheduler* et le constructeur séquentiel mimant la pratique usuelle. Après une présentation des exemplaires utilisés, le choix des paramètres de surcharge utilisés et ses fondements seront exposés et d'ultimes résultats permettront de décrire la performance des deux méthodes générales.

Exemplaires pseudo-réalistes

Les exemplaires générés pour les tests de performances comparées partagent plusieurs caractéristiques. Toutes les causes incluses dans le jeu de données ont été créées en suivant le modèle pseudo-réalistes établi à la section 4.1.2. Initialement, un horizon de planification d'un mois, soit vingt jours ouvrables a été envisagé. Cependant, on peut constater au tableau 4.1 que les durées prévues potentielles peuvent dépasser cet horizon. C'est pourquoi l'horizon utilisé pour les exemplaires est de quarante périodes, soit deux mois si chaque période représente un jour ouvrable. De même, par souci d'avoir un nombre total de salles proche de la taille du bassin de la Cour supérieure, les exemplaires se sont écartés des valeurs découlant du tableau 4.2 pour se porter à sept salles de type I, quinze salles de type II et deux salles de type III. Les proportions pour déterminer les besoins des causes sont demeurées les mêmes.

Le tableau 4.11 dénombre les causes régulières et préassignées des trois exemplaires. Les valeurs sont inspirées du nombre de causes prévues pour les différents mois de l'échantillon fourni par le palais de justice de Montréal. On peut constater dans la figure 4.1 que ce nombre oscille de 166 pour le mois *B* à 211 pour le mois *A*. Pour un horizon de deux mois, un volume allant de 350 à 450 causes apparaît représentatif de la réalité. Les proportions des causes régulières et préassignées est arbitraire et cherche à évoquer une assignation dans

laquelle l’horizon de deux mois a déjà partiellement fait l’objet d’assignations précédentes. L’emplacement au rôle des causes préassignées pour les trois exemplaires a été déterminé en exécutant *CourtCausesScheduler* sur un horizon vierge avec un taux de surcharge $\beta = 1,25$. Cinq paires de conflit entre des causes régulières ont également été déterminées aléatoirement. Chaque cause régulière a reçu trois périodes de disponibilité disposées aléatoirement entre le début de l’horizon et la dernière période permettant à la cause de procéder totalement à l’intérieur de l’horizon. Enfin, la pénalité découlant de ne pas avoir inscrit une cause au rôle, soit la valeur de M_c , est fixée à 200 fois la plus longue durée potentielle pour la cause c .

Tableau 4.11 Caractéristiques des exemplaires pseudo-réalistes

Nom	$ C^A $	$ C^R $
r350	150	200
r400	200	200
r450	250	200

Choix des taux de surcharge

La taille des types de salles dans les exemplaires qu’on souhaite résoudre a un impact sur les taux de surcharge retenus. En effet, pour qu’il y aie effectivement surplanification, le taux doit être au moins de 1,1429 pour le type I ($\frac{8}{7}$), 1,066... pour le type II ($\frac{16}{15}$) et, surtout, 1,5 pour le type III ($\frac{3}{2}$). Pour le constructeur séquentiel, il faut également considérer qu’aucune vérification du risque ne sera entreprise, ce qui impose une certaine réserve sur le taux de surcharge à employer. Les taux de surcharge dont les résultats sont présentés ci-bas vont ainsi de 1,3 à 1,9. Il convient de noter que les premiers taux ne permettront une surcharge que sur les types I et II.

Pour choisir les taux de surcharge utilisés par *CourtCausesScheduler*, le contrôle du risque d’encombrement permet d’utiliser des taux plus élevés. Une surcharge trop ambitieuse peut cependant ralentir la résolution, car un grand nombre de coupes devront être intégrées pour éliminer les solutions trop risquées. Toutefois, en utilisant une valeur du paramètre β tendant vers l’infini, la solution optimale obtenue présenterait forcément une surplanification concrète d’un taux plus faible. Cette solution sera trouvée par toute résolution employant une valeur de β suffisamment grande. Pour les autres valeurs, la solution obtenue devrait théoriquement présenter une certaine « saturation », c’est-à-dire que pour certains couples période-type de salle, la planification proposée sera égale à $\lfloor \beta N_s \rfloor$. La valeur $\beta = 2,2$ utilisée dans les tests

précédents a livré une bonne performance et fut retenue pour éviter d'enliser la résolution avec l'ajout d'un trop grand nombre de coupes.

Approche mathématique

Conformément avec la section précédente, chacun des trois exemplaires fut résolu par *CourtCausesScheduler*. La résolution n'a pu être complétée que pour *r-350*. Le tableau 4.12 rend compte de ces résolutions. On peut remarquer que les exemplaires n'étant que partiellement résolus n'affichent pas un saut d'intégrité plus élevé que celui de *r-350*. En fait, la valeur de la meilleure solution obtenue par *r-450* n'est séparée que par 1 % de celle de la résolution au sommet racine. Par ailleurs, les valeurs des solutions sont peu élevées puisque sur les deux cents causes régulières à assigner dans les trois exemplaires, moins de cinq ont été mises de côté. L'exemplaire *r-400* se distingue par un grand nombre de coupes par rapport au nombre de solutions entières rencontrées durant l'énumération implicite. Néanmoins, il s'agit également de l'exemplaire où le plus de solutions entières réalisables ont été identifiées.

Tableau 4.12 Résultats détaillés en utilisant *CourtCausesScheduler*

Exemplaire	Coupes <i>CPLEX</i>	Valeur de la solution	Saut d'intégrité (10^{-3})	Sommets énumérés	Durée (s)	Temps coupes (s)	Solutions entières	Solutions retenues	Coupes générées
r-350	0	6103	241,46	2050	150,82	147,88	1116	9	42
r-400	18	7283	324,52	227613	18000 ^a	17607,20	11256	23	11053
r-450	6	3318	10,04	200228	18000 ^a	17670,50	100846	18	1844

^aLa résolution n'a pas pu être complétée pour cet exemplaire.

Approche traditionnelle-empirique

Pour chacun des sept taux de surcharge prévus à la section précédente, le constructeur séquentiel a généré des horaires pour les trois exemplaires du problème stochastique d'horaire d'audiences judiciaires. Dans un premier temps, le tableau 4.13 compare la valeur de la solution obtenue par le constructeur séquentiel. Tel qu'appréhendé, les délais sont davantage minimisés avec l'augmentation du taux de surcharge. La valeur élevée des solutions provient

Tableau 4.13 Valeur de la solution du constructeur séquentiel selon le taux de surcharge

Exemplaire	1,3	1,4	1,5	1,6	1,7	1,8	1,9	CCS
r-350	213534	213534	203615	203615	203615	203615	203615	6103
r-400	213732	213732	155313	128186	128186	89530	89530	7283
r-450	195608	195608	163302	110436	110436	99129	99129	3318

des pénalités de non-mise au rôle de certaines causes. La dernière colonne rappelle la valeur obtenue par la meilleure solution de *CourtCausesScheduler*, ce qui permet de constater que le constructeur séquentiel n'arrive pas, dans un horaire déjà partiellement comblé, à offrir des solutions avec des délais moindres et ce, pour les trois exemplaires. En comparant les valeurs obtenues pour un même exemplaire, il est possible de voir qu'à partir de 1,8, la hausse du taux de surcharge ne permet plus d'améliorer la solution. Ce type de plateau survient pour divers taux, lorsque les périodes de départ valides des différentes causes ne permettent plus ou peu de surcharge supplémentaire.

Jusqu'à maintenant, les résultats montrent que l'approche mathématique soutient la comparaison avec le constructeur séquentiel. L'objectif est certes de minimiser les délais judiciaires et d'éviter de reporter indéfiniment la mise au rôle de certaines causes, mais il importe également de maintenir les risques d'encombrement du rôle sous un certain niveau de tolérance. *CourtCausesScheduler* a été exécuté avec un seuil de tolérance $1 - \alpha = 0,05$. Vu leur conservatisme, il serait tentant de supposer que le constructeur séquentiel respecte lui-aussi cette tolérance.

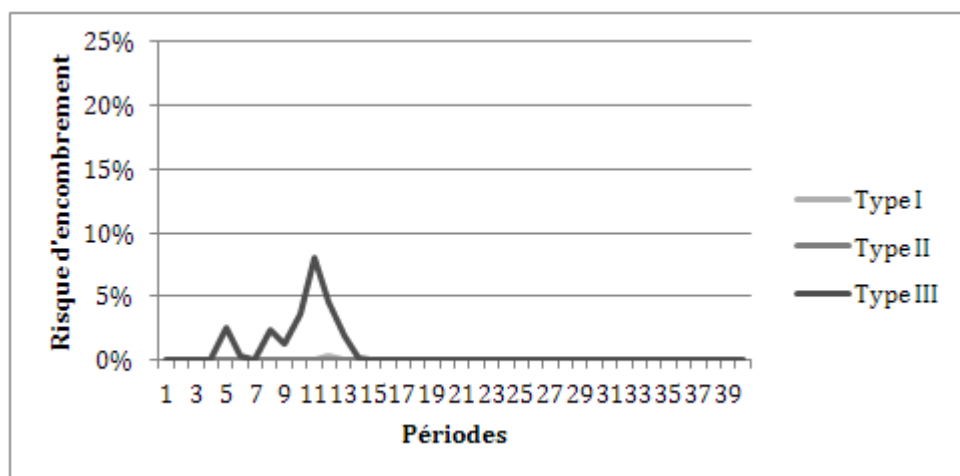


Figure 4.10 Risque d'encombrement pour solution de l'approche séquentielle pour r-350

À partir des meilleures solutions obtenues par assignation séquentielle pour chaque exemplaire, une simulation de Monte-Carlo a été utilisée pour estimer les probabilités d'encombrement du rôle engendrées. Les figures 4.10 à 4.12 présentent sous forme de courbes les résultats obtenus. Toutes les solutions ont été trouvées avec un taux de surcharge de 1,8. D'après les approximations obtenues, le type II ne pose un risque intolérable que dans le cas de la solution de l'exemplaire *r-400*, où la troisième et la quatrième période affichent des risques d'encombrement respectivement de 14,10 % et de 15,65 %. Les deux autres types présentent des risques élevés pour les trois solutions. Dans la figure 4.10, le risque d'encombrement des salles de type III atteint 8,01 % pour la période 8, ce qui est le seul écart pour cette solution.

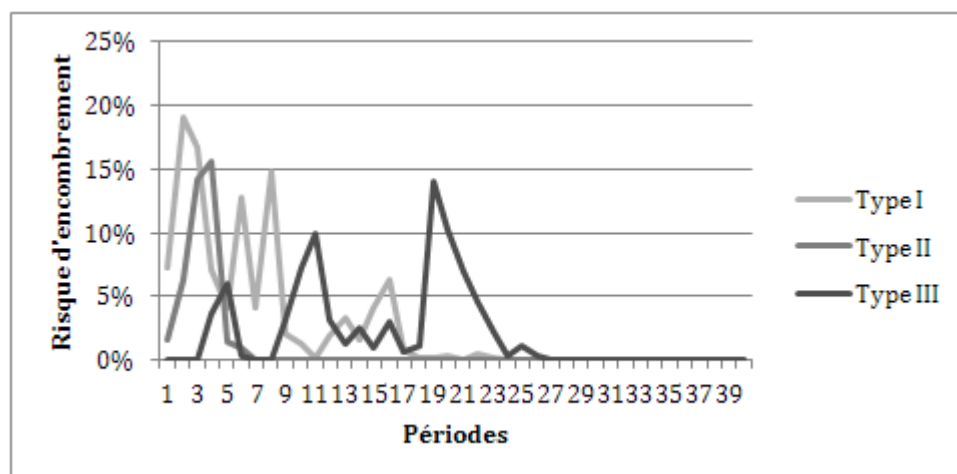


Figure 4.11 Risque d'encombrement pour solution de l'approche séquentielle pour *r-400*

Dans la figure 4.11, le risque le plus élevé est de 19,06 %, survenant dans une série de quatre périodes où les salles de type I présente un risque d'encombrement élevé. Pour ce type de salle, le niveau culmine à nouveau à 12,70 % à la période 6, puis à 14,82 % à la période 8 et, enfin, une poussée à 6,27 % survient à la période 16. Le type III connaît trois épisodes de risque dépassant la tolérance. Un premier pic isolé sur la période 5 amène le risque à 6,00 %. Les périodes 10 et 11 affichent des risques de 7,32 % et de 10,00 % alors que la dernière succession se produit à partir de la période 19 avec un risque de 14,07 %, qui décline à 10,20 % et enfin à 7,01 % avant de repasser sous la barre des 5 %. Mais c'est dans la figure 4.12 qu'on peut observer le plus grand risque, soit 22,85 % pour le premier type de salle à la période 10. Sur les treize premières périodes de l'horizon, neuf d'entre elles affichent un niveau de risque supérieur à 5 % pour un, et parfois deux, types de salles. Le risque pour le type III atteint 14,75 % à la période 3.

La volatilité du risque constatée dans les figures précédentes n'est pas que le fait des solutions analysées. Le tableau 4.14 établit qu'une bonne part des solutions retournées par le

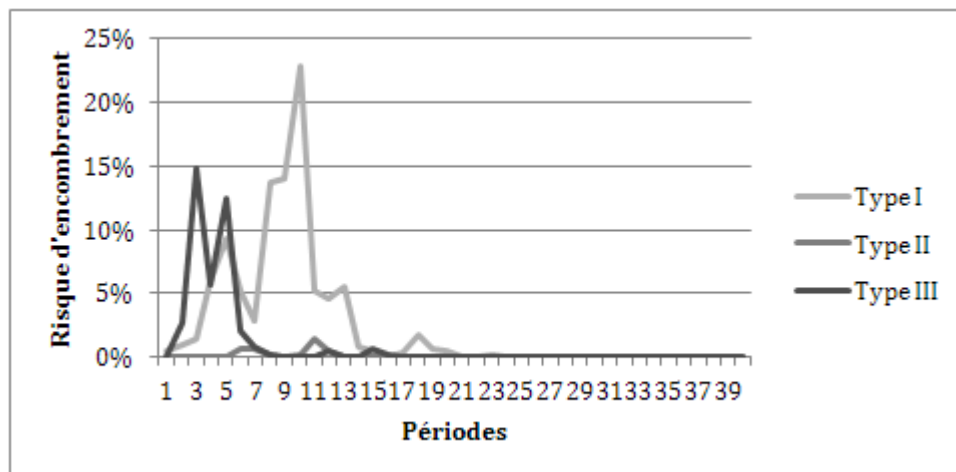


Figure 4.12 Risque d'encombrement pour solution de l'approche séquentielle pour $r-450$

constructeur séquentiel présentent des niveaux de risque d'encombrement du rôle supérieur à une tolérance de 5 %. Pour l'exemplaire $r-350$, une seule période pose problème, même pour le plus élevé des taux de surcharge. Par contre, pour l'exemplaire $r-400$, le nombre de période-type de salle risquées atteint la valeur de seize sur les grands taux de surcharge, après avoir passé à quatre et sept. Ce niveau de risque est légèrement supérieur à celui constaté sur l'exemplaire $r-450$, pour lequel on atteint un dépassement de la tolérance sur environ cinq, puis une dizaine de périodes-types de salle sur les taux de surcharge supérieur à 1,5. Les légères variations d'une période-type de salle ne sont pas nécessairement tributaires d'une faible volatilité du risque d'un taux de surcharge à l'autre, mais peuvent également tirer leur source dans une imprécision de la simulation de Monte-Carlo. Lorsqu'on considère le nombre de type de salles et la taille de l'horizon, ces quantités sont tout de même faibles.

Tableau 4.14 Nombre de périodes-types de salle dont le risque d'encombrement excède 5 %

Surcharge	1,3	1,4	1,5	1,6	1,7	1,8	1,9
r-350	0	0	1	1	1	1	1
r-400	0	0	4	7	7	16	16
r-450	0	0	0	5	4	11	10

4.4 Discussion

Les précédentes sections ont permis de présenter un certain nombre de résultats sur les méthodes d'évaluation de la probabilité d'encombrement du rôle ainsi que sur la performance

de l'approche mathématique développée dans ce mémoire par rapport à un algorithme séquentiel qui imite la pratique actuelle dans la mise au rôle. Pour clore ce chapitre, des constats seront dressés à partir de ces résultats. De façon globale, les postulats sur le comportement des méthodes d'évaluation de la probabilité d'encombrement du rôle seront étayés ou remis en cause. Puis, les mérites d'une approche mathématique plutôt que strictement séquentielle seront discutés.

Conformément à l'ordre des temps de calcul établi à la section 3.2.1 pour la méthode exacte, cette dernière se voit rapidement déclassée par les autres méthodes lorsque la taille des ensembles sur lesquels l'évaluation porte augmente. Le passage brusque d'un temps moyen par solution entière de 2,32 millisecondes à 7,35 secondes, qu'on retrouve dans le tableau 4.6 est à ce sujet très éloquent. On peut par ailleurs souligner qu'entre les deux mêmes catégories d'exemplaires, même s'il se trouve presque à quintupler, le temps moyen par solution entière de la méthode simulée reste bien en deçà de l'explosion combinatoire qui était anticipée pour la méthode exacte.

En présentant le tableau 4.4, la valeur du paramètre γ a été validée par la proximité entre la valeur des solutions obtenues avec la méthode simulée et celles retournées en utilisant la méthode exacte. Au-delà d'un ajustement de paramètre, cette capacité à produire des résultats proches de la méthode exacte démontre que la méthode simulée est un pendant fiable. Cette proximité de résultat s'observe même sur les exemplaires requérant l'évaluation de probabilités sur des grands ensembles, cette fois avec la méthode approchée, qui par théorème, fournit une bonne approximation sur ceux-ci.

Si on se rapporte à la section 3.2.2, les temps de calculs sont présentés comme la principale force de la méthode approchée. Indépendamment de la qualité des résultats, médiocres tels que prévus sur les petits ensembles, le tableau 4.6 montre, lorsque les résultats sont disponibles, que c'est bel et bien la méthode approchée qui évalue le plus rapidement les probabilités d'encombrement du rôle et ce, sur chaque catégorie d'exemplaires. Cette rapidité relative, particulièrement sur les dernières catégories d'exemplaires, permet une économie de temps qui transparait légèrement sur les résultats globaux présentés dans le tableau 4.5. Il s'avère ainsi que les trois méthodes d'évaluation, selon les résultats présentés, ont des comportements conformes à ceux supputés à leur égard au chapitre 3.

On peut aisément constater, à la lecture du tableau 4.5, que les trois méthodes discutées ci-haut ont tour à tour leur intérêt selon la taille des ensembles concernées. Selon l'algorithme 1, la résolution globale ferait bon parti de l'usage de chacune des trois méthodes au gré des ensembles sur lesquels la probabilité d'encombrement du rôle doit être évaluée. Le choix de seuils σ_1 et σ_2 permet de caractériser une méthode de résolution hybride combinant les trois approches. Le tableau 4.10 montre, avec la colonne variation, que les temps moyens de

résolution d'une méthode hybride peuvent se rapprocher de la plupart des temps obtenus par la meilleure des trois méthodes seule, et, parfois, même permettre une résolution légèrement plus rapide.

L'ensemble des méthodes d'évaluation de la probabilité d'encombrement du rôle tire leur importance du fait qu'elles sont ce qui distingue la formulation présentée à la section 3.1.3 de la procédure détaillée dans l'algorithme 2, qui s'en remet entièrement au taux de surcharge pour assurer le contrôle du risque. Celui-ci, les figures 4.10 à 4.12 ainsi que le tableau 4.14 en font foi, est bien mince. Le risque encouru est enclin à varier au cours de l'horizon. Dans le cas des salles de type II, il reste largement contenu, mais il ne peut en être dit autant des différents types de salles pris dans leur ensemble.

L'avantage d'une approche mathématique pour produire un horaire de salles d'audience tient donc de la qualité des solutions mais également dans une garantie de contrôle du niveau de risque. En comparant la qualité des solutions découlant du tableau 4.13, il ressort que le constructeur séquentiel peine à insérer des causes dans un horaire déjà partiellement comblé. Les horaires proposés par *CourtCausesScheduler* bornent le risque, ce que ne peut faire la solution séquentielle. Seul un taux de surcharge plus conservateur permet d'approcher d'une situation où le risque d'encombrement peut paraître contenu, ce qui se fait au prix d'une détérioration notable de la qualité des solutions. L'approche mathématique parvient donc à nettement s'imposer et même, dans une certaine mesure, à transcender l'arbitrage entre réduction des délais et prise de risque.

CHAPITRE 5

CONCLUSION

Après avoir proposé, exposé et étayé une certaine formulation du problème stochastique d'horaire d'audiences judiciaires et la manière de le résoudre, il convient de dresser un bilan récapitulant les travaux de recherche à la source de ce mémoire. Par la suite, un examen de la portée du travail accompli permettra d'en juger la valeur, mais également la pertinence au-delà du simple exercice intellectuel. Enfin, rare sont les productions de recherches qui peuvent être considérées comme totalement finies en elles-mêmes. C'est pourquoi il est nécessaire de reconnaître et d'identifier les pistes sur lesquelles il sera possible de continuer à étoffer la problématique entamée par cet ouvrage.

5.1 Bilan de recherche

Cette section se veut une synthèse de la démarche présente. Tout d'abord, celle-ci sera rappelée pour permettre plus facilement de la considérer dans son ensemble et sa logique. Ce n'est qu'ensuite qu'il sera possible de témoigner des accomplissements de ces travaux de recherche, à la lumière des buts vers lesquels ces derniers devaient mener.

En début d'ouvrage, le contexte de la production des rôles judiciaires a été présenté. La complexité de cette problématique dans le cadre d'un grand district judiciaire en fait un problème sur lequel la recherche opérationnelle peut se pencher. Le caractère original d'une approche mathématique de ce problème dans le domaine judiciaire québécois ainsi que l'intérêt théorique de développer une prise en compte dynamique de l'incertitude sur la durée des procès font du problème stochastique d'horaire d'audiences judiciaires un objet d'étude valable.

De façon plus générale, ce problème s'inscrit dans la famille des problèmes d'horaire, qui ont collectivement suscité une attention appréciable de la communauté scientifique. L'un des enjeux de modélisation de ce groupe de problèmes est la gestion de l'incertitude. Cette dimension est centrale au problème stochastique d'horaire d'audiences judiciaires et a été abordée de manières diverses au sein de la littérature contemporaine. Ainsi, s'il n'y a guère de trace d'un effort mathématique pour produire un rôle judiciaire, il s'avère que ce problème se rapproche d'autres situations d'optimisation qui, elles, ont fait l'objet de l'étude de chercheurs opérationnels dans le passé. Néanmoins, ce mémoire a l'ambition d'apporter son apport au corpus existant par son traitement de l'incertitude sur les durées d'événements.

Pour rendre compte en termes mathématiques du problèmes d’horaire de salles d’audience, deux modèles proches, mais différents, sont formulés. L’un d’entre eux fait usage de contraintes aléatoires non linéaires, ce qui aiguille la réflexion vers une résolution par énumération implicite à plans coupants faisant dynamiquement appel à diverses méthodes d’évaluation des probabilités de risque découlant de la durée incertaine des procès.

En combinant des données obtenues grâce à la collaboration des responsables du rôle judiciaire civil d’une cour provinciale de Montréal et des outils logiciels, des tests expérimentaux ont pu être menés pour calibrer la méthode de résolution mathématique proposée et comparer ses mérites à une approche inspirée de la pratique contemporaine. Il se dégage de ces résultats que la gestion des risques liées à la durée des procès est accomplie par la résolution du modèle mathématique développé et ce, tout en permettant de minimiser les délais avant l’ouverture des divers procès.

5.2 Retombées académiques

D’entrée de jeu, l’une des retombées de nature académique de ce mémoire est l’introduction du problème stochastique d’horaire d’audiences judiciaires, qui est au cœur de l’ensemble des activités de recherche qui ont mené à cet ouvrage. Ce problème est relié à une grande famille, celle des problèmes d’horaire, dont les diverses facettes et les principaux modes de résolution sont décrits dans la revue des sections 2.1 et 2.2.

Dans sa gestion de l’incertitude, la méthode de résolution exposée à la section 2.4.2 propose d’utiliser des méthodes externes à l’énumération implicite pour intégrer dynamiquement des plans coupants qui viennent remplacer des contraintes non linéaires. Cette pratique n’est pas originale, la section 2.2.3 souligne d’autres occurrences comprenant des contraintes aléatoires. Cependant, il s’agit d’une avenue digne d’intérêt pour des problèmes pouvant recevoir une formulation quasi-linéaire.

5.3 Retombées pratiques

Les retombées de ce mémoire ne sont pas uniquement théoriques. Si on se replace dans le cadre de la problématique générale, le simple fait de proposer une approche mathématique pour mener à bien l’élaboration des rôles judiciaires ajoute une perspective à partir de laquelle les autorités compétentes en matière de soutien à l’activité judiciaire peuvent articuler leur prise de décision, voire des intentions de réforme.

Plus capital encore, cette démarche d’optimisation d’un processus de l’administration publique fait sa part pour favoriser encore davantage l’ouverture de cette dernière envers la recherche opérationnelle. Déjà entamée dans des secteurs para-publics et péri-publics, comme

la production d'hydroélectricité et la santé, cette percée disciplinaire est portée encore plus au cœur de l'appareil public avec le soutien à la justice. Dans son ensemble, l'État québécois, ses organismes, ses réseaux et ses entreprises recèlent de vastes champs d'applications potentielles de la recherche opérationnelle. La propagation en son sein des avancées en matière d'optimisation laisse entrevoir des économies tangibles en matière de ressources et une amélioration de l'efficacité des rouages du secteur public au sens large. Ces bénéfices représentent un gain objectif pour la société québécoise et pour les contribuables et prestataires de l'État.

5.4 Avenues de poursuites

Comme toute réalisation en matière d'optimisation, la recherche sous-tendant ce mémoire peut s'inscrire dans une démarche plus large pouvant être constamment améliorée. Ainsi, plusieurs aspects abordés dans ce mémoire pourraient être approfondis et peaufinés dans des travaux ultérieurs. Parmi ceux-ci, la poursuite de la caractérisation théorique du problème stochastique d'horaire d'audiences judiciaires. Par rapport à la typologie générale des problèmes d'horaire, une meilleure définition permettrait de déterminer si l'élaboration d'un rôle judiciaire s'inscrit dans un des groupes déjà établi ou s'il serait pertinent de promouvoir une nouvelle famille de problèmes d'horaire au sein duquel le problème stochastique d'horaire d'audiences judiciaires pourrait être un cas spécifique ou plus général. Dans un angle plus concret, tenter de décrire d'autres contextes que celui de la justice avec ce problème pourrait permettre d'en améliorer la pertinence comme objet théorique et favoriser sa filiation au sein des problèmes classiques.

Dans le chapitre 3, la question de la symétrie a été abordé au sujet des salles en utilisant le concept de multi-salles. La procédure présentée à la section 4.1.2 pour générer les données pseudo-réalistes mène à soulever la question de la symétrie des causes. Avec vingt-deux distributions de durée pour des centaines de causes au sein d'un horizon d'une quarantaine de périodes, la probabilité que plus d'une cause ayant la même distribution de durée partage une ou plusieurs périodes de départ est bien réelle. La formulation actuelle ne distinguant pas ce cas de figure, si ce dernier est susceptible de se poser fréquemment lors de la résolution, il conviendrait d'agir sur la formulation ou de doter l'énumération implicite de règles de branchement qui pourraient agir pour éviter les décisions redondantes.

Dans cet ultime chapitre, il a été rappelé à quelques reprises que l'utilisation de méthodes stochastiques pour procéder à l'évaluation des probabilités d'encombrement du rôle constitue un des traits distinctifs de la résolution centrale à cet ouvrage. Trois méthodes ont été considérées, mais une étude plus poussée pourrait permettre d'identifier d'autres procédés susceptibles d'accélérer la résolution ou de la rendre plus fiable, sans parler des améliorations

potentielles aux méthodes déjà décrites.

Il a été mentionné qu'au niveau des règles de branchement de l'énumération implicite, les rôles avaient été laissés à *CPLEX*. Les résultats obtenus montrent que cette dernière n'améliore que maigrement la borne inférieure en cours de résolution, ce qui est typique d'une énumération privilégiant la descente en profondeur dans l'arbre que son exploration en largeur. Différents modes de sélection des sommets à résoudre, ainsi que des techniques de branchement heuristiques pourraient être expérimentés et conduire sur une réduction des temps de calcul globaux.

En guise de conclusion, il apparaît important de souligner que des efforts ultérieurs pour développer la problématique des horaires de salles d'audience peut se faire en renouvelant la collaboration avec les services du rôle, au palais de justice de Montréal ou dans d'autres administrations intéressées. L'accès à des échantillons de données de plus grandes tailles pourrait permettre de créer des distributions beaucoup plus alignées sur la réalité des causes, ce qui renforce la capacité du modèle à établir des horaires de bonne qualité. Des efforts en ce sens permettront de favoriser des phases d'essai et, à terme, l'adoption par l'administration publique de nouvelles manières de tirer partie d'une approche mathématique de l'aide à la décision.

RÉFÉRENCES

- ABDULLAH, S., TURABIEH, H., MCCOLLUM, B. et MCMULLAN, P. (2010). A hybrid metaheuristic approach to the university course timetabling problem. *Journal of Heuristics*.
- AHMED, A. et ZHOIJUN, L. (2010). A biphasic approach for university timetabling problem. *2010 2nd International Conference on Computer Engineering and Technology*. Chengdu, China, V1–192–V1–197.
- ARJEN VAN WEERT et JAN A.M. SCHREUDER (1997). Construction of basic match schedules for sports competitions by using graph theory. *Practice and Theory of Automated Timetabling II. Second International Conference, PATAT'97*. Springer-Verlag, Toronto, Canada, 201–210.
- BILLINGSLEY, P. (2012). *Probability and Measure*. Wiley Series in Probability and Statistics. Wiley.
- BIRGE, J. et LOUVEAUX, F. (1997). *Introduction to stochastic programming*. Springer Series in Operations Research Series. Springer London, Limited.
- BURKE, E., DE WERRA, D. et KINGSTON, J. (2004). Applications to timetabling. *Handbook of graph theory*, CRC Press, Boca Raton. 444–474.
- BURKE, E. K., MAREČEK, J., PARKES, A. J. et RUDOVÁ, H. (2010). Decomposition, reformulation, and diving in university course timetabling. *Computers & Operations Research*, 37, 582–597.
- CICERONE, S., D'ANGELO, G., STEFANO, G., FRIGIONI, D. et NAVARRA, A. (2009a). Recoverable robust timetabling for single delay : Complexity and polynomial algorithms for special cases. *Journal of Combinatorial Optimization*, 18, 229–257.
- CICERONE, S., D'ANGELO, G., STEFANO, G., FRIGIONI, D., NAVARRA, A., SCHACHTEBECK, M. et SCHÖBEL, A. (2009b). Recoverable robustness in shunting and timetabling. D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, R. K. Ahuja, R. H. Möhring et C. D. Zaroliagis, éditeurs, *Robust and Online Large-Scale Optimization*, Springer Berlin Heidelberg, Berlin, Heidelberg, vol. 5868. 28–60.
- CORDEAU, J.-F., TOTH, P. et VIGO, D. (1998). A survey of optimization models for train routing and scheduling. *Transportation Science*, 32, 380–404.

- CÔTÉ, M.-C., GENDRON, B., QUIMPER, C.-G. et ROUSSEAU, L.-M. (2009). Formal languages for integer programming modeling of shift scheduling problems. *Constraints*, 16, 54–76.
- CSIMA, J. (1965). *Investigations on a Time-Table Problem*. Thèse de doctorat, University of Toronto, Toronto, Canada.
- DE CAUSMAECKER, P. et VANDEN BERGHE, G. (2010). Towards a reference model for timetabling and rostering. *Annals of Operations Research*.
- DE WERRA, D. (1985). An introduction to timetabling. *European Journal of Operational Research*, 19, 151–162.
- DIMITRAKOPOULOS, R. et RAMAZAN, S. (2008). Stochastic integer programming for optimising long term production schedules of open pit mines : methods, application and value of stochastic solutions. *Mining Technology : IMM Transactions section A*, 117, 155–160.
- ÉDITEUR OFFICIEL DU QUÉBEC (2012). Adoption du principe du projet de loi 78 - loi permettant aux étudiants de recevoir l’enseignement dispensé par les établissements de niveau postsecondaire qu’ils fréquentent. *Journal des débats de l’Assemblée*, 6321–6359.
- EMERY, Y. et GIAUQUE, D. (2005). *Paradoxes de la gestion publique*. L’Harmattan, Paris.
- ERNST, A., JIANG, H., KRISHNAMOORTHY, M. et SIER, D. (2004). Staff scheduling and rostering : A review of applications, methods and models. *European Journal of Operational Research*, 153, 3–27.
- FISCHETTI, M., SALVAGNIN, D. et ZANETTE, A. (2009). Fast approaches to improve the robustness of a railway timetable. *Transportation Science*, 43, 321–335.
- GOLAMNEJAD, J., OSANLOO, M. et KARIMI, B. (2006). A chance-constrained programming approach for open pit long-term production scheduling in stochastic environments. *The Journal of The South African Institute of Mining and Metallurgy*, 106, 105 – 114.
- GOTLIEB, C. C. (1963). The construction of class-teacher timetables. *IFIP Congress 1962*. C. M. Popplewell, North-Holland, 73–77.
- HENDERSON, S. et MASON, A. (1998). Rostering by iterating integer programming and simulation. *1998 Winter Simulation Conference. Proceedings (Cat. No.98CH36274)*. Washington, DC, USA, 677–683.
- HERTZ, A. (1991). Tabu search for large scale timetabling problems. *European Journal of Operational Research*, 54, 39–47.

- KLIEWER, N., MELLOULI, T. et SUHL, L. (2006). A time-space network based exact optimization model for multi-depot bus scheduling. *European Journal of Operational Research*, 175, 1616–1627.
- LOMBARDI, M., MILANO, M., RUGGIERO, M. et BENINI, L. (2010). Stochastic allocation and scheduling for conditional task graphs in multi-processor systems-on-chip. *Journal of Scheduling*, 13, 315–345.
- LUTUKSIN, T. et PONGCHAROEN, P. (2010). Best-worst ant colony system parameter investigation by using experimental design and analysis for course timetabling problem. *2010 Second International Conference on Computer and Network Technology*. Bangkok, Thailand, 467–471.
- MCCOLLUM, B. (2006). A perspective on bridging the gap between theory and practice in university timetabling. *Proceedings of the 6th international conference on Practice and theory of automated timetabling VI*. Edmund K. Burke, Brno, vol. VI, 3–24.
- MOONEY, C. (1997). *Monte Carlo Simulation*. No. 116 Quantitative Applications in the Social Sciences. CORWIN PressINC.
- NAJDPOUR, N. et FEIZI-DERAKHSHI, M.-R. (2010). A two-phase evolutionary algorithm for the university course timetabling problem. *2010 2nd International Conference on Software Technology and Engineering*. San Juan, PR, USA.
- NGUYEN, K., NGUYEN, D., TRIEU, K. et TRAN, N. (2010). Automating a real-world university timetabling problem with tabu search algorithm. *2010 IEEE RIVF International Conference on Computing & Communication Technologies, Research, Innovation, and Vision for the Future (RIVF)*. Hanoi, Vietnam, 1–6.
- PEARL, J. (1984). *Heuristics : Intelligent search strategies for computer problem solving*. Addison-Wesley Pub. Co., États-Unis.
- POLLITT, C. et BOUCKAERT, G. (2011). *Public management reform : a comparative analysis : new public management, governance, and the neo-Weberian state*. Oxford University Press, Oxford ; New York.
- QUÉBEC (1964). Loi sur la division territoriale. L.R.Q., c. D-11.
- QUÉBEC (1965). Loi sur le ministère de la justice. L.R.Q., c. M-19.
- QUÉBEC (PROVINCE). MINISTÈRE DE LA JUSTICE (2011). *Plan stratégique 2010-2015 du ministère de la Justice du Québec*. Justice Québec, [Québec, Québec].
- RANSON, D. et AHMADI, S. (2006). An extensible modelling framework for timetabling problems. E. K. Burke et H. Rudová, éditeurs, *Practice and Theory of Automated Timetabling VI*, Springer Berlin Heidelberg, Berlin, Heidelberg, vol. 3867. 383–393.

- RASMUSSEN, R. V. et TRICK, M. A. (2008). Round robin scheduling – a survey. *European Journal of Operational Research*, 188, 617–636.
- RAVI, R. et SINHA, A. (2005). Hedging uncertainty : Approximation algorithms for stochastic optimization problems. *Mathematical Programming*, 108, 97–114.
- ROYAUME-UNI (1982). Canada act 1982. 1982 c. 11 (U.K.).
- SAHINIDIS, N. (2004). Optimization under uncertainty : state-of-the-art and opportunities. *Computers & Chemical Engineering*, 28, 971–983.
- SALWANI ABDULLAH, KHALID SHAKER, BARRY MCCOLLUM et PAUL MCMULLAN (2010). Dual sequence simulated annealing with round-robin approach for university course timetabling. *Lecture Notes in Computer Science*. Springer-Verlag, vol. 6022, 1–10.
- SCHAERF, A. (1999). A survey of automated timetabling. *Artificial Intelligence Review*, 13, 87–127.
- SHIAU, D.-F. (2011). A hybrid particle swarm optimization for a university course scheduling problem with flexible preferences. *Expert Systems with Applications*, 38, 235–248.
- SOCHA, K., KNOWLES, J. et SAMPELS, M. (2002). A MAX-MIN ant system for the university course timetabling problem. M. Dorigo, G. Caro et M. Sampels, éditeurs, *Ant Algorithms*, Springer Berlin Heidelberg, Berlin, Heidelberg, vol. 2463. 1–13.
- VERDERAME, P. M., ELIA, J. A., LI, J. et FLOUDAS, C. A. (2010). Planning and scheduling under uncertainty : A review across multiple sectors. *Industrial & Engineering Chemistry Research*, 49, 3993–4017.
- VILLAGGI, J.-P. (2005). *L'administration publique québécoise et le processus décisionnel : des pouvoirs au contrôle administratif et judiciaire*. Éditions Y. Blais, Cowansville, Québec.
- YANG, M., CAI, L. et SONG, G. (2010). Constraint satisfaction timetabling research based on course-period-template selection and conflict-vector detection. *2010 Fifth International Conference on Frontier of Computer Science and Technology*. Changchun, TBD, China, 582–588.