UNIVERSITÉ DE MONTRÉAL

ORGANISATION ET PLANIFICATION DE LA MAIN-D'ŒUVRE : APPLICATIONS EN SANTÉ ET EN INDUSTRIE

NADIA LAHRICHI
DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIAE DOCTOR (Ph.D.)
(MATHÉMATIQUES DE L'INGÉNIEUR)
JANVIER 2008

©Nadia Lahrichi, 2008.

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-46106-8
Our file Notre référence
ISBN: 978-0-494-46106-8

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée :

ORGANISATION ET PLANIFICATION DE LA MAIN-D'ŒUVRE :
APPLICATIONS EN SANTÉ ET EN INDUSTRIE

présentée par : LAHRICHI Nadia
en vue de l'obtention du diplôme de : Philosophiæ Doctor
a été dûment acceptée par le jury d'examen constitué de :

M. ROUSSEAU Louis-Martin, Ph.D., président
M. HERTZ Alain, Doct. ès Sc., membre et directeur de recherche
M. LAPORTE Gilbert, Ph.D., membre
M. WIESER Philippe, Doctorat, membre externe

REMERCIEMENTS

Mes premiers mots s'adressent à mon directeur, le professeur Alain Hertz. Je vous remercie de tout cœur d'avoir accepté de me diriger, d'avoir investi autant de temps et d'énergie dans nos travaux et enfin d'avoir été présent à tous les instants. Travailler à vos côtés ces quatre dernières années a été un plaisir, mais surtout un honneur. Je remercie le professeur Marino Widmer pour sa collaboration et ses précieux conseils. À mon mentor, le professeur Gilbert Laporte, le doctorat était votre idée, alors merci de m'avoir poussée dans cette voie.

Cette thèse n'aurait sans doute pas été possible sans l'initiative de Sophie Lapierre et la collaboration d'Amina Talib du CLSC Côte-des-Neiges, qui nous a fourni autant les données que ses précieux conseils sur la réalisabilité de notre projet. Une partie de cette thèse a bénéficié du soutien de la Régie régionale de la santé et des services sociaux de Montréal-Centre dans le cadre du programme régional élargi de soutien à l'organisation du travail. Je remercie également les membres du personnel du CRT pour leur disponibilité et leur efficacité. Je pense entre autres à Daniel Charbonneau, Luc Rocheleau, Serge Bisaillon et François Guertin.

Un merci tout particulier à mes amis Jean-François Bérubé, Benoit Crevier et Walter Rei. Votre présence ainsi que tous les moments partagés ont largement contribué à l'avancement de mon travail. Finalement je tiens à remercier Alysson Machado Costa et Marie-Ève Rancourt avec qui j'ai travaillé tous les jours côte à côte.

À mes très chers parents, frère et sœur, votre présence est ce qu'il y a de plus précieux.

RÉSUMÉ

Nous nous intéressons dans cette thèse aux problèmes d'organisation et de planification de la main-d'œuvre dans deux contextes bien particuliers, les services de soins à domicile et les industries de fabrication. Dans le milieu des services de soins à domicile il s'agit de rendre un service de qualité aux patients tout en assurant un niveau de satisfaction élevé au travail des professionnels de la santé, en particulier les infirmiers et infirmières. Les services de soins à domicile reçoivent des demandes de services de la part de patients. Ceux-ci représentent une charge plus ou moins importante suivant la complexité de leur cas. Le problème est de distribuer ces patients au personnel infirmier de façon à répartir le travail équitablement. L'équilibre des charges de travail a été reconnu dans la littérature comme un facteur de satisfaction au travail se traduisant par un service de meilleure qualité. De la même façon, la concurrence oblige les industries à non seulement être à la fine pointe de la technologie, mais également à répondre à la demande de la meilleure façon possible. Ces deux milieux connaissent de grandes pressions dues à l'irrégularité de la demande.

Dans la première partie de cette thèse nous faisons l'analyse de l'approche territoriale en pratique au CLSC Côte-des-Neiges (CLSC CDN) depuis 1980. Le principe est d'associer le personnel infirmier à des secteurs et de leur affecter toute la population de ce secteur sans tenir compte de la lourdeur ou du nombre de cas. En se basant sur la comparaison du nombre de visites rendues entre 1998-1999 et 2002-2003, nous montrons que l'approche territoriale aux services à domicile ne résiste pas aux fluctuations de la demande. Il est difficile d'évaluer le niveau réel des besoins en services car ceux-ci sont davantage déterminés par la disponibilité du personnel, créant ainsi des déséquilibres dans la répartition de la charge de travail. Une approche pour distribuer les patients plus dynamique et basée sur la charge de travail du personnel semble alors plus adaptée pour assurer autant l'équilibre entre les charges des différentes personnes que l'équilibre entre les niveaux de service reçu par les patients.

Dans la seconde partie de cette thèse nous traitons le problème de distribution des patients aux infirmières du CLSC CDN. L'objectif poursuivi est d'équilibrer les charges de travail des différentes infirmières tout en éliminant les longs déplacements. Nous élaborons une nouvelle approche développée en collaboration avec le comité directionnel du CLSC combinant approche par territoire et par charge de travail. L'assouplissement des frontières entre secteurs permet au personnel infirmier de se déplacer vers la demande. Une mesure de la charge de travail basée sur le nombre de visites, les déplacements, le nombre de cas ainsi que la lourdeur des cas a été développée. Le problème d'équilibrage des charges de travail est au centre de notre étude. Le problème est formulé comme un problème d'affectation généralisé avec plusieurs contraintes (MRGAP), problème NP-difficile, où l'objectif est de minimiser l'écart entre les charges de travail des infirmières. Certaines contraintes sont nonlinéaires et la fonction objectif quadratique, encourageant l'utilisation de méthodes heuristiques pour résoudre le problème. Nous avons développé un algorithme basé sur la méthode de recherche avec tabous utilisant des mouvements ainsi que des voisinages complexes. La simplification de la mesure de la charge de travail ramène le problème à un problème linéaire, permettant d'utiliser des solveurs tels que CPLEX.

Le dernier problème que nous abordons s'inspire d'un cas pratique d'une entreprise de fabrication où l'annualisation du temps de travail est en pratique. Cette forme de flexibilité dans la planification de la production permet de mieux répondre à la demande tout en répondant à divers objectifs tels que la satisfaction du personnel. Nous avons développé un modèle linéaire mixte suffisamment flexible pour permettre de tester divers scénarios. Les objectifs principaux sont de minimiser le nombre de semaines travaillé en temps supplémentaire, de minimiser le temps de travail annuel ainsi que minimiser la taille du personnel requis. Nous avons proposé trois stratégies principalement en plus des relaxations de certaines contraintes, telles que l'introduction du temps partiel, de l'engagement graduel et enfin l'introduction de quarts de travail de huit heures au lieu de douze heures. Toutes les variantes de ce problème sont résolues avec CPLEX et montrent que l'introduction de la flexibilité dans la planification de la production fournit de très bons résultats.

ABSTRACT

This thesis studies workforce planning in two contexts : homecare and manufacturing. In homecare, nurses have to visit clients and assure the follow-up of the cases in the best way. Equilibrium between the nurses workload is a good measure of the nurses satisfaction and service quality, pushing the workload assignment to nurses to the first rank. Following the same idea, competition pushes industries to manage workforce planning in the best way not only to fit the demand, but also to prioritize the employees satisfaction. These two contexts are under irregular and variable demand environment.

In the first part of this thesis, we analyze a territorial approach to deliver nursing home care services currently practiced at the CLSC Côte-des-Neiges site (CLSC CDN) since 1980. The territorial approach consists of assigning nurses to districts and assigning them to the population without considering their impact on the workload. Starting with the comparison of the number of visits between 1998-1999 and 2002-2003, we show that the territorial approach to deliver homecare nursing services does not resist to fluctuations in population needs for services. It is difficult to quantify the true needs for services since the availability of nursing services tend to determine the services actually delivered resulting in service delivery inequities across the entire territory. Therefore, a more dynamic assignment of clients to the nurses based on each nurse's workload and case load rather than based on the geographic location of clients is necessary to ensure a more equitable case load attribution between nurses as well as fewer inequities between clients in terms of service delivery considering their needs.

In the second part of the thesis, we consider the problem of assigning patients to nurses for CLSC CDN. The aim is to balance the workload of the nurses while avoiding long trips to visit the patients. We propose a new approach combining territorial approach and workload. The flexibility of the borders between districts is also introduced as well as a workload measure based on the number of visits, traveling, number of cases and the complexity of these cases. We model the problem as a multi-resource generalized assignment

problem, which is known to be NP-hard, where the objective is to minimize the gap between workloads. Some constraints are non-linear and the objective function is quadratic, motivating the use of heuristic methods to solve the problem. We propose a Tabu Search algorithm using complex movements and neighborhoods. A simplification of the workload measure leads to a linear mixed integer program which we optimize using CPLEX.

The last problem we study is the flexibility in multiple-shift workforce planning problem occuring in a real industrial case. The annualized hours approach is one of best ways to fit the fluctuations demand. We develop a mixed integer programming model that seeks to propose different solutions minimizing either the workforce size, the overtime or the workload. Different working arrangements are introduced as part time workers or gradual hiring approach. We finally test the impact of changing the length of shifts from twelve to eight hours. All these versions are solved using CPLEX and show that the use of a form of flexibility in workforce planning gives very good results.

TABLE DES MATIÈRES

REMERCIEMENTS . . . . . **iv**

RÉSUMÉ . . . . . **v**

ABSTRACT . . . . . **vii**

TABLE DES MATIÈRES . . . . . **xi**

LISTE DES TABLEAUX . . . . . **xiii**

LISTE DES FIGURES . . . . . **xv**

INTRODUCTION **1**

CHAPITRE 1 : REVUE DE LA LITTÉRATURE SUR L'ORGANISATION ET LA PLANIFICATION DE LA MAIN-D'ŒUVRE . **12**

1.1 Organisation du personnel et distribution du travail 12

1.1.1 Différenciation du personnel 13

1.1.2 Approches pour la distribution du travail 15

1.2 La charge de travail 17

1.2.1 Évaluation de la charge de travail du personnel infirmier 19

1.2.2 Systèmes de classification des patients (SCP) 23

1.2.3 La dotation en personnel 27

1.2.4 Limites des systèmes d'évaluation de la charge de travail infirmière . 28

1.3 Équilibrage de la charge de travail 30

1.4 La confection des horaires 33

1.5 Le problème de distribution des patients 39

1.5.1 Applications 39

1.5.2 Le problème d'affectation généralisé (GAP) 41
1.5.3 Le problème BGAP 48
1.5.4 Le GAP avec plusieurs contraintes (MCGAP ou MRGAP) 50
1.6 Résumé du chapitre 51

CHAPITRE 2 : ANALYSIS OF A TERRITORIAL APPROACH TO THE DELIVERY OF NURSING HOME CARE SERVICES BASED ON HISTORICAL DATA 53

2.1 The territorial approach 57
2.2 Fluctuations between 1998-1999 and 2002-2003 61
2.3 An in-depth analysis of the 2002-2003 period 66
2.3.1 New patients 67
2.3.2 Outside agencies 69
2.3.3 Surplus team 71
2.4 Conclusion 74

CHAPITRE 3 : A PATIENT ASSIGNMENT ALGORITHM FOR HOME CARE SERVICES 76

3.1 Introduction 79
3.2 The territorial approach at CLSC CDN 81
3.3 Workload measure 84
3.4 Patient assignment model 87
3.5 A Tabu Search algorithm 93
3.6 Experimental results 98
3.7 Dynamic assignment 107
3.8 Conclusion 110

CHAPITRE 4 : A FLEXIBLE MILP MODEL FOR MULTIPLE-SHIFT WORKFORCE PLANNING UNDER ANNUALIZED HOURS 112

4.1 Introduction 115

4.2 Problem description 117
4.2.1 A real life problem 117
4.2.2 Considered variations 118
4.3 A MILP model 119
4.3.1 Constraints 121
4.3.2 Objectives 125
4.4 Computational experiments 126
4.4.1 Solutions for the original problem 126
4.4.2 Restrictions on worktime over T consecutive weeks 132
4.4.3 Restrictions on similar week schedule types 135
4.4.4 Gradual hiring 136
4.4.5 Part time workers 138
4.4.6 Different number of shifts per day 139
4.5 Conclusion 140

DISCUSSION GÉNÉRALE, CONCLUSION ET RECOMMANDATIONS 142

BIBLIOGRAPHIE 162

LISTE DES TABLEAUX

Tableau 1 Formes de flexiblité du travail 2

Tableau 2.1 Number of visits per district for 1998–1999 and 2002–2003 63

Tableau 2.2 Evolution of the number and distribution of nurses 64

Tableau 2.3 Evolution of the number of nursing visits per district 66

Tableau 2.4 Number of different patients per month, per district, and per nurse 67

Tableau 2.5 Number of new clients received by each district in the September 2002 – March 2003 period 68

Tableau 2.6 Distribution of 2002–2003 visits 69

Tableau 2.7 Average number n of visits per day and average visit duration τ . 70

Tableau 2.8 Distribution of the visits delivered by the surplus team during the September 2002 – March 2003 period 72

Tableau 2.9 Distribution of the visits delivered to each district during the September 2002 – March 2003 period 73

Tableau 2.10 Percentage of the visits delivered by the surplus team during the September 2002 – March 2003 period 73

Tableau 2.11 Number of visits, and average visit durations 73

Tableau 3.1 Number of units, nurses and patients per district for the year 2002-2003 84

Tableau 3.2 Comparison between CPLEX and Tabu Search. 102

Tableau 3.3 Results for the case manager nurses on the whole territory. 105

Tableau 3.4 Results for the technician nurses on the whole territory. 106

Tableau 3.5 Dynamic assignment of case management nurses. 108

Tableau 3.6 Dynamic assignment of nurse technicians. 109

Tableau 4.1 Schedule for 1994 obtained with the three phase algorithm of [197]. 127

Tableau 4.2 Schedule for 1995 obtained with the three phase algorithm of [197]. 128

Tableau 4.3 MILP solution for 1994. 130

Tableau 4.4 MILP solution for 1995. 131

Tableau 4.5 Various solutions with different values of parameter λ. 133

Tableau 4.6 Various solutions with $\gamma = 3$. 135

Tableau 4.7 Various solutions with 3 shifts of 8 hours per week. 140

TABLE DES FIGURES

Figure 1 Organisation du réseau québécois de la santé et des services sociaux 5

Figure 1.1 Classification des approches permettant de mesurer la charge de travail infirmière 20

Figure 1.2 Illustration de L'indicateur de charge de cas 23

Figure 1.3 Illustration des différents échanges de tâches utilisés par Amini et Racer [28] 44

Figure 1.4 Illustration des différents échanges de tâches utilisés par Yagiura et al. [199] 46

Figure 2.1 Picture of the different districts constituting the territory of the CLSC Côte-des-Neiges 62

Figure 3.1 The six districts of the CLSC CDN territory 80

Figure 3.2 Graph G associated with the CLSC CDN territory 88

Figure 3.3 Tabu Search for the Patient Assignment Problem 98

Figure 4.1 Solutions with the original constraints. Parameter setting : $\alpha = 1, \beta = 0, \gamma = 1, \lambda = 1, T = 2$ and objective O_1. 132

Figure 4.2 Solutions with an annual bound on the workload. Parameter setting : $\alpha = 1, \beta = 0, \gamma = 1, \lambda = 3, T = 52$ and objective O_2. 134

Figure 4.3 Solutions with possibly different schedule types on each week. Parameter setting : $\alpha = 1, \beta = 0, \gamma = 3, \lambda = 3, T = 52$ and objective O_2. 136

Figure 4.4 Solutions with gradual hiring. Parameter setting : $\alpha = 0, \beta = 0, \gamma = 1, \lambda = 3, T = 52$ and objective O_3. 137

Figure 4.5 Solutions with part time workers. Parameter setting : $\alpha = 1, \beta = 2, \gamma = 1, \lambda = 3, T = 52$ and objective O_4. 138

Figure 4.6 Solutions with 3 shifts of 8 hours per week. Parameter setting : $\alpha = 1, \beta = 0, \gamma = 1, \lambda = 3, T = 52$ and objective O_2. 140

INTRODUCTION

L'organisation du travail a longtemps été synonyme de taylorisme ou de fordisme. Bien que ces modes d'organisation aient suivi la révolution industrielle et permis une accélération importante de la production ainsi que de la consommation, il n'en reste pas moins que d'autres formes ont vu le jour, notamment dans le secteur des services. La flexibilité du travail est une des réponses aux rigidités des anciens modèles d'organisation du travail connus.

Les différentes sources économiques s'accordent à reconnaître la flexibilité du travail comme étant "une politique de gestion de la main-d'œuvre mise en place dans les entreprises pour adapter la production et l'emploi correspondant aux fluctuations rapides de la demande". Cette flexibilité est de plus en plus nécessaire pour répondre autant aux changements du marché, à la concurrence, aux progrès technologiques, aux attentes en qualité de plus en plus élevées qu'à l'augmentation du chômage. L'Organisation de Coopération et de Développement Économiques (OCDE) reconnaît les impacts de cette politique depuis les années 80 et l'utilise pour comparer les performances des plus grands pays.

La flexibilité du travail concerne autant la fixation des salaires, la nature des contrats de travail, le statut du personnel, le temps de travail, les effectifs que la définition des postes de travail. On distingue plusieurs formes de flexibilité : *externe* et *interne*, *quantitative* et *qualitative* et enfin *salariale*. Chacune de ces formes repose sur différents aspects du travail et a donc des effets différents sur l'intégration et la stabilité de l'emploi. La flexibilité externe est en rapport avec les contrats de travail ou encore les effectifs. Elle ne touche pas le travail en tant que tel des employés contrairement à la flexibilité interne qui elle, est directement en rapport avec les activités. La flexibilité externe quantitative peut être atteinte en faisant varier les effectifs. Il existe différentes façons d'y parvenir, que ce soit en faisant appel aux licenciements au besoin ou en instaurant des contrats de travail précaires ou de courte durée. Cette même flexibilité externe peut être atteinte qualitativement en remplaçant les contrats de travail (plutôt qu'en y mettant fin), et ce en faisant appel

au travail différencié (travail temporaire ou contrat de travail à durée déterminée). Cette forme de flexibilité est largement répandue en France [116] où en 2003, sept embauches sur dix se sont faites sous cette forme. La flexibilité interne quant à elle du point de vue quantitatif est atteinte en faisant varier la durée des activités, que ce soit en ayant recours au temps partiel ou en annualisant le temps de travail. En 2007, la flexibilité du temps de travail reste cependant peu développée en Europe [99] où seulement un quart des employés sont concernés. La flexibilité interne qualitative (ou flexibilité fonctionnelle) fait appel à l'utilisation de la polyvalence et la mobilité des travailleurs. Certains peuvent accomplir différentes tâches et sont donc appelés à le faire. On évite ainsi la spécialisation. La dernière forme de flexibilité, la flexibilité salariale (ou flexibilité des rémunérations) recommande la fixation du salaire sur le marché du travail. Le salaire minimum est donc supprimé et dépend de la rentabilité de l'entreprise, des pressions du marché et de la performance des travailleurs. La flexibilité du travail pour être mise en place doit cependant être en accord avec les lois du travail. Le tableau 1 présente le moteur de chacune des formes de flexiblité recensées.

Tableau 1 – Formes de flexiblité du travail

Flexibilité	Externe	Interne
Quantitative	Variation des effectifs : - contrats de travail précaires - licenciements - ...	Variation de la durée des activités : - temps partiel - annualisation du temps de travail - ...
Qualitative	Remplacement du contrat de travail : - intérimaires - ...	Utilisation de la polyvalence des travailleurs : - roulement sur différentes tâches - ...
Salariale	Fixation du salaire sur le marché du travail : - suppression du salaire minimum - ...	

Toutes ces formes de flexibilité ont fait l'objet d'études et ce autant sous l'aspect ressources humaines, économique qu'opérationnel. Les formes de flexibilité externe (qualitative ou quantitative) ainsi que salariale connaissent plusieurs limites, notamment en ce qui a

trait à la stabilité de l'emploi. Le dictionnaire suisse de politique sociale estime que la flexibilité du travail "contribue à multiplier les contraintes qui pèsent sur le personnel" alors que le journal "L'Humanité", tout en titrant en 2004 "La flexibilité du travail favorise l'emploi", considère que "le fort niveau des contrats précaires [...] a alimenté la hausse du chômage". Cependant, et malgré ces côtés négatifs, la flexibilité du travail reste non seulement une solution pour les employeurs, mais est aussi souhaitée par les travailleurs.

Nous allons dans cette thèse nous intéresser à l'aspect opérationnel des formes de flexibilité interne et externe dans deux contextes bien précis où la performance des travailleurs est de plus en plus le seul critère de qualité : le milieu de la santé en soins à domicile et le milieu de l'industrie de la fabrication. L'industrie des soins de la santé *comprend les établissements dont l'activité principale consiste à dispenser des soins des santé en assurant des services diagnostiques et thérapeutiques, à soigner des bénéficiaires internes pour des raisons médicales et sociales, et à offrir à ceux qui en ont besoin une aide sociale. Ce secteur comprend les sous-secteurs de services de soins ambulatoires, les hôpitaux, les établissements de soins infirmiers et de soins pour bénéficiaires internes et l'assistance sociale* (Industrie Canada). Le secteur de l'industrie de la fabrication lui, *comprend les établissements dont l'activité principale est la transformation de matières ou de substances en nouveaux produits par des procédés mécaniques ou physiques* (Industrie Canada). Ces deux secteurs représentent plus de 31% et 6% respectivement du PIB de l'économie canadienne (évalué à 1012,7 milliards de dollars chaînés de 1997 en 2004) [24].

Les formes d'organisation du travail dans le milieu de la santé dépendent très fortement des pays et des systèmes de santé en place. Nous avons recensé trois systèmes de santé : privé, public ou mixte (privé et public). Les systèmes de santé sont financés en majorité par des sources publiques dans les pays de l'OCDE, à l'exception des États-Unis, du Mexique et de la Grèce où les sources de financement sont majoritairement privées. Le Brésil est l'un des pays ayant adopté le système de santé mixte. Il n'existe pas de "meilleur" système dans la mesure où ils diffèrent dans les grands pays de l'OCDE et où, surtout, il devient de plus en plus difficile de quantifier la qualité des services de santé étant donné le vieillissement généralisé de la population.

Au Canada [5], l'accès public aux soins médicaux date de 1957. La *loi canadienne sur la santé* [12] stipule qu'en plus d'avoir accès aux services hospitaliers, médicaux et de chirurgie dentaire médicalement nécessaires, "les services complémentaires de santé" tels que les services de soins de longue durée en établissement et les composantes "santé" des soins à domicile et des soins ambulatoires sont aussi couverts. Chaque province a la responsabilité de gérer son réseau de soins de santé. Au Québec, le Ministère de la Santé et des Services Sociaux (MSSS) a la responsabilité de rendre accessible l'ensemble des services de santé et de services sociaux à la population en vue d'améliorer, restaurer ou maintenir sa santé. Cette responsabilité est coordonnée avec les autorités des 18 régions qui composent le Québec. Ces autorités généralement appelées *agence* ou *réseau régional de services de santé et de services sociaux* ont la charge de coordonner et de mettre en place les différents services de santé et de services sociaux dans leur région que ce soit en gérant les aspects de financement et de répartition des ressources humaines ou en facilitant l'accès à ces services. Chaque territoire d'agence ou de réseau régional est lui même divisé en sous-territoires au cœur desquels opère le Centre de Santé et de Services Sociaux (CSSS). Chacun de ces CSSS est le résultat du regroupement de CLSC (Centre Local de Services Communautaires), de CHSLD (Centre d'Hébergement et de Soins de Longue Durée) et de CHSGS (Centres Hospitaliers de Soins Généraux et Spécialisés). Les soins de santé à domicile sont au Québec coordonnés par les différents CLSC. La figure 1 permet de visualiser plus en détail l'organisation du réseau québécois de la santé et des services sociaux.

L'organisation des soins de santé à domicile diffère selon les pays. En France par exemple, on fait la distinction entre les SSIAD (services de soins infirmiers à domicile) et l'HAD (hospitalisation à domicile). Les SSIAD ont pour mission de contribuer au soutien à domicile des personnes, notamment en prévenant, ou différant, l'entrée à l'hôpital ou dans un établissement d'hébergement et en raccourcissant certains séjours hospitaliers. Ils interviennent au domicile des patients ou dans les établissements non médicalisés prenant en charge des personnes âgées ou des personnes handicapées. L'HAD quant à elle, est une alternative à l'hospitalisation traditionnelle. L'HAD permet au patient de bénéficier de soins médicaux et paramédicaux continus et coordonnés, dans son environnement fa-

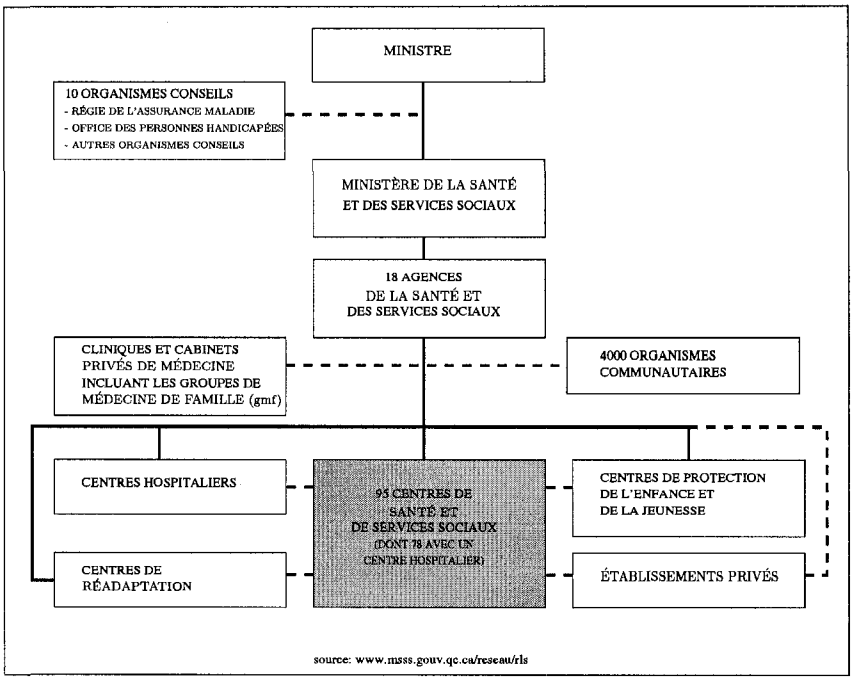

Figure 1 – Organisation du réseau québécois de la santé et des services sociaux

milial et quotidien. Ces soins sont prodigués pour une période théoriquement limitée mais révisable selon l'évolution de l'état de santé du patient. Ils se distinguent des autres soins dispensés à domicile par leur fréquence et leur technicité.

En Italie, les soins à domicile sont aussi sous la responsabilité des municipalités et sont financés par elles. Les services de soins à domicile intégrés (*Integrated Home Care Service*) ont fourni des services à 240 105 personnes d'après un rapport de 2003 [165]. La demande en soins à domicile en Italie a augmenté bien plus que l'offre de services. Le pourcentage de personnes agées de plus de 65 ans qui utilisent ce service est très bas (environ 1% de la population) comparativement au Royaume-Uni (5,5%), à l'Allemagne (6,5%) ou encore à la Scandinavie (presque 10%) [174]. En 1991, 15% de la population en Italie a un âge supérieur à 65 ans, et le segment des au-dessus des 80 ans s'accroît rapidement (ce sont les principaux utilisateurs des soins à domicile à long terme) [150].

Aux États-Unis plus de 2,4 millions de personnes âgées ou handicapées couvertes par l'assurance maladie (*Medicare*) ont reçu des soins à domicile de la part de plus de 8 100

agences de soins à domicile certifiées par *Medicare* en 2006 [2]. Les maladies chroniques comptent pour 75% du budget des dépenses en santé [153]. Les soins à domicile sont couverts par *Medicare*. Les soins consistent en des soins à temps partiel ou médicalement nécessaires prescrits par un médecin, les soins à long terme ne sont par conséquent pas couverts par l'assurance maladie.

Au Luxembourg il existe deux grands réseaux de soins à domicile : Help et Hellef Doheem. Ils ne sont pas des autorités publiques mais des réseaux privés puisqu'il n'existe pas de services de soins à domicile publics. Ces réseaux doivent répondre à deux types de soins à domicile financés différemment : les soins à domicile liés à une maladie et pris en charge par l'assurance maladie et les soins à domicile liés à la dépendance payés par l'"Assurance Dépendance".

En Belgique, l'hospitalisation à domicile est encore en projet de reconnaissance légale comme étant une structure de service de soins. La loi concernant cette alternative de soins était encore à l'étude au Sénat [19].

Le recours aux services de soins à domicile comme alternative à l'hospitalisation ou comme outil de prévention a montré des gains considérables d'après des études faites au Danemark, au Canada et aux États-Unis. En Ontario, un investissement de 300 millions par année dans les services à domicile a permis une économie évaluée à 1800 millions en coûts de capital et 500 autres en coûts opérationnels annuels. Le rapport Romanow de la commission sur l'avenir des soins de santé au Canada [3] décrit les soins à domicile comme étant le prochain service essentiel dès 2002. "Les dépenses totales par habitant engagées par le secteur public pour les services de soins à domicile (soins de santé et soins de soutien) ont augmenté de 54,75$ par habitant en 1994-1995 à 93,60$ en 2003-2004" d'après le rapport sur les dépenses du secteur public et utilisation des services à domicile au Canada [4]. La part des dépenses en services à domicile consacrée aux soins de santé à domicile a grimpé de 43,3% en 1995-1996 à 48,6% en 2003-2004 au Canada. Pour les provinces qui font la distinction entre les deux types de soins (santé et soutien), les chiffres montrent qu'il y a eu une augmentation moyenne de 5,5% des dépenses dans le cas des soins de santé à domicile (contrairement à 2,7% d'augmentation pour le soutien à domicile)

entre 1995-1996 et 2003-2004. La province de Québec a enregistré une hausse de 11,4% des dépenses pour sa part. Cette province enregistre aussi avec le Nouveau-Brunswick le plus grand nombre d'utilisateurs de soins de santé à domicile sur 1000 habitants (19,2 environ). Sur l'île de Montréal en particulier, 571 107 interventions en soins infirmiers à domicile ont été enregistrées pour l'année 2003-2004 et 649 050 pour l'année 2005-2006 Le nombre de personnes desservies aux soins infirmiers à domicile est passé de 46 481 en 2002-2003 à 49 580 en 2005-2006 [16]. L'augmentation attendue du nombre d'interventions due au nombre croissant de personnes âgées (15% de la population en 2001) et aux politiques visant le séjour à court terme dans les hôpitaux pose de plus en plus la question de l'efficacité dans les soins infirmiers à domicile. De plus des études [44] montrent que les soins infirmiers consomment plus de 40% du budget de l'hôpital, et que combinés aux soins d'assistance, ils occupent de 60 à 80% des coûts dans un établissement de santé [95]. Il faut donc porter une attention toute particulière aux employés de ce secteur en leur offrant une organisation adéquate du travail.

Le second milieu que nous proposons de traiter en rendant plus efficace l'organisation du travail, est le milieu de la fabrication. Celui-ci a subi d'importants changements dus à l'introduction des différents contrats de travail visant à diminuer le chômage, ou simplement à introduire plus de flexibilité sur le marché. En France par exemple, les entreprises industrielles privilégient de plus en plus le travail par intérim puisque l'aspect court terme de ce genre d'emploi permet de répondre aux fluctuations de la demande. En vingt ans dans l'industrie, la part des emplois à court terme a été multipliée par trois [10]. Les emplois temporaires sont de plus en plus fréquents. Que ce soit le travail à temps partiel, saisonnier, le travail sur appel, le travail temporaire ou encore l'intérim, tous ces statuts sont reconnus par les lois régissant les contrats de travail dans les pays industrialisés. Par conséquent le recours à ces formes de travail dépend uniquement des entreprises et de l'acceptation du contrat par les travailleurs.

Les différentes formes de contrat de travail ne sont cependant pas appropriées pour tous les contextes d'industrie manufacturière. Dans certains contextes où les travailleurs utilisent une technologie de pointe, il est rare de voir apparaître des contrats par intérim.

La formation des employés est continue et nécessite un suivi constant. Plutôt que d'engager du personnel supplémentaire à la demande, les entreprises préfèrent surévaluer le nombre de travailleurs requis. Lorsque l'entreprise a la possibilité de recourir au stockage ou de reporter sa demande, il est moins critique d'engager des employés à court terme et varier les contrats de travail. Dans l'autre contexte où il est impossible d'accumuler des stocks et où le temps de réaction est court, il est critique de déterminer de façon efficace la taille de l'équipe de travail ainsi que les contrats à mettre en place. Nous proposons d'étudier l'impact de la forme de flexibilité concernant les contrats de travail sur les opérations de l'entreprise. Nous faisons également appel à l'annualisation du temps de travail. Cette forme de flexibilité a déjà fait ses preuves dans plusieurs contextes où la demande est variable. Le principe de l'annualisation du temps de travail est de limiter le nombre d'heures travaillées sur l'année plutôt que sur une période d'une ou deux semaines. Cette annualisation du temps de travail permet de dépasser le nombre d'heures travaillées certaines semaines (équivalent du temps supplémentaire), mais de le compenser par moins d'heures de travail d'autres semaines (récupération du nombre d'heures) plutôt que par une compensation monétaire.

Dans chacun des deux contextes que nous voulons traiter, l'organisation de la main-d'œuvre est au cœur du sujet. Dans le contexte des soins à domicile, il faut non seulement rendre des visites aux patients afin de leur prodiguer le service requis, mais aussi assurer le suivi de leur cas avec les autres professionnels de la santé. Cette forme de prise en charge est connue sous le nom de "gestion de cas" et permet d'assurer la continuité des services [9]. La gestion de cas se définit *comme une stratégie concertée associant le client en vue de la prestation de services de santé et de soutien de qualité au travers de l'utilisation efficace et efficiente des ressources afin d'appuyer l'atteinte des objectifs du client* [8]. La gestion de cas est pratiquée partout au Canada [17] et est définie comme un "service essentiel dans le panier des soins à domicile" par les gouvernements fédéral, provinciaux et territoriaux en 2004 pour consolider les soins de santé [8]. Cette organisation augmente les responsabilités des dispensateurs de services (personnel infirmier et travailleurs sociaux en particulier) ([6, 100, 147]). À chaque membre du personnel infirmier est affecté un certain nombre de patients auquel il doit autant rendre les visites qu'assurer son suivi. Le problème

consiste donc à trouver la meilleure approche pour distribuer les patients à chaque membre du personnel infirmier. Les quatre approches principales sont par territoire, par "charge de travail", par "consensus" ou encore l'approche par rotation. L'approche territoriale consiste à associer chaque membre du personnel à un secteur géographique. Chacun est alors en charge des habitants de son secteur. L'approche par "charge de travail" nécessite d'introduire la signification de la charge de travail. Il n'existe pas de définition formelle de la charge de travail. Globalement elle regroupe l'ensemble des soins directs (en présence du patient) et indirects (coordination du suivi) relatifs à l'ensemble des patients, mais peut inclure d'autres attributs. La méthode par "consensus" nécessite la collaboration de tous les membres du personnel en présence qui doivent décider qui doit prendre en charge les nouveaux patients. Cette collaboration se traduit par des réunions hebdomadaires où se décident le partage et le transfert de cas (certains cas sont pris en charge avant les réunions à cause de l'urgence des soins à rendre et doivent être transférés à la "bonne" personne). La dernière approche que nous avons appelée l'approche par "rotation" est une distribution des patients à tour de rôle. Le personnel reçoit chacun son tour un patient, peu importe la charge qui lui est associée. La majorité des CLSC (responsables des soins à domicile au Québec) ont adopté l'approche populationnelle (qui consiste à gérer la population de leur territoire) en privilégiant le modèle de "soins et services intégrés". Ce modèle prévoit la formation de réseaux intégrés et a recours à la gestion de cas. Cette approche ainsi que ce modèle font partie des recommandations du ministère [18] pour réorganiser les services afin de faciliter le parcours des personnes dans le réseau de la santé et des services sociaux.

Cette thèse par articles est organisée en quatre chapitres. Le premier chapitre est une revue de la littérature traitant des divers aspects de l'organisation et de l'allocation de la main-d'œuvre dans les deux contextes que nous avons choisi de traiter. Dans un premier temps nous présentons diverses organisations de la main-d'œuvre, particulièrement dans le contexte des soins à domicile. Nous traitons ensuite le concept de charge de travail en décrivant divers systèmes permettant de la mesurer. Diverses méthodes pour équilibrer la charge de travail sont ensuite décrites. L'allocation de la main-d'œuvre est finalement abordée, d'abord par l'allocation à des quarts de travail et enfin par l'allocation à une charge de travail. Les travaux reliés au problème d'allocation à des quarts de travail sont

présentés par rapport aux flexibilités qu'ils introduisent. Les différentes méthodes permettant de résoudre le problème d'allocation de la charge de travail sous forme d'un problème d'affectation généralisé avec plusieurs contraintes sont présentées.

Le second chapitre est une analyse de l'approche territoriale en pratique au CLSC côte-des-Neiges (CLSC CDN) de la ville de Montréal. Cette analyse a montré que bien que l'approche territoriale ait de nombreux avantages, elle présente des lacunes reliées principalement aux variations de la demande. Le volume de travail par secteur varie sans être suivi par une redistribution du personnel infirmier. Certaines personnes se retrouvent donc surchargées. Or, l'équilibre de la charge de travail a été reconnu dans [183] comme l'un des éléments clés de la satisfaction du personnel infirmier.

Dans le chapitre trois nous présentons l'algorithme basé sur la méthode de recherche avec tabous de distribution des patients au personnel infirmier. Nous avons, en collaboration avec le comité décisionnel du service de soins à domicile, développé une nouvelle approche, qui consiste à distribuer les cas suivant l'approche territoriale mais en tenant compte de la charge de travail des infirmières (nous utiliserons le féminin tout au long de cette thèse bien qu'il y ait des infirmiers de sexe masculin). Cette charge de travail doit être équilibrée entre les différentes personnes. Nous avons introduit une mesure de la charge de travail au CLSC CDN qui a été validée par les personnes en place. Le problème de distribution des patients au personnel infirmier se formule comme étant un problème d'affectation avec plusieurs contraintes de capacité (MRGAP), où chaque contrainte de capacité est une composante de la charge de travail. Notre objectif est de minimiser l'écart entre les différentes composantes de la charge de travail. La fonction objectif est quadratique et certaines contraintes nonlinéaires. Dans une première approche nous avons simplifié le problème en retirant une composante de la charge infirmière. Le modèle devient linéaire et peut être résolu de façon exacte par CPLEX. Pour résoudre le modèle complet, nous avons utilisé une approche heuristique, la méthode de recherche avec tabous.

Le chapitre quatre présente une application de l'approche par programmation linéaire et en nombres entiers pour résoudre le problème d'annualisation du temps de travail. Cette approche d'annualisation du temps de travail est un exemple typique d'organisa-

tion de la main-d'œuvre dans un contexte où la fluctuation de la demande est importante. Un avantage certain de cette méthode est d'éviter de payer les heures supplémentaires du point de vue des employeurs, et de faire moins d'heures les semaines moins chargées du point de vue des employés (ils compensent le travail supplémentaire en travaillant moins les autres semaines). Le problème consiste à faire la planification de la production en distribuant les horaires de travail à chaque employé. Ces horaires incluent la distribution des semaines de vacances.

Nous avons repris une application dans le contexte d'une entreprise suisse où les opérateurs de machine couvrent des quarts de travail de 12 heures. Tous doivent respecter le même type d'horaire, c'est-à-dire que tous doivent faire du temps supplémentaire en même temps ou encore travailler en temps régulier. Nous avons formulé le problème sous forme d'un programme linéaire et en nombres entiers que nous avons résolu avec CPLEX. Nous avons tout d'abord introduit des relaxations relatives au temps de travail ainsi qu'au type d'horaire (régulier ou temps supplémentaire). Nous avons ensuite introduit deux nouvelles stratégies : l'engagement graduel des employés et le temps partiel. Finalement nous avons évalué l'impact du passage à des quarts de travail de huit heures au lieu de douze. L'objectif de cette étude est de fournir différentes solutions de qualité à la direction pour guider leur choix.

Les chapitres deux, trois et quatre sont la reproduction d'articles en anglais, publiés ou soumis dans des revues scientifiques.

Un chapitre de conclusion clôt cette thèse.

CHAPITRE 1 : REVUE DE LA LITTÉRATURE SUR L'ORGANISATION ET LA PLANIFICATION DE LA MAIN-D'ŒUVRE

La planification de la main-d'œuvre est un processus en plusieurs étapes. Ces étapes diffèrent suivant les contextes, mais on peut les décrire comme suit. La première étape consiste à faire la dotation en personnel. Plus précisément, elle consiste à déterminer le nombre de personnes requises pour répondre à la demande. L'étape de confection des horaires et d'allocation des ressources sont souvent faites en parallèle. Des horaires sont générés respectant les conventions collectives et de façon à satisfaire la demande. Les horaires sont ensuite distribués au personnel. Dans certains contextes, comme celui des soins à domicile, les employés reçoivent une charge de travail, qu'ils sont libres de planifier eux-mêmes plutôt qu'un horaire préétabli.

Ce chapitre débute par la présentation de deux organisations de la main-d'œuvre, par spécialité ou non différenciée ainsi que de trois formes de distribution du travail. Dans la seconde section nous présentons divers systèmes de mesure de la charge de travail qui est un élément clé au moment de la dotation en personnel et de la distribution du travail. La troisième section décrit des méthodes d'équilibrage de la charge de travail. Les deux sections suivantes traitent de l'allocation des ressources, par la confection d'horaires tout d'abord, suivie de la distribution de la charge de travail.

1.1 Organisation du personnel et distribution du travail

Avant de débuter les opérations de planification de la main-d'œuvre, il convient de décider de l'organisation du personnel ainsi que du mode de distribution du travail.

1.1.1 Différenciation du personnel

Le choix de l'organisation du personnel par spécialité ou non différenciée dépend du contexte. L'organisation non différenciée ne fait pas de distinction entre les tâches qui peuvent être accomplies par les différentes personnes. Toute personne est habilitée à accomplir toute tâche. L'organisation par spécialité, par contre, fait cette distinction en particulier lorsque des aptitudes particulières sont demandées. Dans le contexte manufacturier, il est fréquent de faire appel à de la main-d'œuvre hiérarchique, c'est-à-dire qu'un employé A peut accomplir la tâche A, un employé B peut accomplir les tâches A et B, etc. Ce problème se traduit alors par différents niveaux d'efficacité pour accomplir les mêmes tâches.

Dans le domaine de la santé, l'organisation des soins infirmiers dépend de l'infrastructure où ils sont dispensés. Les infrastructures les plus usuelles sont les hôpitaux, les cliniques de jour ou à domicile. Nous ne tiendrons pas compte des cliniques dans notre analyse dans la mesure où le service infirmier rendu est ponctuel et ne nécessite pas de suivi, contrairement aux hôpitaux ou aux services à domicile.

Dans les établissements hospitaliers, les soins infirmiers sont essentiellement organisés autour des unités de soins, qui elles-mêmes, dépendent de services professionnels où sont prodigués les soins et les services aux patients.

C'est à la direction des soins infirmiers dans chaque établissement que revient la gestion de plusieurs aspects entourant la pratique infirmière. En plus de la gestion de la qualité, elle s'occupe de la gestion des activités et définit la philosophie des soins. La gestion des activités consiste à faire la distribution des soins infirmiers ainsi que la planification et évaluation des soins en fonction des besoins identifiés. Elle assure la coordination des activités de l'établissement, et gère les ressources humaines, matérielles et financières sous sa gouverne.

Pour ce qui est des services de soins à domicile, il faut tout d'abord faire la différence entre les organismes qui offrent eux-mêmes les services à domicile, les agences ou les infirmières de profession libérale. Les agences sont des organismes qui mettent leur personnel

à disposition des organismes publics et qui ne gèrent donc pas les horaires ou l'organisation du travail de ces personnes. Les infirmières de profession libérale font aussi leur propre horaire et ne sont sujettes à aucune organisation préétablie. Cette forme d'organisation est particulièrement répandue en France, où les personnes âgées font appel à l'infirmière de leur choix et où leur organisme responsable des soins à domicile règle les frais (exemple du service de Beaumont-sur-Oise qui couvre sept communes sans offrir de soins infirmiers directement). En Suède, une combinaison du service public (le plus commun) et privé (contracteur) est offerte par exemple dans la banlieue de Stockholm.

Lorsque le personnel infirmier dans les services de soins infirmiers est différencié, la durée des soins et la technicité de ces soins sont les principaux critères de séparation. Les patients en phase postopératoire et post-hospitalisation sont considérés comme des patients à court terme. La durée de leur épisode de soins est déterminée à l'avance. Les patients en perte d'autonomie, en phase terminale ou atteints de maladies chroniques ont un épisode de soins de plus longue durée (moyen ou long terme). Certains services de soins à domicile se sont donc pourvus d'équipes à court terme et d'équipes à moyen et long terme. D'autres services de soins à domicile se sont plutôt dotés d'équipes spécialisées telles que les équipes d'évaluation ou encore les équipes responsables des cas en soins palliatifs.

Le CLSC CDN par exemple, fait la distinction par la durée des soins infirmiers : les infirmières "gestionnaires de cas" gèrent les cas à long terme, les cas en perte d'autonomie, les cas complexes de coordination de services et les soins palliatifs tandis que les infirmières soignantes gérent les autres cas qui sont théoriquement du court-terme (postopératoire, post-hospitalisation, intervention ponctuelle en consultation). Le CLSC St-Laurent s'est doté d'une équipe d'infirmières à court terme qui assure les visites de fin de semaine et les visites de grande fréquence. D'autres CLSC enfin privilégient une approche mixte, c'est-à-dire une organisation par équipe pour certaines spécialités et centralisée pour d'autres. Au CLSC Côte-des-Neiges par exemple, des équipes multidisciplinaires sont en place dans chaque district, mais une seule nutritionniste est disponible et se charge de tout le territoire. Le CLSC Verdun s'est doté de dix infirmières gestionnaires de cas dans les équipes de secteur, deux infirmières dans l'équipe d'évaluation ainsi qu'une infirmière responsable.

1.1.2 Approches pour la distribution du travail

Trois modes de distribution du travail dans les unités de soins en hôpital sont recensés : l'affectation à des zones, l'affectation à des patients (lits) ou une combinaison des deux. Lorsque les infirmières sont affectées à des zones, elles prennent en charge tous les patients de ces zones. La taille, le nombre ou la charge de ces zones sont variables et il n'existe pas de norme permettant de les fixer. D'autres critères entrent en compte comme les maladies contagieuses par exemple. Tous les malades sont regroupés et une seule et même infirmière est en charge peu importe le niveau de soins requis (ce nombre est revu à la hausse en cas extrême). Lorsque les infirmières sont affectées à des patients, l'assistante infirmière évalue de son mieux la quantité de soins nécessaire.

Chaque réseau de services de soins à domicile est autonome pour décider de la répartition du travail entre le personnel infirmier. Il n'y a pas d'outil officiel pour gérer cette répartition des activités entre le personnel, que ce soit au Québec ou à l'étranger. L'approche est empirique et intuitive et dépend de chaque établissement. Cependant plusieurs normes sont répandues sans être obligatoires.

Chaque service de soins à domicile gère les soins infirmiers ainsi que leur organisation dans son territoire de façon autonome. Certains organismes privilégient le découpage de ce territoire, où les patients sont affectés à l'infirmière de leur quartier. En effet cette approche est très utilisée en milieu urbain comme dans plusieurs contextes (découpage électoral, découpage de territoires de vente, ...) pour ses avantages connus de simplicité à gérer. Cette approche territoriale est privilégiée par plusieurs CLSC : Côte-des-Neiges (CDN), Notre-Dame-de-Grâce (NDG)/Montréal-Ouest, Lac St-Louis, Mercier-Anjou, Pointe-aux-Trembles/Montréal-Est, Ahunstic, Verdun, etc. La distribution des patients selon l'approche territoriale est pratiquée également, par exemple, à Danderyd en Suède [88], dans une maison de soins infirmiers (*nursing home*) en Alabama aux États-Unis [38] ainsi qu'en France (que ce soit pour les organismes dispensateurs de soins à court terme ou à long terme).

Dans les autres cas cette affectation est soit effectuée suivant la "charge de travail" du personnel, soit "à tour de rôle", ou enfin est le fruit d'un "consensus". La deuxième méthode d'affectation distribue les cas suivant les charges de chaque infirmière. Une réunion est généralement prévue où le (ou la) responsable désigne la personne adaptée. Cette estimation de la charge n'est cependant pas basée sur des outils ou une mesure quantitative, elle est le fruit du jugement professionnel et de l'expérience de la personne en charge. Dans le troisième cas les infirmières prennent en charge à tour de rôle les patients qui entrent dans le système (sans tenir compte de leur position géographique ou de leur charge actuelle). La méthode du consensus enfin nécessite la collaboration de tous les membres du personnel en présence qui doivent décider qui prend en charge les nouveaux patients. Le CLSC Ahunstic par exemple privilégie ces trois formes en même temps. Son territoire est divisé en trois secteurs. L'attribution des cas est faite dans chaque secteur sous la supervision du chef de programme : dans le premier il consulte les membres du personnel infirmier pour arriver à un consensus ; dans le second, les infirmières reçoivent les cas à tour de rôle, alors que dans le troisième il tient compte de la charge des infirmières.

Chaque modèle d'organisation des soins à domicile a ses avantages et ses inconvénients. L'analyse de l'approche territoriale au CLSC CDN, présentée en détail au chapitre deux, montre que bien qu'étant très avantageuse de par sa simplicité et son efficacité à gérer, elle présente un manque de robustesse dans un contexte où la demande est changeante. Malgré les inégalités constatées, le fonctionnement actuel (territorial) donne beaucoup d'autonomie à l'infirmière qui peut réaliser des plans d'intervention de grande qualité avec son client : elle connaît bien son territoire grâce à l'approche communautaire. C'est une approche voulue et encouragée par le CLSC. Nous avons montré que pour pallier à cette variation dans ce contexte, il faut privilégier une approche flexible qui s'adapte aux fluctuations de la demande. La charge des infirmières est le critère le plus important de mesure de la satisfaction des patients (à cause de la qualité du service rendu) et des infirmières elles-mêmes. Il est donc nécessaire de développer des outils qui vont évaluer cette charge et d'adopter une approche qui en tienne compte.

1.2 La charge de travail

Comme nous l'avons souligné précédemment, il n'existe pas de définition formelle de la charge de travail : elle dépend du contexte. Dans le cas des soins infirmiers, la charge de travail inclut généralement la *charge de cas*. Celle-ci désigne l'ensemble des services directs et indirects que le professionnel doit offrir aux clients dans le cadre de sa fonction tandis que la charge de travail consiste en l'ensemble des activités professionnelles et administratives exécutées par le professionnel dans l'exercice de sa fonction. Les *soins directs* sont les actes de soins effectués en présence du malade par le personnel soignant ou aide-soignant. La durée est proportionnelle au nombre de malades et à la gravité de leur état. Les *soins indirects* eux sont les actes de soins effectués en dehors de la présence du patient par les soignants. Par exemple, au CLSC du Vieux Lachine [45], la lourdeur d'une charge de cas est évaluée par la complexité des patients et la fréquence des interventions. Lors du découpage du territoire du CLSC CDN en secteurs en 2002 [42], l'équilibrage de la charge de travail entre chacun des secteurs a été basé sur des données statistiques sur le nombre et la durée des visites effectuées par tous les professionnels. Certains font aussi la distinction entre la charge de soins requis par l'état du patient et ceux réellement offerts. Pour d'autres enfin, la charge de travail ne tient compte que de la charge récurrente ; toutes les tâches temporaires sont donc exclues.

Dans des contextes autres que les soins à domicile, la charge de travail est synonyme de volume de travail. Celui-ci est évalué comme étant le nombre de palettes à décharger dans un terminal d'avions cargos [117], comme le nombre de tâches effectuées par des véhicules autoguidés [137] ou encore par la durée des opérations [135, 203].

L'évaluation de la charge de travail a plusieurs intérêts. Elle est souvent utilisée pour prendre des décisions administratives (telles que le financement de l'établissement de soins) [74], et sert d'indicateur pendant la planification des ressources. Par ailleurs, elle s'inscrit dans le programme du Ministère de la Santé et des Services Sociaux qui a initié en 2002 un programme d'organisation du travail. Plusieurs CLSC ont dans ce cadre entamé une procédure afin de trouver un recours à la surcharge de travail dénoncée par l'ensemble du personnel infirmier. Au CLSC Ahunstic un rapport de 2001 sur la réorganisation des

services de soins infirmiers des services de soutien à domicile [22] montre une réelle prise de conscience du fait qu'il faut réviser les façons de faire et les processus. Au CLSC LaSalle [25], un comité a été formé dès novembre 2000 afin de développer des outils de mesure de la charge de travail des infirmières des soins à domicile. Au CLSC du Vieux Lachine [45] un comité a été mis sur pied en juin 2001 pour trouver des pistes de solutions à la surcharge de travail dénoncée par l'ensemble des infirmières. Les CLSC Verdun [21] et Côte-des-Neiges [46] se sont impliqués aussi dans cette nécessité de réorganisation. Un projet du même type a été mis en place en Nouvelle-Zélande [166] dans un service de soins intensifs en hôpital.

L'évaluation de la charge de travail permet par la même occasion d'identifier les unités (ou les personnes) surchargées. Powers [176] souligne que la surcharge de travail diminue la qualité des soins prodigués. Drennan [81], Catterson [58] ainsi que Farnham et al. [90] soulignent que la charge de travail équitable est essentielle pour une qualité de soins optimale. Dans le cas où cette surcharge ne peut être évitée (en cas de pénurie de main-d'œuvre par exemple), il convient d'atteindre un équilibre entre les charges des différentes infirmières. Du point de vue clinique, la mesure de la charge de travail permet une uniformisation des processus de soins et une surveillance de la qualité des soins.

La méthode la plus commune et la plus ancienne de mesure de charge de travail est le jugement professionnel et la perception de la personne responsable (en général une infirmière) [30]. Celle-ci, en consultation avec son supérieur, détermine le niveau d'effectif nécessaire en se basant sur les spécialités et le nombre de cas, alors que le niveau de soins requis par les patients est mesuré sur une base quotidienne d'après le jugement professionnel. Deux aspects du travail infirmier doivent être quantifiés : le temps direct de soins et le temps indirect. Les études visant à les évaluer sont rares [85]. La difficulté est de s'assurer que la complexité et le niveau d'intensité du travail infirmier soient bien reflétés [30].

Des formes de mesure de la charge de travail plus systématiques ont tout de même été développées. La charge de travail infirmier est soit mesurée directement par des systèmes de mesure de la charge de travail, soit estimée par des systèmes de classification des patients.

1.2.1 Évaluation de la charge de travail du personnel infirmier

Les différentes approches pour mesurer les charges en soins infirmiers peuvent être catégorisées soit en approches implicites ou explicites [188]. La figure 1.1 fait une synthèse de ces approches. Les approches implicites sélectionnent, à partir du patient, les variables pertinentes qui déterminent ainsi son profil. Trois catégories de client sont retenues : complètement déficient, partiellement déficient ou non déficient. Les approches explicites quant à elles utilisent plus de données. Deux types d'approches explicites sont retenues : les approches directes et indirectes. Dans le type d'approche directe, les niveaux de soins d'un patient sont déterminés par l'addition des temps (exprimés en minutes) estimés des actions de soins identifiées et requises. Les approches indirectes elles, ne traduisent pas les actions directement, mais déterminent un forfait de soins en heures par jour pour chaque patient appartenant à une catégorie donnée. Ces catégories dépendent de l'approche indirecte choisie. Il existe plusieurs approches indirectes mais qui se basent toutes sur les variables pertinentes reliées au patient. Une première approche indirecte, dite approche par classification, permet de catégoriser les patients directement par rapport à leur état de santé. Les trois approches suivantes catégorisent chaque patient selon son niveau d'handicap, d'indépendance ou de sévérité de la maladie.

L'approche directe est plus riche en information que l'approche indirecte. Elle permet plus de précision quant à l'évaluation des différents types de charge décrits précédemment.

Une autre manière de classifier les méthodes pour évaluer la charge de travail des infirmières au Royaume-Uni est introduite par Hughes dans [118]. L'auteur les différencie selon qu'elles sont basées sur les activités (de soins requises par un patient) ou sur la dépendance (de ce patient). L'état de dépendance d'un patient est "un état de santé qui nécessite de l'aide pour que soient assurés les actes essentiels de la vie quotidienne" [89]. La méthode basée sur les activités affecte à chaque activité un temps et la somme de ces temps donne la charge du système. La méthode basée sur la dépendance elle, affecte chaque patient à une classification à laquelle est associé un temps exprimé en minutes. Une autre méthode, cette fois statistique, permet en se basant sur les données historiques d'établir des prévisions concernant les besoins.

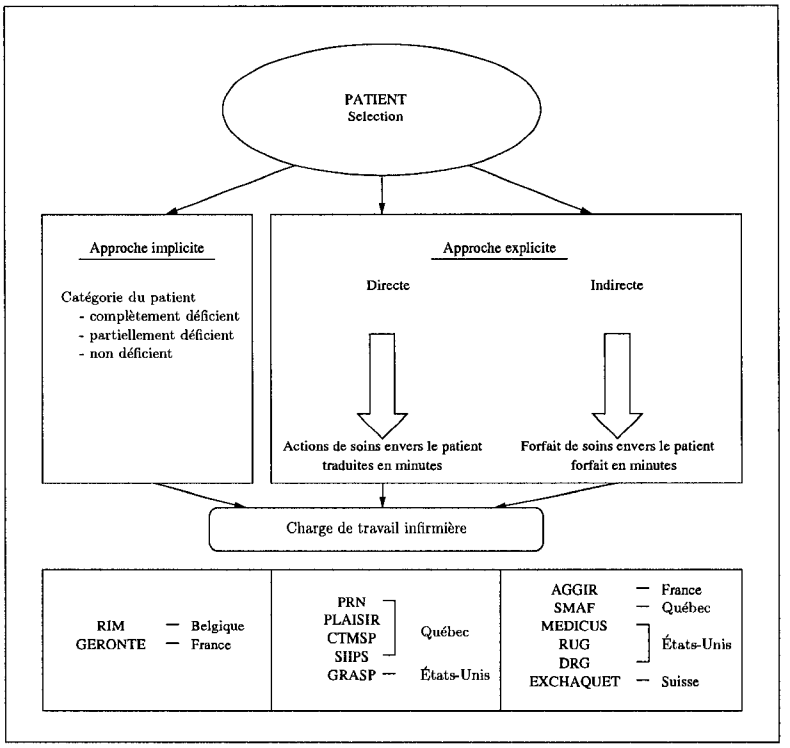

Figure 1.1 – Classification des approches permettant de mesurer la charge de travail infirmière

Avant de présenter les systèmes de mesure de la charge de travail, il convient de remarquer que beaucoup semblent basés sur l'évaluation des activités de soins. La formalisation des actes médicaux et infirmiers est donc une étape nécessaire. Elle permet de définir des normes de qualité mais aussi d'évaluer les prestations d'un établissement ou d'un service (en vue de son accréditation par exemple). Le travail soignant est donc traduit en procédures. Cette standardisation des tâches résulte des nouveaux modes de gestion et d'organisation des hôpitaux et s'inscrit dans une volonté d'uniformisation des différentes tâches [80]. L'un des avantages de cette standardisation est de présenter un caractère rassurant pour les travailleurs. La majorité des systèmes tend à standardiser les ressources nécessaires en termes d'heures de soins pour chaque type d'utilisateur.

Les systèmes développés sont nombreux. Nous n'allons présenter en détail que les plus utilisés.

Le système Projet de Recherche en Nursing (PRN) Le système PRN a été développé au Québec [189] dans les années 70. Il a pour but de déterminer l'effectif en personnel nécessaire par unité de soins pour chacune des trois périodes de travail du lendemain. Ce calcul se base sur les informations planifiées à partir de la démarche de soins. Il tient compte de l'évaluation des soins directs et indirects requis par chaque bénéficiaire pour une période de 24h, chacun des bénéficiaires étant identifié par sa catégorie. Il existe en tout huit catégories divisées en rubriques comprenant un ou plusieurs facteurs numérotés qui s'excluent mutuellement. À chaque facteur est associé un nombre de points défini par des études scientifiques et des consensus internationaux d'experts professionnels.

Le calcul de la charge en soins se fait donc par l'addition des points donnant le temps de soins directs et indirects pour une période de 24 heures, mais aussi du temps de communication au sujet du bénéficiaire* (CSB), des activités administratives et d'entretien (AAE), des déplacements dans et en dehors du service en fonction de sa dimension et du nombre de patients.

*Le temps de communication au sujet du bénéficiaire regroupe les réunions avec d'autres spécialistes au sujet du patient

Le système PRN est utilisé au Canada depuis son développement et a pris de l'envergure dans plusieurs pays d'Europe depuis les années 80 (France, Espagne, Luxembourg, Italie et Suisse).

Le système Soins Infirmiers Individualisés à la Personne Soignée (SIIPS) Les systèmes SIIPS et PRN sont très comparables. Le système SIIPS a été développé au CHU Nancy et est utilisé pour l'instant dans les unités de soins car il donne suffisamment d'informations pertinentes bien que n'étant pas aussi complet que PRN. Une description du système est dans [36]. Ce système a fait aussi l'objet de plusieurs validations quant à l'adéquation de la mesure de la charge de travail infirmier, dont quelques-unes récentes [96, 187]. L'un des avantages de cette méthode est qu'elle est applicable en soins à domicile [185], bien qu'elle ne le soit pas dans les unités de soins intensifs [84].

Le système ENEPCS Malloch et al. [152] reconnaissent la variabilité dans l'offre des soins suivant les compétences et en tiennent compte dans leur système ENEPCS (*Expert Nurse Estimation Patient Classification System*). Ce système permet en se basant sur huit catégories de soins et sur l'expérience des infirmières, de déterminer les soins à prodiguer au patient, le personnel ainsi que les compétences requis. Ce système met l'accent sur les besoins des patients en soins directs et chaque unité doit ensuite identifier sa charge indirecte qui peut varier d'un quart à un autre. Il est actuellement utilisé dans plusieurs hôpitaux aux États-Unis [152].

Dans certains contextes, tels que les unités de soins pour les problèmes mentaux intenses [166], la charge infirmière est particulièrement difficile à mesurer vu que la relation infirmière-patient ne dépend pas uniquement des activités de soins observables ou quantifiables

Plus récemment les CLSC ont entrepris des projets pour évaluer la charge de travail infirmière. Le comité des soins infirmiers du CLSC du Vieux Lachine [45] a identifié comme priorité d'action, l'élaboration d'une grille permettant l'évaluation de la charge de cas.

Indicateur de charge de cas (ICC) Cet indicateur a été développé par Nicole Colette [66] du CLSC Verdun et inclut les caractéristiques de la clientèle, les interventions directes et indirectes planifiées, et la fréquence des visites à domicile pour le mois à venir. Chacun de ces éléments a un score, et la somme de ces scores donne la charge pour chaque client. Cet outil a été utilisé par plusieurs CSLC (Ahunstic, Côte-des-Neiges, Lachine, LaSalle et Verdun). Certains ont revu les scores mais tout le monde semble être d'accord sur les différentes catégories. La figure 1.2 donne une vue d'ensemble de l'ICC.

Caractéristiques du client	Interventions directes et indirectes	
	Spécifiques	Universelles
Guichet unique Déficience intellectuelle Sida/soins palliatifs Santé mentale Aidants épuisés Misère sociale Ressource privée d'hébergement Alcoolisme/toxicomanie etc	Évaluation à domicile Réévaluation à domicile/retour d'hospitalisation Réévaluation régulière Hébergement temporaire Rencontre familiale ou avec des partenaires Intervention téléphonique planifiée	Démarches de suivi médical Fréquence des visites à domicile Mise à jour du programme de soins
⇩	⇩	⇩
Côte caractéristique client	Côte d'interventions spécifiques	Côte d'interventions directes et indirectes universelles

Figure 1.2 – Illustration de L'indicateur de charge de cas

L'analyse de la charge de cas du personnel infirmier à l'aide de l'ICC aux services de soutien à domicile au CLSC Verdun [65] a montré un déséquilibre dans la charge de cas.

Cet ICC est accompagné d'une grille de mesure de la charge mensuelle. Cette grille permet à chaque personne d'avoir une vue globale de sa charge de cas et à l'infirmière responsable d'avoir une vision d'ensemble de la charge de l'ensemble des infirmières afin de distribuer équitablement les nouveaux cas. Un planificateur quotidien permet de dresser les listes des clients à visiter dans la journée en spécifiant l'intervention et la durée.

1.2.2 Systèmes de classification des patients (SCP)

Un système de classification est un processus pour déterminer, valider et suivre de près les services requis par la condition de l'usager. Il est aussi un outil de mesure de la charge de travail : le niveau des besoins des patients est déterminé et traduit en interventions

attendues des soignants, ces interventions sont ensuite exprimées en unités de temps. Les systèmes de classification des patients ont plusieurs intérêts. Ils permettent non seulement de faire la planification des soins, l'allocation des ressources infirmières ou la dotation en personnel, mais aussi de participer à la planification et la défense du budget, au remboursement des coûts reliés aux patients ou au maintien des standards de qualité [85]. Un rapport de l'association des infirmières et infirmiers du Canada [7] rapporte les SCP comme étant les plus courants et largement utilisés par les établissements de santé de nombreux pays afin de déterminer la composition du personnel (combinaison des compétences) infirmier.

Malloch et Conovaloff [151] décrivent le développement des systèmes comme un processus en trois générations. La première génération date d'avant les années 70. Les systèmes de classification des patients (par groupes d'âge) étaient calculés manuellement pour estimer les ratios (patients par infirmière) et étaient basés sur des données historiques pour déterminer le niveau minimal de personnel. La surcharge du personnel était régulière. La seconde génération date des années 80 et a vu l'introduction des *Diagnostic related groups* DRG (décrits au paragraphe suivant) mettant plus d'emphase dans les soins communautaires et utilisant plus les technologies de l'information. Les SCP ont été développés pour donner des réponses plus flexibles aux variations de la charge de travail en employant les agences et les équipes de surplus (ou volantes). Même cette seconde génération n'a pas tenu compte des compétences du personnel infirmier. La troisième génération des années 90 a voulu relever le défi de calculer la charge infirmière en se basant sur chaque quart de travail de façon pratique mais n'y est pas encore parvenue. La quatrième génération enfin a pour objectif d'utiliser la technologie (telle que les moniteurs à distance) pour prédire les besoins en soins infirmiers en heures réelles en tenant compte des compétences et en donnant des informations statistiques à l'équipe de gestion.

Les systèmes DRG et PPS Deux systèmes ont été créés aux États-Unis en vue de contrôler les coûts en santé [170] : *Diagnostic related groups* (DRG) et *Prospective payment systems* (PPS). Ces deux systèmes ont été à la base de plusieurs décisions managériales par les professionnels de la santé dans et à l'extérieur des hôpitaux depuis leur création en 1981. Chaque patient est classifié dans une des 474 catégories DRG et l'appartenance à

un DRG fixe le montant de remboursement de l'hôpital indépendamment des soins reçus. Cette nouvelle politique de remboursement a été un important catalyseur dans la gestion des coûts [150]. Elle a créé une diminution de la durée d'hospitalisation et par conséquent une augmentation de la demande pour les soins qualifiés à domicile de 14%.

Le Système de Mesure de l'Autonomie Fonctionnelle (SMAF) Le système de mesure de l'autonomie fonctionnelle (SMAF) se charge d'évaluer le degré d'autonomie d'un individu. Il "s'appuie sur les notions d'incapacités et de handicaps (ou désavantage) telles que décrites par l'Organisation Mondiale de la Santé dans sa classification des déficiences, incapacités et handicaps. Cette classification se base sur un concept fonctionnel de la maladie qui comporte trois niveaux : la déficience, l'incapacité et le handicap" [108, 109]. Ce système a été développé en 1984 par une équipe de chercheurs et a fait l'objet de plusieurs validations dans les 15 années suivantes [78, 111, 112]. Il évalue 29 fonctions couvrant les activités de la vie quotidienne (AVQ), la mobilité, les communications, les fonctions mentales et les tâches domestiques. Chacune des fonctions est cotée sur une échelle de quatre degrés. Une nouvelle dimension a été ajoutée en 2003 [173] qui tient compte de l'aspect social (réseau d'activités et de relations de l'individu, ses ressources ainsi que sa participation dans la communauté). Cet instrument est utilisé par les cliniciens pour l'évaluation et le suivi des clientèles tant des soins à domicile qu'en institution d'hébergement ou en milieu de réadaptation. Pour faciliter l'utilisation du système, des profils iso-ressources (appelés ISO-SMAF) ont été générés. Chacun de ces profils détermine des besoins homogènes qui correspondent à des plans de services spécifiques [83]. Ce système est aujourd'hui la base des instruments d'évaluation en vigueur au Québec dans les programmes de maintien à domicile et pour l'admission en institution d'hébergement. Les profils ISO-SMAF sont aussi utilisés pour évaluer le niveau de financement des établissements de soins longue durée au Québec [110, 191].

L'outil Resource Utilization Groups (RUG) Les systèmes RUG [57] et RUG-III [95] pour ne citer que ceux-là, utilisent les caractéristiques des clients et l'utilisation qu'ils font du système (en soins infirmiers) en termes de durée pour les classifier. Ces systèmes

ont été implantés en Angleterre et au Pays de Galles [57] et dans 11 états aux États-Unis [95].

Dellefield [77] décrit l'utilisation de l'outil RUG-III dans le cas de trois maisons de soins infirmiers (*nursing homes*) pour vétérans. RUG-III inclut sept groupes de patients qui tiennent compte de l'intensité de leurs besoins. Une des maisons de soins infirmiers a utilisé RUG-III également pour déterminer les besoins en infirmières pour chaque quart de travail.

Le système RAFAELA Le système RAFAELA est un nouveau SCP utilisé dans différents hôpitaux d'Europe [179]. Le but de cette recherche est de montrer l'utilisation de ce système dans la gestion de la taille du personnel et la possibilité de transférer les ressources infirmières d'une salle à une autre en fonction des informations reçues sur la classification de l'intensité des soins infirmiers requis. Le système consiste en trois mesures :

- l'intensité des soins requis,
- les ressources quotidiennes en infirmières,
- la situation optimale d'intensité de soins infirmiers dispensés par infirmière.

Les conclusions de l'étude montrent que RAFAELA est utilisable pour l'allocation des ressources infirmières.

Nous pouvons citer d'autres systèmes tels que celui de Walts et Kapadia [194] où la côte associée à un patient dépend de la fréquence des types de procédures de soins requises.

Classifications internes aux CLSC Les CLSC ont introduit une classification de la clientèle qui a pour but de différencier les patients qui reçoivent une visite sur une base régulière de ceux qui n'en reçoivent que très peu ou pas du tout. Un client est actif s'il reçoit des soins sur une base régulière suivant un plan établi alors qu'un client inactif est un client qui a atteint ses objectifs de soins mais qui est encore à risque, et un client au besoin est un client qui ne reçoit des soins que très ponctuellement. Un client passif est un usager hospitalisé, hébergé, en attente ou sans service. Une première catégorisation est actif/inactif au CLSC CDN[46], actif/passif aux LaSalle [25] et Lachine [45], alors que

d'autres font la différenciation actif/au besoin au CLSC Verdun [145]. D'ailleurs dans une comparaison entreprise entre plusieurs CLSC de l'île de Montréal [147], un biais a été introduit en considérant le nombre de patients par infirmière. L'analyse a montré un écart de 32 patients (les infirmières ont en moyenne entre 41,5 et 73,8 patients) entre les sept CLSC montréalais étudiés, mais sans tenir compte de la définition d'un cas : est-ce un patient qui reçoit une visite ou un patient stable qui nécessite un suivi sans recevoir de visite ?

Un autre exemple de classification utilisé en CLSC à l'interne est le suivant : (P1) pour patient post-opéré post-hospitalisé ; les trois autres profils considèrent les patients présentant une perte d'autonomie modérée à sévère où le réseau social est absent (P2), modérée à sévère avec réseau fiable (P3) ou légère (P4). Le ministère [18] reconnaît aussi quatre profils de clientèles : la population générale en santé ayant des besoins ponctuels et nécessitant des soins aigus, les personnes de la population générale ayant une pathologie chronique généralement stable, les clientèles vulnérables ou des personnes de la population générale présentant des risques particuliers de problèmes sociaux et de santé, et enfin les clientèles particulières présentant une problématique complexe de santé ou sociale exigeant de la réadaptation ou de l'intégration sociale.

Tous ces systèmes de classification des patients ont servi à évaluer le niveau de soins requis pour les patients dans chaque catégorie déterminée. Ils ont donc tous contribué indirectement à évaluer les ressources infirmières nécessaires pour répondre à la demande de soins.

1.2.3 La dotation en personnel

Hurst [133] a recensé plus de 500 ouvrages et articles sur les différents systèmes de dotation en personnel infirmier. Les différents systèmes se basent sur cinq approches :

- l'approche *Telford* est basée sur le jugement professionnel. L'idée est de convertir les décisions et actions qui doivent être prises, en temps équivalent (ou en quarts de travail). Le nombre de ressources nécessaires dépend alors du temps (ou nombre de quarts) que peut exercer une infirmière. Cette méthode est simple, rapide, facile à

mettre à jour et surtout n'est pas chère. Cependant, elle ne montre pas la surcharge et est trop subjective d'après certains gestionnaires ;

- l'approche NPOB (*Nurses per occupied bed*) se base sur le nombre d'infirmières nécessaires par lit. Ce calcul est différent en fonction des groupes de soins. Cette méthode est simple et empirique. De plus elle utilise des données réelles. La lacune principale est qu'elle dépend du service et qu'elle ne tient pas compte de la dépendance du patient ;
- l'approche basée sur le rapport entre l'intensité des besoins du patient et son degré de dépendance utilise pour chaque catégorie de patients, le nombre de patients ainsi que le nombre de minutes de soins nécessaires par jour. Le nombre moyen et total de patients par jour permettent de connaître le nombre d'infirmières nécessaires par jour pour rendre les services demandés ;
- l'approche temps-demandé par activité est basée sur le plan d'intervention des patients qui tient aussi compte des soins indirects. Treize catégories d'interventions sont retenues et la distribution des patients est faite équitablement en se basant sur le temps de soins ;
- l'approche basée sur la régression établit une relation entre le degré d'activité et le nombre d'infirmières ou encore entre le taux d'occupation des lits et le nombre d'infirmières.

Finalement aucun des systèmes permettant de faire la dotation en personnel, qu'il soit basé sur le jugement, sur la classification des patients ou sur la charge des infirmières, ne donne une mesure exacte du nombre de personnes nécessaires. Ils sont cependant tous des bons indicateurs.

1.2.4 Limites des systèmes d'évaluation de la charge de travail infirmière

Plusieurs recherches appuient l'observation faite dans [118] et stipulant que les systèmes de mesure de la charge des infirmières ne permettent d'avoir qu'une estimation de la charge du système et non une mesure exacte. En effet, les études de la dotation en personnel infirmier et l'utilisation des SCP affirment que ces outils soulèvent une méfiance généralisée

et qu'ils ne sont pas suffisants pour déterminer les effectifs dont un service a besoin chaque jour ou pour chaque quart de travail [15], au mieux ils fournissent une prévision moyenne. Une autre limite de ces systèmes a trait à l'exportabilité de ces systèmes à l'extérieur du cadre où ils sont conçus [30] : ils ne permettent pas de faire des comparaisons entre services, alors que la littérature recommande d'établir la robustesse et la validité des systèmes [113]. Cette constatation est encore plus vraie lorsque l'environnement est instable [75, 76] (soins intensifs) ou que les admissions et les congés sont fréquents. Des ajustements ont été faits sur certains systèmes tels que PRN pour tenir compte du changement de l'état du patient. Les auteurs dans [64] incluent au système PRN des événements imprévus dans une unité de réveil postopératoire. Cette étude a bien montré qu'ils nécessitent une augmentation du nombre d'infirmières en place. Les auteurs dans [168] incluent aussi les coûts, le volume de travail des infirmières ainsi que la variabilité de l'utilisation des ressources dans la mesure du système.

Les critiques envers ces systèmes ont aussi trait au manque de rigueur [85, 166, 113] et de validité, que ce soit dans les temps d'intervention estimés ou dans les prévisions en personnel [7, 167]. Le besoin en personnel identifié par RUG-III par exemple est moins précis que celui obtenu par l'expérience des infirmières en place [159]. Une autre étude entreprise [57] montre que les temps de soins associés à chaque classification de patients varient suivant les pays.

D'autres études ont par ailleurs montré que les infirmières ne surestiment pas le temps nécessaire pour les soins aux patients [166] et plusieurs systèmes donnent le même niveau de dotation en personnel en comparant plusieurs systèmes [15, 171].

Les solutions proposées pour remédier à ces problèmes sont nombreuses. La plus intéressante est la dotation "ponctuelle" en personnel répondant aux variations non prévues du volume des patients ou de l'intensité des besoins. Cette dotation peut être le fruit d'une surévaluation des effectifs, en donnant une formation plus large aux infirmières qui leur permet de travailler dans d'autres services au besoin, ou encore en recourant à des infirmières d'agence. L'excédent de personnel n'est pas considéré comme du gaspillage, cette marge de manoeuvre permet au personnel de suivre les progrès de la technologie et de l'in-

formation. D'ailleurs un rapport américain récent [14] concernant la dotation en personnel infirmier et la qualité des soins aux patients, montre que l'augmentation de la taille du personnel infirmier en hôpital est associée à une baisse de la mortalité reliée au séjour à l'hôpital sans que cette association soit causale.

Malgré les différents reproches faits aux systèmes de mesure de la charge de travail ou de classification des patients, un recensement en 1990 montre que sur 496 établissements hospitaliers en France ils restent largement utilisés : 21% utilisent le système SIIPS, 13% utilisent le système PRN, 9% utilisent la méthode EXCHAQUET, 44% ont une grille de dépendance, 20% utilisent GERONTE ou ISIS (moyens et longs séjours), 7 l'IPSIS et 179 n'ont pas de méthode d'évaluation [82].

Le reproche principal adressé à l'ICC développé en CLSC [65], est en rapport avec le processus de compilation des données. Il est complexe et prend une demi-journée à compléter excluant l'analyse par les chefs de programme. Il a été utilisé au CLSC Ahunstic pendant une courte période ainsi qu'au CLSC CDN après modification mais uniquement pour une évaluation ponctuelle et non régulière. Une mesure de la charge de cas pour être efficace en pratique doit être facile et rapide à évaluer.

Les différents systèmes permettant d'évaluer la charge de travail infirmière, qu'ils soient directs ou non, doivent considérer des éléments clés. Il faut tenir compte des patients (nombre d'utilisateurs, caractéristiques démographiques, niveaux d'handicap et de dépendance, présence du support familial), des interventions requises (traitements et durée de ces traitements) ainsi que de la compétence du personnel.

1.3 Équilibrage de la charge de travail

Équilibrer la charge de travail revient à trouver la façon la plus équitable possible pour distribuer le travail. La distribution du travail en tant que telle sera présentée dans les deux prochaines sections alors que dans cette section nous faisons une description d'applications où l'équilibre de la charge de travail est recherché.

Il y a deux façons fondamentales d'équilibrer la charge de travail : soit en ayant une charge équitable pour tout le monde (la surcharge et la sous-charge sont distribuées équitablement), soit en ne tenant compte que de la surcharge. Plus formellement, équilibrer la charge de travail revient à minimiser la surcharge maximale (problème bottleneck GAP présenté à la section 1.5), minimiser la somme des surcharges et sous-charges ou minimiser l'écart entre la charge minimale et maximale. La mesure de variance (entre la charge moyenne et les charges observées) est une bonne mesure de l'équilibre, mais est difficile à implanter à cause de sa nature non linéaire. Guerro et al. [104] estiment que la mesure la plus efficace d'équilibrage de la charge de travail est la différence entre la charge minimale et la charge maximale. Kumar et Shanker [140] concluent que le meilleur objectif d'équilibrage est de minimiser cette différence. Ces mesures sont particulièrement adaptées et utilisées dans un contexte manufacturier. D'ailleurs Riezebos et al. [180] soulignent l'importance du tampon pour absorber les surcharges en contexte manufacturier. Ils font aussi la distinction entre les agents décentralisés qui reçoivent leur tâches et les autres qui sont regroupés (en *pool*) et qui ont pour rôle d'absorber le surplus. Ils reconnaissent aussi le temps d'attente comme une façon de limiter la charge.

Dans un contexte d'affectation de tâches à des opérateurs sur des machines, Hajri-Gabouj dans [107] tente d'obtenir un équilibre parfait entre les charges des opérateurs. Pour cela, la charge moyenne de travail des opérateurs est évaluée et la distribution de la charge est calculée comme étant la valeur absolue de la différence entre la charge des opérateurs et la moyenne. Pour atteindre l'équilibre entre les charges, une contrainte est ajoutée au modèle stipulant que la distribution est inférieure à un seuil ϵ. Comme cette contrainte est non linéaire, elle est pénalisée dans la fonction objectif. La solution recherchée doit respecter d'autres critères. Une hiérarchisation de ces critères est incluse dans la fonction objectif en leur donnant des poids plus ou moins importants.

Dans un contexte de déchargement d'avions de fret dans deux terminaux, Huang et al. [117] tentent d'équilibrer la charge de travail entre les terminaux. Le but poursuivi est d'augmenter l'efficacité opérationnelle. La charge w est mesurée par le nombre total de palettes et conteneurs qui doivent être déchargées. Cette charge est variable dans le

temps, et les auteurs tentent d'obtenir une charge équilibrée sur l'ensemble des périodes j considérées. Pour cela ils utilisent deux mesures :

- la surcharge maximale $\max\{w_j^1, w_j^2 : \forall j\}$ où w_j^1 et w_j^2 représentent les charges au terminal 1 et 2 respectivement à la période j,
- la somme des écarts de charge entre les deux terminaux pour chaque période $\sum_{j \in J} |w_j^1 - w_j^2|$.

Le premier objectif est de minimiser une fonction qui tient compte de ces deux critères, chacun ayant un poids différent.

Dans un environnement manufacturier, Kim et al. dans [137] résolvent le problème de distribution de véhicules autoguidés. Un index de priorité pour chaque tâche est calculé représentant la différence entre le nombre de tâches à une station et la suivante. Un index d'urgence permet de différencier deux tâches de même index de priorité (qui tient compte du temps de transport dans le système). La distribution se fait par la priorité des tâches.

L'équilibre des charges est aussi particulièrement important dans la distribution de mémoire aux processeurs informatiques. L'équilibre de la charge est considérée comme la mesure de l'efficacité du système. Dans [63] la charge d'un processeur est calculée comme étant le ratio d'utilisation de la mémoire. Dans [203] la charge de chaque noeud est évaluée par le temps passé pour accomplir le travail en ce noeud. Dans [172], la flexibilité de l'affectation des tâches est exploitée pour améliorer l'équilibre sur les processeurs. Le problème peut être formulé comme une minimisation de la norme infinie des charges des processeurs, et peut se réduire à résoudre le problème de la minimisation de la norme carrée.

Dans une application en contexte hospitalier [135], la charge des infirmières est évaluée par le nombre d'heures qu'elles ont travaillé. Cette charge doit se trouver dans un intervalle. Les autres ressources telles que le nombre de fin de semaine de congé, le nombre de changement de quart, le nombre de quarts consécutifs travaillés, la combinaison des différents quarts de travail et les vacances sont considérés de cette même façon. La résolution du problème ainsi formulé bien que menant à une solution de bonne qualité du point de vue de la minimisation du personnel, ne garantit pas l'équilibre des charges entre les personnes. Dans une autre application [31], la charge de travail est considérée équilibrée lorsque le

nombre de nuits travaillées, le nombre de fins de semaines en congé ainsi que le nombre de jours travaillés ou en congé isolés sont comparables entre les différentes personnes.

Dans [196], Weigel et Cao développent un système qui permet de résoudre le problème de répartition des livraisons et des techniciens chez Sears. Une phase d'amélio-ration lors de la construction des routes permet de modifier les routes qui présentent une charge de travail inacceptable (en termes de nombre d'arrêts ou de temps supplémentaire). Les arrêts sont alors transférés vers d'autres routes de façon à créer des routes plus équilibrées.

Dans leur souci de réorganiser le travail infirmier, plusieurs CLSC ont tenté d'établir des normes à respecter qui fixe la charge maximale de chaque ressource infirmière. Trois normes en particulier ont été étudiées : le nombre de visites à domicile par jour, la durée des visites aux patients ainsi que le nombre de patients maximal à prendre en charge. Le temps disponible pour les visites à domicile n'est pas suffisant pour déterminer un nombre moyen de visites quotidien idéal. En effet, les patients du service à domicile diffèrent de par la complexité de leur cas et ne peuvent pas être considérés comme étant équivalents. Cette limite là n'a donc pas pu être quantifiée. Cependant la moyenne du nombre de visites à domicile par semaine par infirmière gestionnaire de cas a été fixée à 25 au CLSC Verdun [145]. La charge de cas (nombre de cas) normale au CLSC Verdun [145] se compose de 65 cas de clientèles variées alors qu'elle est de 50 dossiers actifs (avec un maximum fixé par catégorie de client) à Ahunstic [66]. Des temps moyens pour les visites ont été définis tels que trente minutes pour une visite à domicile, trente minutes pour une prise en charge d'un nouveau dossier, etc. Des balises de fréquence fixées selon les profils des clients ont été définies au CLSC Ahunstic [103] ainsi qu'un nombre de visites minimal par profil.

1.4 La confection des horaires

Quarts de travail La confection des horaires des employés se fait fréquemment par la distribution de quarts de travail. Lorsque l'environnement de l'entreprise se prête à un quart de travail unique, le problème revient à distribuer les journées de travail aux employés. Dans

l'autre cas où l'environnement se prête mieux à des quarts de travail multiples (quart de jour, quart de soir et quart de nuit par exemple), le problème revient à distribuer aux employés les journées de travail ainsi que le quart de travail qu'ils doivent couvrir. Tous les contextes où le travail est en continu (24 heures) tels que les établissements hospitaliers où les entreprises manufacturières s'inscrivent dans ce dernier cas.

Mode de construction Divers modes de construction d'horaires sont recensés dans [47] : centralisé, par unité ou personnel. Lorsque le mode est centralisé, un département est responsable des horaires de tout l'établissement. Lorsque le mode est par unité, la personne en charge de l'unité (ou son assistante) est responsable de cette planification. Finalement chaque personne construit son propre horaire dans le dernier mode. Les études ont démontré que cette dernière méthode a de sérieux inconvénients tels que la création de surplus ou de manque en personnel ou encore le déséquilibre de la charge de travail entre les différentes personnes dans les unités et à l'extérieur. D'autres encore montrent que cette méthode de planification donne des rotations plus efficaces puisque les infirmières connaissent déjà leurs préférences et coopèrent donc plus facilement. Finalement, une étude faite en 2000 montre que tout dépend du contexte opérationnel : la planification centralisée semble plus efficace lorsque les salles sont grandes et les problèmes de planification complexes, la planification par unité est plus efficace lorsque les tailles des équipes sont moyennes avec des problèmes relativement faciles alors que la planification personnelle est plus efficace lorsque les salles sont petites avec des problèmes directs et précis. Les modes les plus répandus dans les établissements hospitaliers sont la planification manuelle et décentralisée des horaires du personnel [170].

Le cas de l'unité de chirurgie de l'hôpital Royal Victoria à Montréal est présenté dans [135]. L'infirmière en chef fait la confection des horaires des autres membres du personnel infirmier toutes les six semaines (les horaires couvrent des périodes de deux semaines à la fois). Chaque journée se divise en quatre quarts de travail et l'objectif est de minimiser le nombre d'infirmières en service (à l'exception de l'infirmière en chef et de son assistante). La confection de l'horaire doit respecter une contrainte de couverture minimale (nombre minimal d'infirmières en service) pendant la semaine (et la fin de semaine). Lorsque les

horaires sont complétés mais qu'il y a pénurie de personnel, les infirmières qui acceptent de faire du temps supplémentaire sont sollicitées par priorité d'ancienneté.

Le personnel infirmier établit lui-même son horaire quotidien dans les CLSC. Chaque personne est responsable du choix des clients à visiter et de l'ordre des visites. Les membres des équipes de surplus (qui absorbent le surplus de travail) n'ont par contre pas le choix des clients à visiter. Ces clients leur sont imposés par le personnel régulier. Dans le cas des services de soins à domicile en France par exemple, la cohérence des tournées générées par les visites à rendre à domicile est plus centralisée.

Dans le contexte manufacturier, la confection des horaires est centralisée. Une planification de la production pour répondre à la demande est accomplie, les horaires sont ensuite distribués aux employés.

Type d'horaires Plusieurs types d'horaires sont identifiés : les horaires cycliques et les horaires acycliques. Les horaires cycliques couvrent un horizon de n périodes qui se répète. Dans les horaires acycliques, il n'y a pas de répétition (du moins pas systématique).

Plusieurs revues de littérature traitant des problèmes d'horaires de personnel ont été publiées, parmi celles-ci citons celles de Burke et al. [47] et de Ernst et al. [87]. Les auteurs ont fait la classification des contraintes en fonction des contraintes de couverture (nombre de personnes requis par quart par exemple) et des contraintes reliées au temps. Ces deux types de contraintes sont, suivant les méthodes de résolution, dures ou souples. Les coûts et les paramètres, qu'ils soient dans les contraintes ou l'objectif sont tantôt fixés tantôt flexibles. Les problèmes qui ont été résolus comprennent une flexibilité à différents niveaux :

- pour fixer les contrats ou arrangements de travail,
- pour définir les contraintes de couverture (en quarts de travail, en heures, en jours, etc),
- pour définir la longueur des quarts de travail : de longueur définie, à définir, plusieurs quarts de différentes longueur, etc,
- pour définir les périodes de planification possibles (4 jours, 1-2 ou 3 semaines, 2-4 ou 2-6 semaines, 1 mois, 1 an ou définissable par l'usager),

- pour définir la capacité des quarts de travail pour chaque personne (nombre maximum, temps supplémentaire, maximum d'affectations à un quart particulier, nombre maximum de quarts par semaine),
- pour tenir compte des préférences personnelles (jours de congé, séquence et alternance des quarts de travail, jours travaillés, etc),
- pour définir les contraintes de consécutivité (jours travaillés consécutifs, quarts travaillés consécutifs, jours de congé consécutifs, etc),
- pour équilibrer la charge de travail (à définir par l'usager ou fixée d'avance),
- pour déterminer la préférence des fins de semaines de congé,
- pour définir la taille de l'équipe (définie ou à minimiser),
- pour définir le choix des horaires cycliques ou acycliques,
- pour définir le choix du nombre de catégories d'employés spécialisés et de possibilités de substitution (d'une personne par une autre).

Certains chercheurs ont abordé ces problèmes de façon formelle. Les méthodes d'optimisation utilisées pour les résoudre sont autant des méthodes exactes (qui garantissent l'optimalité des solutions si elles existent) qu'heuristiques (méthodes de résolution approchées). Les métaheuristiques (méthodes de résolution approchées possédant des outils permettant de s'extirper des optima locaux) sont largement privilégiées pour résoudre ces problèmes, dans la mesure où ils sont plus complexes et nécessitent des outils efficaces. Les différentes méthodes ainsi que leur contexte d'application sont résumées dans l'annexe B de [47].

Plusieurs applications en construction d'horaires de personnel infirmier sont plus récentes. Dans [202] les auteurs simulent le flux entrant des patients à l'unité d'urgence d'un hôpital et tentent de minimiser le temps d'attente des patients. Les horaires de personnel sont générés respectant le nombre de personnes requises par quart de travail. Dans [35] les auteurs tiennent compte des préférences du personnel infirmier en minimisant les horaires indésirables dans un contexte d'horaires cycliques. Bard et Purnomo [34] résolvent le problème de gestion des déséquilibres en hôpital en temps réel en réallouant les quarts de travail lorsque la demande est supérieure ou en annulant les quarts de travail lorsque la demande est plus basse que prévu. Les différentes options pour réagir aux fluctuations de

la demande sont le temps supplémentaire, l'appel du personnel en congé, l'utilisation de ressources externes ou finalement l'absence de réaction.

L'approche par programmation multi-objectifs *goal programming* a été utilisée dans [31, 92, 141] pour construire les horaires du personnel infirmier et [43] pour déterminer le volume et composition des patients à accepter en hôpital, de façon à générer suffisamment de revenus (que ce soit pour l'établissement hospitalier ou pour les professionnels de la santé). Cette variation de la programmation linéaire est particulièrement intéressante dans les contextes où plusieurs objectifs sont à considérer, chacun ayant un niveau de priorité. Dans le contexte de services de soins à domicile que nous traitons, plusieurs composantes forment la charge de travail. Du point de vue de la direction, il est intéressant de connaître l'impact de chacune de ces composantes sur la distribution des patients. Cette approche par multi-objectifs semble être adaptée. Dans l'application de Ferland et al. [92], chaque infirmière doit être affectée à un certain nombre de jours de travail durant la période considérée. Les différents objectifs à atteindre sont classés par ordre d'importance et correspondent aux contraintes à respecter. Le problème est résolu par recherche avec tabous où diverses stratégies de diversifications sont utilisées.

Dans le contexte industriel où l'évolution de la demande ainsi que les conditions et contrats de travail introduisent beaucoup de flexibilité, les problèmes de planification de la main-d'œuvre peuvent être classés en quatre catégories [33] : les horaires réguliers, les horaires compressés, la main-d'œuvre hiérarchique et l'annualisation du temps de travail. Le type d'horaires le plus connu est sans doute l'horaire régulier, que ce soit dans le contexte de quarts de travail unique [48] ou dans le contexte de quarts de travail multiples [49, 128, 129].

Les horaires de travail compressés par contre, permettent de travailler le même nombre d'heures par semaine mais en diminuant le nombre de jours travaillés. La longueur des quarts de travail s'en trouve bien sur augmentée. Les horaires compressés consistent à travailler des semaines de trois, quatre ou trois/quatre jours. Plusieurs applications en contexte de quarts de travail unique [52, 54, 120, 127] et en contexte de quarts de travail multiples [121, 122, 123, 124] sont recensées. Alfares [26] traite ce problème en imposant

trois jours de repos consécutifs après chaque quatre jours de travail alors que dans [27] il fait varier le nombre de jours de repos consécutifs entre deux et trois pour obtenir plus de flexibilité.

Dans le contexte où la main-d'œuvre est hiérarchique, les employés d'une classe A sont capables d'accomplir les tâches A, les employés de classe B sont capables d'accomplir les tâches A et B, et ainsi de suite. Des exemples d'applications en contexte de quart de travail unique sont décrits dans [86, 125] et en contexte de quarts de travail multiples dans [119, 162]. Billionnet [41] a étudié le cas de la main-d'œuvre hiérarchique en profondeur.

Plusieurs auteurs enfin ont étudié des combinaisons de ces types d'horaires, comme les horaires hiérarchiques sous horaires compréssés [126, 132] par exemple.

Le dernier problème d'horaire à étudier, et non le moindre, est le temps de travail annualisé. Le nombre d'heures de travail par semaine n'est pas fixé d'avance tandis que le nombre maximal par année est déterminé à l'avance. Ce type d'horaire est de plus en plus utilisé vu qu'il permet une grand flexibilité pour répondre aux fluctuations de la demande, surtout en contexte manufacturier. Ce type d'horaire peut encore une fois être appliqué autant dans un contexte de quart de travail unique [32, 33, 131] que dans le contexte de quarts de travail multiples [67, 73, 130]. Corominas et al. [71] ont étudié ce problème en considérant des employés qui appartiennent à différentes catégories pouvant accomplir différentes tâches à des niveaux d'efficacité différents.

Corominas et al. [68, 70] montrent que cette approche est efficace pour atteindre la flexibilité dans la planification de la production. Ils proposent une classification des problèmes d'annualisation du temps de travail, que ce soit dans le secteur des services ou manufacturier. L'approche par programmation linéaire et en nombres entiers semble donner de bons résultats. Résoudre le problème de planification de la production tient aussi compte de la distribution des semaines de vacances [67, 69]. Dans les applications, l'objectif est de minimiser les coûts dus au paiement du temps supplémentaires aux employés et à l'engagement de travailleurs temporaires. Les auteurs tentent de plus de régulariser le temps de travail. Les contrats temporaires sont aussi intégrés dans [72]. Dans [71], la main-d'œuvre est hiérarchique et les horaires sont choisis dans un ensemble préétabli. Dans [73], ils

considèrent la version du problème où tous les employés sont en vacances simultanément. La chaîne de production s'arrête pendant cette période.

Corominas et al. [68] dans leur classification des problèmes d'annualisation du temps de travail considèrent que le problème suivant est l'un des plus complexes à résoudre :

(i) les employés ne prennent pas leur vacances simultanément ;

(ii) les semaines de vacances doivent être distribuées ;

(iii) la production ne s'arrête pas ;

(iv) les produits sont périssables.

L'application que nous étudions au chapitre cinq fait partie de cette catégorie.

1.5 Le problème de distribution des patients

Le problème de distribution des patients au personnel infirmier des services à domicile consiste à affecter chaque patient à exactement un membre du personnel infirmier. Chaque personne peut prendre en charge plusieurs patients. Cette affectation peut avoir différents objectifs suivant le contexte, que ce soit une distribution équitable des patients ou encore la maximisation des préférences du personnel.

Le problème d'affectation de cas (ou patients) a déjà été traité à quelques reprises dans la littérature, même si le problème d'affectation des visites semble avoir été plus souvent étudié. Nous présentons trois applications, la première en hôpital et les deux suivantes dans un contexte de soins à domicile, avant de présenter les méthodes appropriées pour résoudre ce problème.

1.5.1 Applications

Mullinax et Lawley dans [160] font l'affectation du personnel infirmier dans des unités de soins intensifs néonatales. Au début de chaque quart de travail, la personne en chef doit affecter un groupe d'enfants à chaque infirmière. Ce problème d'affectation est complexe vu que les enfants ne sont pas tous dans la même condition. De plus, pour chaque

quart de travail, une personne chargée des admissions est obligatoire et doit donc être moins chargée que les autres. Chaque infirmière doit être relativement proche physiquement du groupe d'enfants dont elle a la charge, ce qui élimine l'affectation interzones. L'objectif de l'affectation est d'équilibrer le travail entre le personnel infirmier tout en respectant les contraintes. Tout le monde devrait recevoir le même niveau d'intensité de soins (le tout mesuré en fonction des soins dont a besoin le patient). Le but du modèle développé est de minimiser l'écart entre l'intensité de soins des différents groupes de façon à minimiser le déséquilibre entre les charges de travail. L'approche adoptée par les auteurs consiste à décomposer le problème en deux : le premier problème consiste à affecter le personnel infirmier à des zones et ensuite à déterminer les patients qui font partie de chaque zone. Ils comparent quatre méthodes de résolution : résolution exacte du modèle, résolution de la décomposition du problème, affectations générées aléatoirement et enfin la pratique courante en place. Le temps de résolution a été limité à 30 minutes puisque c'est le temps mis en pratique pour faire l'affectation. La méthode de décomposition a donné des améliorations du point de vue de l'équilibre des charges de travail comparativement à la méthode manuelle en pratique.

L'approche utilisée pour résoudre le problème à Danderyd dans la banlieue de Stockholm [88] de services de soins à domiciles est le *repeated matching*. Le système développé LAPS CARE a été testé dans une unité géographique où opèrent 28 employés qui visitent 150 clients. Le secteur est divisé en 3 sous-secteurs et chaque employé est affecté à exactement un sous-secteur. Le système développé permet d'introduire de la flexibilité dans les frontières entre les sous-secteurs. Les anciens employés restent confinés dans leur sous-secteur tandis que les nouveaux sont autorisés à traverser de façon à aller chercher un gain en efficacité. Le système permet de construire les routes (ordre des visites) des infirmières. Pour cela, le système complète la planification partielle des visites cycliques, qui est généralement préparée quelques jours à l'avance.

Dans [39], Bertels et Fahle résolvent le problème de distribution des visites aux patients dans un contexte de soins à domicile. Ils incluent les qualifications requises, les limites en temps de travail ainsi que les préférences : les patients préfèrent êre servis à une heure

particulière, la "chimie" entre le patient et la personne qui le sert est souhaitée, les changements d'infirmières soignantes doivent être évités et la satisfaction du personnel maximisée. La méthode proposée pour résoudre ce problème est une méthode hybride combinant programmation par contraintes et recherche avec tabous.

1.5.2 Le problème d'affectation généralisé (GAP)

Le problème de distribution des patients au personnel infirmier consiste à affecter chaque patient à exactement une infirmière. Chaque infirmière a une capacité maximale en termes de nombre de patients à ne pas dépasser. Ce problème est un problème d'affectation généralisé où m patients doivent être affectées à n infirmières de façon à minimiser le coût d'affectation d'exactement une infirmière par patient. Chaque infirmière a une capacité limitée b_j représentant par exemple le nombre maximal de patients dont elle peut avoir la charge. Le problème d'affectation généralisé se formule comme suit :

$$\text{Min} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

$$\sum_{j=1}^{n} x_{ij} \quad = 1 \quad i = 1, ..., m \tag{1.1}$$

$$\sum_{i=1}^{m} a_{ij} x_{ij} \quad \leq b_j \quad j = 1, ..., n \tag{1.2}$$

$$x_{ij} \quad = 0 \text{ ou } 1 \quad i = 1, ..., m, j = 1, ..., n \tag{1.3}$$

où x_{ij} vaut 1 lorsque le patient i est affecté à l'infirmière j et 0 sinon. Ce problème est connu comme étant un problème NP-difficile. Les applications du GAP sont nombreuses : localisation d'installations, ordonnancement, confection d'horaires, routing et chargement pour des systèmes manufacturiers flexibles, etc [164].

Dans ce qui suit, nous allons plutôt que d'utiliser le contexte de patients à distribuer au personnel infirmier, utiliser la notion de tâches à distribuer aux agents, peu importe le contexte réel de l'application présentée.

Cattrysse et Wassenhove [61] présentent une revue complète des différents algorithmes autant exacts qu'heuristiques utilisés pour résoudre ce problème. Les méthodes exactes de résolution sont généralement basées sur le "branch-and-bound" en profondeur d'abord. Plusieurs formes de relaxations ont été utilisées pour résoudre le problème.

La relaxation linéaire consiste à relaxer la contrainte (1.3) qui devient $x_{ij} \geq 0$. La relaxation lagrangienne a surtout été appliquée à la relaxation de la contrainte (1.2). Les solutions recherchées ne respectent donc plus les contraintes de capacité du problème. La relaxation surrogate enfin remplace un ensemble de contraintes d'inégalités par un autre. Par exemple, l'ensemble des contraintes de capacité (1.2) devient :

$$\sum_{j=1}^{n}\sum_{i=1}^{m}\mu_j a_{ij} x_{ij} \leq \sum_{j=1}^{n}\mu_j b_j$$

où μ_j est le dual optimal du problème relaxé. Cette dernière relaxation transforme le problème en un problème de sac-à-dos où la capacité est représentée par la somme des capacités b_j. La relaxation de la contrainte (1.1) finalement ramène le problème à un problème de sac-à-dos lorsque le problème est reformulé comme un problème de maximisation.

Martello et Toth [155] font état des différents algorithmes pour résoudre le GAP présenté comme un problème de sac-à-dos multiples (*knapsack*). Ce problème consiste à remplir des sacs de capacité finie d'un certain nombre d'éléments appartenant à M auxquels sont associés une taille et un coût (une utilité). Dans le contexte de l'affectation des patients, cela revient à distribuer un sous-ensemble des patients (c'est la contrainte d'affectation de tous les patients qui est relaxée) à chaque infirmière.

Deux algorithmes heuristiques sont particulièrement reconnus. Ils ont souvent été repris dans la littérature : l'heuristique HGAP de Martello et Toth [154] et l'algorithme de Fisher et Jaikumar [94].

L'algorithme HGAP de Martello et Toth [154] se déroule en deux phases, où la première génère une affectation réalisable en affectant un objet (patient) à la fois dans le "meilleur" sac (une infirmière) avec pour objectif de maximiser une fonction f. Plusieurs fonctions de pseudo-coût sont définies qui évaluent la différence entre le choix du meilleur et du second meilleur objet (patient) à affecter par rapport :

- au coût c_{ij},
- à la consommation de la ressource a_{ij},
- au ratio entre la consommation de la ressource et la capacité $\frac{a_{ij}}{b_j}$,
- et enfin au ratio entre le coût et la consommation de la ressource $\frac{-c_{ij}}{a_{ij}}$.

La seconde phase est une phase d'amélioration qui échange un objet i s'il y a amélioration. Les auteurs ont aussi développé une procédure de réduction de la taille des problèmes qu'ils résolvent encore avec le même algorithme.

La relaxation de la contrainte (1.1) est utilisée par Fisher et Jaikumar [94] où le problème devient :

$$\max \sum_i \sum_j (c_{ij} - \mu_i) x_{ij}$$

sous les contraintes (1.2) et (1.3) et où μ_i est fixé au second coût c_{ij} le plus élevé. Les tâches (patients) affectées en priorité sont celles ayant une différence positive entre le coût c_{ij} et μ_i, les suivantes sont celles ayant un écart nul (dans l'ordre décroissant de coût). Les μ_i sont ensuite ajustés pour renforcer la borne.

Cario et al. [55] font dans leur article une comparaison entre quatre méthodes pour évaluer leur performance sur des problèmes générés de deux façons basées sur la corrélation entre le coût pour effectuer chaque tâche et sa consommation de ressource. Les quatre méthodes testées sont : CPLEX et LINDO (deux solveurs exacts disponibles sur le marché), les heuristiques développées par Trick [192] et Martello et Toth [154]. Ces deux heuristiques ont été modifiées pour être plus efficaces. L'heuristique de Trick commence par une phase de prétraitement qui consiste à fixer les variables lorsque des tâches ne peuvent être affectées à des agents à cause de la consommation de la ressource (consommation supérieure à la capacité). Dans l'heuristique HGAP de Martello et Toth, ce sont les fonctions de pseudo-coût qui ont été modifiées.

L'approche par ajout de coupes valides a été utilisée par Nauss [164] et Cattrysse et al. [59] et la décomposition lagrangienne appliquée dans [60, 105, 106]. Dans [105] en particulier, la relaxation utilise les mêmes principes que les heuristiques HGAP et celle de Mazzola [157]. Lorena et Narciso [149] suggèrent de combiner les relaxations lagrangienne et surrogate pour résoudre le problème GAP.

Toutes ces méthodes sont basées sur la résolution de la relaxation du modèle mathématique. Dans ce qui suit nous présentons quelques méthodes de résolution du GAP qui n'utilisent pas directement la résolution du modèle par un solveur. Elles sont toutes heuristiques.

Amini and Racer [28] ont développé l'algorithme heuristique "Variable-Depth-Search" (VDSH) et testé sa performance par une analyse statistique. Ils comparent la performance de leur algorithme avec l'algorithme HGAP de Martello et Toth [154]. La première phase consiste à affecter chaque tâche (les tâches sont classées dans un ordre aléatoire) au premier agent disponible (sinon à un agent fictif). La solution obtenue est ensuite améliorée en utilisant des mouvements de transfert (*flip*) d'une tâche d'un agent à un autre, d'échange (*swap*) d'affectation entre deux tâches ou une chaîne de transferts et d'échanges. Une illustration de ces mouvements est donnée à la figure 1.3. En 1995, ils introduisent une heuristique hybride [29] qui est une combinaison de ces deux méthodes. Celle-ci utilise la meilleure solution fournie par HGAP après quatre appels et fait appel à VDSH si celle-ci n'est pas optimale.

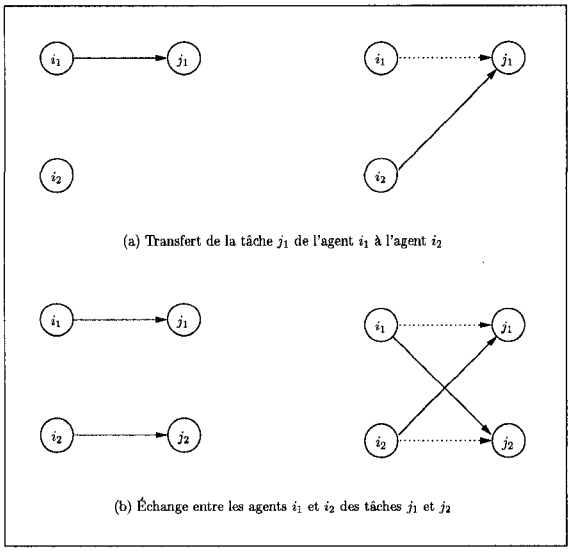

Figure 1.3 – Illustration des différents échanges de tâches utilisés par Amini et Racer [28]

En 2000, Romeijn et Morales [181] présentent des algorithmes gloutons pour résoudre le GAP. Une fonction de pénalité est définie qui évalue le pseudo-coût d'affecter une tâche à

un agent. Les tâches sont ensuite distribuées dans l'ordre décroissant de coûts. Les fonctions utilisées sont celles de Martello et Toth [154] auxquelles est ajoutée une fonction qui permet de tenir compte en même temps du coût c_{ij} et de la consommation de la ressource a_{ij} sous la forme $c_{ij} + \lambda_i a_{ij}$. Il est donc possible de faire varier l'importance du coût par rapport à la consommation de la ressource.

Yagiura et al. [199] ont développé un algorithme tabou qui reprend les composants de VDSH mais inclut le concept de chaîne d'éjection. Cette chaîne d'éjection est composée d'une succession de mouvements de transfert d'une tâche d'un agent à un autre. La longueur de la chaîne dépend du nombre de transferts effectués. Trois voisinages sont considérés, la chaîne simple (un transfert unique), la chaîne double (deux transferts ayant un agent en commun), et la chaîne de longueur variable. Une illustration de ces mouvements est présentée à la figure 1.4. Ces chaînes d'éjection sont construites à l'aide de l'information fournie par la relaxation lagrangienne (le coût pour choisir un voisin tient compte du multiplicateur de Lagrange obtenu par cette résolution). Ces coûts permettent de réduire la taille du voisinage à explorer. L'exploration du voisinage de solutions non admissibles est autorisé par le biais d'une pénalisation de la non-admissibilité de la solution dans la fonction objectif. Ce calcul est ajusté automatiquement par l'algorithme. L'exploration commence par une solution initiale générée aléatoirement et explore le voisinage des transferts simples, passe aux transferts doubles et enfin à la longue chaîne.

Higgins [114] présente un algorithme de recherche avec tabous dynamique pour résoudre des problèmes de grande taille. Il compare quatre versions d'algorithmes basés sur la méthode de recherche avec tabous. Les mouvements privilégiés dans chacun de ces algorithmes sont les transferts et les échanges de tâches. Les particularités des algorithmes concernent l'exploration de la région non admissible : exploration interdite, exploration autorisée suivant des règles [143], exploration autorisée par pénalisation (constante) de la violation dans la fonction objectif, et enfin exploration autorisée par pénalisation dynamique. À ce dernier algorithme est combinée la taille dynamique du voisinage (à cause de la grande taille du problème il est inefficace d'évaluer le voisinage au complet).

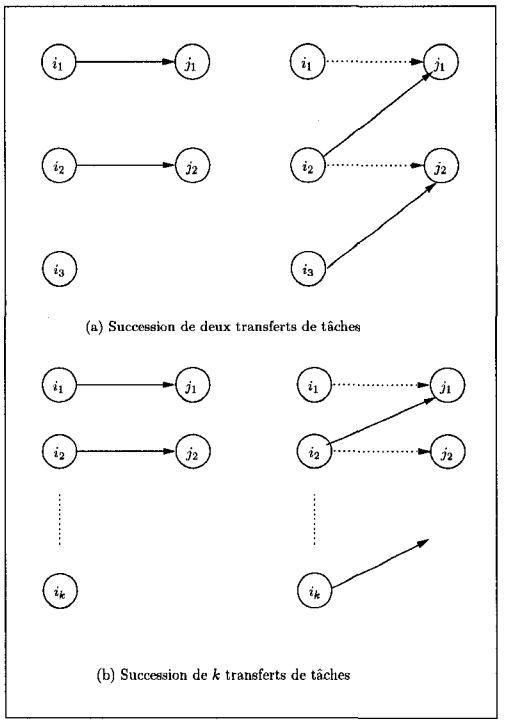

Figure 1.4 – Illustration des différents échanges de tâches utilisés par Yagiura et al. [199]

Diaz et Fernandez [79] proposent un algorithme de recherche avec tabous qui se distingue par sa simplicité et sa flexibilité. Sa flexibilité provient du fait que non seulement l'espace de recherche est élargi en tenant compte des solutions non admissibles, mais que la fonction objectif est adaptée et modifiée au cours de la recherche en tenant compte d'un terme de pénalité qui est ajusté dynamiquement. Pour cela l'algorithme utilise l'historique de la recherche à court et moyen terme. La structure de voisinage retenue est simple et les deux mouvements privilégiés sont l'échange et le transfert (échange entre deux paires (agent, tâche), et transfert d'une tâche d'un agent à un autre). Les auteurs résolvent une version relaxée du problème où les contraintes de capacité sont pénalisées dans la fonction objectif. Cet algorithme a permis d'obtenir des solutions de bonne qualité (en le comparant à plusieurs heuristiques connues) en des temps de calcul raisonnables.

Les algorithmes génétiques [115] ont suscité un réel intérêt pour la résolution du GAP. Chu and Beasley [62] ont testé cette approche. Pour chaque solution du problème, une mesure de non-admissibilité est calculée qui tient compte du dépassement de la capacité. Chaque solution représente une affectation de toutes les tâches aux agents. Le croisement entre deux solutions choisies aléatoirement se fait en seul point.

Une version plus efficace des algorithmes génétiques est proposé par Feltl et Raidl [91]. La phase d'initialisation aléatoire est remplacée par une initialisation basée sur l'heuristique *constraint-ratio* qui tient compte de la contrainte de capacité pour générer des solutions initiales de meilleure qualité qu'ils comparent à la solution de la relaxation linéaire. Une phase d'amélioration suit qui choisit un nombre de tâches aléatoire et les réaffecte en utilisant les fonctions de pseudo-coût de Martello et Toth [154].

L'algorithme utilisé par Wilson [198] se déroule en deux phases. Dans la première, les tâches sont affectées sans tenir compte de la deuxième contrainte qui fixe une limite par agent. Dans la deuxième phase, une mesure de la violation de la contrainte de capacité est calculée. Deux agents sont choisis, l'un enregistrant un dépassement de capacité et l'autre non. Chaque tâche est réaffectée du premier vers le second si une telle réaffectation conduit vers une solution réalisable et si elle n'a pas déjà été visitée. Une phase d'amélioration se déclenche alors basée sur l'échange des agents en charge des tâches pour toutes les paires

de tâches qui ne sont pas affectées à la même personne. Comparé à l'algorithme de Chu et Beasley [62] basé lui aussi sur les algorithmes génétiques, restaurer la réalisabilité en faisant des échanges semble plus prometteur qu'utiliser les échanges à des fins d'amélioration uniquement.

Osman [169] présente une revue des méthodes hybrides combinant des algorithmes exacts et heuristiques pour le problème GAP, notamment une méthode hybride combinant la méthode du recuit simulé [138] et de recherche avec tabous [101].

Yagiura et al. [200] proposent un nouvel algorithme qui semble supérieur en terme d'efficacité aux autres. Ils utilisent l'approche chemin reliant (*path relinking*) qui permet de générer de nouvelles solutions en combinant deux solutions (ou plus). Les mouvements privilégiés pour passer d'une solution à une autre sont le transfert (d'une tâche à un autre agent), l'échange (entre deux paires (tâches, agents)) et enfin la succession de plusieurs transferts et/ou échanges qui s'inspire des principes de chaîne d'éjection.

Kogan et al. ont considéré une capacité stochastique [139]. Romeijn et Piersma [182] étudient la réalisabilité du GAP lorsque les coûts et consommations de ressources sont générés par un modèle stochastique.

D'autres approches ont été proposées telles que les réseaux de neurones [186, 195].

1.5.3 Le problème BGAP

Dans le contexte des soins à domicile, minimiser le coût d'affectation des patients aux infirmières (GAP) peut être remplacé par un autre objectif. Plutôt que de considérer le coût global d'affectation des patients aux infirmières, il pourrait être judicieux de comparer le coût d'affectation des patients pour chaque personne et tenter de minimiser le coût le plus élevé. En d'autres mots, cela revient à minimiser la surcharge (ou la charge) maximale du personnel infirmier. Ce problème est la version min-max du GAP ou version bottleneck du GAP. Il se formule comme suit :

$$\text{Min } z = \max_{j \in N} \sum_{i=1}^{m} c_{ij} x_{ij}$$

sous les contraintes (1.1)-(1.3). La différence avec le GAP se situe au niveau de la fonction objectif, les contraintes étant identiques dans les deux problèmes.

Les auteurs dans [156] présentent un algorithme exact ainsi qu'un algorithme approximatif en introduisant plusieurs bornes inférieures. Des bornes inférieures sont calculées en relaxant les contraintes de capacité (1.2) et en utilisant la relaxation surrogate. Cette dernière relaxation revient à résoudre un cas particulier du *bottleneck multiple-choice knapsack problem* (BMCKP) où des éléments doivent être mis dans un sac j sans dépasser sa capacité et où la valeur maximale est minimisée. L'algorithme développé considère les objets et les met dans un sac dans l'ordre décroissant de coût. La borne obtenue de cette façon est meilleure que la borne de la première relaxation. L'algorithme développé pour BGAP considère les objets dans l'ordre décroissant de la valeur de regret (entre la meilleure et la seconde meilleure) et les insère dans l'ordre de la moindre consommation de ressource (FEAS). Plusieurs algorithmes d'approximation peuvent être obtenus en faisant varier les différents critères utilisés dans FEAS. Ces différentes solutions sont finalement utilisées pour résoudre le problème par branch-and-bound. Le BAP (*bottleneck assignement problem*) est un cas particulier du BGAP où une seule ressource est impliquée. Ce problème aussi est NP-difficile. Il a été résolu par les algorithmes BASS de Carpaneto et Toth [56] et BOTJV de Jonker et Volgenant [136].

Un exemple d'application du BGAP est présenté par Potts et Whitehead [175]. Le problème d'équilibrage de la charge de travail sur les machines en imposant une contrainte de charge maximale est décrit. Cette charge est évaluée par le temps total de traitement sur chaque machine. La charge maximale est le maximum de tous les temps de traitement pour chacun des produits. Cette valeur doit être minimisée dans une deuxième étape. Les auteurs tiennent également compte du temps de transfert entre les machines et tentent de le minimiser dans la deuxième phase de leur méthode. À cette étape, la contrainte de charge maximale est prise en compte en ajoutant un seuil de tolérance de 5% : la charge maximale peut être dépassée pour chaque machine tant que ce n'est pas de plus de 5% de cette charge.

1.5.4 Le GAP avec plusieurs contraintes (MCGAP ou MRGAP)

La formulation du problème de distribution des patients aux infirmières sous la forme du GAP ne considère qu'une seule contrainte de capacité. Il est cependant plus réaliste de considérer que le problème devrait tenir compte de plusieurs contraintes de capacité. En effet le nombre de visites que peuvent rendre les infirmières n'est pas infini, de même que le nombre de patients dont elles peuvent avoir la charge. La reformulation du problème de distribution des patients aux infirmières en problème GAP avec K ressources (ou contraintes) s'exprime comme suit :

$$\text{Min} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

$$\sum_{j=1}^{n} x_{ij} \quad = 1 \quad i = 1, ..., m \tag{1.4}$$

$$\sum_{i=1}^{m} a_{ijk} x_{ij} \quad \leq b_{jk} \quad j = 1, ..., n, k = 1, ..., K \tag{1.5}$$

$$x_{ij} = \quad 0 \text{ ou } 1 \quad i = 1, ..., m, j = 1, ..., n \tag{1.6}$$

Gavish et Pirkul [97, 98] étudient différentes relaxations de ce problème et présentent diverses heuristiques. La première heuristique est basée sur le principe du regret, c'est-à-dire que les tâches affectées en priorité sont celles qui provoquent la plus grande hausse de coût si elles étaient affectées à un autre agent. Les deux autres heuristiques utilisent les solutions obtenues par la relaxation des contraintes. La seconde heuristique développée utilise la solution de la relaxation des contraintes (1.5) et réaffecte les tâches des agents en dépassement de capacité. La troisième heuristique utilise la relaxation de la contrainte d'affectation(1.4), certaines tâches sont affectées plus d'une fois, alors que d'autres ne le sont pas. Le problème d'allocation de ces tâches (non affectées ou affectées plus d'une fois) est résolu de nouveau. Cette dernière heuristique semble donner les meilleurs résultats.

Privault et Hérault [178] approchent le MRGAP par la méthode du recuit simulé [138] pour résoudre un problème de grande taille. Les auteurs avaient déjà tenté une approche par réseaux de neurones dans [177]. Une solution initiale est générée à partir de la solution

de la relaxation linéaire du problème résolue par l'algorithme du simplexe. Les contraintes actives sont utilisées pour limiter l'espace de recherche. La procédure d'échanges est ensuite déclenchée, encore une fois basée sur des mouvements de transferts et d'échanges de tâches. Les auteurs font appel à une procédure d'intensification pour réduire l'espace de recherche qui utilise le coût réduit des variables de base à chaque fois que l'algorithme trouve une solution réalisable.

Yagiura et al. [201] développent une méthode de recherche avec tabous pour résoudre ce problème. Les mouvements privilégiés sont le transfert et l'échange de tâches entre agents ainsi que la séquence de transferts (principe des chaînes d'éjection). Les solutions sont évaluées suivant leur coût ainsi qu'une pénalisation de la non-admissibilité de cette solution (écart entre les capacités et la valeur de l'affectation). L'algorithme débute par des mouvements de transferts, lorsqu'il n'y a plus d'amélioration possible, les mouvements d'échange sont implantés jusqu'à amélioration, et ainsi de suite.

Mazzola et Wilcox [158] introduisent diverses procédures heuristiques pour résoudre le MRGAP. Ils ont tout d'abord développé TPH, une heuristique en trois phases. Les deux premières phases ont pour but de générer une solution réalisable (en utilisant une mesure de regret et la consommation des ressources), la troisième tente de l'améliorer (en réaffectant des tâches). Cette heuristique s'inspire de celles proposées par Gavish et Pirkul [98], où plusieurs relaxations sont résolues par la méthode du sous-gradient et ensuite réparées en réaffectant, par exemple, les tâches de agents enregistrant un dépassement de capacité.

Des versions dynamiques de ce problème sont étudiées dans [184] où la demande pour les tâches change dans le temps et par Toktas et al. [190] où les capacités ne sont pas connues d'avance.

1.6 Résumé du chapitre

Dans ce chapitre nous avons introduit les diverses organisations de la main-d'œuvre dans les deux contextes de soins à domicile et d'industrie de la fabrication que nous avons choisi de traiter dans cette thèse. Nous avons ensuite présenté divers travaux traitant du

concept de charge de travail, et ce dans divers contextes. Nous avons finalement décrit les différentes approches pour distribuer la charge de travail, que ce soit par l'affectation de la charge de travail ou par la construction d'horaires. Les trois prochains chapitres sont la reproduction d'articles publiés ou soumis traitant de l'analyse de l'approche territoriale en vigueur au CLSC CDN, de l'algorithme développé pour mieux distribuer la charge de travail aux infirmières et enfin de l'approche intégrant la flexibilité de l'annualisation du temps de travail pour faire la planification de la production.

CHAPITRE 2 : ANALYSIS OF A TERRITORIAL APPROACH TO THE DELIVERY OF NURSING HOME CARE SERVICES BASED ON HISTORICAL DATA

Nous avons fait une analyse détaillée de la distribution des patients aux infirmières selon l'approche territoriale en se basant sur le cas du CLSC CDN. Le CLSC CDN a divisé son territoire en six secteurs, chacun géré par une équipe multidisciplinaire. Une équipe d'infirmières supplémentaire appelée équipe de surplus est disponible et a pour objectif de servir de tampon et d'absorber le surplus de travail du personnel de secteur régulier. Deux catégories d'infirmières sont en place dans chacun des secteurs, une équipe d'infirmières bachelières en charge des cas à long terme et une équipe d'infirmières soignantes en charge des cas à court terme. Un standard reçoit toutes les demandes de services et transfère chacune des demandes reçue au professionnel approprié en se basant sur l'adresse du patient. En se basant sur des données historiques sur le nombre de visites ainsi que la durée de ces visites (pour les équipes d'infirmières bachelières, soignantes et de surplus), nous avons établi que le découpage du territoire en six secteurs conduit à un déséquilibre des charges de travail entre les infirmières. La raison principale tient du fait que le découpage du territoire date de plusieurs années (quatre ans) et qu'il a lui-même été basé sur des données historiques datant de deux ans. L'évolution hétérogène de la demande dans les différents secteurs, sans ajustement du nombre d'infirmières, a conduit à une augmentation excessive du nombre de cas ainsi que du nombre de visites dans certains secteurs (alors que certains ont même enregistré une baisse). L'analyse du recours à l'équipe de surplus, qui a pour rôle premier, rappelons-le, d'absorber le surplus, a montré que la notion de surplus est variable d'un secteur à l'autre, ainsi que d'une infirmière à l'autre. Finalement cette analyse a montré qu'une distribution des patients basée sur l'appartenance au secteur uniquement n'est pas optimale en termes d'équilibre des charges des infirmières. La solution envisagée en collaboration avec le comité directeur du CLSC est une approche plus dynamique, toujours basée sur le territoire, mais qui permet de tenir compte de la charge de chaque infirmière. Afin

de pouvoir mettre en place ce type d'approche il faut rendre les frontières plus souples et permettre aux infirmières de se déplacer d'un secteur à l'autre pour répondre à la demande.

Analysis of a territorial approach to the delivery of nursing home care services based on historical data

N. Lahrichi[1,2], S.D. Lapierre[1,2], A. Hertz[2,3],
A. Talib[4], L. Bouvier[4]

[1]Centre de recherche sur les transports, C.P. 6128,
succ. Centre-ville, Montréal, Canada H3C 3J7
[2]École Polytechnique, C.P. 6079 succ. Centre-Ville,
Montreal, Canada H3C 3A7
[3]GERAD - École des HEC,
3000, chemin de la Côte-Sainte-Catherine, Montréal, Canada H3T 2A7
[4]CSSS Côte-des-Neiges, Métro et Parc Extension
5700, chemin de la Côte-des- Neiges, Montreal, Canada H3T 2A6

Article paru dans
Journal of Medical Systems, Volume 30, 2006. Pages 283-291

Abstract

We analyze a territorial approach to deliver nursing home care services to a territory public health. We present the case of the CSSS assigned to Côte-des-Neiges, Métro center and Parc Extension, specifically the case of the Côte-des-Neiges site (CLSC CDN), where a territorial approach is used since 1980. We first give an historical comparison of patient visits delivered in 1998-1999 and in 2002-2003. We follow with an in-depth analysis of the home services delivered in 2002-2003 to determine whether or not the territorial approach can well support the changing needs of the population. We conclude that the territorial approach to deliver homecare nursing services does not sufficiently support fluctuations in population needs for services. Not only is it difficult to predict these fluctuations, but it is difficult to accurately quantify the true needs for services since the availability of nursing services tend to determine the services actually delivered. In sectors of the territory where resources are more scarce (based on previous population needs analyses) or demand for services is greater, the result is work overload for the nursing staff. In addition, this results in service delivery inequities across the entire territory. Therefore, a more dynamic assignment of clients to the nurses based on each nurse's workload and case load rather than based on the geographic location of clients is worth the extra administrative time in case assignment to ensure a more equitable case load attribution between nurses as well as less inequities between clients in terms of service delivery considering their needs.

2.1 The territorial approach

In Canada, as in many Western countries, home care services is taking a greater place in the health care system. Health Canada reports that total public funding for home care has increased substantially over the last seven years (from $1.028 billion in 1990-91 to $2.096 billion in 1997-98), an average annual rate of increase of almost 11% [20]. This is due in part to the ageing of the population, but also to the need to reduce or at least contain health expenditures. It is difficult to estimate the homecare expenses since they are not computed separately. In 1998, Canadian provinces have allocated 2 to 6% of their yearly health care budget to homecare services, or 34$ to 124$ per capita [20]. In the United States, annual expenditures for home health care are projected to be $41.3 billion in 2001 [1] or about 145$ per capita.

In the province of Quebec, homecare services fall under the responsibility of local community health centres (or CLSC for Centre Local de Services Communautaires) now merged into CSSS (for Centre de Santé et de Services Sociaux) since 2005. In 2003, there were 398 CLSC's in the province of Quebec [23]. The province has spent $19,1 billions in 2003-2004 for health care services delivered to the population, about 2639$ per capita and 8% of its Gross Domestic Product (GDP) (this does not include the health care expenditures made by the private sector not covered by the national health insurance) [13, 23]. CSSSs are responsible for delivering homecare services to the population living in the *territory* assigned to them. In fact, contrary to hospitals which cater to the population beyond a specific territory, CSSS's are responsible for the well-being of the population on a given territory. Homecare services may be required for acute illness, post-hospitalization and post-operation treatment, long-term health conditions and/or chronic conditions, permanent disability, including physical and mental disability, or terminal illness. Homecare services are provided by health care professionals that include nurses, occupational therapists, physiotherapists, nutritionists, homemakers, social workers and physicians.

Funding of Quebec health care institutions have been traditionally based on historical budgets. Over the years, efforts have been made to take demographic data as well as data

on resource utilization into account to determine future budgets. Needs based budgeting is however still in its embryonic stage.

This study originates from a practical case encountered at CLSC CDN in Montreal which caters to 125 000 inhabitants, among which 5200 are regular homecare service users. The territorial approach to manage homecare services has been used since 1980. Given the size of the territory, the management team decided to partition the territory into districts, with each district being assigned to a multidisciplinary team of professionals. This has allowed for increased efficiency in terms of case assignment since the geographic location of the client determines which team will be responsible for the care of that client. Furthermore, this allows for a reduction in transport time and therefore allows for more time for direct patient care.

Since September 2000, the territory has been divided into six districts, after having been divided into four districts from 1993 to 2000 in order to absorb the increase in the number of patients and staff.

As indicated, the use of districts facilitates the dispatching of homecare services. When a request for homecare services is sent to the CLSC from a hospital or a physician's office, or when a request is made by a patient or a family member or a friend, the intake nurse identifies the district associated with the patient's address and forwards the request to the manager of the team responsible for that particular district, after having analyzed the nature and urgency of the request for services. The analysis by the intake nurse will yield a decision to assign the case to one professional so that a patient requiring nursing care as a major component of his care plan will be assigned to a nurse. This decision is confirmed by the manager of the multidisciplinary team who receives every new request for homecare services and who may decide to assign the case to a different professional, based on the information on the patient's condition. Of course, the professional responsible for the case will, in most cases, involve other professionals in the care of the patient.

A distinction is made between *case manager nurses* (who typically hold a Bachelor's degree in nursing) and *nurse technicians* (who typically hold a community college degree

in nursing). The nurse technician will be assigned the short-term cases or the ones not requiring case management.

For instance, a case requiring short-term and specific nursing care such as a wound dressing or a home based antibiotic-therapy treatment post-operation or post-hospitalization, will typically be assigned to the nurse technician of the team (there is one nurse technician per team).

Conversely, a case requiring the organization of a more complex service plan such as organizing activities of daily living, coordinating visits to and ensuring links with doctors and specialists as well as with the pharmacist, consulting with and arranging for evaluation by other professionals (occupational therapist, physiotherapist, social worker, dietician, etc.) will be assigned to a case manager nurse. Such cases typically include frail elderly patients with a great loss of autonomy, palliative patients, patients with cancer, patients suffering from degenerative diseases or chronic illnesses and patients with serious mental health problems.

There are about two to four case manager nurses per team. In addition to the nurses (case managers and nurse technicians) who are part of the six multidisciplinary teams and to whom cases get assigned to, there are three to four nurses who make up the surplus team. These nurses are not assigned any cases nor a particular district. They are asked to deliver specific nursing care treatments by the professionals responsible for the case and are not responsible for the global care plan of the client. This team will handle nursing visits that the team nurses are unable to absorb, given the number of visits they have already scheduled for themselves. Furthermore, since the working hours of the surplus team nurses are extended until late in the evening (11 p.m.) as well as on week-ends, these nurses are able to absorb visits that are needed outside regular working hours. For instance, a wound dressing that would be required three times a day, seven days a week, will be taken care of by the homecare department and, even though one nurse will be assigned that particular case, she cannot be required to perform all the nursing visits needed. However, the surplus team nurses will be able to absorb some, if not most, of the visits required.

Another feature of the surplus team is that it can serve as a "buffer" team to absorb temporary increases in demands for nursing services thereby contributing to the reduction in work overload. It should be noted however that despite the fact that the surplus team can serve as a buffer, it is difficult to absorb increases in demands in all situations due to, on the one hand, the difficulty to predict these peaks in demands and, on the other hand because the surplus team is often used to compensate for the shortage of nurses during absences of regular nurses. Whether the absence is planned or not (sick day), it is indeed often difficult to find a nurse to replace the one that is absent.

When an increase in workload is more sustainable, the homecare department may also resort to outside agencies. However, such agencies are usually used for long term replacement of nurses.

Since client assignment to nurses is performed according to territory, i.e., the territorial origin of the demand, and not according to actual nurse workload, partitioning of the territory and assignment of nurses to each district must be carried out carefully in order that the nurses end up with similar case loads. The partitioning is performed on the basis of historical data on number of patients and number of nursing visits. Since the population make-up changes over time thereby bringing about changes in the demands for services originating in each district, districting exercises must be performed on a regular basis to compensate for the fluctuations in demands over time. But since reorganizing districts requires time and resources not to mention that it can be very disruptive for patients as well as for the nurses because of all the changes involved, such an operation would make sense if the demands for services were to be relatively stable for at least five to ten years. Otherwise, a more dynamic method should be considered to assign clients to nurses. By that, we mean that each case assignment should be based on the actual case load of every nurse at the time of the demand for services. While the previous division of the territory into districts at CSSS CDN lasted 7 years, there is now a sense that this approach can no longer sustain now more rapid changes in population needs. In addition to workload inequities between nurses that such fluctuations tend to create, it has been observed that the availability of nursing services tend to determine the services actually delivered. This

in turn leads to inequities on how patients are being serviced depending on their place of residence.

Before switching to an alternative method of assigning new cases, an in-depth analysis and assessment of the territorial approach should be performed. This assessment should also include an assessment of the significance of the perceived inequities or differences between case loads as well as differences between nursing services provided for similar patient needs. This assessment will help to determine whether or not the territorial approach is still suitable for the CLSC CDN.

The present paper meets this objective by providing a description of the districting exercise as well as a quantitative assessment of service fluctuations per district using 2002-2003 and 1998-1999 data, the latter having been used to design the current districts. This is followed by an analysis of the use of the surplus team as well as outside agencies in the different districts for the 2002-2003 period. We conclude with a discussion on the advantages and disadvantages of the territorial approach.

2.2 Fluctuations between 1998-1999 and 2002-2003

The territory of CLSC CDN is currently divided into six *districts*, each one being constituted by several *basic units* which are the census tracts used by Statistics Canada. The current division into districts results from an analysis performed in 1998-1999, and summarized in [42]. Five criteria have been used for this districting exercise :

- indivisibility of the basic units : each one has to be assigned to only one district
- respect of borough boundaries : the territory of CLSC CDN intersects three different cities (Mont-Royal, Outremont and CDN/Snowdon), and it is preferred that a district not span two different cities for more efficient work with community agencies, which traditionally cater to a specific borough.
- connectivity : no enclaves are authorized

- mobility : since it is encouraged that the visiting staff use public transportation, the districts should be designed so that travel by bus is easy. Also, travel should not be constrained by major barriers such as railway lines or motorways.
- workload equilibrium : the total workload (measured by the sum of the time spent for the visits and for traveling) of each district should be roughly the same.

The model used in [42] focused primarily on the three first criteria and tried to reach the two last ones in the best possible way. A picture of the resulting partition into districts is given in Figure 2.1. In 2000, it was considered as optimal in terms of the satisfaction of the professionals, the team managers and the head managers. In the next sections, we check whether the optimality was still reached in 2002-2003.

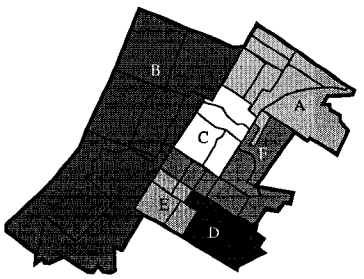

Figure 2.1 – Picture of the different districts constituting the territory of the CLSC Côte-des-Neiges

Comparison between 1998-1999 and 2002-2003

Since complete data are available for the 2002-2003 period, they will be compared with those of 1998-1999 that served for the partition into districts. Such a comparison will help us to determine whether or not the territorial approach supports the changing needs of the population over time.

In Table 2.1, we compare the total number of visits in the six districts. In 1998-1999 each district had between 12% and 20% of the CLSC's visits, and the situation is almost the same in 2002–2003 since each district has from 12 to 21% of the visits. However, in 2002–2003 the total number of visits has increased by 20% compared to 1998–1999, and

Table 2.1 – Number of visits per district for 1998–1999 and 2002–2003

District	Number of visits		Distribution of the visits (%)		% change
	1998-1999	2002-2003	1998-1999	2002-2003	
A	5014	6892	16	18	+34
B	3739	4655	12	12	+24
C	5652	7832	18	21	+39
D	4538	5285	15	14	+16
E	6206	5333	20	14	-14
F	6059	7555	19	20	+25
Total	31208	37552	100	100	—
Average	5201	6259	—	—	+20

this increase is not the same for each district, as shown in the last column. For example, district E has a decrease in terms of annual number of visits, while districts A and C have seen the largest increase, followed by districts B and F. District D has seen an increase of 16%, but since the average number of visits per district has increased by 20%, it is equivalent to a decrease of relative needs.

We can also observe a big change in the percentage of the demands supported by each district. For example, district E was taking care of about 20% of the population in 1998-1999, and this number has decreased to 14%. There is a smaller decrease of 1% for district D. Conversely, nurses of district C took care of 3% more of the population, and this increase is of 2% for district A and 1% for district F.

In conclusion, this first measure of equilibrium shows that four years are sufficient to create a wide change in the total number of visits and the distribution of these visits among the districts.

When the population needs change, the CLSC has the possibility of modifying the assignment of nurses (of the permanent staff) to districts. As shown in Table 2.2, the assignment in 2002-2003 is not the same as the one recommended by [42] in 1998-1999.

Table 2.2 – Evolution of the number and distribution of nurses

District	Number of nurses		
	1998-1999	2000-2001	2001- 2005
A	4	5	5
B	3	3.5	4
C	5	4.5	5
D	4	5	3
E	5	4	4
F	5	4	5
Total	26	26	26

- In the column labeled 1998-1999, we indicate the assignment recommended in [42]. This assignment was never implemented because changes had occurred between the districting exercise and its implementation in September 2000.
- In the column labeled 2000-2001, we indicate the assignment that was valid from September 2000 to June 2001. Notice that one nurse was working half time in B and the other half in C.
- The last column corresponds to the current situation that is valid since June 2001. The nurse that was working half time in B and half time in C has been assigned full time to B. Notice also that Sector D has lost 2 nurses : one was assigned to district C and the other one to district F.

If we compare the 1998-1999 column with the last one of Table 2.2, we can observe that the number of nurses was decreased by one unit in districts D and E and increased by one unit in districts A and B. This is coherent with Table 2.1 which indicates a decrease of relative needs in districts D and E, and an increase in districts A and B. We can therefore point out that the CLSC CDN has adapted the assignment of nurses to the changing needs of the population.

In Table 2.3, we report the average number of visits per nurse. We can observe a big variance between the districts. For example, in 1998-1999, a nurse in district A had to perform an average of 1254 visits while the average number of visits per nurse in the

whole territory was 1203. This represents an excess of 4.2%. Districts B, E and F had also more visits than the average, the excess being of 3.6%, 3.2% and 0.7%, respectively. Conversely, districts C and D had 6% and 5.7% less visits than the average. Hence, there was a difference of 10.2% between the least and the most busy districts. The situation is totally different in 2002-2003. The most significant differences occur in districts B and D : district B was 3.6% above the average and is now 19.9% below the average ; district D was 5.7% below the average and is now 21.3% above the average. The difference between the least and the most busy districts is 41.2% in 2002-2003, which points to a large imbalance.

Another way of studying the change in the distribution of the number of visits is to simply measure the increase or decrease of this number in each district. For example, the number of visist has changed from 1254 to 1378 in district A, which represents an increase of 10%. As shown in Table 2.3, there was in fact an increase in each district except in district B where the decrease is of 6.6%. The most critical situation seems to be in district D where the number of nurses is 3 while the number of visits has increased by more than 50%, followed by district C where the number of nurses is kept equal to 5 while there is an increase by about 39% in the number of visits.

When looking at Table 2.3, one could conclude that districts C, D and F are the most productive ones, followed by A and E. District B seems to be the less productive. However, in order to get a sense of the productivity of a district, one must also take into account the duration of the visits. Each visit has a duration that strongly depends on the type of care. For example, a visit for a blood drawing is more likely to be of a short duration as compared to a visit to a frail elderly patient needing the organization of a more complex service plan. However, in this case, it is also possible for a nurse to visit that particular patient more often and for short visits when the service plan has been implemented and the situation stabilized.

To conclude this section, we can observe that the CLSC CDN has modified the staff assignment (without changing the district zoning) in order to adapt to the changes in the population needs. But this reassignment appears to be imperfect since districts A, B and E seem to be less productive than the others. We can however not state that the other

Table 2.3 – Evolution of the number of nursing visits per district

District	average number of visits per nurse		% difference with the average		% change over time
	1998-1999	2002-2003	1998-1999	2002-2003	
A	1254	1378	+4.2	-5.1	+10
B	1246	1164	+3.6	-19.9	-6.6
C	1130	1566	-6.0	+7.8	+38.6
D	1135	1762	-5.7	+21.3	+55.4
E	1241	1333	+3.2	-8.2	+7.4
F	1212	1511	+0.7	+4.0	+24.7
Average	1203	1452			

districts (C, D and F) are the most productive since the larger number of visits may be the result of shorter visits and less complex cases . A more in-depth analysis of the 2002-2003 situation is carried out in the next section.

2.3 An in-depth analysis of the 2002-2003 period

To better understand inequities during the 2002-2003 period, we include three important factors in our analysis.

- A first visit to a patient is usually longer than subsequent ones, the reason being that data collection and elaboration of the care plan are done during the first visit to the patient. A district with many new cases can therefore appear as less productive than the other districts.
- During the absence of permanent CLSC staff due to illness, disabilities, maternity leaves, training, etc, nurses from *outside agencies* are called for replacement.
- The *surplus team* is called for the visits that need to be performed during regular hours (that the regular staff is unable to absorb due to work overload), evening and weekend hours. Also, some patients need to receive more than one visit per day, and these supplementary visits are handled by the surplus team.

Table 2.4 – Number of different patients per month, per district, and per nurse

District	Apr.	May	June	July	Aug.	Sep.	Oct.
A	245	230	202	206	213	207	212
B	158	144	122	133	130	100	125
C	228	230	213	219	220	210	208
D	171	161	162	152	157	142	165
E	192	184	148	184	154	135	152
F	228	215	177	202	193	185	189

District	Nov.	Dec.	Jan.	Feb.	Mar.	Total	ρ
A	254	222	189	200	206	712	11.87
B	168	181	135	123	155	477	9.94
C	259	199	227	210	209	732	12.20
D	159	146	150	120	142	550	15.28
E	141	125	129	178	114	508	10.58
F	238	185	224	241	250	647	10.78

2.3.1 New patients

In Table 2.4, we show the number of different patients per month receiving care in each district. In the last two columns, we give the total number of different patients seen during the whole period as well as the average number ρ of different patients per month visited by a nurse. This value is obtained by dividing the total number of different patients in a year by 12 (for an average per month), and by the number of nurses. For example, a nurse in district A will care for an average of 11.87 different patients per month, this number being the result of the division of 712 by 60.

Districts B, E and F have the lowest average ratio of different patients per nurse per month while D has the highest average ratio.

On the basis of detailed data for the seven last months of the 2002-2003 period, we have been able to count the number of new patients per month in each district. These numbers are reported in Table 2.5. We can observe a peak in November. The column

Table 2.5 – Number of new clients received by each district in the September 2002 – March 2003 period

Dist.	Sept	Oct	Nov	Dec	Jan	Feb	March	Average	Average by nurse
A	49	51	67	47	30	45	48	48	9.6
B	28	37	57	39	43	36	28	38	9.5
C	49	47	62	47	43	42	49	48	9.6
D	35	49	49	34	25	39	40	39	13
E	27	41	40	27	27	33	19	31	7.75
F	38	51	75	44	40	45	33	47	9.4

labeled "Average" indicates the average number of new clients per month in each district. We can observe big differences since, for example, there are in average 31 new patients per month in district E, while this number is equal to 48 in districts A and C. The last column indicates the average number of new patients per month and per nurse. This average is the highest in district D, and the lowest in district E.

Putting together Tables 2.3, 2.4 and 2.5, we can produce the three following "productivity" rankings of the districts. It is interesting to note that these rankings are similar (all rankings are shown from the lowest to the largest value) :

- ranking according to the average number of visits per nurse (see Table 2.3) : B, E, A, F, C, D
- ranking according to the average number ρ of different patients per nurse and per month (see Table 2.4) : B, E, F, A, C, D
- ranking according to the average number of new patients per month and per nurse (see Table 2.5) : E, F, B, A, C, D

Districts C and D are the two most productive ones according to the three rankings, while district E is always among the two less productive. This is perhaps due to a lower number of different patients. District B appears as not very productive according to the first two rankings, but this can be due to a larger number of new patients, as shown by the third position of district B in the third ranking, lower number of different patients and

Table 2.6 – Distribution of 2002–2003 visits

District	Total number of visits	by nurses from the regular staff	by nurses from outside agencies	% by nurses from outside agencies
A	6892	5608	1284	19
B	4655	2587	2068	44
C	7832	6005	1827	23
D	5285	4476	809	15
E	5333	3761	1572	20
F	7555	5779	1776	23
Total	37552	28216	9336	
Average				25

possibly the instability of the replacement (absenteeism observed in B). This confirms the need to analyze the load of each nurse in more details.

2.3.2 Outside agencies

In Table 2.6, we show how the total number of visits in each district (which can be read in Table 2.1) is distributed among nurses from the regular staff and those from outside agencies.

On average, nurses from agencies deliver 25% of the visits. District B made the largest use of this outside help (44%), while district D has the highest percentage of visits performed by the permanent staff (85%). This is consistent with Table 2.3, since we have observed that district B has the lowest number of visits per nurse while this number is the highest in district D. Districts A and E look similar in both tables, and the same observation holds for districts C and F. It has been observed that during long-term absences of nurses, the total number of visits over a long period of time carried out by outside agencies is consistently lower. This is explained by several factors :

- it is not possible to find replacement for every day of absence and therefore non-urgent visits are cancelled when replacement is not found;

Table 2.7 – Average number n of visits per day and average visit duration τ

District	Nurses from agencies		Regular staff			
			case manager nurses		nurse technicians	
	n	τ	n	τ	n	τ
A	11.02	0.48	4.82	0.54	6.52	0.41
B	7.91	0.52	3.61	0.67	4.78	0.54
C	11.02	0.51	4.71	0.54	4.7	0.48
D	7.68	0.69	4.68	0.52	6.09	0.37
E	10.59	0.50	4.17	0.59	5.23	0.58
F	9.62	0.47	4.01	0.64	5.07	0.52

- it takes more time for an agency nurse at the beginning of the replacement to get acquainted with the patients and therefore she will not be able to visit as many patients as the regular nurse;
- furthermore, the stability of the replacement (one agency nurse versus several agency nurses for short periods of time each time) will also affect the total number of visits.

Therefore, one can conclude that problems with absenteeism and possibly replacement in district B may explain the low number of visits.

With the help of detailed data for the seven last months of the 2002-2003 period, we have been able to split the nurses into three categories :

- nurses from outside agencies,
- nurses from the regular staff with a Bachelor degree, called *case manager nurses*,
- nurses from the regular staff with a community college degree, called *nurse technicians.*

We have computed the number n of visits performed in average by each type of nurse in each district, as well as the average duration τ of a visit. These numbers are reported in Table 2.7. We can observe that *case manager nurses* usually do 4 to 5 visits per day while this number can go up to 7 for the *nurse technicians.*

We observe from Table 2.7 that the ranking of the districts according to the number of visits per day (from the lowest to the largest number) for case manager nurses is B, F,

E, D, C, A, while this ranking is C, B, F, E, D, A for nurse technicians. Hence, district A seems to be the most productive district, and B appears to belong to the set of not very productive districts. However, the average duration of a visit is the longest in district B. This can be explained by the fact that district B is more extended than the others which implies longer travel times for the nurses. District A has many visits for blood drawing. Since such visits are of short duration, this can explain the large number of visits per day and the small average visit duration.

As already mentioned, a first contact with a new client lasts longer than a regular visit. Notice however that a long average visit duration is not only due to a larger number of new patients. Indeed, for example, we observe from Table 2.7 that the regular staff in district E has relatively long average visit durations while this district has the smallest average number of new clients per nurse (see Table 2.5).

2.3.3 Surplus team

We report in Table 2.8 the number of visits made by the surplus team during the seven last months of the 2002-2003 period. We can clearly observe large disparities among the districts. An important use of the surplus team in a district is either a sign that the nurses are overloaded in this district or that many patients need more than one visit. Districts C and F make the largest use of the surplus team, followed by A and B, and then by D an E.

To further analyze the use of the surplus team during the seven last months of the 2002-2003 period, we give in Table 2.9 the total number of visits per month in each district, and we then report in Table 2.10 the percentage of visits delivered by the surplus team. We can observe large fluctuations in the total number of visits per month, with a peak in November. Conversely, the use of the surplus team does not vary a lot. For example, if we compare October with November in district D, we can see that the total number of visits does not vary a lot (from 427 down to 413) while the percentage of visits delivered by the surplus team increases from 12% to 25%.

Table 2.8 – Distribution of the visits delivered by the surplus team during the September 2002 – March 2003 period

District	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Total	Average
A	118	115	137	126	99	105	161	861	123
B	118	164	125	90	112	68	179	856	122
C	197	159	224	178	154	105	157	1174	168
D	119	50	103	71	60	47	23	473	67
E	67	54	77	72	51	77	90	488	70
F	205	203	246	190	169	170	138	1321	189
Total	824	745	912	727	645	572	748	5173	—

The last column of Table 2.10 shows that in average, the surplus team takes care of 31% of the visits in district F, while this percentage goes down to 17% and 18% in districts D and E respectively. This points to an important imbalance.

As already mentioned, the first contact with a new patient lasts longer than subsequent visits. Hence, a large number of new patients can explain a larger use of the surplus team. However, by comparing Table 2.5 with Table 2.10, it clearly appears that the number of new patients is not correlated with the use of the surplus team. For example, by comparing October and December in district E, one can read in Table 2.5 that the number of new patients has decreased from 41 down to 27, while Table 2.10 reports an increase of the use of the surplus team from 13% up to 16%.

We finally analyze if the workload fluctuations have an impact on the duration of the visits. For this purpose, we report in Table 2.11, the total number of visits and their average durations for three different months : September and December 2002 and March 2003. We put in bold the highest visit durations and in italic the lowest ones for each month. We can see that district D has systematically the lowest durations, while districts B, E and F have the highest. It appears that the nurses adjust themselves : the number of clients to visit tends to determine the duration of the visit.

Table 2.9 – Distribution of the visits delivered to each district during the September 2002 – March 2003 period

District	Sep.	Oct .	Nov.	Dec.	Jan.	Feb.	Mar.	Total
A	531	553	665	615	453	510	614	3941
B	312	447	488	370	398	314	529	2858
C	655	669	719	622	621	485	564	4335
D	474	427	413	416	372	309	320	2731
E	334	423	422	439	377	391	366	2752
F	634	630	746	555	532	593	521	4211
Total	2940	3149	3453	3017	2753	2602	2914	20828

Table 2.10 – Percentage of the visits delivered by the surplus team during the September 2002 – March 2003 period

District	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Average
A	22	21	21	20	22	21	26	22
B	38	37	26	24	28	22	34	30
C	30	24	31	29	25	22	28	27
D	25	12	25	17	16	15	7	17
E	20	13	18	16	14	20	25	18
F	32	32	33	34	32	29	26	31
Total	28	24	26	24	23	22	26	**25**

Table 2.11 – Number of visits, and average visit durations

District	September 2002		December 2002		March 2003	
	n	*t*	n	*t*	n	*t*
A	531	0,50	615	0,50	614	0,50
B	312	0,53	370	**0,56**	529	0,50
C	655	0,51	622	*0,49*	564	0,56
D	474	*0,48*	416	*0,49*	320	*0,44*
E	334	**0,55**	439	0,53	366	**0,62**
F	634	0,53	555	**0,56**	521	0,52

2.4 Conclusion

The territorial approach is with no contest a good way for providing homecare services. The most important advantage of this approach is that case assignment is very straightforward since it is dependent on the geographic location of the client. This in turn shortens the delay between the time the intake nurse receives the request and the actual assignment of the case.

There are however some disadvantages that have to be taken into account since they can create workload inequities between nurses. The first problem is due to the fact that the districting exercice can not forecast the demand in each district. We have observed that the current partition into districts was optimal in 1998-1999, while disparities are apparent between nurses 4 years later. Since the population needs are fluctuating from a district to another, one always has the possibility to reassign nurses to different districts in order to maintain a balance of the workloads. However, even if two such re-assignment have been performed between 1998-1999 and 2002-2003, the difference between the most busy district and the least one has become very large. Another disadvantage of the territorial approach is that it implies that the nurse assigned to one district need not be concerned with the increased workload for her colleague in another district. This approach does not encourage collaboration between nurses. Furthermore, each nurse will adjust her practice according to her workload. This does not lead to uniformity of practice.

Reorganizing the districts is time consuming, and the CLSCs are therefore not interested in modifying the partition into districts too often. If the partition has to remain stable while the population needs are changing, one could think about using a more dynamic assignment approach where the nurses are not assigned to a fixed district (while the patients are). When a request for homecare services is sent to the CLSC, the intake service would identify the district associated with the address and transfer the demand to the team manager. The manager would then choose a nurse for this new patient, taking into account the workload of each nurse as well as the location of the current patients of each nurse. This however implies the use of an efficient information system. Such a system will increase the availability of information to all nurses, which will in turn lead to more transparency and uniformity in

nursing practice. Another possibility that would lie between the current situation and the above proposed one, is to divide the nurses into two groups : the nurses from the first group are assigned to a specific district (like in the current situation), while each nurse of the second group can work on all the territory, or in a fixed subset of districts. This second option is in fact equivalent to increasing the number of nurses in the surplus team. Such solutions imply major changes in the way cases get assigned at the CLSC CDN but can lead to optimized situations in terms of reducing imbalance and inequities.

CHAPITRE 3 : A PATIENT ASSIGNMENT ALGORITHM FOR HOME CARE SERVICES

Nous avons étudié le problème d'affectation des patients aux infirmières au service de soins à domicile du CSLC CDN. L'objectif est de distribuer les patients (visites ainsi que la prise en charge du suivi du dossier) tout en trouvant un équilibre entre les charges de travail des différentes infirmières. Nous avons développé une mesure de la charge de travail qui se base sur les visites, la charge de cas et les déplacements. Les visites et la charge de cas tiennent compte de la "lourdeur" des patients puisque le niveau de soins requis par chaque patient est différent. Nous avons défini cinq catégories de patients, chacune ayant un poids différent reflétant le niveau de "lourdeur". Deux catégories d'infirmières sont considérées, des infirmières bachelières et des infirmières soignantes. Les premières prennent en charge les patients à moyen et long terme nécessitant de la gestion de cas tandis que les secondes prennent en charge les patients à court terme, généralement des post-hospitalisés et des post-opérés. Chacune des infirmières est affectée à un ensemble d'unités d'un secteur donné. L'approche que nous avons développée tient toujours compte des unités d'affectation des infirmières, mais encourage les déplacements vers les unités surchargées de travail.

Nous avons modélisé ce problème sous la forme d'un problème d'affectation généralisé avec plusieurs contraintes (MRGAP) où certaines contraintes ainsi que la fonction objectif sont non linéaires. Ce problème est reconnu comme NP-difficile. Nous avons proposé un algorithme basé sur la méthode de recherche avec tabous pour résoudre ce problème. Nous avons testé notre algorithme sur les données disponibles du CLSC CDN. Nous avons comparé différents scénarios en donnant des poids différents aux différentes composantes de la charge de travail des infirmières. Notre approche s'appuie aussi sur l'assouplissement des frontières entre secteurs. Nous avons donc introduit un paramètre permettant de pénaliser les déplacements et de tester différentes possibilités de rendre les frontières plus flexibles à la demande.

Lorsque la composante de charge de cas n'est pas considérée dans la charge de travail, le problème d'affectation est plus simple à résoudre puisque toutes les contraintes deviennent linéaires. Dans ce cas de figure il est possible de résoudre le problème directement avec un solveur exact tel que CPLEX. Nous comparons les solutions obtenues par CPLEX avec celles obtenues avec notre algorithme pour cette version simplifiée du problème afin de tester son efficacité.

A Patient Assignment Algorithm for Home Care Services

Alain Hertz **Nadia Lahrichi**
alain.hertz@gerad.ca nadial@crt.umontreal.ca

Département de mathématiques et de génie industriel
École Polytechnique de Montréal
C.P. 6079, succ. Centre-ville
Montréal, Québec H3C 3A7

Article à paraître dans
Journal of the Operational Research Society

Abstract

We consider the problem of assigning patients to nurses for home care services. The aim is to balance the workload of the nurses while avoiding long travels to visit the patients. We analyze the case of the CSSS Côte-des-Neiges, Métro and Parc Extension for which a previous analysis has shown that demand fluctuations may create work overload for the nursing staff. We propose a mixed integer programming model with some non linear constraints and a non linear objective which we solve using a Tabu Search algorithm. A simplification of the workload measure leads to a linear mixed integer program which we optimize using CPLEX.

Keywords : health service ; heuristics ; assignment problem

3.1 Introduction

The "Ministère de la Santé et des Services Sociaux" (MSSS) and its network offer health and social services to the entire population of Québec to ensure the welfare of its residents. In 2004-2005, 37% of the overall budget of the government of Québec was allocated to health and social services. Orientations, budgetary resources and results assessment obtained in the entire health care network are established at the central level. At the regional level, the "Health and Social Services Agencies" are charged with regional planning, resource management and budget allocation to institutions. At the local level, the 95 "Health and Social Services Centres" (CSSS) established in June 2004 and their partners in local services networks share a collective responsibility for the population on their local territories.

The CSSSs were created by merging existing local community health centres (CLSCs), residential and long-term care centres (CHSLDs) and general and specialized hospital

centres (CHSGSs). Each CSSS ensures the population on its territory has access to health and social services. The local network of services thus created within a single territory has many objectives such as to promote health and well-being, and offer a cohesive set of services to the public. This enables people to move through the health and social services network and ensure better patient management, particularly of the most vulnerable users [11].

Home care services are a great part of the services managed by the CSSSs. They are provided by health care professionals and are required for acute illness, post-hospitalization and post-operation treatment, long-term health conditions and/or chronic conditions, permanent disability, including physical and mental disability, or terminal illness.

The territorial approach to manage home care services has been used since 1980 in the specific CLSC Côte-des-Neiges site in Montreal (CLSC CDN for short), which catered for 130 000 inhabitants in 2004, among which 5200 are regular home care service users. Given the size of the territory, the management team partitioned the territory into 6 districts [42], with each district being assigned to a multidisciplinary team of professionals. This has allowed for increased efficiency in terms of patient assignment (the geographic location of the patient determines which team will be responsible for the care of that patient), reduced transportation time, and therefore allowed for more time for direct patient care.

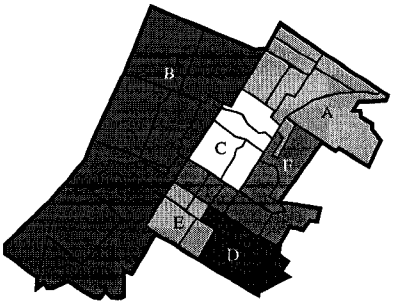

Figure 3.1 – The six districts of the CLSC CDN territory

Since patient assignment to nurses is performed according to a territorial approach (i.e., the assignment is based on the territorial origin of the demand, and not according to the actual workload of the nurses), partitioning of the territory and assignment of nurses to each

district must be carried out carefully so that the nurses can end up with similar workloads. The partitioning in [42] was performed on the basis of historical data on number of patients and number of nursing visits. Since the population changes over time, thereby bringing about changes in the demands for services originating in each district, districting exercises must be performed on a regular basis to counterbalance these fluctuations in demands over time. In addition to workload inequities between nurses that such fluctuations tend to create, it has been observed that the availability of nursing services tends to determine the services actually delivered. This in turn leads to inequities on level of service depending on districts. Since reorganizing districts is time and resource consuming and can cause important changes in patients follow-up (by changing the case holder), a more dynamic method should be considered to assign patients to nurses [144]. By this, we mean that the patient assignment to nurses should be based on the actual workload of every nurse at the time of the demand for services.

In this paper, we propose a model and an algorithm for assigning patients to nurses that take into account not only the geographic location of the patients, but also each nurse's workload. We first give a brief description of the territorial approach used at CLSC CDN, and present measures that help comparing the workload of the nurses. The patient assignment model is then described with the proposed solution method. Experimental results are then reported, and we conclude with final remarks.

3.2 The territorial approach at CLSC CDN

The territory of CLSC CDN is currently divided into six *districts*, each one being constituted by several *basic units* which are the census tracts used by Statistics Canada. Requests for home care services arrive at CLSC CDN from a hospital or a physician's office, directly from the patient, a family member or a friend. The intake nurse identifies the district associated with the patient's address and forwards the request to the manager of the team responsible for that particular district. At the same time the nurse performs an analysis of the nature and the urgency of the request for services and then takes a

decision to assign the patient to one professional. A patient requiring nursing care as a major component of his care plan will be assigned to a nurse. This decision is confirmed by the manager of the multidisciplinary team who receives every new request for home care services. The professional responsible for the patient will, in most cases, involve other professionals in the care of the patient.

A distinction is made between *case manager nurses* (who typically hold a Bachelor's degree in nursing) and *nurse technicians* (who typically hold a community college degree in nursing). The nurse technicians will be assigned the short-term patients or long-term patients needing punctual nursing care. For instance, a patient requiring short-term and specific nursing care such as a wound dressing or a post-hospitalization home based antibiotic-therapy treatment, will typically be assigned to a nurse technician.

Conversely, a patient requiring the organization of a more complex service plan such as organizing the activities of daily living, coordinating visits and ensuring links with doctors and specialists as well as with the pharmacist, consulting with and arranging evaluations by other professionals (occupational therapist, physiotherapist, social worker, dietician, etc.) will be assigned to a case manager nurse. Such patients typically include frail elderly patients with a great loss of autonomy, palliative patients, patients with cancer, patients suffering from degenerative diseases or chronic illnesses and patients with serious mental health problems.

In addition to these nurses who are part of the six multidisciplinary teams and to whom patients get assigned to, there are three to four nurses who make up the "surplus" team. These nurses are not assigned any patient nor a particular district. They are asked to deliver specific nursing care treatments by the professionals responsible for the patient and are not responsible for the global care plan of the patient.

Typically case manager and nurse technicians deliver the nursing visits to the patients they are assigned to. The surplus team will handle nursing visits that the team nurses are unable to absorb, given the number of visits they have already scheduled for themselves. Furthermore, since the working hours of the surplus team nurses are extended until late in the evening (11 p.m.) as well as on week-ends, these nurses are able to absorb visits that

are needed outside regular working hours. Some specific cares (such as wound dressing) are required several times a day, seven days a week. Although one nurse will be assigned that particular patient, she cannot be required to perform all the nursing visits needed. However, the surplus team nurses will be able to absorb some, if not most, of the visits required. Another feature of the surplus team is that it can serve as a "buffer" team to absorb temporary increases in demands for nursing services thereby contributing to the reduction in work overload. It should be noted however that despite the fact that the surplus team can serve as a buffer, it is difficult to absorb increases in demands in all situations due to, on the one hand, the difficulty to predict these peaks in demands and, on the other hand, because the surplus team is often used to compensate for the shortage of nurses during absences of regular nurses. Whether the absence is planned or not (sick day), it is indeed often difficult to find a nurse to replace the one that is absent. For long term replacement of nurses, the home care department usually resort to outside agencies.

Table 3.1 shows the number of basic units in each district, as well as the repartition of the case manager nurses, the nurse technicians, and the patients over the CLSC CDN territory for the year 2002-2003. We observe that the number of case manager nurses in each district varies between 2 and 4 while the number of nurse technicians varies between 1 and 2. Each nurse is associated with a set of basic units in her district and a patient from a basic unit is preferably assigned to a nurse associated with that unit. Each basic unit is assigned to exactly one nurse technician, but can be assigned to more than one case manager nurse. The current division of the CLSC CDN territory results from an analysis performed in 1998-1999, and summarized in [42]. In 2000, it was considered as optimal in terms of the satisfaction of the professionals, the team managers and the head managers.

Table 3.1 – Number of units, nurses and patients per district for the year 2002-2003

District	Number of			
	units	case manager nurses	nurse technicians	patients
A	8	4	1	712
B	7	2	2	477
C	4	4	1	732
D	6	2	1	550
E	4	3	1	508
F	7	4	1	647

3.3 Workload measure

Workload balance is known as being a good measure of the performance of systems such as multiprocessor systems [63], software distributed shared memory systems [203], planning production [180] or health care systems [118].

In [144], we analyzed the impact of the demand fluctuations on the workload of the nurses in each district, and thereby concluded that in order to reduce imbalance and inequities, one should consider the possibility of assigning patients from a basic unit to nurses that are not associated with that unit.

Actually at the CLSC CDN there is no fixed measure to evaluate the workload of the nurses. However, in the event of overload, the manager of every team is usually able to designate which nurses are concerned. In case of imbalance, the manager has the latitude to reassign only new requests (since follow-up of the patients already in the system has to be held) to another nurse even if the address of the patient does not correspond to the nurse's set of basic units. Meanwhile, managers do not necessarily have the same information of the workloads of the nurses, since each one works with his own team. The profile of the patients varies a lot from one district to another for demographic and socio-economic regards. By considering the possibility of assigning patients to nurses from a different district, one may better respond to demand fluctuations without creating too much imbalance among the

nurses, and this may also result in a closer collaboration between the six team managers. Since CLSC CDN overlaps 6 districts, there is a need for uniformization of the case load evaluation to promote clarity and efficiency.

Workload measures used in the UK are analyzed in [118]. The author compares activity based measures (where the workload of each nurse mainly depends on the time spent with the patients) with dependency based measures (where the patients are grouped into categories, and a workload is associated with each category). He concludes that these two types of measures are neither valid nor reliable. We however believe that a mix of these measures can provide a good picture of the nurses' workload at the CLSC CDN. In cooperation with the CLSC committee, we have defined a workload measure which we now define.

The activities of a nurse can be divided into *direct* work and *indirect* work. Direct work includes every task related to patients, as visits and case management, while indirect work encompass tasks related to the nursing job itself as meetings, syndicate and associations activities and trainings. We consider here only the direct work which we aim to balance.

In the previous section we highlighted the difference between case manager nurses and nurse technicians. This difference is related to the educational background as much as to the kind of patients they are usually assigned to. In practice, even if nurse technicians are not assigned long term patients, deterioration of a short term case or overload of a case manager nurse can lead to this situation. The workload of a nurse depends on the type of patients she is assigned to. For budget purposes, the CLSCs are used to classify the patients according to 16 well defined categories [110]. With the help of the CLSC CDN committee, we have simplified this classification to obtain only 5 categories which we use to measure the workload of each nurse. These five categories can be summarized as follows :

- category 1 : short term patients that do not require case management ;
- category 2 : short term patients that need post-hospitalization or post-surgery care ;
- category 3 : long term patients needing punctual nursing care ;
- category 4 : patients with loss of autonomy ;
- category 5 : palliative patients.

In a previous work [144], we have shown that the duration of a visit at CLSC CDN is independent of the patient category and lasts in average 30 minutes, with a very small standard deviation. Care procedures have also been used to evaluate nurses workload [160], but in the context of CLSCs it is considered fastidious by the nurses. Hence, instead of using the duration of a visit or the care procedures to evaluate the workload due to a patient, we prefer to use the *heaviness* of the case which depends on the category of the patient. If the patient is a complex case, we assume that he represents a heavy case for the nurse, while a short term patient is considered less complex usually. To evaluate the heaviness of a case, we have defined a witness visit. The load of the witness visit includes not only the actual nursing care which has to be provided, but also the clerical work and the management associated to the case follow-up. Each visit required by a patient has a weight that indicates how much heavier it is when compared to the witness visit. The CLSC CDN board has fixed these weights to 0.75, 1, 1, 2 and 4 for categories $1, \ldots, 5$, respectively.

The number of patients in each category is also part of the workload. Obviously, a high number of heavy cases is not suitable. It is even more significant for the fifth category of patients which represent palliative cases. These cases usually require very complex nursing cares and heavy case management.

Finally, when a nurse has to travel to a basic unit she is not associated with, inside or outside her district, this creates an additional workload that depends as much on the distance traveled as on the number of visits required by the patients. This consideration is not so important with the actual approach since the nurses typically take care of patients located in the basic units they are associated with, but finds its relevance within the new assignment policy proposed in this paper.

In summary, the workload of every nurse depends on three components :

- the *visit load* which is equal to the weighted sum of the visits that the nurse has to perform, the weight of a visit being defined according to its heaviness when compared to a witness visit ;

- the *case load* which depends on the number of patients assigned to the nurse in each category ;
- the *travel load* which depends on the distance that the nurse has to travel outside her basic units to visit patients, and the number of visits required by these patients.

To balance the workload of the nurses, we will proceed as follows. We first solve the patient assignment problem with no regards to patients follow-up. This may occur, for example, when a new partition of the territory into districts is defined, or when the total imbalance is so high that a complete reassignment is needed. We then perform a dynamic assignment where new patients have to be assigned to nurses knowing their respective workload.

Note that nurses in most CLSCs practice "self-scheduling" and the boards are not willing to change this work organization. Thus we do not concentrate on the rostering problem which has been well studied in the literature (see for example [47]).

3.4 Patient assignment model

We start this section by giving some basic notations which will be used in the proposed mathematical model.

- I is the set of nurses working at the CLSC CDN ;
- P is the set of patients ;
- $J = \{1, \ldots, 5\}$ is the set of patient categories (see previous section) ;
- $K = \{m, t\}$ is the set of nurses types, where m stands for case manager nurses and t for nurse technicians ;
- I_k is the set of nurses of type $k \in K$;
- P_{jk} is the set of patients of category $j \in J$ that can be assigned to nurses of type $k \in K$, and $\mathcal{P}_k = \bigcup_{j \in J} P_{jk}$;
- k_i is the type of nurse i ;
- U_i is the set of basic units to which nurse i is associated ;

- v_p is the number of visits required by patient p. This value is obtained by considering a period of one month and by multiplying the number of visits needed per week by the number of weeks in the caring plan of the considered month ;
- j_p is the category of patient p ;
- u_p is the basic unit where patient p is located ;
- h_{jk} is the heaviness of a visit to a patient of category j if assigned to a nurse of type k.

To determine the distance between two basic units, we define a graph G in which each vertex is associated with a basic unit, and two vertices are linked by an edge if the corresponding basic units share a common frontier. The graph associated with the CLSC CDN territory is represented in Figure 3.2. The length of an edge is equal to 1 if it connects two basic units of the same district, and λ otherwise, where λ helps penalizing the move of a nurse from a district to another. The distance ℓ_{ip} that nurse i has to travel to take care of patient p is defined as the length of the shortest chain in G linking the vertex associated with u_p to a vertex associated with a basic unit in U_i. The travel load t_{ip} of patient p for nurse i is defined as $v_p \cdot e^{\ell_{ip}}$. It is proportional to v_p to take into account the number of times i will have to move to the basic unit u_p to take care of p. The exponential term is to discourage too long travels.

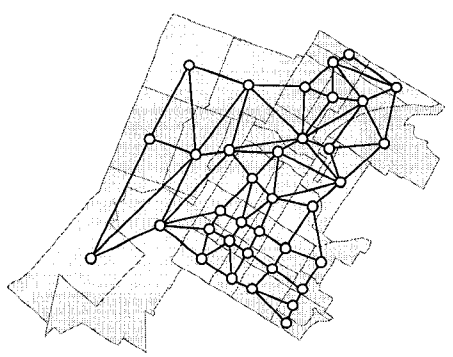

Figure 3.2 – Graph G associated with the CLSC CDN territory

In the following mathematical formulation, we denote x_{ip} the boolean variable that equals 1 if patient p is assigned to nurse i, and 0 otherwise. The initial idea of the CLSC

CDN board was to determine an assignment $\boldsymbol{x}$ with balanced visit loads, case loads, and travel loads. More precisely, the average visit load $\overline{V}_k(\boldsymbol{x})$ of the nurses of type k is equal to

$$\overline{V}_k(\boldsymbol{x}) = \frac{\sum_{p \in \mathcal{P}_k} \sum_{i \in I_k} v_p \cdot h_{j_p k} \cdot x_{ip}}{|I_k|}$$

and the visit loads are considered as balanced if

$$\sum_{p \in \mathcal{P}_{k_i}} v_p \cdot h_{j_p k_i} \cdot x_{ip} = \overline{V}_{k_i}(\boldsymbol{x}) \qquad \forall i \in I.$$

Similarly, the average number $\overline{N}_{jk}(\boldsymbol{x})$ of patients of category j assigned to nurses of type k is equal to

$$\overline{N}_{jk}(\boldsymbol{x}) = \frac{\sum_{p \in P_{jk}} \sum_{i \in I_k} x_{ip}}{|I_k|}$$

and since the number of patients that every nurse has in every category is an integer, the case loads are considered as balanced if

$$\sum_{p \in P_{jk_i}} x_{ip} \leq \lceil \overline{N}_{jk_i}(\boldsymbol{x}) \rceil \qquad \forall i \in I,\ \forall j \in J.$$

Finally, the average travel load $\overline{T}_k(\boldsymbol{x})$ of the nurses of type k is equal to

$$\overline{T}_k(\boldsymbol{x}) = \frac{\sum_{p \in \mathcal{P}_k} \sum_{i \in I_k} t_{ip} \cdot x_{ip}}{|I_k|}$$

and the visit loads are considered as balanced if

$$\sum_{p \in \mathcal{P}_{k_i}} t_{ip} \cdot x_{ip} = \overline{T}_{k_i}(\boldsymbol{x}) \qquad \forall i \in I.$$

Since a perfect balance of the case load, visit load and travel load is typically out of reach, the first objective fixed in collaboration with the board of the CLSC CDN was to determine a patient assignment to the nurses respecting all the following constraints, where ϵ_v, ϵ_c and ϵ_t in constraints (3.1)-(3.3) are fixed acceptable deviations from the ideal average visit load, case load and travel load, respectively, and constraints (3.4) and (3.5) impose that each patient p is assigned to exactly one nurse of type k with $p \in P_{j_p k}$.

$$\sum_{p \in \mathcal{P}_{k_i}} v_p \cdot h_{j_p k_i} \cdot x_{ip} \leq \overline{V}_{k_i}(\boldsymbol{x}) + \epsilon_v \qquad \forall i \in I \tag{3.1}$$

$$\sum_{p \in P_{jk_i}} x_{ip} \leq \overline{N}_{jk_i}(\boldsymbol{x}) + \epsilon_c \qquad \forall i \in I, \forall j \in J \tag{3.2}$$

$$\sum_{p \in \mathcal{P}_{k_i}} t_{ip} \cdot x_{ip} \leq \overline{T}_{k_i}(\boldsymbol{x}) + \epsilon_t \qquad \forall i \in I \tag{3.3}$$

$$\sum_{k \in K | p \in P_{j_p k}} \sum_{i \in I_k} x_{ip} = 1 \qquad \forall p \in P \tag{3.4}$$

$$x_{ip} \in \{0, 1\} \qquad \forall i \in I, \forall p \in P \tag{3.5}$$

If every patient $p \in P$ can only be assigned to one type of nurse (i.e., there is a unique $k \in K$ with $p \in P_{j_p k}$), then the average visit load $\overline{V}_k(\boldsymbol{x})$ and case load $\overline{N}_{jk}(\boldsymbol{x})$ do not depend on the assignment $\boldsymbol{x}$, and constraints (3.1), (3.2), (3.4) and (3.5) are then typical of a multi-resource generalized assignment problem (MRGAP for short), where jobs (patients) have to be assigned to agents (nurses) subject to resource constraints [98]. Finding a feasible solution to an MRGAP is an NP-hard problem since, as shown in [155], it is already NP-hard with one resource constraint. Heuristic solution methods have been developed to solve the MRGAP with a linear objective function (e.g., [98, 158, 201]).

Choosing the most appropriate values for ϵ_v, ϵ_c and ϵ_t is a difficult task. Too high values create workload imbalance while too small values make the problem infeasible. An alternate strategy consists in defining a measure for the imbalance and to minimize this measure rather than imposing an almost perfect balanced workload. A discussion with the CLSC CDN board has resulted in the definition of the three following functions which respectively measure the case load, visit load and travel load imbalance.

The visit load of nurse i is defined as

$$V_i(\boldsymbol{x}) = \sum_{p \in \mathcal{P}_{k_i}} v_p \cdot h_{j_p k_i} \cdot x_{ip}$$

while the visit load imbalance of nurse i is defined as

$$C_{i1}(\boldsymbol{x}) = \max\{0, V_i(\boldsymbol{x}) - \overline{V}_{k_i}(\boldsymbol{x})\}. \tag{3.6}$$

The first objective is to minimize

$$f_1(\boldsymbol{x}) = \sum_{i \in I} (C_{i1}(\boldsymbol{x}))^2 .$$

The second objective takes into account the difference between the number of patients of category j assigned to nurse i and the ideal number $\lceil \overline{N}_{jk_i}(\boldsymbol{x}) \rceil$. For every patient in excess, a penalty is added which is equal to the average number $\overline{v}_{ij}(\boldsymbol{x})$ of visits performed by nurse i to patients of category j, multiplied by the weight h_{jk_i} of such patients. More precisely, we have

$$\overline{v}_{ij}(\boldsymbol{x}) \cdot \sum_{p \in P_{jk_i}} x_{ip} = \sum_{p \in P_{jk_i}} v_p \cdot x_{ip}$$

and

$$C_{i2}(\boldsymbol{x}) = \sum_{j \in J} \max \left\{ 0, \sum_{p \in P_{jk_i}} x_{ip} - \lceil \overline{N}_{jk_i}(\boldsymbol{x}) \rceil \right\} \cdot \overline{v}_{ij}(\boldsymbol{x}) \cdot h_{jk_i}. \tag{3.7}$$

The second objective is to minimize

$$f_2(\boldsymbol{x}) = \sum_{i \in I} (C_{i2}(\boldsymbol{x}))^2 .$$

The travel load of nurse i is defined as

$$T_i(\boldsymbol{x}) = \sum_{p \in \mathcal{P}_{k_i}} t_{ip} \cdot x_{ip}$$

and the travel load imbalance of nurse i is defined as

$$C_{i3}(\boldsymbol{x}) = \max\{0, T_i(\boldsymbol{x}) - \overline{T}_{k_i}(\boldsymbol{x})\}. \tag{3.8}$$

The third objective is to minimize

$$f_3(\boldsymbol{x}) = \alpha \cdot \sum_{i \in I} (C_{i3}(\boldsymbol{x}))^2 + \sum_{k \in K} \left(\overline{T}_k(\boldsymbol{x})\right)^2$$

where the first component of $f_3(\boldsymbol{x})$ aims to reduce imbalance in the travel loads, while the second term minimizes the total distance traveled. Parameter α helps giving more or less importance to one of the two components of $f_3(\boldsymbol{x})$.

The patient assignment problem we solve has for objective to minimize the weighted sum $f(\boldsymbol{x}) = \omega_1 \cdot f_1(\boldsymbol{x}) + \omega_2 \cdot f_2(\boldsymbol{x}) + \omega_3 \cdot f_3(\boldsymbol{x})$, where ω_i $(i = 1, 2, 3)$ are parameters that

give more or less importance to each component of $f(\boldsymbol{x})$.

The mathematical formulation of the patient assignment problem we solve can now be summarized as follows.

$$\text{Minimize} \quad \sum_{i\in I}\left(\omega_1\cdot(C_{i1})^2+\omega_2\cdot(C_{i2})^2+\omega_3\cdot\alpha\cdot(C_{i3})^2\right)+\omega_3\cdot\sum_{k\in K}\left(\overline{T}_k\right)^2$$

subject to

$$\overline{V}_k\cdot|I_k| = \sum_{p\in\mathcal{P}_k}\sum_{i\in I_k} v_p\cdot h_{j_pk}\cdot x_{ip} \quad \forall k\in K \tag{3.9}$$

$$\sum_{p\in\mathcal{P}_{k_i}} v_p\cdot h_{j_pk_i}\cdot x_{ip}-\overline{V}_{k_i} \leq C_{i1} \quad \forall i\in I \tag{3.10}$$

$$\overline{N}_{jk}\cdot|I_k| = \sum_{p\in P_{jk}}\sum_{i\in I_k} x_{ip} \quad \forall j\in J,\forall k\in K \tag{3.11}$$

$$\sum_{p\in P_{jk_i}} x_{ip}-\overline{N}_{jk_i} \leq s_{ij} \quad \forall i\in I,\forall j\in J \tag{3.12}$$

$$\overline{v}_{ij}\cdot\sum_{p\in P_{jk_i}} x_{ip} = \sum_{p\in P_{jk_i}} v_p\cdot x_{ip} \quad \forall i\in I,\forall j\in J \tag{3.13}$$

$$\sum_{j\in J} s_{ij}\cdot\overline{v}_{ij}\cdot h_{jk_i} \leq C_{i2} \quad \forall i\in I \tag{3.14}$$

$$\overline{T}_k\cdot|I_k| = \sum_{p\in\mathcal{P}_k}\sum_{i\in I_k} t_{ip}\cdot x_{ip} \quad \forall k\in K \tag{3.15}$$

$$\sum_{p\in\mathcal{P}_{k_i}} t_{ip}\cdot x_{ip}-\overline{T}_{k_i} \leq C_{i3} \quad \forall i\in I \tag{3.16}$$

$$\sum_{k\in K|p\in P_{j_pk}}\sum_{i\in I_k} x_{ip} = 1 \quad \forall p\in P \tag{3.17}$$

$$x_{ip} \in \{0,1\} \quad \forall i\in I,\forall p\in P \tag{3.18}$$

$$C_{i1},C_{i2},C_{i3} \geq 0 \quad \forall i\in I \tag{3.19}$$

$$s_{ij} \geq 0 \text{ and integer} \quad \forall i\in I,\forall j\in J \tag{3.20}$$

Constraints (3.9) and (3.10) coupled with the positivity constraint (3.19) on C_{i1} imply that $C_{i1}(\boldsymbol{x}) = \max\{0, V_i(\boldsymbol{x})-\overline{V}_{k_i}(\boldsymbol{x})\}$. Similarly, constraints (3.11), (3.12) and the integrality and positivity constraint (3.20) on s_{ij} imply that $s_{ij} = \max\{0, \sum_{p\in P_{jk_i}} x_{ip} - \lceil \overline{N}_{jk_i}(\boldsymbol{x})\rceil\}$, which combined with constraints (3.13), (3.14) and (3.19) give definition (3.7) for C_{i2}. Constraints (3.15), (3.16) and (3.19) define C_{i3} according to (3.8), and constraints

(3.17) and (3.18) impose that each patient p is assigned to exactly one nurse of type k with $p \in P_{j_p k}$.

In the next section, we propose a Tabu Search algorithm for solving this problem. Note that all constraints are linear, except constraints (3.13) and (3.14), which means that if we remove the constraints on the case load, the considered patient assignment problem has linear constraints and a quadratic objective.

3.5 A Tabu Search algorithm

Given a solution space X and a function f that measures the value $f(\boldsymbol{x})$ of every solution $\boldsymbol{x} \in S$, Tabu Search is an algorithm which objective is to determine a solution $\boldsymbol{x}^*$ with minimum value $f(\boldsymbol{x}^*)$ over X. For this purpose, a neighborhood $N(\boldsymbol{x})$ is defined for every $\boldsymbol{x} \in X$. It corresponds to the set of *neighbor solutions* that can be obtained from s by performing a *local move*. Tabu Search generates a sequence $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_r$ of solutions such that $\boldsymbol{x}_0$ is an initial solution and $\boldsymbol{x}_i \in N(\boldsymbol{x}_{i-1})$ for $i = 1, \ldots, r$. In order to avoid cycling, a tabu list is created that contains forbidden local moves. Hence, a move m from $\boldsymbol{x}_i$ to $\boldsymbol{x}_{i+1}$ can only be performed if m does not belong to the tabu list, unless $f(\boldsymbol{x}_{i+1}) < f(\boldsymbol{x}^*)$, where $\boldsymbol{x}^*$ is the best solution encountered so far. For more details on Tabu Search, the reader may refer to [102].

For our problem, we define X as the set of assignments such that each patient p is assigned to a nurse i with $p \in P_{j_p k_i}$. Each solution $\boldsymbol{x} \in X$ is measured using the objective function $f(\boldsymbol{x})$ defined in the previous section. The initial solution is generated randomly by assigning a nurse of the right type to each patient. When moving from a solution $\boldsymbol{x}$ to a neighbor one $\boldsymbol{x}'$, we will change the assignment of several patients. If a patient p is transferred from nurse i to nurse i', then we put the pair (i, p) in the tabu list T, with the meaning that it is forbidden for θ iterations to reassign p to i.

Two patients in P are considered as equivalent if they are of the same category and require the same number of visits. By analyzing the data set of the CLSC CDN, we have observed that it contains many equivalent patients. Hence, when transferring a patient p

from a nurse i to a nurse i', there is a danger that the next move will consist in moving a patient equivalent to p from i' to i, which would create cycles in the algorithm. To avoid such a situation, when moving a patient p from i to i', we also introduce all pairs (i, p') into T, where p' is any patient equivalent to p and assigned to i'.

As mentioned above, when moving from a solution $\boldsymbol{x}$ to a neighbor one $\boldsymbol{x}'$, we possibly change the assignment of several patients. If all these changes belong to the tabu list T, then the move from $\boldsymbol{x}$ to $\boldsymbol{x}'$ is designated *tabu*, while there is no restriction if at least one of the changes is not in T.

For moving from a solution $\boldsymbol{x}$ to a solution $\boldsymbol{x}'$, we first choose a patient p. This is done according to one of the eight following rules, where i denotes the nurse currently assigned to p in $\boldsymbol{x}$, and j is the category of p (i.e., $j = j_p$) :

(a) p is any patient in P;

(b) p is any patient such that $V_i(\boldsymbol{x}) > \overline{V}_{k_i}(\boldsymbol{x})$;

(c) p is any patient such that more than $\lceil \overline{N}_{jk_i}(\boldsymbol{x}) \rceil$ patients of category j are currently assigned to i;

(d) p is any patient such that $\ell_{ip} > 0$;

(e) p is any patient such that $\ell_{ip} > 1$;

(f) p is any patient such that $T_i(\boldsymbol{x}) > \overline{T}_{k_i}(\boldsymbol{x})$ and $\ell_{ip} > 0$;

(g) p is any patient chosen according to (b), (c) or (d);

(h) p is any patient such that $v_p > \frac{\sum_{p' \in P_{jk_i}} v_{p'}}{|P_{jk_i}|}$.

Rule (a) ensures that every patient gets a chance to be moved. Rule (b) helps reducing $f_1(\boldsymbol{x})$, while rule (c) does the same job for $f_2(\boldsymbol{x})$, and rules (d), (e), and (f) for $f_3(\boldsymbol{x})$. Rule (g) is for trying to reduce at least one of the three components $f_i(\boldsymbol{x})$. The transfer of patient p from nurse i to nurse i' may create an important overload for i' when v_p is large. Hence such patients are eventually never moved. However, since the initial solution is randomly generated, it may be necessary to change the assignment of such patients, and rule (h) helps doing it.

Once p is chosen, a move from $\boldsymbol{x}$ to a neighbor $\boldsymbol{x}'$ is performed according to one of the five following procedures (where, as above, i is the nurse assigned to p in $\boldsymbol{x}$)) :

(1) A *flip* consists in choosing a nurse $i' \neq i$ with $p \in \mathcal{P}_{k_{i'}}$, and assigning p to i' instead of i;

(2) A *2-swap* consists in choosing a nurse $i' \neq i$ with $p \in \mathcal{P}_{k_{i'}}$ and a patient p' assigned to i' with $p' \in \mathcal{P}_{k_i}$, and then exchanging patients p and p' between nurses i and i';

(3) A *3-swap* consists in choosing two nurses i' and i'' distinct from i and two patients p' and p'' assigned to i' and i'', respectively, such that $p \in \mathcal{P}_{k_{i'}}$, $p' \in \mathcal{P}_{k_{i''}}$ and $p'' \in \mathcal{P}_{k_i}$, and then assigning p to i', p' to i'', and p'' to i;

(4) A *2-mswap* consists in choosing a nurse $i' \neq i$ with $p \in \mathcal{P}_{k_{i'}}$ and a set P' of patients assigned to i' with $P' \subseteq \mathcal{P}_{k_i}$, and then assigning p to i' and all patients in P' to i;

(5) A *3-mswap* consists in choosing two nurses i' and i'' distinct from i, and two sets P' and P'' of patients assigned to i' and i'', respectively, such that $p \in \mathcal{P}_{k_{i'}}$, $P' \subseteq \mathcal{P}_{k_{i''}}$ and $P'' \subseteq \mathcal{P}_{k_i}$, and then assigning p to i', all patients in P' to i'', and all patients in P'' to i.

Procedures (1) and (2) are standard moves which are typically used in assignment problems [93]. The three other procedures are inspired by ejection chain techniques [143] where after a flip of patients p from i to i', subsequent moves directly depend on the first one. Moves of type (4) and (5) are especially important in our context where patients may have very different required number of visits. For example, by analyzing the data set of the CLSC CDN, we have noticed that the number of required visits in a given patient category ranges from 1 to 57. If we want to move a patient p with $v_p = 57$ from i to i', without creating a big visit load for i', it may be necessary to remove more than one patient from i'.

We now explain how sets P' and P'' are determined in moves of types (4) and (5). For two nurses i' and i'', let $A_{i'i''}(\boldsymbol{x})$ denote the set of patients currently assigned to i' which can be assigned to i''. More precisely, define $A_{i'i''}(\boldsymbol{x})$ as the set of patients p' with $x_{i'p'} = 1$ and $p' \in \mathcal{P}_{k_{i''}}$. Let $m_{i' \to i''} = \min_{p' \in A_{i'i''}(\boldsymbol{x})} \{t_{i'p'} - t_{i''p'}\} - 1$. Given a patient p assigned to a nurse i and given any two nurses i' and i'' of type k_i with $i' \neq i$, we consider a knapsack

problem, denoted $P_{i'\to i''}(p)$, that determines a set of patients in $A_{i'i''}(\boldsymbol{x})$ to be moved from nurse i' to nurse i'' :

$$P_{i'\to i''}(p)\begin{cases} \text{Maximize} \quad \sum_{p'\in A_{i'i''}(\boldsymbol{x})}(t_{i'p'}-t_{i''p'}-m_{i'\to i''})\cdot y_{p'} \\ \text{subject to} \\ \displaystyle\sum_{p'\in A_{i'i''}(\boldsymbol{x})} v_{p'}\cdot h_{j_{p'}k_i}\cdot y_{p'} \leq v_p\cdot h_{j_pk_i} \\ \qquad\qquad y_{p'} \in \{0,1\} \qquad \forall p'\in A_{i'i''}(\boldsymbol{x}) \end{cases}$$

The objective of $P_{i'\to i''}(p)$ is to gain as much as possible in the travel loads of nurses i' and i'' when moving patients from i' to i''. The term $m_{i'\to i''}$ is to ensure that each patient in $A_{i'i''}(\boldsymbol{x})$ gets a chance to be moved. The constraint of the knapsack problem ensures that the total visit load of the patients moved from i' to i'' is not larger than the visit load of patient p for nurse i. Let $Q_{i'\to i''}(p)$ be the subset of patients $p'\in A_{i'i''}(\boldsymbol{x})$ such that $y_{p'}=1$ in the optimal solution of the above problem. For moves of type (4), we define $P'=Q_{i'\to i}(p)$ and for moves of type (5), we define $P'=Q_{i'\to i''}(p)$ and $P''=Q_{i''\to i}(p)$.

We solve the knapsack problems using the implementation of [40] of Martello and Toth's algorithm [155].

In what follows, we denote $N_{q,r}(\boldsymbol{x})$ the set of solutions that can be obtained from $\boldsymbol{x}$ by choosing patient p according to rule (q) with $q\in\{a,\ldots,h\}$ and then applying procedure (r) with $r\in\{1,\ldots,5\}$. Also, we denote $N_{q,6}(s)=\bigcup_{r=1}^{5}N_{q,r}(\boldsymbol{x})$. Every neighborhood $N_{q,r}(\boldsymbol{x})$ is explored until M_I iterations (where M_I is a parameter) have been performed without improvement of $\boldsymbol{x}^*$.

As *intensification* strategy, we check, at each iteration, whether $N_{a,1}(\boldsymbol{x})\cup N_{a,2}(\boldsymbol{x})$ contains a solution which is better than $\boldsymbol{x}^*$, in which case we determine such a solution and update $\boldsymbol{x}^*$. When all neighborhoods have been tested, we use a *diversification* strategy which consists in performing M_D moves (where M_D is a parameter) using neighborhoods $N_{d,4}(\boldsymbol{x})$ and $N_{d,5}(\boldsymbol{x})$, but using a different objective function in the knapsack problem. More precisely, we maximize $\sum_{p'\in A_{i'i''}(\boldsymbol{x})}(t_{i'p'}-t_{i''p'})\cdot y_{p'}$ (i.e., the term $-m_{i'\to i''}$ is removed), the consequence being that the patients which are closer to i' than to i'' will not be moved

since the increase of the travel load for i'' would be larger than the decrease of the travel load for i'. The output of the knapsack problem is then typically a set of patients with total visit load much smaller than the visit load $v_p \cdot h_{j_p k_i}$ of p for i. This means that the visit load of nurse i' will probably increase with such a move in $N_{d,4}(\boldsymbol{x})$ and $N_{d,5}(\boldsymbol{x})$, while the total travel load of the nurses involved in the move will eventually decrease. In summary, the proposed diversification put the emphasis on the decrease of the travel load, even if this induces a large increase in the visit load of some nurses. The process of testing all neighborhoods $N_{q,r}(\boldsymbol{x})$ followed by a diversification is called a *loop*. We apply M_L such loops (where M_L is our last parameter) before stopping the algorithm, each new loop starting from the solution produced by the diversification strategy. The proposed algorithm is summarized in Figure 3.3.

The parameters of our Tabu Search have been fixed on the basis of some preliminary experiments. The following choices have been implemented. For a solution $\boldsymbol{x}$, let $\mu_q(\boldsymbol{x})$ denote the number of patients p which can be chosen according to rule (q) in $\boldsymbol{x}$. There are at most $(|I|-1)$ nurses to which patient p can be reassigned. Hence, $|N_{q,1}(\boldsymbol{x})| \leq \mu_q(\boldsymbol{x}) \cdot (|I|-1)$. When moving from $\boldsymbol{x}$ to $\boldsymbol{x}' \in N_{q,r}(\boldsymbol{x})$, all pairs (i,p) introduced in the tabu list remain in the list for $2 \cdot \sqrt{\mu_q(\boldsymbol{x}) \cdot (|I|-1)}$ iterations. We have chosen the same value for parameter M_I, which means that every neighborhood $N_{q,r}(\boldsymbol{x})$ is used until $2 \cdot \sqrt{\mu_q(\boldsymbol{x}) \cdot (|I|-1)}$ iterations have been performed without improvement of $\boldsymbol{x}^*$. The diversification strategy is used for $M_D = \sqrt{\mu_q(\boldsymbol{x}) \cdot (|I|-1)}$, and the number M_L of loops is set equal to 10. Parameter α in $f_3(\boldsymbol{x})$ helps giving more or less importance to balanced travel loads in comparison with the total travel load of the nurses. Since the CLSC CDN aims to avoid too much travelling, we set $\alpha = \frac{1}{|I|}$ which means that journeys have to be minimized in priority before being balanced between nurses.

```
Generate an initial solution x ∈ X at random and set x ← x* and T ← ∅;
for loop = 1 to M_L do
  for q = a to g do
    for r = 1 to 6 do
      counter ← 0
      while counter < M_I do
        while N_{a,1}(x) ∪ N_{a,2}(x) contains a solution x' such that f(x') < f(x*) do
          Select such a solution x';
          Set x ← x', x* ← x and counter ← 0;
        end while
        Determine the solution x' ∈ N_{q,r}(x) with minimum value f(x') such that the move from x to
        x' is not tabu or f(x') < f(x*);
        Set x ← x' and update T;
        if f(x') < f(x*) then set x* ← x and counter ← 0;
        else set counter ← counter + 1;
      end while
      Set x ← x*;
    end for
  end for
  for diversification =1 to M_D do
    Determine the best solution x' in N_{d,4}(x) ∪ N_{d,5}(x), using the modified objective function for the
    knapsack problem, such that the move from x to x' is not tabu or f(x') < f(x*);
    Set x ← x';
  end for
end for
```

Figure 3.3 – Tabu Search for the Patient Assignment Problem

3.6 Experimental results

Since there is no algorithm to which our tabu search can be compared when all constraints (3.9)-(3.20) are to be taken into account, we have decided to perform our first experiments on instances without constraints on the case load. As already observed, the resulting patient assignment problem has linear constraints and a quadratic objective and can therefore

be solved using CPLEX. In order to evaluate the efficiency of the proposed Tabu Search algorithm, we therefore compare the solutions it provides to those obtained using CPLEX version 9.1.3 on the proposed mixed integer program, where ω_2 is set equal to 0 and constraints (3.11)-(3.14) are removed.

We solve three different problems. The first one, denoted $ABCDEF$, consists in solving the patient assignment problem for the whole territory. We have considered real data from June 2002. The problem contains 19 case manager nurses, 7 nurse technicians, 1413 patients and 36 basic units. Since the CLSC CDN board is not convinced that the six team managers in the districts will easily accept to collaborate, we have also considered a problem, denoted $AB, CDEF$, in which we solve two patient assignment problems, one for districts A and B, and the other one for the four other districts. We can then merge the two assignments and compare them with the solution obtained by solving $ABCDEF$. Such a solution requires a collaboration between the two team managers in A and B, and another collaboration between the four other team managers. The patient assignment problem for districts A and B contains 9 nurses (6 case managers and 3 technicians), 440 patients and 15 basic units, while there are 17 nurses (13 case managers and 4 technicians), 992 patients and 21 basic units in districts C, D, E and F. For comparison, we also solve the patient assignment problem in each district separately, and then merge the assignments. The solution thus obtained corresponds to the current situation at CLSC CDN. This last problem will be denoted A, B, C, D, E, F.

Table 3.2 reports the results obtained with CPLEX and Tabu Search. Instead of reporting the values of each component $f_i(\boldsymbol{x})$ of the objective function which do not clearly indicate the various overloads, we report the average visit overload of the nurses, the average number of visits performed by the nurses in basic units that do not belong to the set of basic units where there are located, and the average number of patients in each category that the nurses have above the ideal average. More precisely, Table 3.2 can be read as follows :

– The first column indicates the values of the weights ω_i.

- The second column indicates the problem solved. Since the visit, case and travel loads are typically very different when comparing case manager with nurse technicians, we have decided to split the results into two parts, one for the case manager nurses (we add a m at the end of the instance name), and the other one for the nurse technicians (we add a t at the end of the instance name). So, for example, the line with label $AB, CDEF - t$ means that we report results for the nurse technicians after having solved $AB, CDEF$.
- For each instance, except A, B, C, D, E, F, we give two lines of results. The first line was obtained using $\lambda = 3$, which means that travels outside the district are discouraged, while the second line considers $\lambda = 1$ and therefore makes no difference between adjacent basic units of the same district or of neighbor districts. This parameter is not relevant for instance A, B, C, D, E, F since the nurses are not allowed to move to another district, and this explains why we have only one line for this instance.
- The next nine columns contain the results obtained using CPLEX.
  - Column labeled O_1 indicates the average visit overload for the considered type k of nurses. Hence,
  $$O_1 = \frac{\sum_{i \in I_k} \max\{0, V_i(\boldsymbol{x}) - \overline{V}_k(\boldsymbol{x})\}}{|I_k|}$$
  - The columns labeled O_{2j} indicate the average number of patients of category j that the nurses have above the ideal average, multiplied by the average number of visits that these nurses have to perform to patients of category j. More precisely, for a type k of nurses, we have
  $$O_{2j} = \frac{\sum_{i \in I_k} \max\{0, \sum_{p \in P_{jk_i}} x_{ip} - \lceil \overline{N}_{jk}(\boldsymbol{x}) \rceil\} \cdot \overline{v}_{ij}(\boldsymbol{x})}{|I_k|}$$
  Notice that CPLEX does not optimize these values since $\omega_2 = 0$, but we report them for comparison with solutions obtained using Tabu Search with $\omega_2 > 0$ (see Tables 3.3 and 3.4).
  - Column labeled O_{31} reports the average number of visits that the nurses have to do in basic units at distance $\ell_{ip} = 1$ from where they are located. More precisely, for a nurse i, let L_i denote the set of patients p with $x_{ip} \cdot \ell_{ip} = 1$. For a type k of

nurses, we have

$$O_{31} = \frac{\sum_{i \in I_k} \sum_{p \in L_i} v_p}{|I_k|}$$

Notice that when $\lambda = 3$, O_{31} does not take into account travels to adjacent basic units in different districts. Columns O_{32} and O_{3+} give the same information but for travels to basic units at distance 2 or more.

- The last four columns indicate the values obtained with Tabu Search. The Δ in front of each parameter means that we report differences with the CPLEX solutions. So, for example, the value 0.05 in column ΔO_{31} for the instance $ABCDEF - m$ with $\omega_1 = \omega_3 = 1$ and $\lambda = 3$ means that Tabu Search has produced a solution where the nurses make in average $2.53 + 0.05 = 2.58$ visits in basic units at distance 1 from the basic units where they are located.

We find it important to mention that a positive value in a column for Tabu Search does not necessarily mean that Tabu Search was not able to find the optimal solution. For example, for instance $A, B, C, D, E, F - m$ with $\omega_1 = 1$ and $\omega_3 = 100$, the solutions found by CPLEX and Tabu Search are equal except for district A where the four nurses have a visit load of 192, 260, 172 and 181 in the CPLEX solution, while these values are 192, 260, 160 and 193 for the Tabu Search solution. Since the average visit load in district A is 201.25, the unique nurse with an excess in the visit load is the second one and both solutions have therefore the same value for $f_1(\boldsymbol{x})$. However, by merging the solutions obtained from the 6 districts, the average visit load of the nurses becomes equal to 186.7, which means that the fourth nurse of district A has no overload in the CPLEX solution, while its overload is equal to $193 - 186.7$ in the Tabu Search solution. This induces an increase of $\frac{193-186.7}{19} = 0.33$ of the average visit overload of the nurses, and this is the reported value in Table 3.2.

Notice also that if three nurses have a visit load of $x - 4$, $x + 2$ and $x + 2$, the average visit overload O_1 is equal to $\frac{4}{3}$ while $f_1(s) = 8$. If the same three nurses have a visit load of $x - 1$, $x - 2$ and $x + 3$, the average visit overload O_1 is equal to $1 < \frac{4}{3}$ while $f_1(s) = 9 > 8$. Hence, a positive value under column ΔO_1 does not mean that the f_1 component of the objective function is larger for Tabu Search. On the opposite, a negative value does not mean that CPLEX has not found the optimal solution.

Table 3.2 – Comparison between CPLEX and Tabu Search.

ω_1-ω_2-ω_3	instance	CPLEX									TABOU			
		O_1	O_{21}	O_{22}	O_{23}	O_{24}	O_{25}	O_{31}	O_{32}	O_{3+}	ΔO_1	ΔO_{31}	ΔO_{32}	ΔO_{3+}
	ABCDEF-m	15.49	1.58	0	0	6.25	2.90	2.53	0	0.32	0	0.05	0	0
		5.46	1.58	0	0	7.16	2.52	8.00	0.05	0	0.84	-1.26	0.27	0
	AB,CDEF-m	16.75	1.82	0	0	4.23	2.32	2.21	0	0	0	0	0	0
		10.25	1.58	0	0	6.06	2.31	6.11	0.53	0	1.58	-0.58	-0.06	0
1-0-1	A,B,C,D,E,F-m	16.75	1.58	0	0	5.75	2.26	2.95	0	0	0	0	0	0
	ABCDEF-t	40.21	5.37	11.35	30.68	0	0	0	0	0	0	0	0	0
		28.08	7.14	10.85	25.93	0	0	11.29	0	0	0	1.00	0	0
	AB,CDEF-t	40.21	5.37	11.35	30.68	0	0	0	0	0	0	0	0	0
		37.17	5.95	11.35	30.68	0	0	1.71	0	0	0.14	2.00	0.29	0
	A,B,C,D,E,F-t	40.21	5.03	11.35	30.68	0	0	0.14	0	0	0	1.29	0	0
	ABCDEF-m	4.23	1.58	0	0	6.32	3.30	2.89	0	3.26	0.21	7.06	0.37	-0.10
		0.22	1.50	0	0	7.56	2.66	12.16	0.37	0	-0.04	1.63	1.10	0
	AB,CDEF-m	6.73	1.82	0	0	8.33	3.40	3.42	0	2.68	0.06	2.69	0	-0.05
		3.81	1.82	0	0	6.81	2.88	13.11	1.84	0	-0.15	3.36	0.32	0
100-0-1	A,B,C,D,E,F-m	16.75	1.82	0	0	6.40	2.58	4.16	0	0	0	1.68	0	0
	ABCDEF-t	33.24	7.19	11.20	30.51	0	0	0	0	5.29	0.04	0	0	0
		2.33	7.63	8.00	9.81	0	0	37.14	0	0	-0.01	0.72	0	0
	AB,CDEF-t	36.88	6.20	11.35	30.68	0	0	0	0	0	0.04	0	0	3.00
		27.54	7.92	8.36	20.99	0	0	39.71	2.00	0	0	2.29	0	0
	A,B,C,D,E,F-t	40.21	4.66	11.35	30.68	0	0	0.86	0	0	0	-0.72	0	0
	ABCDEF-m	17.39	1.58	0	0	4.54	2.25	0.68	0	0	0	0	0	0
		16.76	1.82	0	0	4.46	1.77	0.89	0	0	0	0	0	0
	AB,CDEF-m	17.60	1.82	0	0	5.25	2.25	0.37	0	0	0.05	0	0	0
		17.52	1.58	0	0	4.20	1.69	0.53	0	0	0	0	0	0
1-0-100	A,B,C,D,E,F-m	17.94	1.82	0	0	5.81	2.69	0.05	0	0	0.33	0	0	0
	ABCDEF-t	40.21	5.37	11.35	30.68	0	0	0	0	0	0	0	0	0
		39.92	5.69	11.35	30.68	0	0	0.29	0	0	0	0	0	0
	AB,CDEF-t	40.21	5.37	11.35	30.68	0	0	0	0	0	0	0	0	0
		40.21	5.37	11.35	30.68	0	0	0	0	0	0	0	0	0
	A,B,C,D,E,F-t	40.21	5.37	11.35	30.68	0	0	0	0	0	0	0	0	0

It clearly appears in Table 3.2 that the differences between the solutions produced by Tabu Search and CPLEX are very small. The largest difference for O_1 is 1.58, while, as a counterpart, both O_{31} and O_{32} have a lower value in the Tabu Search solution on the same instance. We also observe that the largest gap is 7.06 for O_{31} meaning that nurses perform on average 7.06 more visits to adjacent units. The counterpart for this instance is a decrease of O_{3+} which means that nurses have less visits to perform at patients located very far (i.e., at distance > 2) from their basic units. We also observe that Tabu Search and CPLEX solutions are very similar when the traveling component is important in the optimization process ($\omega_3 = 100$) while differences are more apparent when $\omega_1 = 100$.

These results lead to important observations. Problem $AB, CDEF$ is a partition of the real problem $ABCDEF$ into two subsets, and it seems that the CLSC CDN board finds it easier to implement since it does not require collaboration of the six team managers. When compared to A, B, C, D, E, F (i.e., the actual situation) we observe that the visit overload can be drastically reduced. For example, with $\omega_1 = 100, \omega_3 = 1$ and $\lambda = 1$, the average visit overload decreases from 16.75 for $A, B, C, D, E, F - m$ to 3.81 for $AB, CDEF - m$. This value can even reduced to 0.22 (i.e., an almost perfect balanced visit load) if the six team managers are ready to collaborate.

We also observe that the choice for parameter λ makes a big difference. For example, for the instance $ABCDEF - m$ with $\omega_1 = \omega_3 = 1$, the average visit overload decreases from 15.49 with $\lambda = 3$ to 5.46 with $\lambda = 1$. Such a decrease is achieved at the expense of an increase of O_{31} from 2.53 to 8.00, and of O_{32} from 0 to 0.05, but with a decrease of O_{3+} from 0.32 to 0. The same phenomenon can be observed for the other instances, and the CLSC CDN board should therefore consider traveling as a good opportunity to reduce imbalance in the visit load. In summary, even though districting simplifies the work of each team managers, authorizing travels to other districts helps obtaining more balanced visit loads.

Notice also that it often happens when $\lambda = 3$ that $O_{32} = 0$ while $O_{3+} > 0$. This is for example the case for $ABCDEF - m$ with weights 100-0-1 where $O_{3+} = 3.26$. This simply

means that the nurses make in average 3.26 visits to adjacent basic units of other districts, but do not perform any visit is basic units at distance 2 in their district.

We now report results that include the case load (i.e., where $\omega_2 > 0$). We have performed tests for the instance $ABCDEF$ with all ω_i equal to 1, with two of them equal to 1 and one equal to 100, with two of them equal to 100 and one to 1, and finally with three different values 1, 100 and 10000. Each test with $\omega_1 = \omega_3$ is compared with the solution reported in Table 3.2 for $\omega_1 = \omega_3 = 1$, while the tests with $\omega_1 > \omega_3$ are compared with the solution with $\omega_1 = 100$ and $\omega_3 = 1$ in Table 3.2, and the tests with $\omega_1 < \omega_3$ are compared with the solution with $\omega_1 = 1$ and $\omega_3 = 100$ in Table 3.2. The results are given in Tables 3.3 and 3.4. The columns are labeled as in Table 3.2 and we also give two lines of results for each instance, the first one with parameter $\lambda = 3$, and the second one with $\lambda = 1$. Table 3.3 contains the results for the case manager nurses while Table 3.4 contains those for the nurse technicians.

We observe that the solutions obtained with $\lambda = 1$ have systematically lower average visit overloads O_1 than those obtained with $\lambda = 3$, except when $\omega_1 - \omega_2 - \omega_3$ is equal to $1 - 1 - 100$, $10000 - 1 - 100$ and $1 - 10000 - 100$, and even in those cases, the increase in the visit overload is only equal to 0.5, 4.6 and 0.2, respectively. The decrease is particularly impressive for nurse technicians where the average visit overload drops from 33.13 to 2.01 for $\omega_1 = 100$ and $\omega_2 = \omega_3 = 1$. For the same instance, O_{3+} is reduced to 0, and the O_{2j} values decrease, for example from 30.68 to 8.45 for O_{23}. However, this is done at the expense of more travels to adjacent basic units from different districts, since O_{31} increase from 0 to 45.57.

We can also observe that Tabu Search is able to balance the case load since all O_{2j} are almost equal to 0 when ω_2 is larger than the two other weights, especially when $\lambda = 1$. If we compare the solution 1-0-1 (obtained with CPLEX) with the solution 1-100-1 (obtained with Tabu Search) for the case manager nurses, we can observe that O_1 does not change and the travel load slightly increases, while the case load is reduced to almost 0. Hence balanced case loads can be obtained without inducing too much increase in the visit and travel overloads.

Table 3.3 – Results for the case manager nurses on the whole territory.

ω_1-ω_2-ω_3	O_1	O_{21}	O_{24}	O_{25}	O_{31}	O_{32}	O_{3+}
1-0-1	15.49	1.58	6.25	2.90	2.53	0	0.32
	5.46	1.58	7.16	2.52	8.00	0.05	0
1-1-1	16.33	0.08	2.23	0.43	5.05	0.16	0.11
	6.98	0.08	0.74	0.06	9.32	0.37	0
100-1-100	15.49	1.58	4.93	2.50	2.63	0.05	0.32
	7.11	0.08	2.93	2.04	9.26	0.16	0
1-100-1	15.49	0	0.08	0	4.42	0.16	1.00
	8.52	0	0	0	9.74	0.53	0
1-0-100	17.39	1.58	4.54	2.25	0.68	0	0
	16.76	1.82	4.46	1.77	0.89	0	0
1-1-100	17.84	0.08	3.18	1.62	0.63	0	0
	18.32	1.74	2.02	1.51	2.47	0.05	0
100-1-10000	17.39	1.58	3.99	1.63	0.68	0	0
	16.76	1.58	4.22	1.79	0.89	0	0
1-100-100	19.15	0.08	1.47	0.08	2.32	0.11	0
	17.85	0.08	0.28	0.05	2.32	0.32	0
1-100-10000	20.23	0.08	2.95	1.14	0.37	0	0
	19.76	0.08	2.35	0.62	0.63	0	0
1-10000-100	18.48	0	0.08	0	1.00	0.16	1.00
	18.69	0	0	0	2.53	0.42	0
100-0-1	4.23	1.58	6.32	3.30	2.89	0	3.26
	0.22	1.50	7.56	2.66	12.16	0.37	0
100-1-1	4.23	0.08	3.15	2.43	10.84	0.37	3.26
	0.37	0.08	2.25	0.39	22.32	1.95	0.16
10000-1-100	4.23	1.58	3.97	2.52	10.68	0.37	3.21
	8.83	1.82	2.23	1.62	24.21	12.05	17.11
100-100-1	4.23	0.08	0.24	0	9.84	0.84	4.16
	3.27	0.47	1.25	0	25.47	12.47	4.79
10000-100-1	0.13	0	0.94	0.47	9.74	3.42	12.68
	0.05	0	0.51	0	20.21	11.00	0
100-10000-1	4.86	0	0	0	7.84	0	4.21
	0.61	0	0	0	21.42	5.84	0

Table 3.4 – Results for the technician nurses on the whole territory.

ω_1-ω_2-ω_3	O_1	O_{21}	O_{22}	O_{23}	O_{31}	O_{32}	O_{3+}
1-0-1	40.21	5.37	11.35	30.68	0	0	0
	28.08	7.14	10.85	25.93	11.26	0	0
1-1-1	39.86	5.37	10.24	30.68	0	0	0.43
	27.66	3.42	10.63	7.08	13.43	0	0
100-1-100	40.21	5.37	11.35	30.68	0	0	0
	28.08	5.62	9.62	26.51	11.71	0	0
1-100-1	36.55	1.61	11.35	16.61	0.14	0	7.14
	27.55	0.37	0.69	0.29	21.29	0.29	0
1-0-100	40.21	5.37	11.35	30.68	0	0	0
	39.92	5.69	11.35	30.68	0.29	0	0
1-1-100	40.21	5.37	11.35	30.68	0	0	0
	39.98	5.37	10.35	30.68	0.57	0	0
100-1-10000	40.21	5.37	11.35	30.68	0	0	0
	39.92	5.69	11.35	30.68	0.29	0	0
1-100-100	40.21	5.03	11.35	30.68	0.14	0	0
	34.61	1.08	10.34	9.46	10.14	0	0
1-100-10000	40.21	5.37	11.35	30.68	0	0	0
	39.98	5.37	10.35	30.68	0.57	0	0
1-10000-100	36.55	1.61	11.35	16.61	0.14	0	7.14
	29.21	0.18	0.69	0.29	18.43	0.57	0
100-0-1	33.24	7.63	8	9.81	0	0	5.29
	2.33	7.19	11.2	30.51	37.14	0	0
100-1-1	33.13	6.4	9.81	30.68	0	0	5.43
	2.01	4.29	3.6	8.45	45.57	0	0
10000-1-100	33.24	7.4	11.35	30.5	0	0	5.29
	2.44	6.02	6.97	14.46	43.29	0	0
100-100-1	33.51	5.24	11.35	9.46	0	0	9.71
	2.02	0.93	0.78	0.35	48	0.14	0
10000-100-1	3.98	5.51	8.43	22.13	0	0	31
	0.05	0.78	1.26	0.34	70.14	3.43	0
100-10000-1	27.52	0.21	0.7	0.35	0.29	0	17.14
	2.23	0	0	0	53.43	0	0

If we put the emphasis on the travel load (i.e., ω_3 is the largest weight), we observe that O_{32} and O_{3+} are reduced to almost 0 while the O_{2j} values typically decrease and O_1 slightly increase when we compare the CPLEX solution (obtained with $\omega_2 = 0$) to the Tabu Search solutions (obtained with $\omega_2 > 0$). On the opposite, if we put the emphasis on the visit load (i.e., ω_1 is the largest weight), Tabu Search is able to reduce the visit overload to very low values while decreasing the O_{2j} values simultaneously. For example weights 10000-100-1 for the case manager nurses produce a solution with $O_1 = 0.13$ and all O_{2j} smaller than 1, while CPLEX obtains $O_1 = 4.23$, $O_{24} = 1.58$, $O_{24} = 6.32$ and $O_{25} = 3.30$. This is done at the expense of the travel load.

3.7 Dynamic assignment

In practice, the list of patients is not known in advance, with the exception of long term patients who are already assigned to a nurse and can not be reassigned to a different nurse. When a new request arrives at the CLSC CDN, it is typically immediately assigned to a nurse, although the team managers can consider making the assignment a few days later in the week. By waiting a little, the team managers have the possibility to perform the assignment of several patients at the same time and can thus better control the balance of the workload of the nurses.

In order to analyze the gain that can be obtained by not assigning the new requests on a daily basis, we compare five strategies using historical data from June and July 2002. We have first produced an assignment using Tabu Search for the patients of June, and we have then removed the patients not needing any home care in July. This gives an initial workload for each nurse which can not be modified. We have then considered the new requests in July and these have been assigned to the nurses on a regular basis. We have assigned the new patients to the nurses every τ days, where $\tau = 1$ is a daily basis, $\tau = 3$ is twice a week, $\tau = 7$ is once per week, $\tau = 15$ is twice a month, and $\tau = 31$ is once per month. When assigning new requests every τ days, we consider that the assignment generated in the previous days can not be modified. Hence, the solutions typically improve

Table 3.5 – Dynamic assignment of case management nurses.

ω_1-ω_2-ω_3		$\lambda = 3$					$\lambda = 1$				
		$\tau = 1$	$\tau = 3$	$\tau = 7$	$\tau = 15$	$\tau = 31$	$\tau = 1$	$\tau = 3$	$\tau = 7$	$\tau = 15$	$\tau = 31$
1-1-1	O_1	11.74	12.05	11.77	11.77	11.43	4.49	5.81	4.23	5.91	5.43
	O_{21}	0	0	0	0	0	0	0	0	0	0
	O_{24}	2.58	2.51	2.06	2.57	2.45	1.29	1.75	1.01	1.74	1.43
	O_{25}	0.49	0.49	0.49	0.49	0.32	0.07	0.09	0	0	0
	O_{31}	5.95	6	5.95	5.84	6.47	9.79	9.63	9.63	9.58	9.74
	O_{32}	0.21	0.32	0.21	0.11	0.05	0.53	0.68	0.37	0.37	0.26
	O_{3+}	0.37	0.32	0.21	0.32	0.26	0	0	0	0	0
1-1-100	O_1	15.2	15.63	16.83	16.98	16.92	17.11	16.13	16.47	16.92	16.99
	O_{21}	0	0	0	0	0	0	0	0	0	0
	O_{24}	3.78	3.95	3.43	4.02	3.89	4.04	3.97	3.47	3.55	3.73
	O_{25}	1.24	1.32	1.82	1.03	1.03	0.69	0.76	0.96	0.75	0.75
	O_{31}	1.42	1.32	1.05	1	1.05	1.16	1.47	1.37	1.32	1.21
	O_{32}	0	0	0	0	0	0	0	0	0	0
	O_{3+}	0.26	0.16	0.32	0.26	0.21	0	0	0	0	0
10000-1-100	O_1	1.77	1.64	2.06	3.95	3.32	0.38	0.26	0.26	0.52	0.18
	O_{21}	0	0	0	0	0	0.24	0	0	0.24	0
	O_{24}	3.96	5.4	4.99	5.1	5.29	4.88	4.78	3.25	2.79	4.45
	O_{25}	1.74	2.25	1.92	2.25	2.19	1.9	1.56	1.93	2.04	2.29
	O_{31}	12.16	11.95	11.74	12.16	11.58	21.95	22.47	22.84	25.37	23.21
	O_{32}	0.79	0.79	1.16	1.21	1.37	5.37	4.84	5	5.11	4.42
	O_{3+}	4.16	3.89	3.79	3.74	2.84	0	0	0	0.11	0.05
10000-100-1	O_1	0.3	0.25	0.2	0.15	0.19	0.18	0.19	0.14	0.16	0.16
	O_{21}	0	0	0	0	0	0	0	0	0	0
	O_{24}	2.09	1.06	1.71	1.69	2.1	0.96	1.29	1.23	1.19	0.73
	O_{25}	0.11	0	0	0	0	0.08	0	0	0	0
	O_{31}	10.95	14.47	12.84	12	11.95	24.53	26.47	25	22.63	21.58
	O_{32}	4.11	0.63	3.95	4.74	4.26	17.37	16.11	13.26	11.63	13.16
	O_{3+}	22.26	20.37	18.89	15.63	13.47	3.89	2.68	2.58	2.32	1.58

when τ increases. In Tables 3.5 and 3.6, we report the results obtained with $\lambda = 3$ and 1, and with weightings 1-1-1, 1-1-100, 10000-1-100 and 10000-100-1 (which are considered as the most realistic by the nurses and the CLSC CDN board). Table 3.5 contains the results for the case manager nurses while Table 3.6 contains those for the nurse technicians.

A first observation is that there is no big difference between the solutions obtained with $\tau = 1$ and those with $\tau = 31$. A decrease of one overload is often obtained at the expense of an increase of another overload. For example, the instance 1-1-100 puts the emphasis on the travel load. For the case manager nurses with $\lambda = 3$, we observe that O_{31} and O_{3+} can be decreased from 1.42 and 0.26 to 1.05 and 0.21 when τ increases from 1 to 31. This is obtained at the expense of a slight increase of O_1 and O_{24}, but with a decrease of O_{25} from 1.24 to 1.03.

Table 3.6 – Dynamic assignment of nurse technicians.

		$\lambda = 3$					$\lambda = 1$				
ω_1-ω_2-ω_3		$\tau = 1$	$\tau = 3$	$\tau = 7$	$\tau = 15$	$\tau = 31$	$\tau = 1$	$\tau = 3$	$\tau = 7$	$\tau = 15$	$\tau = 31$
1-1-1	O_1	31.62	31.4	31.11	31.8	31.6	2.01	2.24	2.47	2.36	2.33
	O_{21}	8.37	8.66	5.76	7.88	8.58	5.38	4.72	6.08	6.62	4.61
	O_{22}	5.91	5.13	8.67	7.63	8.67	5.32	6.99	6.27	7.22	7.53
	O_{23}	24.77	24.88	24.36	21.19	23.73	7.61	9.84	6.36	4.85	8.58
	O_{31}	0	0	0	0	0	52.43	50.57	51.43	53.29	51
	O_{32}	0	0	0	0	0	0.14	1	0	0	0
	O_{3+}	16.29	16.43	16.29	16.43	15	0	0	0	0	0
1-1-100	O_1	40.3	40.3	40.3	40.3	40.3	39.79	39.59	39.76	39.82	39.67
	O_{21}	9.67	9.67	9.67	9.67	9.67	8.82	7.78	8.36	9.26	9.03
	O_{22}	8.67	8.67	8.67	8.67	8.67	8.43	7.56	8.36	8.36	8.36
	O_{23}	27.81	27.45	27.81	27.81	27.81	27.01	27.09	27.45	26.99	26.99
	O_{31}	0	0	0	0	0	1	1.57	0.71	0.71	0.86
	O_{32}	0	0	0	0	0	0.14	0	0	0	0
	O_{3+}	0	0.29	0	0	0	0	0	0	0	0
10000-1-100	O_1	31.8	31	31.17	31.62	31.86	2.38	2.13	2.58	2.41	2.64
	O_{21}	8.17	3.55	5.45	8.92	8.36	5.05	7.09	5.7	4.95	6.45
	O_{22}	5.74	5.11	8.18	7.67	8.67	5.57	2.83	2.09	6.94	2.46
	O_{23}	24.71	24.37	18.87	23.14	21.63	11.36	10.77	9.56	11.28	11.3
	O_{31}	0	0	0	0	0	51.71	51.57	54.71	52.43	55
	O_{32}	0	0	0	0	0	0	1	0	0	0
	O_{3+}	16.14	18.71	18.71	15.43	15.71	0	0	0	0	0
10000-100-1	O_1	4.93	4.76	4.87	4.67	4.61	0.13	0.18	0.06	0.	0.07
	O_{21}	7.04	6.89	5.49	7.02	8.39	1.13	1.07	1.22	0.84	0.82
	O_{22}	7.53	6.06	8.08	8.18	5.77	0.91	1.35	1.16	0.84	1.07
	O_{23}	15.24	14.96	14.04	11.78	16.51	1.37	1.98	2.17	1.12	2.15
	O_{31}	0	0	0	0	0	62.29	59	52.14	58.71	49.86
	O_{32}	0	0	0	0	0	12.29	20	17	10	11.71
	O_{3+}	57.14	58.71	56.29	56.43	56.57	0	0	0	0	3.86

Parameters 10000-100-1 produce interesting results for the case manager nurses. Such a weighting gives a higher priority to the visit load, but without neglecting the case load. We can observe that by increasing τ from 1 to 31, one can reduce all overloads since O_1, O_{31}, O_{32}, O_{3+}, O_{21}, O_{24}, and O_{25} decrease from 0.18, 24.53, 17.37, 3.89, 0, 0.96 and 0.08 to 0.16, 21.58, 13.16, 1.58, 0, 0.73 and 0, respectively.

All solutions reported in Tables 3.5 and 3.6 show again that fixing λ equal to 1 (i.e., encouraging travels to adjacent basic units of different districts) helps obtaining much lower visit and case overloads, but at the expense of a higher travel load. For example, for the nurse technicians with parameters 1-1-1, O_1 decreases from 31.6 to 2.33, O_{21}, O_{22} and O_{23} decrease from 8.58, 8.67 and 23.73 to 4.61, 7.53 and 8.58, O_{3+} is reduced from 15 to 0, while O_{31} augments from 0 to 51.

3.8 Conclusion

We have considered the problem of assigning patients to nurses for home care services. A previous work has shown that when the nurses offer their services only in the district where they are located, demand fluctuations may create imbalance and inequities among them, and one should therefore consider the possibility of assigning them patients from basic unit in other districts.

For this purpose, we have developed a measure of the workload of the nurses which takes into account the number of visits performed by the nurses, the heaviness of each patient, the number of patients that the nurses have in each category, and the travels needed to visit the patients. We have then modeled the patient assignment problem as a mixed integer program with some non linear constraints and a non linear objective. When the case load is not taken into account, while the objective is to minimize the travel and visit overloads, we have shown that the model contains only linear constraints and can therefore be solved using CPLEX. By adding the objective of minimizing the case overload, non linear constraints must be taken into account and we solve the problem using a Tabu Search algorithm with various neighborhoods.

The effectiveness of the Tabu Search algorithm has been confirmed by making comparisons with CPLEX on instances where the case load is not considered. The tests performed on real historical data have shown that it is possible to drastically reduce the visit and case loads of the nurses if they accept to move to basic units that are not too far from where they are located, but possibly in another district. Giving the opportunity to nurses to leave their district is comparable to make borders between districts more flexible. This is an interesting alternative when compared to reorganizing districts which is time and resource consuming and can cause important changes in patients follow-up.

It was shown that the proposed Tabu Search algorithm performed well for the addressed problem. However, some more complex and possibly more performing heuristics, such as genetic and memetic algorithms, scatter search and variable neighborhood search (e.g., see [134]) may be investigated as an extension to this paper.

CHAPITRE 4 : A FLEXIBLE MILP MODEL FOR MULTIPLE-SHIFT WORKFORCE PLANNING UNDER ANNUALIZED HOURS

Nous abordons dans ce chapitre un problème qui s'inspire d'un cas pratique d'une entreprise de fabrication. L'annualisation du temps de travail est utilisée pour mieux adapter la production à l'irrégularité de la demande. Le contexte de cette entreprise est particulier dans la mesure où la demande ne peut être reportée (délai de 24 heures pour répondre à la demande), et aucun stockage ne peut être fait (les produits sont uniques).

Les opérateurs de machine travaillent des quarts de travail d'une longueur de 12 heures. Le nombre de quarts de travail couvert par chaque opérateur chaque semaine varie entre deux et trois lorsque les opérateurs travaillent en temps régulier, et entre quatre et cinq lorsque les opérateurs font du temps supplémentaire. L'objectif de l'entreprise est de minimiser le nombre de semaines travaillé en temps supplémentaire.

Nous avons développé un modèle linéaire mixte flexible qui non seulement répond au problème de l'entreprise, mais permet de tester divers scénarios. Les objectifs principaux sont de minimiser le nombre de semaines travaillé en temps supplémentaire ainsi que le temps de travail annuel. Nous avons tout d'abord proposé la relaxation des contraintes relatives au temps de travail. Au lieu de fixer ce temps sur deux semaines, nous avons fait varier cet horizon. Nous avons aussi, en plus de cette variation, testé l'impact de la relaxation de cette contrainte. Le temps de travail peut être relaxé ou borné, permettant ainsi plus de flexibilité. Nous avons ensuite proposé trois stratégies permettant d'introduire plus de flexibilité. L'introduction du temps partiel et de l'engagement graduel permettent tout en continuant à répondre à toutes les demandes de l'entreprise, d'économiser du temps de travail et d'évaluer le besoin exact en personnel. Une dernière variation qui touche à la longueur des quarts de travail permet d'obtenir les meilleures solutions.

Toutes les variantes de ce problème sont résolues avec CPLEX et montrent que l'introduction de la flexibilité dans la planification de la production fournit de très bons résultats.

A flexible MILP model for multiple-shift workforce planning under annualized hours

Alain Hertz **Nadia Lahrichi**

alain.hertz@gerad.ca nadial@crt.umontreal.ca

Département de mathématiques et de génie industriel

École Polytechnique de Montréal

C.P. 6079, succ. Centre-ville

Montréal, Québec H3C 3A7, Canada

and

GERAD

3000, chemin de la Côte-Sainte-Catherine

Montréal, Québec H3T 2A7, Canada

Marino Widmer

marino.widmer@unifr.ch

Département d'informatique

Université de Fribourg

Bd de Prérolles 90

1700 Fribourg, Switzerland

Article soumis à

European Journal of Operational Research

Abstract

Flexibility in workforce planning is one of the best ways to respond to fluctuations of the demand. This paper proposes a flexible mixed integer linear programming (MILP) model to solve a multiple-shift workforce planning problem under annualized working hours. The model takes into account laws and collective agreements that impose constraints on overtime and holidays. We consider possible gradual hiring of full time and partial time workers. Several objectives are pursued such as balancing the workload of the employees or minimizing the workforce size. Computational experiments on a real life problem demonstrate the effectiveness of the model.

4.1 Introduction

Demand fluctuations are one of the most significant and potentially serious external factors that industrial companies have to face. Indeed, operation planning is typically demand based while the market is in constant evolution. Different strategies are offered to the companies to overcome these changes at each operation level : supply chain management, inventory management, demand forecasting or workforce planning.

Workforce planning is a well studied problem that deals not only with workforce scheduling but also with flexible working arrangements. Studies on this topic are divided into *single-shift* scheduling and *multiple-shift* scheduling. In multiple-shift scheduling, each day is divided into several shifts (for example, day, evening and night), and scheduling involves the determination of both work hours of the day and work days of the week for each employee. Azmat and Widmer [33] divide these two scheduling types into four sub-categories. *Regular work schedules* have five work days per week while *compressed work schedules* have three, four or three-four work days per week. *Hierarchical workforce scheduling* occurs when different classes of workers are considered (class A worker is able to perform job A, class

B worker is able to perform jobs A and B, and so on). *Annualized hours scheduling* occurs when the number of work hours per week are not fixed, but the annual amount is fixed or limited.

Regular and compressed work are very common in different contexts such as manufacturing, health care or services. Constraints are typically imposed on the number of weekends off over a certain period, and on the number of consecutive work days. The single-shift case is studied for example in [48, 120, 127, 53] while multiple-shift is considered in [49, 50, 121, 122, 123, 124, 128, 129]. Different solution approaches are proposed such as constraint programming [146], mixed integer programming [37], and heuristic algorithms [49, 142]. NP-hardness is proved in Lau in [148], which justifies the use of heuristic algorithms. Hierarchical workforce scheduling is particularly relevant when different jobs have to be performed and some employees can perform more than one. This kind of problem was studied under the single-shift [86, 125, 161, 163] and the multiple-shift [51, 119, 162] environments. An integer programming model of hierarchical workforce scheduling is presented in [41].

This paper considers annualized hours scheduling, which introduces flexibility to respond to demand fluctuations. The use of annualized working hours, can drastically reduce overtime, as shown for example in [193], since overtime during certain weeks can be compensated by less working hours during other weeks. Corominas et al. [68, 70] show that annualized hours are an efficient means to achieve production flexibility. They propose a MILP model to minimize the cost due to overtime and temporary workers [67, 69]. Temporary workers were also introduced in [72]. In [73], they consider the version of the problem where all employees take their holidays simultaneously and the production stops when workers are on holidays. Azmat et al. [32] propose a MILP model to generate a workforce schedule that minimizes overtime and balances the workload of the employees over a year, satisfying weekly customer demands. Heuristic algorithms for annualized hours scheduling problems can be found in [33, 131, 130].

The problem studied in this paper is a multiple-shift workforce planning problem under annualized hours. It comes from a real life scheduling problem in a Swiss company. It differs

from standard workforce scheduling problems mainly because each product is unique and customized for its client, and every customer demand has to be fulfilled within 24 hours. This means that there is no storage option and the production cannot be delayed. The next section contains a precise description of the considered real life problem. While Swiss laws and collective agreements impose strong constraints on the schedule, we analyze several options which, under different laws and agreements, could drastically diminish overtime or the workforce size. The MILP model presented in Section 3 is flexible enough to deal with all these options and can therefore be considered as a decision management tool. Computational experiments are reported in Section 4 and we conclude with final remarks.

4.2 Problem description

In this section, we give a detailed description of the considered real life problem and describe several options which can help to reduce overtime or the workforce size.

4.2.1 A real life problem

The considered Swiss company is running 24 hours every day of the year, which means that production never stops. Every day is divided into two shifts of 12 hours. Customer demands can arrive at any time and have to be fulfilled within 24 hours. Since each product is unique and customized for its client, there is no possibility to accumulate stocks or to delay the production of a product. There must be at least one worker on each shift every day. There are 11 available employees which are all full time workers (i.e., part time is not considered). Temporary workers are not an option since sharp knowledge is needed to work in the company. A *regular* week schedule contains 2 or 3 shifts (i.e., 24 or 36 hours), while an *irregular* week schedule can have 4 or 5 shifts (i.e., 48 or 60 hours). The following constraints must be satisfied :

- Each worker must have 3 holiday weeks per year, among which at least two are consecutive.

- When a worker has two consecutive regular week schedules, the total amount of working hours for these weeks must be 60, which means that he works 2 shifts in one week and 3 shifts in the other. Also, for two consecutive irregular week schedules, the total amount must be 108 hours, which means one week having 4 shifts and the other week having 5 shifts. There is no constraint for two week schedules of different type.
- Every week, all workers not on holiday must have the same schedule type, either regular or irregular. We will therefore speak about regular and irregular weeks.
- It is forbidden to have 12 or more consecutive irregular weeks.
- The annual workload (i.e., the total number of hours worked) of an employee cannot exceed 2136 hours (i.e., 178 shifts).

The aim of the company is to fulfill the demand with a minimum number of irregular weeks.

4.2.2 Considered variations

The above constraints are very restrictive. We will therefore consider several variations which should help reducing the overtime or the workforce size.

The first considered option is to soften the condition on worktime over consecutive weeks with the same schedule type. Instead of imposing 60 hours work for two consecutive regular week schedules and 108 for two consecutive irregular week schedules, we consider 60 and 108 as upper bounds, which means that two consecutive weeks with 2 shifts (for a total of 48 hours) are allowed, as well as two consecutive weeks with 4 shifts (for a total of 96 hours). Two consecutive weeks with 3 shifts or with 5 shifts are however still forbidden since they give a total of 72 hours for two regular week schedules and 120 hours for two irregular week schedules. The next variation will then consist on totally removing the constraints on two consecutive week schedules, which means that consecutive weeks with 3 or 5 shifts will also be permitted. We will also consider similar restrictions on more than two consecutive weeks. For example, instead of imposing different working hours on two consecutive weeks, we will impose a maximum of $\frac{T}{2}$ weeks with the same number of

working hours on T consecutive weeks. We will consider $T = 2, 4$ and 52, which corresponds to impose constraints every two weeks, every month or over the whole year.

The rule that imposes that all workers not on holiday must have the same schedule type is very constraining. We will therefore relax this constraint. We will then say that a particular week k has an irregular schedule if and only if at least one employee has an irregular schedule on week k. Since such a relaxation may create large workload imbalance among the employees, we will analyze the impact of fixing an upper bound on such an imbalance. In particular, we will try to generate perfectly balanced workloads.

Another option is to consider gradual hiring, which means that the company can hire an employee any time during the year. Once an employee starts his work, he remains in the company for the rest of the year.

The last variation we consider is to change the status of some employees from full time to part time and to analyze the impact on the number of irregular week schedules. We assume that part time employees have a 50% workload. Hence, a regular week schedule for a part time employee has 12 or 18 hours, while an irregular one has 24 or 30 hours. While we assume that full time employees keep their status during the whole year, we will consider the possibility of modifying the status of part time employees to full time during the year.

4.3 A MILP model

We now present a mixed integer linear programming model for the problem described in Section 4.2.1 and for its variations in Section 4.2.2.

We consider a period of K weeks and assume that each day of each week is divided into s shifts of h hours. A set I of operators are available to satisfy the demand. Each operator can work as full time or part time. We denote W_f (resp. W_p) the set of possible week loads for a full time (resp. part time) employee. In our real life problem we have $|I| = 11$, $K = 52$ weeks with $s = 2$ shifts of $h = 12$ hours, and $W_f = \{0, 24, 36, 48, 60\}$ while $W_p = \{0, 12, 18, 24, 30\}$.

A *configuration* is a pair (x, y) with $x \in W_f$ and $y = f$ (where "f" stands for full time) or $x \in W_p$ and $y = p$ (where "p" stands for part time). The assignment of a configuration (x, y) to an employee $i \in I$ on week $k \in \{1, \cdots, K\}$ means that i works x hours with status y at week k. We denote J the set of all possible configurations. We consider the partition of J into two subsets J_f and J_p, where J_f is the set of configurations (x, y) with $y = f$ and J_p is the set of those with $y = p$. We also consider a second partition of J into three sets J_{hol}, J_{reg} and J_{irr}, where J_{hol} denotes the set of configurations (x, y) with $x = 0$ (i.e., holidays), J_{reg} stands for the set of configurations (x, y) with a regular week schedule, and J_{irr} stands for those with an irregular week schedule. In our real life problem, we have

- $J_f = \{(0, f), (24, f), (36, f), (48, f), (60, f)\}$,
- $J_p = \{(0, p), (12, p), (18, p), (24, p), (30, p)\}$,
- $J_{hol} = \{(0, f), (0, p)\}$,
- $J_{reg} = \{(24, f), (36, f), (12, p), (18, p)\}$,
- $J_{irr} = \{(48, f), (60, f), (24, p), (30, p)\}$,
- $J = J_f \cup J_p = J_{hol} \cup J_{reg} \cup J_{irr}$.

To simplify notations, we denote a_j the number x of hours in configuration $j = (x, y)$. For example, we have $a_j = 24$ for configuration $j = (24, p)$.

The demand at week k is denoted d_k and corresponds to the number of hours that employees have to work to satisfy it. The company estimates that only $t\%$ of the hours spent by an employee in the company are dedicated to satisfy a demand. For the rest of the time, the employee is busy with other tasks such as cleaning the machines or doing administrative tasks. Hence, a demand d_k at week k means that there is a need for $\frac{100 d_k}{t}$ working hours. The company has estimated $t = 79$.

We denote w_{reg} (resp. w_{irr}) the imposed total number of working hours for two consecutive regular (resp. irregular) week schedules for a full time employee. These values have to be divided by 2 when considering part time workers. Moreover,

- A is the maximum annual workload of each employee,
- C is the maximum number of consecutive irregular weeks,
- Q is the minimum number of employees requested on each shift every day,

- H is the number of holiday weeks that each employee must receive considering he works for the company during K weeks.

In our problem we have $w_{reg} = 60$, $w_{irr} = 108$, $A = 2136$, $C = 12$ and $Q = 1$ and $H = 3$.

In the following MILP model, we consider the three following sets of boolean variables :

$$x_{ijk} = \begin{cases} 1 & \text{if configuration } j \text{ is assigned to employee } i \text{ at week } k \\ 0 & \text{otherwise} \end{cases}$$

$$y_k = \begin{cases} 1 & \text{if week } k \text{ is irregular} \\ 0 & \text{otherwise} \end{cases}$$

$$z_{ik} = \begin{cases} 1 & \text{if weeks } k \text{ and } k+1 \text{ are holiday weeks for employee } i \\ 0 & \text{otherwise} \end{cases}$$

We can now give a mathematical formulation of each constraint. Greek letters will correspond to parameters that help considering the variations mentioned in Section 4.2.2.

4.3.1 Constraints

Constraints (4.1) impose that the demand has to be fulfilled every week :

$$\sum_{i \in I} \sum_{j \in J} \frac{t}{100} a_j \, x_{ijk} \geq d_k \qquad \forall k = 1, \cdots, K \tag{4.1}$$

Constraints (4.2) bound the number of consecutive irregular weeks :

$$\sum_{k}^{k+C} y_k \leq C \qquad \forall k = 1, \ldots, K - C \tag{4.2}$$

Constraints (4.3) require at least Q persons on each shift every day, for a total of $7sQ$ persons per week (where $\frac{a_j}{h}$ corresponds to the number of shifts in configuration j) :

$$\sum_{i \in I} \sum_{j \in J} \frac{a_j}{h} \, x_{ijk} \geq 7 \, s \, Q \qquad \forall k = 1, \cdots, K \tag{4.3}$$

We use parameter $\alpha = 1$ to indicate that gradual hiring is not permitted. In such a case, every employee has to be assigned to exactly one configuration every week of the year,

which means that for every $i \in I$ and $k = 1, \cdots, K$, there must exist a unique $j \in J$ with $x_{ijk} = 1$. This is imposed with constraints (4.4). We set $\alpha = 0$ to allow gradual hiring, and constraints (4.4) then impose that at most one configuration can be assigned to employee i on week k.

$$\alpha \leq \sum_{j \in J} x_{ijk} \leq 1 \qquad \forall i \in I, \forall k = 1, \cdots, K \tag{4.4}$$

Parameter $\beta = 0$ indicates that part time employees are not considered, which is imposed by constraint (4.5). With $\beta = 2$, we consider possible hiring of part time employees but we then keep their part time status for the rest of the year. With $\beta = 1$, it is possible to switch the status of an employee from part time to full time. This is imposed with constraints (4.6). Constraints (4.7) impose that no full time status can change to part time.

$$\sum_{i \in I} \sum_{j \in J_p} \sum_{k=1}^{K} x_{ijk} \leq \beta \, K \, |I| \tag{4.5}$$

$$\sum_{j \in J_p} x_{ijk} \leq \sum_{j \in J_p} x_{ijk+1} - \beta + 2 \qquad \forall i \in I, \forall k = 1, \ldots, K-1 \tag{4.6}$$

$$\sum_{j \in J_f} x_{ijk} \leq \sum_{j \in J_f} x_{ijk+1} \qquad \forall i \in I, \forall k = 1, \ldots, K-1 \tag{4.7}$$

The number of holiday weeks of employee i is equal to $\sum_{j \in J_{hol}} \sum_{k=1}^{K} x_{ijk}$. This number should be equal to H if i works K weeks for the company. For the employees hired during the year and working only $K' \leq K$ weeks, the number of holiday weeks is reduced to $\lfloor \frac{H}{K} \, K' \rfloor$. Since the number of weeks worked by employee i is equal to $\sum_{j \in J} \sum_{k=1}^{K} x_{ijk}$, employee i should have $\lfloor \frac{H}{K} \sum_{j \in J} \sum_{k=1}^{K} x_{ijk} \rfloor$ holiday weeks. This can be written with the following linear constraints (4.8) :

$$\frac{H \sum_{j \in J} \sum_{k=1}^{K} x_{ijk} - (K-1)}{K} \leq \sum_{j \in J_{hol}} \sum_{k=1}^{K} x_{ijk} \leq \frac{H \sum_{j \in J} \sum_{k=1}^{K} x_{ijk}}{K} \qquad \forall i \in I \tag{4.8}$$

When an employee has a minimum of two holiday weeks, at least two of them must be consecutive. To impose this constraint, we use variable z_{ik}. Constraints (4.9) and (4.10) impose that z_{ik} equals 1 if weeks k and $k+1$ are holiday weeks for employee i, and 0 otherwise. The number of holiday weeks of employee i is equal to $\sum_{j \in J_{hol}} \sum_{k=1}^{K} x_{ijk}$, and

when this number is at least equal to 2, constraints (4.11) require $\sum_{k=1}^{K-1} z_{ik} \geq 1$ (i.e., i must have at least two consecutive holiday weeks).

$$\sum_{j \in J_{hol}} (x_{ijk} + x_{ijk+1}) \geq 2z_{ik} \qquad \forall i \in I, \forall k = 1, \ldots, K-1 \tag{4.9}$$

$$\sum_{j \in J_{hol}} (x_{ijk} + x_{ijk+1}) - 1 \leq z_{ik} \qquad \forall i \in I, \forall k = 1, \ldots, K-1 \tag{4.10}$$

$$\sum_{j \in J_{hol}} \sum_{k=1}^{K} x_{ijk} - 1 \leq K \sum_{k=1}^{K-1} z_{ik} \qquad \forall i \in I \tag{4.11}$$

Constraints (4.12) define y_k as equal to 1 if at least one employee i has an irregular schedule at week k, and 0 otherwise. The next parameter γ can be set equal to 1, 2 or 3. In order to impose the same schedule type (i.e., regular or irregular) each week on every employee, we set $\gamma = 1$, and constraints (4.13) then impose that no employee can have a regular schedule at week k if $y_k = 1$ (i.e., if another employee has an irregular schedule at week k). When $\gamma = 2$ or 3, there is no such restriction.

$$y_k \leq \sum_{i \in I} \sum_{j \in J_{irr}} x_{ijk} \leq |I|\, y_k \qquad \forall k = 1, \cdots, K \tag{4.12}$$

$$y_k + \sum_{j \in J_{reg}} x_{ijk} \leq \gamma \qquad \forall i \in I, \forall k = 1, \cdots, K \tag{4.13}$$

The total number of irregular week schedules of employee i is equal to $\sum_{j \in J_{irr}} \sum_{k=1}^{K} x_{ijk}$. When $\gamma \neq 1$, this number can vary a lot among the employees. To impose the same number of irregular week schedules to each employee, we introduce a variable U and use constraints (4.14) with $\gamma = 3$ to impose that all employees must have the same number U of irregular week schedules. Notice that U is possibly strictly smaller than $\sum_{k=1}^{K} y_k$ since it may happen that each employee is on holiday during an irregular week. With $\gamma = 2$ there is no such restriction, which means that all workers do not necessarily have the same schedule type every week, and the total number or irregular week schedules can vary a lot among the employees.

$$\sum_{j \in J_{irr}} \sum_{k=1}^{K} x_{ijk} \leq U \leq \sum_{j \in J_{irr}} \sum_{k=1}^{K} x_{ijk} + (3-\gamma)K \qquad \forall i \in I \tag{4.14}$$

We now impose restrictions on worktime over T consecutive weeks. We consider $T = 2, 4$ and 52 which corresponds to impose constraints every two weeks, every month or over the whole year. By setting parameter $\lambda = 1$, constraints (4.15) impose that at most $\frac{T}{2}$ of T consecutive weeks can have the same number of working hours. There is no restriction when $\lambda \geq 2$.

$$\sum_{k'=k}^{k+(T-1)} x_{ijk'} \leq \lambda \frac{T}{2} \qquad \forall i \in I, \forall j \in J_{irr} \cup J_{reg}, \forall k = 1, \ldots, K-(T-1) \tag{4.15}$$

A full time employee i possibly works with a regular schedule during a subset of T consecutive weeks. Constraints (4.16) with $\lambda = 3$ impose that the total number of working hours of i during these regular weeks cannot exceed $w_{reg}\ \frac{T}{2}$. For example, when $T = 2$, this gives an upper bound of w_{reg} hours (60 in our real life problem) for two consecutive regular weeks. For $T = 4$ and $w_{reg} = 60$, employee i can have 3 consecutive weeks with 36 hours and one with 60 hours, even if this gives a total of $168 > w_{reg} \cdot \frac{T}{2} = 120$ hours, the reason being that only $108 < 120$ hours are worked during regular weeks; however, 3 consecutive weeks with 36 hours and one with 24 hours is forbidden since this gives a total of $132 > 120$ hours during regular weeks. No constraint is imposed when $\lambda = 2$.

Constraints (4.17), (4.18) and (4.19) impose similar restrictions on irregular weeks and part time employees.

$$\sum_{j \in J_{reg} \cap J_f} a_j \sum_{k'=k}^{k+(T-1)} x_{ijk'} \leq (4-\lambda) w_{reg} \frac{T}{2} \qquad \forall i \in I, \forall k = 1, \ldots, K-(T-1) \tag{4.16}$$

$$\sum_{j \in J_{irr} \cap J_f} a_j \sum_{k'=k}^{k+(T-1)} x_{ijk'} \leq (4-\lambda) w_{irr} \frac{T}{2} \qquad \forall i \in I, \forall k = 1, \ldots, K-(T-1) \tag{4.17}$$

$$\sum_{j \in J_{reg} \cap J_p} a_j \sum_{k'=k}^{k+(T-1)} x_{ijk'} \leq (4-\lambda) w_{reg} \frac{T}{2} \qquad \forall i \in I, \forall k = 1, \ldots, K-(T-1) \tag{4.18}$$

$$\sum_{j \in J_{irr} \cap J_p} a_j \sum_{k'=k}^{k+(T-1)} x_{ijk'} \leq (4-\lambda) w_{irr} \frac{T}{2} \qquad \forall i \in I, \forall k = 1, \ldots, K-(T-1) \tag{4.19}$$

We finally limit the number of hours worked during the year. A full time employee cannot work more that A hours per K weeks. For full time employees that are hired during

the year and work only $K' \leq K$ weeks, this maximum number of hours is reduced to $\frac{A\,K'}{K}$. Similarly, part time employees who work $K' \leq K$ weeks have a limited number of working hours equal to $\frac{A\,K'}{2\,K}$. Constraints (4.20) impose these bounds, taking into account that an employee can have a part time status during some period of time, and a full time status later.

$$\sum_{j \in J} \sum_{k=1}^{K} a_j x_{ijk} \leq \frac{A}{K} \left(\sum_{j \in J_f} \sum_{k=1}^{K} x_{ijk} + \frac{1}{2} \sum_{j \in J_p} \sum_{k=1}^{K} x_{ijk} \right) \qquad \forall i \in I \tag{4.20}$$

4.3.2 Objectives

The objective of the company is to minimize the total number of irregular weeks. This is achieved by minimizing

$$O_1 = \sum_{k=1}^{K} y_k$$

The second considered objective is to minimize the workload of the most loaded employee. This is achieved by minimizing a variable V defined as the largest workload of an employee. More precisely, we add the following constraints to assure that V is larger or equal to the workload of every employee,

$$\sum_{j \in J_f} \sum_{k=1}^{K} a_j \; x_{ijk} + 2 \sum_{j \in J_p} \sum_{k=1}^{K} a_j \; x_{ijk} \leq V \qquad \forall i \in I \tag{4.21}$$

and we then minimize

$$O_2 = V$$

Notice that a working hour as part time employee is considered as equivalent to two working hours as full time employee.

In order to favor gradual hiring, we count the number of weeks during which each employee is working for the company, and minimize the sum of these numbers over all employees. This gives our third objective

$$O_3 = \sum_{i \in I} \sum_{j \in J} \sum_{k=1}^{K} x_{ijk}$$

Finally, in order to favor part time employees, we count the number of weeks during which each employee is working full time for the company, and we minimize the sum of these numbers over all employees. This gives our fourth and last objective

$$O_4 = \sum_{i \in I} \sum_{j \in J_f} \sum_{k=1}^{K} x_{ijk}$$

4.4 Computational experiments

We first report experiments performed on the original real life problem (i.e., with $\alpha = 1, \beta = 0, \gamma = 1, \lambda = 1$) and compare the solution of our MILP model with the solutions obtained using the algorithm described in [197]. We then analyze variations of the original problem by using different parameter settings. All results of our MILP model are obtained using CPLEX 10.0.1 for 3 hours, and we report the best encountered feasible solution.

4.4.1 Solutions for the original problem

We first report the solutions obtained with the three phase algorithm described in [197]. We used real demand data for years 1994 and 1995. The algorithm in [197] solves a relaxed version of the original problem : it imposes constraints (4.15) only on odd weeks k, which means that an employee can have two consecutive weeks k and $k+1$ of the same type and with the same workload if k is even ; also, constraints (4.2) which give an upper bound on the number of consecutive irregular weeks and constraints (4.11) which impose at least two consecutive holiday weeks are relaxed ; finally, it may happen that the demand is not satisfied, thus violating constraints (4.1). In the first phase of the algorithm, regular week schedules (24 or 36 hours per week) are assigned to each employee in order to satisfy the demand. In the second phase, when a lack of capacity is detected, some regular week schedules are changed to irregular ones. Finally, since operators need holidays, these weeks are chosen among weeks with an excess capacity. The solutions obtained on this relaxed problem appear in Tables 4.1 and 4.2.

Table 4.1 – Schedule for 1994 obtained with the three phase algorithm of [197].

Week	1	2	3	4	5	6	7	8	9	10	11	Offer	Demand	Excess
1	24	24	24	24	24	24	36	0	0	0	0	142.20	38.22	103.98
2	0	0	36	36	36	36	24	24	24	24	24	208.56	105.62	102.94
3	36	36	0	0	36	36	24	24	24	24	24	208.56	102.78	105.78
4	24	24	24	24	24	24	36	36	36	36	36	255.96	145.82	110.14
5	24	24	24	24	24	24	0	36	36	36	36	227.52	91.47	136.05
6	36	36	36	36	36	36	24	24	24	24	24	265.44	92.07	173.37
7	36	36	36	36	36	36	24	24	24	24	24	265.44	115.60	149.84
8	24	24	24	24	24	24	36	36	36	36	36	255.96	113.78	142.18
9	24	24	24	24	24	24	36	36	36	36	36	255.96	228.83	27.13
10	36	36	36	36	36	36	24	24	24	24	24	265.44	191.03	74.41
11	36	36	36	36	36	36	24	24	24	24	24	265.44	231.08	34.36
12	24	24	24	24	24	24	36	36	36	36	36	255.96	144.08	111.88
13	24	24	24	24	24	24	36	36	36	36	36	255.96	182.70	73.26
14	36	36	36	36	36	36	24	24	24	24	24	265.44	248.55	16.89
15	36	36	36	36	36	36	24	24	24	24	24	265.44	226.02	39.42
16	24	24	24	24	24	24	36	36	36	36	36	255.96	193.05	62.91
17	24	24	24	24	24	24	36	36	36	36	36	255.96	173.35	82.61
18	36	36	36	36	36	36	24	24	24	24	24	265.44	157.95	107.49
19	36	36	36	36	36	36	24	24	24	24	24	265.44	161.53	103.91
20	48	48	48	48	48	48	60	60	60	60	60	464.52	168.62	295.9
21	48	48	48	48	0	0	60	60	60	60	60	388.68	262.55	126.13
22	60	60	60	60	60	60	48	0	0	0	48	360.24	158.03	202.21
23	60	60	60	60	60	60	0	48	48	48	48	436.08	336.00	100.08
24	24	24	24	24	24	24	36	36	36	36	36	255.96	207.78	48.18
25	24	24	24	24	24	24	36	36	36	36	36	255.96	163.15	92.81
26	36	36	36	36	36	36	24	24	24	24	24	265.44	113.42	152.02
27	36	36	36	36	36	36	24	24	24	24	24	265.44	193.30	72.14
28	24	24	24	24	24	24	36	36	36	36	36	255.96	191.48	64.48
29	24	24	24	24	24	24	36	36	36	36	36	255.96	157.60	98.36
30	36	36	36	36	36	36	24	24	24	24	24	265.44	169.75	95.69
31	36	36	36	36	36	36	24	24	24	24	24	265.44	164.82	100.62
32	24	24	24	24	24	24	36	36	36	36	36	255.96	120.40	135.56
33	24	24	24	24	24	24	36	36	36	36	36	255.96	189.82	66.14
34	36	36	36	36	36	36	24	24	24	24	24	265.44	131.55	133.89
35	36	36	36	36	36	36	24	24	24	24	24	265.44	155.05	110.39
36	24	24	24	24	24	24	36	36	36	36	36	255.96	122.40	133.56
37	24	24	24	24	24	24	36	36	36	36	36	255.96	205.62	50.34
38	0	0	0	60	60	60	48	48	48	48	48	331.80	203.07	128.73
39	60	60	60	0	60	60	48	48	48	48	0	388.68	263.23	125.45
40	48	48	48	48	0	48	60	60	60	60	60	426.60	304.05	122.55
41	48	48	48	48	48	0	60	60	60	60	60	426.60	264.45	162.15
42	60	60	60	0	0	60	48	48	48	48	48	379.20	281.08	98.12
43	60	60	60	60	60	0	48	48	48	48	48	426.60	300.33	126.27
44	48	48	48	48	48	48	60	60	60	60	60	464.52	356.73	107.79
45	48	48	48	48	48	48	60	60	60	60	0	417.12	305.53	111.59
46	0	60	0	60	60	60	48	48	48	48	48	379.20	284.17	95.03
47	60	0	60	60	60	60	48	48	48	48	48	426.60	305.93	120.67
48	48	48	48	48	48	48	0	60	60	0	60	369.72	243.85	125.87
49	48	48	48	48	48	48	60	0	0	60	60	369.72	226.20	143.52
50	36	36	36	36	36	36	24	24	24	24	24	265.44	217.83	47.61
51	36	36	36	36	36	36	24	24	24	24	24	265.44	209.83	55.61
52	24	24	24	24	24	24	36	36	36	36	36	255.96	0.00	255.96
Total	1788	1788	1788	1788	1788	1788	1800	1800	1800	1800	1800			

Table 4.2 – Schedule for 1995 obtained with the three phase algorithm of [197].

Week	1	2	3	4	5	6	7	8	9	10	11	Offer	Demand	Excess
1	24	24	24	24	24	24	36	36	36	36	36	255.96	0.00	255.96
2	36	36	36	36	36	36	24	24	24	24	24	265.44	200.28	65.16
3	36	36	36	36	36	36	24	24	24	24	24	265.44	199.95	65.49
4	24	24	24	24	24	24	36	36	36	36	0	227.52	125.78	101.74
5	24	24	24	24	24	24	36	36	36	36	36	255.96	150.77	105.19
6	36	36	36	36	36	36	24	24	24	24	24	265.44	206.22	59.22
7	36	36	36	36	36	36	24	24	24	24	24	265.44	225.18	40.26
8	24	24	24	24	24	24	36	36	36	36	36	255.96	202.43	53.53
9	24	24	24	24	24	24	36	36	36	36	36	255.96	164.88	91.08
10	36	36	36	36	36	36	24	24	24	24	24	265.44	169.70	95.74
11	36	36	36	36	36	36	24	24	24	24	24	265.44	244.05	21.39
12	0	48	48	48	48	48	60	60	60	60	0	379.20	240.18	139.02
13	48	48	48	48	48	48	60	60	60	0	60	417.12	299.37	117.75
14	0	60	60	60	60	60	48	48	48	48	48	426.60	304.12	122.48
15	60	0	60	60	0	60	48	48	48	48	48	379.20	276.38	102.82
16	48	0	0	48	48	48	0	60	0	60	60	293.88	184.45	109.43
17	48	48	48	0	48	48	60	0	60	60	60	379.20	267.73	111.47
18	60	60	0	60	60	0	48	48	48	48	48	379.20	293.22	85.98
19	60	60	60	0	60	60	48	48	48	48	48	426.60	323.80	102.80
20	24	24	24	24	24	24	36	36	36	36	36	255.96	173.82	82.14
21	24	24	24	24	24	24	36	36	36	0	36	227.52	121.35	106.17
22	36	36	36	36	36	0	24	24	24	24	24	237.00	151.37	85.63
23	36	36	36	36	36	36	24	24	24	24	24	265.44	159.33	106.11
24	24	24	24	24	24	24	36	36	36	36	36	255.96	233.45	22.51
25	24	24	24	24	24	24	36	36	36	36	36	255.96	304.87	-48.91
26	36	36	36	36	36	36	24	24	24	24	24	265.44	186.07	79.37
27	0	0	36	36	0	36	24	24	24	24	24	180.12	90.02	90.10
28	24	24	24	24	24	24	36	36	36	36	36	255.96	214.80	41.16
29	24	24	24	24	24	24	36	36	0	0	0	170.64	88.22	82.42
30	36	36	0	36	36	36	24	24	24	24	24	237.00	134.05	102.95
31	36	36	36	0	36	36	24	24	24	24	24	237.00	128.75	108.25
32	24	24	24	24	24	24	36	0	0	36	36	199.08	100.72	98.36
33	24	24	24	24	24	24	0	0	36	36	36	199.08	103.33	95.75
34	36	36	36	36	36	36	24	24	24	24	24	265.44	176.73	88.71
35	36	36	36	36	0	36	24	24	24	24	24	237.00	127.95	109.05
36	24	24	24	24	24	24	36	36	36	36	36	255.96	239.67	16.29
37	24	24	24	24	24	24	36	36	36	36	36	255.96	238.88	17.08
38	60	60	60	60	60	60	48	48	48	48	48	474.00	391.38	82.62
39	60	60	60	60	60	60	0	48	48	48	48	436.08	343.35	92.73
40	48	48	48	48	48	0	60	60	60	60	60	426.60	339.37	87.23
41	48	48	48	48	48	48	60	60	60	60	60	464.52	330.80	133.72
42	60	60	60	60	60	60	48	48	48	48	48	474.00	354.70	119.30
43	60	60	60	60	60	60	48	48	48	48	48	474.00	353.98	120.02
44	48	48	48	48	48	48	60	60	60	60	60	464.52	344.78	119.74
45	48	48	48	48	48	48	60	60	60	60	60	464.52	365.30	99.22
46	60	60	60	60	60	60	48	48	48	48	48	474.00	366.65	107.35
47	60	60	60	60	60	60	48	48	48	48	48	474.00	327.95	146.05
48	48	48	48	48	48	48	60	60	60	60	60	464.52	355.70	108.82
49	48	48	48	48	48	48	60	60	60	60	60	464.52	274.00	190.52
50	60	60	60	60	60	60	48	48	48	48	48	474.00	293.42	180.58
51	60	60	60	60	60	60	48	48	48	48	48	474.00	321.78	152.22
52	24	24	24	24	24	24	36	36	36	36	36	255.96	104.42	151.54
Total	1944	1944	1944	1944	1956	1944	1944	1956	1956	1956	1956			

The first column indicates the week number and the next 11 columns represent the workload of the 11 available employees. Irregular week schedules are represented with bold characters. The next three columns indicate the offer, the demand and the excess capacity. For example, during the first week of 1994, 6 employees have a 24 hours workload, 1 works for 36 hours, and the 4 other employees are on holiday. This gives a total of 180 working hours. Since only 79% of the hours spent by an employee in the company are dedicated to satisfy a demand, we get an offer of $0.79 \cdot 180 = 142.2$, which exceeds the demand $d_1 = 38.22$ by 103.98. Notice that the demand is not satisfied on week 25 in 1995.

The solutions obtained with our MILP model, but without the relaxations on constraints (4.1), (4.2),(4.11) and (4.15) are shown in Tables 4.3 and 4.4.

For 1994, we observe from Tables 4.1 and 4.3 that we have been able to reduce the number of irregular weeks from 16 to 6, the maximum number of consecutive irregular weeks from 12 (weeks 38-49) to 1, and the maximum workload from 1800 to 1656. Hence, constraints (4.2) and (4.20) do not seem to be very restrictive. The workload is however better balanced in Table 4.1 since the maximum annual difference between two employees is 12 hours (one shift), while the MILP solution has a difference of 60 hours (5 shifts) for employees 10 and 11.

For 1995, a comparison between Tables 4.2 and 4.4 shows a reduction of the number of irregular weeks from 22 to 18. Also, the maximum number of consecutive irregular weeks is reduced from 14 (which is too much according to constraint (4.2)) to 11. The maximum workload is reduced from 1956 to 1932, but a better balance appears in Table 4.2 with a maximum difference of 12 hours (1 shift), while this difference grows up to 108 hours (9 shifts) in Table 4.4.

The offer and demand of the solutions in Tables 4.3 and 4.4 are shown on Figure 4.1. The dark area represents the demand while the excess capacity appears in grey.

Table 4.3 – MILP solution for 1994.

Week	1	2	3	4	5	6	7	8	9	10	11	Offer	Demand	Excess
1	36	36	36	24	24	36	24	24	36	0	24	237.00	38.22	198.78
2	24	24	0	36	36	24	36	36	24	0	36	218.04	105.62	112.42
3	36	36	0	24	24	36	24	24	36	24	24	227.52	102.78	124.74
4	0	0	36	0	36	24	36	36	24	36	36	208.56	145.82	62.74
5	0	0	24	0	24	36	24	24	36	24	24	170.64	91.47	79.17
6	24	24	36	24	36	24	36	36	24	36	36	265.44	92.07	173.37
7	36	36	24	36	24	36	24	24	36	24	24	255.96	115.60	140.36
8	24	24	36	24	0	24	36	36	24	36	36	237.00	113.78	123.22
9	36	36	24	36	0	36	24	24	36	24	24	237.00	228.83	8.17
10	24	24	36	24	36	24	36	36	24	36	36	265.44	191.03	74.41
11	36	36	24	36	24	36	24	24	36	24	24	255.96	231.08	24.88
12	24	24	36	24	36	24	36	36	24	36	36	265.44	144.08	121.36
13	36	36	24	36	24	36	24	24	36	24	24	255.96	182.70	73.26
14	24	24	36	24	36	24	36	36	24	36	36	265.44	248.55	16.89
15	36	36	24	36	24	36	24	24	36	24	24	255.96	226.02	29.94
16	24	24	36	24	36	24	36	36	24	36	36	265.44	193.05	72.39
17	36	36	24	36	24	36	24	24	36	24	24	255.96	173.35	82.61
18	24	0	36	24	36	24	36	36	24	36	36	246.48	157.95	88.53
19	36	36	24	36	24	36	24	24	36	0	24	237.00	161.53	75.47
20	24	24	36	24	36	24	36	36	24	24	36	255.96	168.62	87.34
21	36	36	24	36	24	36	24	24	36	36	24	265.44	262.55	2.89
22	24	24	36	24	36	24	36	36	24	24	36	255.96	158.03	97.93
23	**60**	**60**	**60**	**60**	**0**	**0**	**60**	**60**	**60**	**48**	**60**	417.12	336.00	81.12
24	24	24	24	36	36	24	24	36	24	24	24	237.00	207.78	29.22
25	36	36	36	24	24	36	36	24	36	36	36	284.40	163.15	121.25
26	24	24	24	36	36	24	0	36	24	24	24	218.04	113.42	104.62
27	36	36	36	24	24	36	0	24	36	36	36	255.96	193.30	62.66
28	24	24	24	36	36	24	24	36	24	24	24	237.00	191.48	45.52
29	36	36	36	24	24	36	36	24	36	36	36	284.40	157.60	126.80
30	24	24	24	36	36	24	24	36	0	24	24	218.04	169.75	48.29
31	36	36	36	24	24	36	36	24	0	36	36	255.96	164.82	91.14
32	24	24	24	36	36	24	24	36	24	24	24	237.00	120.40	116.60
33	36	36	36	24	24	36	36	24	36	36	36	284.40	189.82	94.58
34	24	24	24	36	36	24	24	36	24	24	24	237.00	131.55	105.45
35	36	36	36	24	24	36	36	24	36	36	36	284.40	155.05	129.35
36	24	24	24	36	36	24	24	36	24	24	24	237.00	122.40	114.60
37	36	36	36	24	24	36	36	24	36	36	36	284.40	205.62	78.78
38	24	24	24	36	36	24	24	0	24	24	24	208.56	203.07	5.49
39	36	36	36	24	24	36	36	0	36	36	36	265.44	263.23	2.21
40	**60**	**60**	**60**	**60**	**60**	**60**	**60**	**60**	**60**	**60**	**60**	521.40	304.05	217.35
41	36	36	36	36	36	36	36	36	36	36	36	312.84	264.45	48.39
42	**60**	**48**	**0**	**0**	**60**	**60**	**0**	**60**	**60**	**60**	**60**	369.72	281.08	88.64
43	36	36	36	36	36	36	36	36	36	36	36	312.84	300.33	12.51
44	**60**	**60**	**48**	**60**	**60**	**60**	**48**	**60**	**60**	**60**	**48**	492.96	356.73	136.23
45	36	36	36	36	36	36	36	36	36	36	36	312.84	305.53	7.31
46	**0**	**48**	**60**	**48**	**48**	**48**	**60**	**48**	**0**	**60**	**0**	331.80	284.17	47.63
47	36	36	36	36	36	36	36	36	36	36	36	312.84	305.93	6.91
48	**60**	**60**	**60**	**60**	**60**	**0**	**48**	**0**	**48**	**60**	**0**	360.24	243.85	116.39
49	36	24	36	36	36	0	36	24	24	36	0	227.52	226.20	1.32
50	24	36	24	24	24	36	24	36	36	24	24	246.48	217.83	28.65
51	36	24	36	36	36	24	36	24	24	36	36	274.92	209.83	65.09
52	24	36	24	24	24	36	24	36	36	24	24	246.48	0.00	246.48
Total	1644	1656	1644	1620	1632	1608	1620	1632	1632	1656	1596			

Table 4.4 – MILP solution for 1995.

Week	1	2	3	4	5	6	7	8	9	10	11	Offer	Demand	Excess
1	36	36	36	36	24	36	36	24	36	24	36	284.40	0.00	284.40
2	24	24	24	24	36	24	24	36	24	36	24	237.00	200.28	36.72
3	36	36	36	36	24	36	36	24	36	24	36	284.40	199.95	84.45
4	24	24	24	24	36	24	24	36	24	36	24	237.00	125.78	111.22
5	36	36	36	36	24	36	36	24	0	24	36	255.96	150.77	105.19
6	24	24	24	24	36	24	24	36	0	36	24	218.04	206.22	11.82
7	36	36	36	36	24	36	36	0	36	24	36	265.44	225.18	40.26
8	24	24	24	24	36	24	24	0	24	36	24	208.56	202.43	6.13
9	36	36	0	36	24	36	36	24	36	24	36	255.96	164.88	91.08
10	24	24	0	24	36	24	24	36	24	36	0	199.08	169.70	29.38
11	36	36	24	36	24	36	36	24	36	24	0	246.48	244.05	2.43
12	24	24	36	24	36	24	24	36	24	36	36	255.96	240.18	15.78
13	48	48	60	0	48	60	60	60	60	48	48	426.60	299.37	127.23
14	36	36	36	36	36	36	36	36	36	36	36	312.84	304.12	8.72
15	60	48	48	0	60	48	60	60	60	60	48	436.08	276.38	159.70
16	24	36	24	0	36	24	24	24	24	24	24	208.56	184.45	24.11
17	36	24	36	24	24	36	36	36	36	36	36	284.40	267.73	16.67
18	60	60	60	60	60	60	60	60	48	48	48	492.96	293.22	199.74
19	48	48	48	48	48	48	48	48	60	60	60	445.56	323.80	121.76
20	24	24	24	24	24	36	24	36	24	36	36	246.48	173.82	72.66
21	36	36	36	36	36	24	36	24	36	24	24	274.92	121.35	153.57
22	24	24	24	24	24	36	24	36	24	36	36	246.48	151.37	95.11
23	36	36	36	36	36	24	36	24	36	24	24	274.92	159.33	115.59
24	24	24	24	24	24	36	24	36	24	36	36	246.48	233.45	13.03
25	60	60	60	60	60	48	60	60	60	60	60	511.92	304.87	207.05
26	24	24	24	36	36	36	24	24	24	36	24	246.48	186.07	60.41
27	36	36	36	24	24	24	36	36	36	24	36	274.92	90.02	184.90
28	24	24	24	36	36	36	24	24	24	36	24	246.48	214.80	31.68
29	36	36	36	24	24	24	36	36	36	24	36	274.92	88.22	186.70
30	24	24	24	36	36	36	24	24	24	36	24	246.48	134.05	112.43
31	36	36	36	24	24	24	36	36	36	24	36	274.92	128.75	146.17
32	24	24	24	36	36	36	24	24	24	36	24	246.48	100.72	145.76
33	36	36	36	24	24	24	36	36	36	24	36	274.92	103.33	171.59
34	24	24	24	36	36	36	24	24	24	36	24	246.48	176.73	69.75
35	36	36	36	24	24	24	36	36	36	24	36	274.92	127.95	146.97
36	24	24	24	36	36	36	24	24	24	36	24	246.48	239.67	6.81
37	36	36	36	24	24	24	36	36	36	24	36	274.92	238.88	36.04
38	48	48	48	60	60	60	48	60	48	48	48	455.04	391.38	63.66
39	60	60	0	48	48	48	60	48	60	60	60	436.08	343.35	92.73
40	0	48	60	60	60	60	48	60	48	48	48	426.60	339.37	87.23
41	0	60	48	48	0	48	60	48	60	60	60	388.68	330.80	57.88
42	60	48	60	60	60	0	48	60	48	48	48	426.60	354.70	71.90
43	48	60	48	48	48	60	60	0	60	60	60	436.08	353.98	82.10
44	60	0	60	60	60	48	0	60	48	48	0	350.76	344.78	5.98
45	48	48	48	48	48	0	48	48	60	60	48	398.16	365.30	32.86
46	60	60	60	60	60	0	60	60	0	0	60	379.20	366.65	12.55
47	48	48	48	48	48	60	0	48	48	0	48	350.76	327.95	22.81
48	60	60	60	60	60	48	0	60	60	60	60	464.52	355.70	108.82
49	36	36	36	36	0	24	36	36	36	36	36	274.92	274.00	0.92
50	0	0	48	48	0	60	60	60	48	0	48	293.88	293.42	0.46
51	48	0	60	60	48	48	48	48	60	60	60	426.60	321.78	104.82
52	36	24	36	36	36	24	24	36	24	24	24	255.96	104.42	151.54
Total	1848	1824	1896	1872	1872	1824	1848	1932	1896	1860	1896			

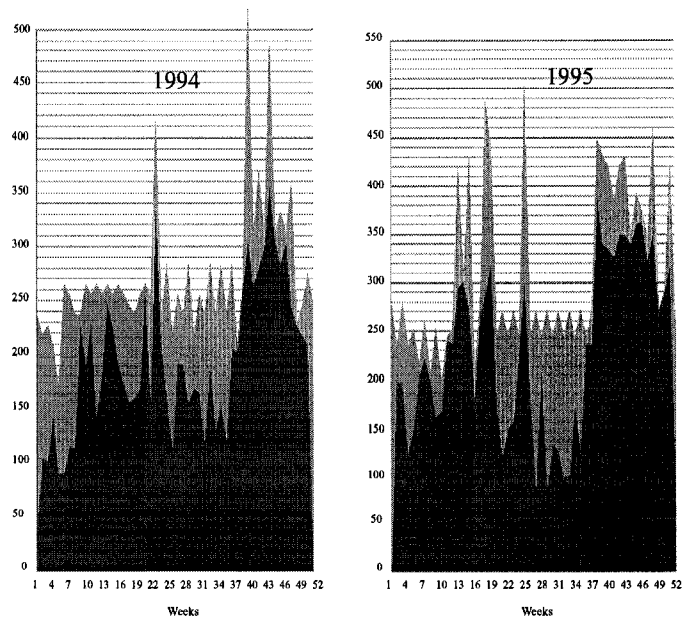

Figure 4.1 – Solutions with the original constraints. Parameter setting : $\alpha = 1, \beta = 0, \gamma = 1, \lambda = 1, T = 2$ and objective O_1.

4.4.2 Restrictions on worktime over T consecutive weeks

We first analyze the impact of setting λ equal to 2 or 3 instead of 1. As explained in Section 3, by setting $\lambda = 3$ with $T = 2$ in constraints (4.16)–(4.19), we impose that the total number of working hours for two consecutive regular week schedules cannot exceed 60 (while it must be equal to 60 when $\lambda = 1$). Hence, we still forbid two consecutive regular week schedules with 3 shifts and two consecutive irregular week schedules with 5 shifts. However a 24-24 or 48-48 sequence is permitted. Similarly, when $T = 4$, sequences like 36-36-36-48 or 48-60-60-36 are allowed, while 36-36-24-36, for example, is forbidden. With $T = 52$ we impose, for example, that the total number of working hours of an employee during regular week schedules cannot exceed 60 $\frac{T}{2} = 1560$ hours. When setting $\lambda = 2$, we do not impose any restriction on worktime over T consecutive weeks.

In Table 4.5 we compare results with $T = 2, 4, 52$ and $\lambda = 1, 2, 3$ and two different objectives O_1 and O_2. In column labelled "$\sum_{k=1}^{K} y_k$", we report the total number of irregular

Table 4.5 – Various solutions with different values of parameter λ.

			1994					1995				
λ	Objective	T	$\sum_{k=1}^{K} y_k$	U_{min}	U_{max}	V_{min}	V_{max}	$\sum_{k=1}^{K} y_k$	U_{min}	U_{max}	V_{min}	V_{max}
1	O_1	2	6	4	6	1620	1656	18	15	17	1896	1932
		4	6	4	6	1596	1656	15	12	14	1728	1836
		52	2	2	2	1500	1548	13	12	13	1740	1836
	O_2	2	6	3	5	1548	1560	18	15	15	1776	1812
		4	6	3	5	1548	1560	15	12	13	1728	1740
		52	2	1	2	1464	1476	13	10	12	1596	1596
3	O_1	2	6	5	5	1512	1608	18	15	17	1740	1860
		4	4	3	4	1476	1596	15	12	14	1740	1788
		52	2	1	2	1452	1632	13	11	13	1788	1956
	O_2	2	6	4	5	1356	1356	18	15	15	1584	1596
		4	4	3	3	1332	1344	15	12	13	1560	1572
		52	2	1	2	1320	1332	13	10	12	1536	1560
2	O_1	-	2	2	2	1476	1572	13	11	13	1812	1884
	O_2	-	2	1	2	1332	1332	13	10	12	1560	1560

weeks. This is an upper bound on the number $U_i = \sum_{j \in J_{irr}} \sum_{k=1}^{K} x_{ijk}$ of irregular week schedules for employee i. The two next columns labelled "U_{min}" and "U_{max}" indicate the smallest and largest number of irregular week schedules among the employees (i.e., $U_{min} = \min_{i \in I} U_i$ and $U_{max} = \max_{i \in I} U_i$). Finally, in columns labelled "V_{min}" and "V_{max}", we report the workload of the less loaded and of the most loaded employee.

With $\lambda = 1$ and objective O_1, we observe that $\sum_{k=1}^{K} y_k$ drops from 6 to 2 with $T = 52$ in 1994 and from 18 to 13 with $T = 52$ in 1995. When using the second objective O_2, the maximum load V_{max} is decreased from 1560 hours to 1476 in 1994 and from 1812 to 1596 in 1995. The gains when varying T from 2 to 4 are less marked in 1994 than in 1995.

When comparing the solutions obtained with $\lambda = 1$ and 3, we observe a big decrease of the maximum workload V_{max} since it drops from 1560 to 1356 in 1994, and from 1812 to 1596 in 1995 with $T = 2$ and objective O_2. When increasing T to 52, the maximum workload is even reduced to 1332 in 1994 and 1560 in 1995. On the opposite, the number $\sum_{k=1}^{K} y_k$ of irregular weeks does not vary a lot since it remains unchanged except for $T = 4$

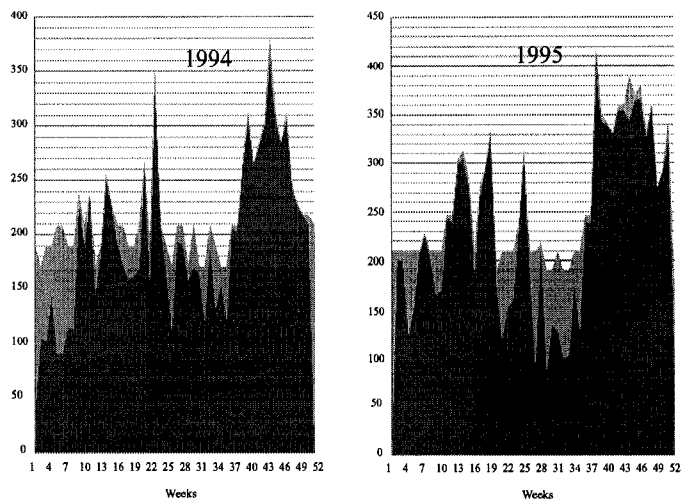

Figure 4.2 – Solutions with an annual bound on the workload. Parameter setting : $\alpha = 1, \beta = 0, \gamma = 1, \lambda = 3, T = 52$ and objective O_2.

where it drops from 6 to 4 in 1994. Looking at the solutions with $T = 52$ we believe that the reduction of the maximum workload is important enough to balance the fact that operators might work more than 60 hours per two consecutive regular weeks or more than 108 hours per two consecutive irregular weeks.

The solutions obtained with $\lambda = 2$, without any restriction on worktime over T consecutive weeks are comparable to those obtained with $\lambda = 3$ and $T = 52$ except that the workloads are better balanced with objective O_2. Indeed $V_{max} = V_{min}$ with $\lambda = 2$ while a difference of 1 or 2 shifts appears with $\lambda = 3$.

The major observation is that switching from objective O_1 to objective O_2 reduces the maximum workload V_{max} without augmenting the number of irregular weeks, which means that objective O_2 seems to dominate O_1.

The offer and demand of the solutions in Table 4.5 with $T = 52$, $\lambda = 3$ and objective O_2 are represented in Figure 4.2. When comparing it to Figure 4.1, we clearly observe a decrease of the excess capacity.

Table 4.6 – Various solutions with $\gamma = 3$.

			1994			1995		
λ	Objective	T	$\sum_{k=1}^{K} y_k$	V_{min}	V_{max}	$\sum_{k=1}^{K} y_k$	V_{min}	V_{max}
1	O_1	2	2	1524	1548	5	1608	1644
		4	2	1512	1548	5	1596	1656
		52	1	1476	1524	3	1584	1596
	O_2	2	2	1500	1500	5	1620	1620
		4	2	1500	1500	5	1608	1608
		52	1	1464	1464	3	1512	1512
3	O_1	2	2	1356	1416	5	1500	1596
		4	2	1524	1428	4	1572	1620
		52	1	1464	1584	3	1536	1692
	O_2	2	2	1308	1320	5	1488	1500
		4	2	1308	1320	4	1500	1500
		52	1	1320	1320	3	1488	1500
2	O_1	-	1	1440	1536	3	1548	1680
	O_2	-	1	1308	1320	3	1500	1500

4.4.3 Restrictions on similar week schedule types

We observe in Table 4.5 that U_{min} is often strictly smaller than U_{max}. We have therefore performed the same tests as in the previous section, but with parameter $\gamma = 3$, which means that we do not impose the same type of schedule (regular or irregular) each week for every employee, but we impose the same number of irregular week schedules for every employee. The results appear in Table 4.6.

We observe much better results when it is compared to Table 4.5. With $T = 2$ and $\lambda = 1$, the number of irregular weeks falls from 6 to 2 in 1994 and from 18 to 5 in 1995 (when using objective O_1), and the maximum workload is reduced from 1560 hours to 1500 in 1994 and from 1812 to 1620 in 1995 (when using objective O_2). Clearly, this relaxation seems profitable and the enterprise should give this option highest priority.

Again, better results are obtained with $\lambda = 3$ and by raising T to 4 or 52, and objective O_2 seems to dominate objective O_1. The best solution in 1994 has only 1 irregular week and a perfect balance workload of 1320 hours among the employees, while for 1995, the

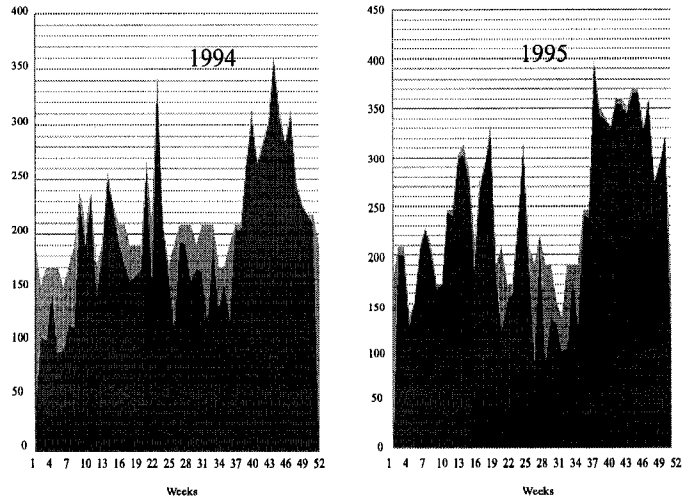

Figure 4.3 – Solutions with possibly different schedule types on each week. Parameter setting : $\alpha = 1, \beta = 0, \gamma = 3, \lambda = 3, T = 52$ and objective O_2.

number or irregular weeks is reduced to 3 (instead of 13) with almost perfectly balanced workloads ($V_{max} - V_{min}$=12 hours = 1 shift).

The offer and demand of the solutions in Table 4.6 with $T = 52$, $\lambda = 3$ and objective O_2 are represented in Figure 4.3. Again, in comparison with Figures 4.1 and 4.2, we notice an important decrease of the excess capacity.

4.4.4 Gradual hiring

Looking at Figures 4.1, 4.2 and 4.3, we observe that the excess capacity is rather small during periods with a large demand, while it is more important whenever the demand is smaller. One way to sidestep the excess capacity at the beginning of the year for both 1994 and 1995 is to consider a gradual hiring strategy. An excess capacity later during the year is more difficult to deal with since once an employee starts his work, he remains in the company for the rest of the year. Temporary workers would have been an option if the kind of work did not require very specific knowledge from the employees.

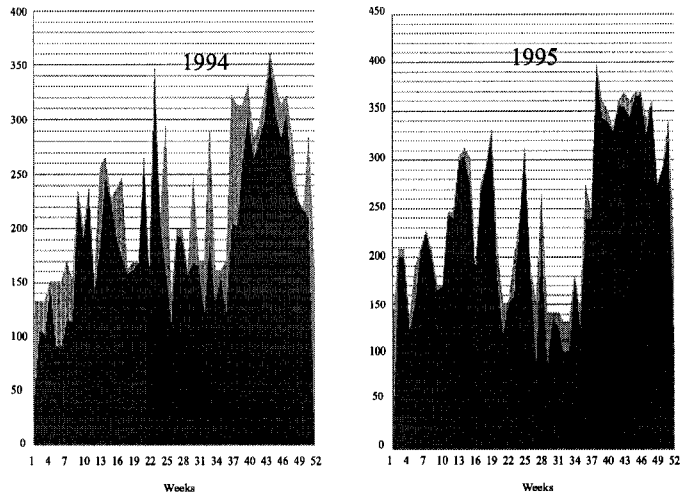

Figure 4.4 – Solutions with gradual hiring. Parameter setting : $\alpha = 0, \beta = 0, \gamma = 1, \lambda = 3, T = 52$ and objective O_3.

Gradual hiring is obtained by setting $\alpha = 0$ and using the third objective O_3 which corresponds to the total number of weeks worked by the employees during the year. An operator is thus hired only when there is no other way to satisfy the demand. Solving our MILP model with $T = 52$, $\lambda = 3$ and $\gamma = 1$, we have obtained for 1994 a solution in which $\sum_{j \in J} \sum_{k=1}^{K} x_{ijk} = 0$ for three employees, which means that only 8 employees work for the company. For 1995, only 10 of the 11 are needed. The offer and demand of this solution are shown in Figure 4.4.

We observe that the excess capacity is much smaller when compared to the previous figures. While in all previous solutions every employee worked during 52 weeks for a total of $52 \cdot 11 = 572$ weeks worked, this number drops down to 350 in 1994 and 376 in 1995. However, such a gradual hiring has induced a very high number of irregular weeks which has raised up to 37 in 1994 and 1995 (instead of 1 and 3, as reported in Table 4.6), while the maximum workload has reached the limit of $A = 2136$ hours.

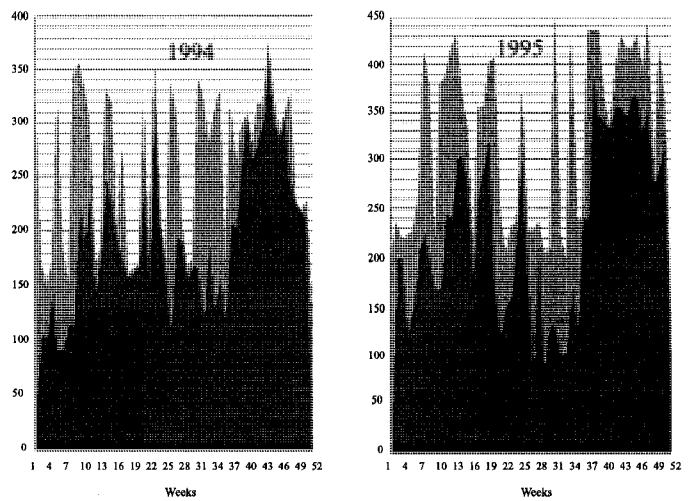

Figure 4.5 – Solutions with part time workers. Parameter setting : $\alpha = 1, \beta = 2, \gamma = 1, \lambda = 3, T = 52$ and objective O_4.

4.4.5 Part time workers

We now analyze the impact of allowing part time workers, without the possibility of switching the status of an employee from part time to full time. This is achieved by setting $\beta = 2$.

We use objective O_4 which is equivalent to minimize the number of full time employees since there is no possibility to change the status of an employee. By setting $\gamma = 1$, $\lambda = 3$ and $T = 52$, we have obtained for 1994 a solution with 5 full time (and 6 part time) employees. Again, the number of irregular weeks is very big for a total of 24 and the maximum workload is equal to $A = 2136$ hours. For 1995, the best solution has 9 full time (and 2 part time) employees, a total of 23 irregular weeks, and a maximum workload of 2010 hours. The offer and demand of these solutions are shown in Figure 4.5.

We observe a very big excess capacity at certain periods. This can be explained by the fact that the objective O_4 does not aim to minimize such an excess. Also, consider for example week 48 in 1994 where the demand in equal to 243.85. If the company works with 6 full time and 5 part time employees and the week is regular, there is a maximum

of $36 \cdot 6 + 18 \cdot 5 = 306$ working hours from which only $79\% = 241.7$ hours are offered to satisfy the demand. Since this is not sufficient, the week is necessarily irregular which gives a minimal offer of $79\%(48 \cdot 6 + 24 \cdot 5 = 408) = 322.3$ if no employee is on holiday. Hence excess capacity is often due to the fact that 11 full time employees are just enough to satisfy the demand on certain regular weeks, and if some of them work part time, the week becomes irregular which increases the workload (and the offer) during this week.

4.4.6 Different number of shifts per day

In our last experiments, we have considered $s = 3$ shifts of $h = 8$ hours instead of 2 shifts of 12 hours per day. Moreover, we have considered the limits $w_{reg} = 64$ hours $= 8$ shifts for two consecutive regular weeks, and $w_{irr} = 96$ hours $= 12$ shifts for two consecutive irregular weeks. A regular week schedule contains 3 or 5 shifts (i.e., 24 or 40 hours), while an irregular week schedule has 6 shifts, which corresponds to 48 hours. We have not considered week schedules with 7 shifts for a total of 56 hours in order to make sure that every employee has at least one day free every week. Constraints (4.15) with $\lambda = 1$ impose that at most $\frac{T}{2}$ of T consecutive weeks can have the same number of working hours. Since all irregular week schedules necessarily have 48 hours, we have removed these constraints for $j \in J_{irr}$. Hence, when $T = 2$, an employee can have two consecutive irregular week schedules with 6 shifts each, but when he has two consecutive regular week schedules, he works 3 shifts in one week and 5 shifts in the other.

The solutions obtained by setting $\alpha = 1$, $\beta = 0$ and $\gamma = 1$ are reported in Table 4.7. They have a very small number of irregular weeks when compared to Table 4.5. For example, with $\lambda = 3$, $T = 52$ and objective O_2, there is only one irregular week instead of 2 in 1994, while for 1995, the number of irregular weeks is reduced from 13 to 7. The offer and demand of these two solutions are shown in Figure 4.6. We observe that the excess capacity is very small when compared to Figure 4.2 and it seems that 3 shifts of 8 hours is much more appropriate than 2 shifts of 12 hours to cover the demand without too much excess.

Table 4.7 – Various solutions with 3 shifts of 8 hours per week.

			1994					1995				
λ	Objective	T	$\sum_{k=1}^{K} y_k$	U_{min}	U_{max}	V_{min}	V_{max}	$\sum_{k=1}^{K} y_k$	U_{min}	U_{max}	V_{min}	V_{max}
1	O_1	2	4	3	4	1616	1672	10	8	10	1736	1776
		4	3	2	3	1616	1648	9	8	9	1712	1744
		52	1	1	1	1552	1616	6	6	6	1592	1736
	O_2	2	5	3	5	1624	1632	10	9	10	1720	1744
		4	3	2	3	1592	1600	9	7	9	1712	1720
		52	1	0	1	1544	1552	6	4	6	1576	1592
3	O_1	2	4	3	4	1496	1576	10	9	10	1632	1736
		4	3	2	3	1448	1568	7	5	7	1584	1696
		52	1	1	1	1488	1696	6	6	6	1624	1768
	O_2	2	5	3	5	1352	1360	10	7	10	1544	1552
		4	3	2	3	1336	1344	7	5	7	1520	1528
		52	1	0	1	1312	1328	7	6	7	1520	1528
2	O_1	-	1	1	1	1520	1696	6	5	6	1568	1752
	O_2	-	1	0	1	1320	1328	6	4	6	1512	1528

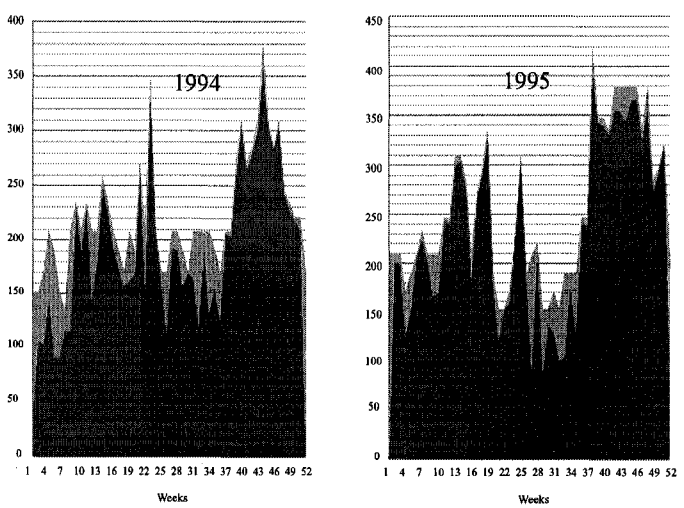

Figure 4.6 – Solutions with 3 shifts of 8 hours per week. Parameter setting : $\alpha = 1, \beta = 0, \gamma = 1, \lambda = 3, T = 52$ and objective O_2.

4.5 Conclusion

As written in [71], annualizing working hours is a good economical way of adjusting productive capacity to seasonal demand and obtaining flexibility in the distribution of annual

working hours. We have presented a flexible mixed integer linear programming (MILP) model to solve a multiple-shift workforce planning problem under annualized working hours and without capacity shortage. The model includes many options such as gradual hiring, and part time work. Different restrictions on sets of consecutive weeks are considered such as imposing the same type of week schedule each week on each employee, or bounding the workload on consecutive weeks. While our model does not take into account temporary workers, it would not be difficult to introduce them in the model, for example by removing constraints (4.6) and (4.7) and setting $\alpha = 0$ in constraints (4.4) for all temporary employees $i \in I$.

We have tested the model on a real life industrial problem and compared different solutions obtained with various parameter settings. We have observed, for example, that while the considered company has 11 employees, 10 would be sufficient, but the counterpart of such a reduction would be a big increase of the workload of the employees, leading possibly to an important overtime. The experiments reported in Section 4 clearly demonstrate that the proposed model is an interesting tool for managers facing demand fluctuations. They can use our model to generate several solutions corresponding to different options and then negotiate with the workers to choose the most appropriate solution for both.

DISCUSSION GÉNÉRALE, CONCLUSION ET RECOMMANDATIONS

L'introduction de la flexibilité dans l'organisation de la main-d'oeuvre est sans aucun doute un excellent moyen pour faire face à l'irrégularité de la demande. Nous avons, dans cette thèse, étudié l'impact de cette flexibilité dans les milieux des services de soins à domicile et de l'industrie manufacturière.

Les problématiques dans le milieu de la santé sont nombreuses. Le vieillissement de la population ainsi que l'augmentation de la demande imposent un poids croissant sur les organismes dispensateurs de services de soins. Chaque service de soins à domicile est responsable de la population de son territoire. Les deux rôles principaux des services de soins à domicile (dispensés par les CLSC au Québec) sont de diminuer la longueur des séjours dans les établissements hospitaliers et de prévenir et maintenir la santé. Afin de remplir ce dernier rôle, la méthode de "gestion de cas" est largement répandue. Le suivi de l'état du patient et la coordination avec les autres professionnels de la santé sont aussi importants que les visites à lui rendre.

L'organisation des services des soins infirmiers à domicile est propre à chaque organisme dispensateur de services, que ce soit au niveau de l'organisation de la main-d'œuvre (différenciée ou non) ou au niveau de la distribution du travail (ou distribution des patients). Nous avons recensé différentes méthodes de distribution des patients : l'approche par territoire, l'approche par "charge de travail", l'approche par "consensus" et enfin l'approche par "rotation". Nous avons, dans cette thèse, débuté par l'analyse de l'approche par territoire pratiquée au CLSC CDN. Chaque infirmière est responsable d'un territoire donné et prend en charge tous ses habitants, que la demande soit en augmentation ou non. Bien que cette approche ait un avantage certain de simplicité à gérer et à implanter, elle présente des lacunes importantes dans un contexte où la demande évolue et où les ressources ne sont pas relocalisées.

Pour s'adapter aux fluctuations de la demande sans recommencer le découpage du territoire desservi par le CLSC, nous avons développé une nouvelle méthode de distribution des patients qui tout en se basant sur le territoire, tient compte de la charge des infirmières. Nous avons développé une mesure de la charge infirmière au CLSC CDN qui tient compte des quatre composantes suivantes : le déplacement pour aller rendre visite aux patients, le nombre de patients dans chaque profil, le nombre de visites rendues pondérées par leur charge et enfin l'excès de charge comparativement aux autres infirmières. Ces différentes composantes sont pondérées dans la charge infirmière de façon à être adaptable dans différents contextes. Il est possible de tenir compte d'une seule de ces composantes, par exemple dans un contexte où les déplacements sont très peu importants. Dans notre application nous avons introduit la flexibilité des frontières entre les secteurs desservis par les différentes infirmières. Les infirmières les moins chargées se déplacent vers les secteurs les plus chargés. Nous avons évalué plusieurs scénarios qui donnent plus ou moins d'importance aux déplacements. Le problème de distribution des patients lorsque l'écart entre les charges des infirmières doit être minimisé se formule comme un problème d'affectation généralisé avec plusieurs ressources et contraintes additionnelles. Ce problème est NP-difficile d'autant plus que la fonction objectif est quadratique et certaines contraintes non linéaires. Nous avons développé une méthode de recherche avec tabous pour le résoudre faisant appel à des voisinages et mouvements complexes. La version simplifiée du problème se résout directement avec un solveur exact tel que CPLEX. Ce même modèle est assez flexible pour évaluer l'impact de l'introduction de différentes catégories d'infirmières et analyser différentes organisations du travail.

La flexibilité dans le milieu industriel intervient dans la planification de la production. L'annualisation du temps de travail est depuis plusieurs années introduite dans les entreprises pour s'adapter à la fluctuation de la demande. Le temps de travail annuel est limité, mais est variable de semaine en semaine. Nous avons appliqué cette annualisation du temps de travail dans une manufacture suisse. À chaque opérateur de machine est affecté un nombre de quarts de travail hebdomadaire à combler. Deux types d'horaires sont en vigueur dans cette entreprise, l'horaire minimal (ou temps régulier) et l'horaire maximal (temps supplémentaire). Dans la situation actuelle, les opérateurs respectent le même type

d'horaire, c'est-à-dire qu'ils sont tous sous horaire minimal (ou horaire maximal) en même temps. L'objectif de l'entreprise est de minimiser le nombre de semaines en horaire maximal. Nous avons introduit plusieurs flexibilités et relaxations de façon à toujours respecter les contraintes de l'entreprise mais en fournissant différentes solutions qui permettent par exemple de réduire le déséquilibre entre les charges des opérateurs. Parmi les scénarios étudiés nous retrouvons l'introduction du temps partiel et de l'engagement graduel. Nous avons aussi analysé l'impact de l'introduction des quarts de travail de huit heures. Nous avons formulé le problème sous forme de modèle linéaire mixte que nous avons résolu avec CPLEX.

Que ce soit dans le secteur des services de soins à domicile ou dans l'industrie de la fabrication, l'introduction des différentes formes de flexibilité a montré un gain substantiel dans l'équilibre de la charge de travail. Cet équilibre est important à atteindre surtout dans les milieux où la demande est en hausse sans être talonnée par l'augmentation des ressources.

La charge de travail est un sujet particulièrement sensible dans le milieu de la santé. Nous avons dans cette thèse étudié le cas particulier des infirmières gestionnaires de cas dans le milieu des soins à domicile. Une extension de nos travaux à l'équilibrage de la charge de travail au niveau des tournées pour rendre les visites nous semble particulièrement intéressante. En effet dans certains CLSC, par exemple ceux en milieu rural, les infirmières ont une composante déplacement bien plus importante qu'en milieu urbain, donnant alors autant de poids à la gestion des cas qu'aux visites. D'autres problèmatiques telles que la confection des horaires et des tournées, revêtent alors plus d'importance. Ces différents problèmes de recherche opérationnelle sont donc toujours d'actualité et montrent l'importance de cette discipline dans la gestion des systèmes de santé dans la mesure où elle constitue un outil efficace.

BIBLIOGRAPHIE

[1] «Basic statistics about home care». National Association for Home Care and Hospice, United States, 2003.

[2] «Centers for medicare & medicaid services». U.S. Department of Health & Human Services, 2006.

[3] «Commission sur l'avenir des soins de santé au canada». Santé Canada, 2003.

[4] «Dépenses du secteur public et utilisation des services à domicile : Examen des données». Institut canadien d'information sur la santé, 2007.

[5] «Événements économiques clés». Gouvernement du Canada, 2003.

[6] «Impacts du guichet unique en CLSC : Utilisation des services et vécu des intervenants». René-Cassin, Institut de gérontologie sociale du Québec, 2000.

[7] «La composition du personnel infirmier : Une recherche documentaire». Association des infirmières et infirmiers du Canada, 2005.

[8] «La gestion de cas à domicile - Table ronde sur invitation, le 3 à 4 mars 2005 : Résumé des déliberations». Association Canadienne de Soins et Services à Domicile, Santé Canada, 2005.

[9] «Le centre de santé et de services sociaux : Au cœur de la nouvelle organisation des services». Agence de développement de réseaux locaux de services de santé et de services sociaux, 2004.

[10] «L'emploi dans l'industrie entre 1992 et 2002 : Le poids croissant de l'intérim». In *Premières Synthèses*, Ministère du Travail des Relations Sociales et de la Solidarité, DARES, 2005.

[11] «Local services network». Ministère de la Santé et des Services Sociaux, 2004.

[12] «Loi canadienne sur la santé». Santé Canada, 2003.

[13] «Macro-level health spending». Canadian Institute for Health Information, 2003.

[14] «Nurse staffing and quality of patient care». Agency for Healthcare Research and Quality, 2007.

[15] «Nursing staff mix : A literature review». Canadian Nurses Association, 2004.

[16] «Portrait de la région de Montréal». Agence de développement de réseaux locaux de services de santé et de services sociaux, 2004.

[17] «Portraits des soins à domicile 2003». Association Canadienne de Soins et Services à Domicile, Santé Canada, 2003.

[18] «Projet clinique : Cadre de référence pour les réseaux locaux de services de santé et de services sociaux». La direction des communications du ministère de la santé et des services sociaux, 2004.

[19] «Proposition de loi relative à l'hospitalisation à domicile». Session extraordinaire, Sénat de Belgique, 2007.

[20] «Provincial and territorial home care programs : A synthesis for canada». Health Canada, 1999.

[21] «Rapport d'analyse - Organisation des services de soins à domicile aux personnes âgées en perte d'autonomie». CLSC Verdun, 1998.

[22] «Réorganisation des services de soins infirmiers des services de soutien à domicile - Critères de priorisation des demandes de soins infirmiers». CLSC Ahunstic, 2001.

[23] «Statistique sur les ressources institutionnelles». Ministère de la Santé et des Services Sociaux, 2003.

[24] «Statistiques relatives à l'industrie canadienne». Industrie Canada, 2004.

[25] «Évaluation de la charge de travail des infirmières des soins à domicile». CLSC LaSalle, 2001.

[26] ALFARES, H.K. (2000). Dual-based optimization of cyclic four-day workweek scheduling, *IMA Journal of Mathematics Applied in Business and Industry*, vol. 11, pp. 269–283.

[27] ALFARES, H.K. (2003). Four-day workweek scheduling with two or three consecutive days off, *Journal of Mathematical Modeling and Algorithms*, vol. 2, pp. 67–80.

[28] AMINI, M.M. et RACER, M. (1994). A rigorous computational comparison of alternative solution methods for the generalized assignment problem, *Management Science*, vol. 40, pp. 868–890.

[29] AMINI, M.M. et RACER, M. (1995). A hybrid heuristic for the generalized assignment problem, *European Journal of Operational Research*, vol. 87, pp. 343–348.

[30] ARTHUR, T. et JAMES, N. (1994). Determining nurse staffing levels : A critical review of the literature, *Journal of Advanced Nursing*, vol. 19, pp. 558–565.

[31] AZAIEZ, M.N. et AL SHARIF, S.S. (2005). A $0-1$ goal programming approach for nurse scheduling, *Computers & Operations Research*, vol. 32, pp. 491–507.

[32] AZMAT, C.S., HÜRLIMANN, T. et WIDMER, M. (2004). Mixed integer programming to schedule a single-shift workforce under annualized hours, *Annals of Operations Research*, vol. 128, pp. 199–215.

[33] AZMAT, C.S. et WIDMER, M. (2004). A case study of single shift planning and scheduling under annualized hours : A simple three-step approach, *European Journal of Operational Research*, vol. 153, pp. 148–175.

[34] BARD, J.F. et PURNOMO, H.W. (2005). Hospital-wide reactive scheduling of nurses with preference considerations, *IIE Transactions*, vol. 37, pp. 589–608.

[35] BARD, J.F. et PURNOMO, H.W. (2007). Cyclic preference scheduling of nurses using a lagrangien-based heuristic, *Journal of Scheduling*, vol. 10, pp. 5–23.

[36] BEAUGHON, M.C. (1999). «La méthode SIIPS indicateurs d'activité en soins infirmiers», Lamarre, Paris.

[37] BEAUMONT, N. (1997). Scheduling staff using mixed integer programming, *European Journal of Operational Research*, vol. 98, pp. 473–484.

[38] BEGUR, S.V., MILLER, D.M. et WEAVER, J.R. (1997). An integrated spatial DSS for scheduling and routing home-health-care nurses, *Interfaces*, vol. 27, pp. 35–48.

[39] BERTELS, S. et FAHLE, T. (2006). A hybrid setup for a hybrid scenario : Combining heuristics for the home health care problem, *Computers & Operations Research*, vol. 33, pp. 2866–2890.

[40] BÉRUBÉ, J.-F., GENDREAU, M. et POTVIN, J.-Y. (2006). A branch-and-cut algorithm for the undirected prize collecting traveling salesman problem. Cahier du CRT 2006-30, Centre de recherche sur les transports, Montreal, Canada.

[41] BILLIONNET, A. (1999). Integer programming to schedule a hierarchical workforce with variable demands, *European Journal of Operational Research*, vol. 114, pp. 105–114.

[42] BLAIS, M., LAPIERRE, S.D. et LAPORTE, G. (2003). Solving a home care districting problem in an urban setting, *Journal of the Operational Research Society*, vol. 54, pp. 1141–1147.

[43] BLAKE, J.T. et CARTER, M.W. (2002). A goal programmimg approach to strategic resource allocation in acute care hospitals, *European Journal of Operational Research*, vol. 140, pp. 541–561.

[44] BOST, D. et LAWLER, T. (1989). Measuring nursing resource consumption, *Nursing Management*, vol. 20, pp. 34–36.

[45] BOUCLIN, S., COURSOL, A-A., GLOUTNAY, F. et LAURIER, L.C. (2002). Mesure de la charge de travail en soins infirmiers. CLSC du Vieux Lachine.

[46] BOUVIER, L (2002). Bilan mi-parcours : Étude de la charge de cas en soins infirmiers. CLSC Côte-des-Neiges.

[47] BURKE, E.K., DE CAUSMAECKER, P., VANDEN BERGHE, G. et VAN LANDEGEM, H. (2004). The state of the art of nurse rostering, *Journal of Scheduling*, vol. 7, pp. 441–499.

[48] BURNS, R.N. et CARTER, M.W. (1985). Work force size and single shift schedules with variable demands, *Management Science*, vol. 31, pp. 599–607.

[49] BURNS, R.N. et KOOP, G.J. (1987). A modular approach to optimal multiple-shift manpower scheduling, *Operations Research*, vol. 35, pp. 100–110.

[50] BURNS, R.N. et NAMIMHAN, R. (1994). 10-Hour multiple shift scheduling. Working paper 94–36, Queen's School of Business.

[51] BURNS, R.N. et NAMIMHAN, R. (1995). 8-Hour multiple shift scheduling of hierarchical workforce. Working paper 95–02, Queen's School of Business.

[52] BURNS, R.N. et NARASIMHAN, R. (1999). Multiple shift scheduling of workforce on four-day workweeks, *Journal of the Operational Research Society*, vol. 50, pp. 979–981.

[53] BURNS, R.N., NARASIMHAN, R., et SMITH, L.D. (1995). An algorithm for scheduling a single-category workforce on four-day work weeks. Working paper 95–11, Queen's School of Business.

[54] BURNS, R.N., NARASIMHAN, R. et SMITH, L.D. (1998). A set processing algorithm for scheduling staff on 4-day or 3-day work weeks, *Naval Research Logistics*, vol. 45, pp. 839–853.

[55] CARIO, M.C., CLIFFORD, J.J., HILL, R.R., YANG, J., YANG, K. et REILLY, C.H. (2002). An investigation of the relationship between problem characteristics and algorithm performance : A case study of the generalized assignment problems, *IIE Transactions on Operations Engineering*, vol. 34, pp. 297–312.

[56] CARPANETO, G. et TOTH, P. (1981). Algorithm for the solution of the bottleneck assignment problem, *Computing*, vol. 27, pp. 179–187.

[57] CARPENTER, G.I., IKAGAMI, N., LJUNGGREN, G. et AL., (1997). RUG-III and resource allocation : Comparing the relationship of direct care time with patient characteristics in five countries, *Age and Ageing*, vol. 26, pp. 61–65.

[58] CATTERSON, J. (1988). How busy are you ?, *Nursing Times*, vol. 11, pp. 28–31.

[59] CATTRYSSE, D.G., DEGRAEVE, Z. et TISTAERT, J. (1998). Solving the generalized assignment problem using polyhedral results, *European Journal of Operational Research*, vol. 108, pp. 618–628.

[60] CATTRYSSE, D.G., SALOMON, M. et VAN WASSENHOVE, L.N. (1994). A set partitioning heuristic for the generalized assignment problem, *European Journal of Operational Research*, vol. 72, pp. 167–174.

[61] CATTRYSSE, D.G. et VAN WASSENHOVE, L.N. (1992). A survey of algorithms for the generalized assignment problem, *European Journal of Operational Research*, vol. 60, pp. 260–272.

[62] CHU, P.C. et BEASLEY, J.E. (1997). A genetic algorithm for the generalized assignment problem, *Computers & Operations Research*, vol. 24, pp. 17–23.

[63] CHU, S., HUANG, T. et LEE, L. (2004). Improving workload balance and code optimization on processor-in-memory systems, *The Journal of Systems and Software*, vol. 71, pp. 71–82.

[64] COHEN, M.M., O'BRIEN-PALLAS, L.L. , COPPLESTONE, C., WALL, R., PORTER, J. et ROSE, D.K. (1999). Nursing workload associated with adverse events in the postanesthesia care unit, *Anesthesiology*, vol. 91, pp. 1882–1890.

[65] COLETTE, N. (2001). Bilan de l'expérimentation de l'outil ILOT concernant l'analyse de la charge de cas des infirmières aux services de soutien à domicile et recommandations déposées au comité soins infirmiers - Document de travail. CLSC Ahunstic.

[66] COLETTE, N. (1999). Guide d'analyse de la charge de cas et de la charge de travail. In *Indicateurs de la charge de cas (I.C.C.) et de la charge de travail (I.C.T.) universels et spécifiques aux intervenants sociaux, infirmières, ergothérapeutes, physiothérapeutes et nutritionniste des services aux personnes en perte d'autonomie au CLSC Verdun/ Côte Saint-Paul*, CLSC Verdun/ Côte Saint-Paul.

[67] COROMINAS, A., LUSA, A. et PASTOR, R. (2002). Using MILP to plan annualised working hours, *European Journal of Operational Research*, vol. 53, pp. 1101–1108.

[68] COROMINAS, A., LUSA, A. et PASTOR, R. (2004). Characteristics and classification of the annualised working hours planning problems, *International Journal of Services Technology and Management*, vol. 5, pp. 435–447.

[69] COROMINAS, A., LUSA, A. et PASTOR, R. (2004). Planning annualised hours with a finite set of weekly working hours and joint holidays, *Annals of Operations Research*, vol. 128, pp. 217–233.

[70] COROMINAS, A., LUSA, A. et PASTOR, R. (2005). Annualised hours : A real flexibility tool, *OR Insight*, vol. 18, pp. 10–14.

[71] COROMINAS, A., LUSA, A. et PASTOR, R.(2007). Planning annualised hours with a finite set of weekly working hours and cross-trained workers, *European Journal of Operational Research*, vol. 176, pp. 230–239.

[72] COROMINAS, A., LUSA, A. et PASTOR, R. (2007). Planning production and working time within an annualised hours scheme framework, *Annals of Operations Research*, vol. 155, pp. 5–23.

[73] COROMINAS, A., LUSA, A. et PASTOR, R. (2007). Using a MILP model to establish a framework for an annualised hours agreement, *European Journal of Operational Research*, vol. 177, pp. 1495–1506.

[74] DE VRIES, G. (1987). Nursing workload measurement as management information, *European Journal of Operational Research*, vol. 29, pp. 199–208.

[75] DEGROOT, H.A. (1994). Patient classification systems and staffing - Part 1, problems and promise, *Journal of Nursing Administration*, vol. 24, pp. 43–51.

[76] DEGROOT, H.A. (1994). Patient classification systems and staffing - Part 2, practice and process, *Journal of Nursing Administration*, vol. 24, pp. 17–23.

[77] DELLEFIELD, M.E. (2006). Using the resource utilization groups (RUG-III) system as a staffing tool in nursing homes, *Geriatric Nursing*, vol. 27, pp. 160–165.

[78] DESROSIERS, J., BRAVO, G., HÉBERT, R. et DUBUC, N. (1995). Reliability of the revised functional autonomy measurement system (SMAF) for epidemiological research, *Age and Ageing*, vol. 24, pp. 402–406.

[79] DIAZ, J.A. et FERNANDEZ, E. (2001). A tabu search heuristic for the generalized assignment problem, *European Journal of Operational Research*, vol. 132, pp. 22–38.

[80] DOUGUET, F. et MUNÕZ, J. (2005). Les effets de l'accréditation et des mesures d'amélioration sur la qualité des soins sur l'activité des personne soignants. Ministère de l'emploi, de la cohésion sociale et du logement, Ministère de la santé et des solidarités.

[81] DRENNAN, V. (1990). Striving for fairer workloads, *Nursing Times*, vol. 10, pp. 12–14.

[82] DUBOYS FRESNEY, C. (1993). Les indicateurs infirmiers, *Informatique et Santé*, vol. 6, pp. 175–184.

[83] DUBUC, N. (1999). Système de classification basé sur le profil d'autonomie fonctionnelle. In *Autonomie et Vieillissement*, pp. 86–101.

[84] DUMAY, M.F., MISSET, B., OLIVA, F., VENAILLE, B. et CARLET, J. (1999). The SIIPS scoring system does not measure nursing care workload in the intensive care unit : A comparison with PRN 87, *Réanimation Urgences*, vol. 8, pp. 639–644.

[85] EDWARDSON, S.R. et GIOVANNETI, P.B. (1994). Nursing workload measurement systems, *Annual Review of Nursing Research*, vol. 12, pp. 95–123.

[86] EMMONS, H. et BURNS, R.N. (1991). Off-day scheduling with hierarchical worker categories, *Operations Research*, vol. 39, pp. 484–495.

[87] ERNST, A.T., JIANG, H., KRISHNAMOORTHY, M. et SIER, D. (2004). Staff scheduling and rostering : A review of applications, methods and models, *European Journal of Operational Research*, vol. 153, pp. 3–27.

[88] EVEBORN, P., FLISBERG, P. et RONNQVIST, M. (2006). LAPS CARE - An operational system for staff planning of home care, *European Journal of Operational Research*, vol. 171, pp. 962–976.

[89] EXCHAQUET, N. et ZUBLIN, L. (1980). «Effectif et organisation pour des soins infirmiers adéquats». Centurion, Paris.

[90] FARNHAM, J.A., MAEZ-RAUZI, V. et CONWAY, K. (1992). Balancing assignments : A patient classification system for a step-down unit, *Nursing Management*, vol. 23, pp. 49–54.

[91] FELTL, H. et RAIDL, G.R. (2004). An improved hybrid genetic algorithm for the generalized assignment problem. In *Proceedings of the 2004 ACM Symposium on Applied Computing*.

[92] FERLAND, J.A., BERRADA, I., NABLI, I., AHIOD, B., MICHELON, P., GASCON, V. et GAGNÉ, É. (2001). Generalized assignment type goal programming problem : Application to nurse scheduling, *Journal of Heuristics*, vol. 7, pp. 391–413.

[93] FERLAND, J.A., HERTZ, A. et LAVOIE, A. (1996). An object-oriented methodology for solving assignment-type problems with neighborhood search techniques, *Operations Research*, vol. 44, pp. 347–359.

[94] FISHER, M.L. et JAIKUMAR, R. (1981). A generalized assignment heuristic for vehicle routing, *Networks*, vol. 11, pp. 109–124.

[95] FRIES, B.E., SCHNEIDER, D.P., FOLEY, W.J., GAVAZZI, M., BURKE, R. et CORNELIUS, E. (1994). Refining a casemix measure for nursing homes : Resource utilization groups (RUG-III), *Medical Care*, vol. 32, pp. 668–685.

[96] GARNIER, G., CALLEJA, M. et LESTEL, A.-M. (2006). Les SIIPS, un outil de management qualitatif et quantitatif au quotidien des pôles. L'expérience du centre hospitalier du Val d'Ariège, *Gestions Hospitalières*, vol. 459.

[97] GAVISH, B. et PIRKUL, H. (1985). Efficient algorithms for solving multiconstraint generalized assignment problem. Working paper QM 8523, University of Rochester.

[98] GAVISH, B. et PIRKUL, H. (1991). Algorithms for the multi-resource generalized assignment problem, *Management Science*, vol. 37, pp. 695–713.

[99] GERMAIN, P.-L. (2007). La flexibilité du temps de travail encore peu répandue en europe, *Journal du Net*.

[100] GIANNOPOULOS, D. (1998). Impacts du guichet unique en CLSC : Utilisation des services, coûts et vécu des intervenant, *Bulletin Bien-vieillir*.

[101] GLOVER, F. (1986). Future paths for integer programming and links to artificial intelligence, *Computers & Operations Research*, vol. 13, pp. 533–549.

[102] GLOVER, F. et LAGUNA, M. (1997). «Tabu search». Kluwer, Boston.

[103] GRÉGOIRE, M. et LAPIERRE, M. (2002). Orientation des balises en soins infirmiers au soutien à domicile - Document de travail. CLSC Ahunstic.

[104] GUERRO, F., LOZANO, S. et KOLTAI, T.J. (1999). Machine loading and part type selection in flexible manufacturing systems, *International Journal of Production Research*, vol. 37, pp. 1303–1317.

[105] HADDADI, S. (1999). Lagrangian decomposition based heuristic for the generalized assignment problem, *INFOR*, vol. 37, pp. 392–402.

[106] HADDADI, S. et OUZIA, H. (2001). An effective lagrangian heuristic for the generalized assignment problem, *INFOR*, vol. 39, pp. 351–356.

[107] HAJRI-GABOUJ, S. (2003). A fuzzy genetic multiobjective optimization algorithm for a multilevel generalized assignment problem, *IEEE Transactions on Systems, Man, and Cybernetics - Part C : Applications and Reviews*, vol. 33, pp. 214–224.

[108] HÉBERT, R., CARRIER, R. et BILODEAU, A. (1988). The functional autonomy measurement system (SMAF) : Description and validation of an instrument for the measurement of handicaps, *Age and Ageing*, vol. 17, pp. 293–302.

[109] HÉBERT, R., DESROSIERS, J., DUBUC, N., TOUSIGNANT, M., GUILBEAULT, G. et PINSONNAULT, E. (1988). Le système de mesure de l'autonomie fonctionnelle (SMAF), *La Revue de Gériatrie*, vol. 28, pp. 323–336.

[110] HÉBERT, R. et DUBUC, N. (2001). Les profils ISO-SMAF : Un système de gestion clinico-administratif pour la planification des services de longue durée dans un système de soins intégrés. Centre d'expertise en santé de Sherbrooke.

[111] HÉBERT, R., GUILBAULT, J., DESROSIERS, J. et DUBUC, N. (1999). Le système de mesure de l'autonomie fonctionnelle (SMAF). In *Autonomie et Vieillissement*, pp. 86–101.

[112] HÉBERT, R., GUILBAULT, J., DESROSIERS, J. et DUBUC, N. (2001). The functional autonomy measurement system (SMAF) : A clinical-based instrument for measuring disabilities and handicaps in older people, *Geriatrics Today : Journal of the Canadian Geriatrics Society*, vol. 4, pp. 1–7.

[113] HERNANDEZ, C.A. et O'BRIEN-PALLAS, L.L. (1996). Validity and reliability of nursing workload measurement systems : Strategies for nursing administrators, *Canadian Journal of Nursing Administration,*, vol. 9, pp. 33–52.

[114] HIGGINS, A.J. (2001). A dynamic tabu search for large-scale generalized assignment problems, *Computers & Operations Research*, vol. 28, pp. 1039–1048.

[115] HOLLAND, J.H. (1975). «Adaptation in natural and artificial systems», *University of Michigan Press*.

[116] HOUSSON, Y. (2004). La flexibilité du travail favorise l'emploi, *L'Humanité*.

[117] HUANG, H.C., LEE, C. et XU, Z. (2006). The workload balancing problem at air cargo terminals, *OR Spectrum*, vol. 28, pp. 705–727.

[118] HUGHES, M. (1999). Nursing workload : An unquantifiable entity, *Journal of Nursing Management*, vol. 7, pp. 317–322.

[119] HUNG, R. (1991). A cyclical schedule of 10-hour, four-day workweeks, *Nursing Management*, vol. 22, pp. 30–33.

[120] HUNG, R. (1991). Single-shift workforce scheduling under a compressed workweek, *OMEGA*, vol. 19, pp. 494–497.

[121] HUNG, R. (1993). A three-day workweek multiple-shift scheduling model, *Journal of the Operational Research Society*, vol. 44, pp. 141–146.

[122] HUNG, R. (1994). Managing compressed workweeks : A comparison of 4-day and 3-4 day workweeks, *International Journal of Technology Management*, vol. 9, pp. 261–266.

[123] HUNG, R. (1994). Multiple-shift scheduling under the 3-4 workweek with different weekday and weekend labor requirements, *Management Science*, vol. 40, pp. 280–284.

[124] HUNG, R. (1994). A multiple-shift workforce scheduling model under the 4-day workweek with weekday and weekend labour demands, *Journal of the Operational Research Society*, vol. 45, pp. 1088–1092.

[125] HUNG, R. (1994). Single-shift off-day scheduling of a hierarchical workforce with variable demands, *European Journal of Operational Research*, vol. 78, pp. 49–57.

[126] HUNG, R. (1995). Compressed work schedules in a police force : A survey of applications, *Optimum*, vol. 26, pp. 32–36.

[127] HUNG, R. (1996). An annotated bibliography of compressed workweeks, *International Journal of Manpower*, vol. 17, pp. 43–53.

[128] HUNG, R. (1997). Scheduling for continuous operations : The Baylor plan, *International Journal of Materials and Product Management*, vol. 12, pp. 37–42.

[129] HUNG, R. (1997). Shiftwork scheduling with phase-delay feature, *International Journal of Production Research*, vol. 35, pp. 1961–1968.

[130] HUNG, R. (1999). A multiple-shift workforce scheduling model under annualized hours, *Naval Research Logistics*, vol. 46, pp. 726–736.

[131] HUNG, R. (1999). Scheduling a workforce under annualized hours, *International Journal of Production Research*, vol. 37, pp. 2419–2427.

[132] HUNG, R. et EMMONS, H. (1993). Multiple-shift workforce scheduling under the 3-4 compressed workweek with a hirearchical workforce, *IIE Transactions*, vol. 25, pp. 82–89.

[133] HURST, K. (2002). Selecting and applying methods for estimating the size and mix of nursing teams. Nuffield Institute for Health.

[134] IBARAKI, T., NONOBE, K. et YAGIURA, M. (2005). «Metaheuristics : Progress as real problem solvers». Springer, New York.

[135] JAUMARD, B., SEMET, F. et VOVOR, T. (1998). A generalized linear programming model for nurse scheduling, *European Journal of Operational Research*, vol. 107, pp. 1–18.

[136] JONKER, R. et VOLGENANT, A. (1995). A shortest augmenting path algorithm for dense and spare linear assignment problems, *European Journal of Operational Research*, vol. 83, pp. 621–638.

[137] KIM, C.W., TANCHOCO, J.M.A. et KOO, P.-H. (1999). AGV dispatching based on workload balancing, *International Journal of Production Research*, vol. 37, pp. 4053–4066.

[138] KIRKPATRICK, S., GELATT, C.D. et VECCHI, M.P. (1983). Optimizing by simulated annealing, *Science*, vol. 220, pp. 671–680.

[139] KOGAN, K., KHMELNITSKY, E. et IBARAKI, T. (2005). Dynamic generalized assignment problem with stochastic demands and multiple agent-task relationships, *Journal of Global Optimization*, vol. 31, pp. 17–43.

[140] KUMAR, N. et SHANKER, K. (2001). Comparing the effectiveness of workload balancing objectives in fms loading, *International Journal of Production Research*, vol. 39, pp. 843–871.

[141] KWAK, N.K. et LEE, C. (1997). A linear goal programming model for human resource allocation in a health-care organization, *Journal of Medical Systems*, vol. 21, pp. 129–140.

[142] LAGODIMAS, A.G. et LEOPOULOS, V. (2000). Greedy heuristic algorithms for manpower shift planning, *International Journal Production Economics*, vol. 68, pp. 95–106.

[143] LAGUNA, M., KELLY, J.P., GONZLEZ-VELARDE, J.L. et GLOVER, F. (1995). Tabu search for the multilevel generalized assignment problem., *European Journal of Operational Research*, vol. 82, pp. 176–189.

[144] LAHRICHI, N., LAPIERRE, S.D., HERTZ, A., TALIB, A. et BOUVIER, L. (2006). Analysis of a territorial approach to the delivery of nursing home care services based on historical data, *Journal of Medical Systems*, vol. 30, pp. 283–291.

[145] LANDRY, G. (1998). Rapport d'analyse : Organisation des services de soins à domicile aux services aux personnes en perte d'autonomie. CLSC Verdun.

[146] LAPORTE, G. et PESANT, G. (2004). A general multi-shift scheduling system, *Journal of the Operational Research Society*, vol. 55, pp. 1208–1217.

[147] LARIVIÈRE, C., HOCKENSTEIN, E. et MONTEJO, M.-E. (2001). Guichet unique en maintien à domicile : L'expérimentation par sept CLSC montréalais. Institut de Gérontologie Sociale du Québec.

[148] LAU, H.C. (1996). On the complexity of manpower shift scheduling, *Computers & Operations Research*, vol. 23, pp. 93–102.

[149] LORENA, L.A.N. et NARCISO, M.G. (1996). Relaxation heuristics for a generalized assignment problem, *European Journal of Operational Research*, vol. 91, pp. 600–610.

[150] MAGGI, S. et STEEL, K. (1996). Planning home care services : Which data do we need ?, *Archives of Gerontology and Geriatrics*, vol. 5, pp. 493–498.

[151] MALLOCH, K. et CONOVALOFF, A.J. (1999). Patient classification systems - Part 1 : The third generation, *Journal of Nursing Administration*, vol. 29, pp. 49–56.

[152] MALLOCH, K., NEELD, A, MCMURRAY, C, MEEKS, L., WALLACH, M., WILLIAMS, S. et CONOVALOFF, A.J. (2002). Classification systems - Part 2 : The third generation, *Journal of Nursing Administration*, vol. 8, pp. 33–42.

[153] MAREK, K.D., POPEJOY, L., PETROSKI, G. et RANTZ, M. (2006). Nurse care coordination in community-based long-term care, *Journal of Nursing Scholarship*, vol. 38, pp. 80–86.

[154] MARTELLO, S. et TOTH, P. (1981). An algorithm for the generalized assignment problem. J. Brans, Ed., Operational research, IFORS, North-Holland, Amsterdam, pp. 589–603.

[155] MARTELLO, S. et TOTH, P. (1990). «Knapsack problems : Algorithms and computer implementations». Wiley, Chichester.

[156] MARTELLO, S. et TOTH, P. (1995). The bottleneck generalized assignment problem, *European Journal of Operational Research*, vol. 83, pp. 621–638.

[157] MAZZOLA, J.B. (1989). Generalized assignement with nonlinear capacity interaction, *Management Science*, vol. 35, pp. 923–941.

[158] MAZZOLA, J.B. et WILCOX, S.P. (2001). Heuristics for the multi-resource generalized assignment problem, *Naval Research Logistics*, vol. 48, pp. 468–483.

[159] MUELLER, C. (2000). The RUG-III case mix classification system for long-term care nursing facilities : is it adequate for nurse staffing ?, *Journal of Nursing Administration*, vol. 30, pp. 535–543.

[160] MULLINAX, C. et LAWLEY, M. (2002). Assigning patient to nurses in neonatal intensive care, *The Journal of the Operational Research Society*, vol. 53, pp. 25–35.

[161] NAMIMHAN, R., (1995). An algorithm for single shift scheduling of hierarchical workforce. Working paper 95–04, Queen's School of Business.

[162] NAMIMHAN, R. (2000). An algorithm for multiple shift shceduling of hierarchical workforce on four days or three days workweek, *INFOR*, vol. 38, pp. 14–32.

[163] NAMIMHAN, R. et BURNS, R.N. (1994). 10-Hours single shift scheduling of hierarchical workforce. Working paper 94–37, Queen's School of Business.

[164] NAUSS, R.M. (2003). Solving the generalized assignment problem : An optimizing and heuristic approach, *INFORMS Journal of Computing*, vol. 15, pp. 249–266.

[165] NESTI, G., CAMPOSTRINI, S., GARBIN, S., PIVA, P., DI SANTO, P. et TUNZI, F. (2003). Providing integrated health and social care for older persons in Italy. PROCARE.

[166] O'BRIEN, A.J., ABAS, M., CHRISTENSEN, J., NICHOLLS, T.H., LE PROU, T., HEKAU, A. et VANDERPYL, J. (2002). Nursing workload measurement in acute mental health

inpatient units. In *A report for the mental health research and development strategy*, Health research council of New-Zealand, Mental Health Publications.

[167] O'BRIEN-PALLAS, L., COCKERILL, R. et LEATT, P. (1992). Different systems, different costs ? An examination of the comparability of workload measurement systems, *Journal of Nursing Administration*, vol. 22, pp. 17–22.

[168] O'BRIEN-PALLAS, L., IRVINE, D., PEEREBOOM, E. et MURRAY, M. (1997). Measuring nursing workload : Understanding the variability, *Nursing Economics*, vol. 15, pp. 171–82.

[169] OSMAN, I.H. (1995). Heuristics for the generalized assignment problem : Simulated annealing and tabu search approaches, *OR Spektrum*, vol. 17, pp. 211–225.

[170] PERROTT SIFERD, S. et BENTON, W.C. (1992). Workforce staffing and scheduling : Hospital nursing specific models, *European Journal of Operational Research*, vol. 60, pp. 233–246.

[171] PHILLIPS, C.Y., CASTORR, A., PRESCOTT, P.A. et AL. (1992). Nursing intensity. Going beyond patient classification, *Journal of Nursing Administration*, vol. 22, pp. 46–52.

[172] PINAR, A. (2005). Improving load balance with flexibly assignable tasks, *IEEE Transactions on Parallel and Distributed Systems*, vol. 16, pp. 956–965.

[173] PINSONNAULT, E., DESROSIERS, J., DUBUC, N., KALFAT, H., COLVEZ, A. et DELLI-COLLI, N. (2003). Functional autonomy measurement system : Development of a social subscale, *Archives of Gerontology and Geriatrics*, vol. 37, pp. 223–233.

[174] POLVERINI, F., PRINCIPI, A., BALDUCCI, C., MELCHIORRE, G., QUATTRINI, M.S., GIANELLI, V. et LAMURA, G. (2004). EUROFAMCARE - national background report for Italy. Istituto Nazionale di Riposo e Cura Anziani.

[175] POTTS, C.N. et WHITEHEAD, J.D. (2001). Workload balancing and loop layout in the design of a flexible manufacturing system, *European Journal of Operational Research*, vol. 129, pp. 326–336.

[176] POWERS, J.L. (1993). Accepting and refusing assignments, *Nursing Management*, vol. 24, pp. 64–73.

[177] PRIVAULT, C. et HÉRAULT, L. (1998). Constraints satisfaction through recursive neural networks with mixed penalties : A case study, *Neural Processing Letters*, vol. 8, pp. 15–26.

[178] PRIVAULT, C. et HÉRAULT, L. (1998). Solving a real world assignment problem with a metaheuristic, *Journal of Heuristics*, vol. 4, pp. 383–398.

[179] RAINIO, A.-K. et OHINMAA, A.E. (2005). Assessment of nursing management and utilization of nursing resources with the RAFAELA patient classification system - Case study from the general wards of one central hospital, *Journal of Clinical Nursing*, vol. 14, pp. 674–684.

[180] RIEZEBOS, J., KORTE, G.J. et LAND, M.J. (2003). Improving a practical buffering approach using workload control, *International Journal of Production Research*, vol. 41, pp. 699–712.

[181] ROMEIJN, H.E. et MORALES, D.M. (2000). A class of greedy algorithms for the generalized assignment problem, *Discrete Applied Mathematics*, vol. 103, pp. 209–235.

[182] ROMEIJN, H.E. et PIERSMA, N. (2000). A probabilistic feasibility and value analysis of the generalized assignment problem, *Journal of Combinatorial Optimization*, vol. 4, pp. 325–355.

[183] RUGGIERO, J.S. (2005). Health, work variables, and job satisfaction among nurses, *Journal of Nursing Administration*, vol. 35, pp. 254–263.

[184] SHTUB, A. et KOGAN, K. (1998). Capacity planning by the dynamic multi-resource generalized assignment problem (DMRGAP), *European Journal of Operational Research*, vol. 105, pp. 91–99.

[185] SIOUFFI, P., MALLET, M., PASCAL, L., DURIF, L., HARLE, J.-R., SAMBUC, R., BONGRAND, M.-C. et TIMON-DAVID, P. (2001). Évaluation de la charge en soins en hospitalisation à domicile, *Gestions Hospitalières*, vol. 402.

[186] SIQUEIRA, P.H., SCHEER, S. et STEINER, M.T.A. (2005). Application of the "winner takes all" principle in Wang's recurrent neural network for the assignment problem. In *Advances in neural networks - ISNN 2005, PT 1, Lecture Notes in Computer Science 3496*, pp. 731–738.

[187] TARIGHT, N., HUGUENET, D., COUTUREAU, F., MESTRAND, M.-J. et NOIRÉ, M.-G. (2003). Indicateur SIIPS et PMSI en court séjour, *Gestions Hospitalières*, vol. 427.

[188] TILQUIN, C (2003). «Stratégies et outils de mesure des charges en soins infirmiers».

[189] TILQUIN, C. et AL. (1978). Determining nursing team size and composition, *Dimensions in Health Service*, vol. 55, pp. 12–16.

[190] TOKTAS, B., YEN, J.W. et ZABINSKY, Z.B. (2006). Addressing capacity uncertainty in resource-constrained assignment problems, *Computers & Operations Research*, vol. 33, pp. 724–745.

[191] TOUSIGNANT, M., HÉBERT, R., DUBUC, N., SIMONEAU, F. et DIELEMAN, L. (2004). Application de la classification des profils ISO-SMAF à la détermination des profils d'accueil et à la répartition budgétaire à l'intérieur de la méthode d'équité - volet hébergement de la région de l'Estrie. Centre de recherche sur le vieillissement.

[192] TRICK, M.A. (1992). A linear relaxation heuristic for the generalized assignment problem, *Naval Research Logistics*, vol. 39, pp. 137–151.

[193] VILA, G.F.E. et ASTORINO, J.M. (2001). Annualized hours as a capacity planning tool in make-to-order or assemble-to-order environment : An agricultural implements company case, *Production Planning & Control*, vol. 12, pp. 388–398.

[194] WALTS, M.L. et KAPADIA, A.S. (1996). Patient classification system : An optimization approach, *Health Care Management Review*, vol. 21, pp. 75–82.

[195] WANG, J. (1992). Analog neural network for solving the assignment problem, *Electronics Letters*, vol. 28, pp. 1047–1050.

[196] WEIGEL, D. et CAO, B. (1999). Applying GIS and OR techniques to solve Sears technician-dispatching and home-delivery problems, *Interfaces*, vol. 29, pp. 112–130.

[197] WIDMER, M (1997). Flexible timetabling : A solution against the irregularity of the demand ?. In *EURO XV Conference, Barcelona, Spain.*

[198] WILSON, J.M. (1997). A genetic algorithm for the generalised assignment problem, *Journal of the Operational Research Society*, vol. 48, pp. 804–809.

[199] YAGIURA, M., IBARAKI, T. et GLOVER, F. (2004). An ejection chain approach for the generalized assignment problem, *INFORMS Journal of Computing*, vol. 16, pp. 133–151.

[200] YAGIURA, M., IBARAKI, T. et GLOVER, F. (2006). A path relinking approach with ejection chains for the generalized assignement problem, *European Journal of Operational Research*, vol. 169, pp. 548–569.

[201] YAGIURA, M., IWASAKI, S. et IBARAKI, T. (2004). A very large-scale neighborhood search algorithm for the multi-resource generalized assignment problem, *Discrete Optimization*, vol. 1, pp. 87–98.

[202] YEH, J.-Y. et LIN, W.-S. (2007). Using simulation technique and genetic algorithm to improve the quality care of a hospital emergency department, *Expert Systems with Applications*, vol. 32, pp. 1073–1083.

[203] ZHUANG, Y., LIANG, T.Y., SHIEH, C., LEE, J. et YANG, L. (2004). A group-based load balance scheme for software distributed shared memory systems, *The Journal of Supercomputing*, vol. 28, pp. 295–309.