



HAL
open science

**Modélisation et optimisation bi-objectif et multi-période
avec anticipation d'une place de marché de prospects
Internet : adéquation offre/demande**

Manel Maamar

► **To cite this version:**

Manel Maamar. Modélisation et optimisation bi-objectif et multi-période avec anticipation d'une place de marché de prospects Internet : adéquation offre/demande. Autre. Université Paris Saclay (COmUE), 2015. Français. NNT : 2015SACL025 . tel-01275017

HAL Id: tel-01275017

<https://theses.hal.science/tel-01275017>

Submitted on 16 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT : 2015SACL025

Place des
Leads
Affiliation, lead generation & monetisation

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PARIS-SACLAY
PRÉPARÉE À
"CENTRALESUPÉLEC"

ÉCOLE DOCTORALE N° 573
INTERFACES: approches interdisciplinaires / fondements, applications et innovation

Spécialité de doctorat : INFORMATIQUE

Par

Manel MAAMAR

**Modélisation et optimisation bi-objectif et multi-période avec
anticipation d'une place de marché de prospects Internet :
Adéquation offre/demande**

Thèse présentée et soutenue à « CentraleSupélec », le « 7 Décembre 2015 » :

Composition du Jury :

M. Michel MINOUX, Université Paris-6	Président
M. Roberto WOLFLER CALVO, Université Paris-13	Rapporteur
Mme. Safia KEDAD-SIDHOUM, Université Paris-6	Rapporteur
M. Oualid JOUINI, CentraleSupélec	Examineur
M. Vincent MOUSSEAU, CentraleSupélec	Directeur de thèse
Mme. Wassila OUERDANE, CentraleSupélec	Co-encadrante de thèse
M. Alexandre AUBRY, Place des leads	Co-encadrant de thèse
M. Ralph RUIMY, Place des leads	Invité

Remerciements

Il me sera très difficile de remercier tout le monde car c'est grâce à l'aide de nombreuses personnes que j'ai pu mener cette thèse à son terme. En effet, je n'aurais jamais pu réaliser ce travail doctoral sans le soutien d'un grand nombre de personnes dont la générosité, la bonne humeur et l'intérêt manifestés à l'égard de ma recherche m'ont permis de progresser pendant ces trois années de thèse.

Je voudrais tout d'abord remercier grandement mon directeur de thèse, monsieur Vincent Mousseau, pour la confiance qu'il m'a accordée en acceptant d'encadrer ce travail doctoral et pour son aide tout au long de cette recherche. Je suis ravie d'avoir travaillé en sa compagnie car outre son appui scientifique, il a toujours été là pour me soutenir et me conseiller au cours de l'élaboration de cette thèse.

Ce travail doctoral doit également beaucoup à madame Wassila Ouerdane, co-encadrante de cette thèse, qui a mis ses compétences et son écoute au service de cette recherche. Je la remercie également pour son accueil chaleureux à chaque fois que j'ai sollicité son aide, ainsi que pour ses multiples encouragements.

Mes remerciements vont également à mon tuteur à "Place des Leads", monsieur Alexandre Aubry, pour m'avoir donné l'opportunité de réaliser ma thèse en entreprise, pour toutes les heures qu'il a consacrées afin de pouvoir réaliser les tests sur le serveur de l'entreprise et pour toute son aide malgré ses nombreuses occupations. Je remercie également monsieur Ralph Ruimy, directeur de l'entreprise, pour le financement de cette thèse et son implication tout au

long du travail.

Ces remerciements seraient incomplets si je n'en adressais pas à l'ensemble des membres du Laboratoire Génie Industriel (LGI) à CentraleSupélec pour leur soutien logistique et moral ainsi que pour la très bonne ambiance. Je souhaite aussi exprimer ma gratitude à l'ensemble des membres de "Place des Leads" qui m'ont beaucoup facilité l'intégration au sein de l'entreprise. Ma reconnaissance va à ceux qui ont plus particulièrement assuré le soutien affectif de ce travail doctoral : ma famille ainsi que mes amis.

Enfin, un grand merci à toutes les personnes ayant participé, de près comme de loin, à la réalisation de ce travail.

Résumé

Le travail que nous présentons dans cette thèse porte sur le problème d'affectation dans une place de marché de prospects Internet. Plus précisément, ce travail a pour ambition de répondre à la problématique de l'adéquation de l'offre et de la demande, dans un contexte caractérisé par des flux continus faisant évoluer en temps réel l'ensemble des offres disponibles et les demandes à satisfaire. Pour ce faire, nous proposons dans un premier temps un modèle mono-période qui optimise le problème d'affectation à un instant donné et en considérant une seule période de temps. Ce modèle mono-période permet également la prise en compte instantanée des nouvelles offres et demandes et leur adéquation en temps réel. Deux objectifs sont à optimiser, à savoir : le chiffre d'affaires et la satisfaction des clients.

Par la suite nous proposons d'étendre ce modèle sur plusieurs périodes de temps futures afin de prendre en compte l'aspect temps réel de l'activité de la place de marché et donc le fait que des flux continus font évoluer en temps réel l'ensemble des offres et des demandes. L'objectif étant de tirer profit de la connaissance concernant cette évolution, par le biais de l'intégration d'un modèle de prévision dans un modèle d'optimisation multi-période. Ainsi, nous proposons un modèle d'optimisation multi-période permettant d'envisager à un instant donné des affectations sur plusieurs périodes de temps futures afin de réaliser les meilleures affectations possibles. Aussi, nous proposons un modèle de prévision des nouveaux flux tout en considérant les caractéristiques du modèle d'optimisation multi-période.

Construire un modèle de prévision nécessite de définir les données à prévoir avant d'envisager

toute méthode de prévision. En d'autres termes, nous devons choisir les paramètres du modèle de prévision, à savoir : les données historiques appropriées, le pas de temps de la prévision ainsi que l'horizon de la prévision. Le défi consiste donc à définir les paramètres du modèle de prévision qui conviendront au fonctionnement du modèle d'optimisation multi-période.

Par ailleurs, une des caractéristiques de la place de marché est la temporalité de son système. Ainsi, nous proposons un algorithme assurant l'aspect temps réel et donc le fait que les affectations s'effectuent toutes les minutes. L'algorithme que nous proposons fonctionne de manière continue à longueur de journée en optimisant, à chaque instant, l'adéquation offre/demande de prospects Internet. Cet algorithme permet également de considérer instantanément les flux continus de prospects Internet ainsi que la mise à jour régulière de la demande.

Enfin, pour mettre en évidence l'efficacité et les bénéfices que la place de marché peut en tirer par l'utilisation des modèles d'optimisation et de l'algorithme proposés, nous avons mené des tests et différentes expérimentations sur des données réelles de la place de marché de prospects Internet. Ces tests nous ont permis de valider nos travaux et d'évaluer la qualité des résultats obtenus.

L'objectif de ce travail est double, d'une part, donner un cadre solide et formel pour répondre à la problématique de la place de marché de prospects Internet. D'autre part, le cadre proposé devrait être aussi générique que possible afin de pouvoir résoudre les problèmes analogues à celui de la place de marché de prospects Internet.

Mots-clés : problème d'affectation, optimisation bi-objectif, optimisation multi-période, modèle de prévision.

Abstract

The work that we present in this thesis focuses on the assignment problem in a marketplace of Internet prospects. More precisely, this work aims to address the problem of matching offers and demands in a context characterized by a continuous flows. These latter evolve in real time the set of available offers and demands to satisfy. To do this, we propose initially a mono-period model which optimizes the assignment problem at a given instant and taking into account a single period of time while allowing the instantaneous consideration of new offers and demands and their adequacy in real time. This model considers two objectives to optimize, namely : maximization of turnover as well as clients satisfaction.

Thereafter, we propose to extend this model over several future time periods in order to take into account the real time aspect of the marketplace activity and so the fact that a continuous flows evolve in real time the set of offers and demands. The objective is to take advantage of knowledge about this evolution, through the integration of a forecasting model in a multi-period optimization model. Thus, we propose a multi-period optimization model for considering at a given instant assignments over several future time periods. Also, we propose a forecasting model for new flows while considering the characteristics of the multi-period optimization model.

Building a forecasting model requires defining the data before considering any forecasting method. In other words, we have to choose the parameters of the forecasting model, namely the appropriate historical data, the forecasting time step and the forecasting horizon. The challenge

is to define the parameters of the forecasting model which agree with the functioning the multi-period optimization model.

Furthermore, a feature of the marketplace is the temporality of its system. Thus, we propose an algorithm ensuring real-time aspect and so the fact that assignments are made every minute. The proposed algorithm works continuously all day long while optimizing every instant the offer/demand adequacy of Internet prospects and instantly considering the continuous flux of Internet prospects as well as the regular updating demand.

Finally, in order to show the efficiency and the benefits that the marketplace can reap by the use of the proposed models, we conducted tests and various experiments on real data. These tests have allowed us to validate the proposed models and evaluate the quality of the results.

The aim is twofold, giving a strong and formal framework to address the issue of the marketplace of Internet prospects but also proposing a generic framework to solve any problem similar to that of the marketplace of Internet prospects.

Keywords : Assignment problem, bi-objective optimization, multi-period optimization, forecasting model.

Table des matières

Introduction générale	9
1 Place de marché de prospects Internet	19
1.1 Introduction	19
1.2 Caractéristiques de la place de marché	21
1.2.1 Prospects Internet et commandes	21
1.2.2 Fonctionnement de la place de marché	24
1.3 Problématique et enjeux	26
1.3.1 Problématique	26
1.3.2 Enjeux	26
1.4 Positionnement et objectifs de la thèse	28
1.5 Conclusion	30
2 Etat de l’art	31
2.1 Introduction	31
2.2 Modèles d’optimisation	31
2.2.1 Problèmes d’optimisation	32
2.2.2 Problèmes d’optimisation discrète	33

2.2.3	Problèmes d'optimisation combinatoire	33
2.2.4	Problèmes d'optimisation combinatoire multi-objectif	34
2.2.4.1	Solutions de bon compromis	35
2.2.4.2	Approche de résolution.	36
2.2.5	Problèmes d'affectation	39
2.2.5.1	Formulations	39
2.2.5.2	Quelques algorithmes pour le problème d'affectation	40
2.3	Modèles de prévision	42
2.3.1	Pourquoi élaborer des prévisions pour le processus d'affectation ?	44
2.3.2	Ce qui caractérise un problème de prévision	44
2.3.3	Typologie des problèmes de prévision	45
2.3.4	Ce qui caractérise une solution de prévision	47
2.3.4.1	Les données à prévoir	47
2.3.4.2	Horizon et préodicité de la prévision	47
2.3.4.3	Quelle(s) méthode(s) choisir ?	48
2.3.5	Principales méthodes de prévision	48
2.3.5.1	Les méthodes fondées sur le lissage exponentiel	48
2.3.5.2	Box-Jenkins et modèle causal	49
2.3.6	L'ensemble des modèles <i>ETS</i> au sein du logiciel <i>R</i>	49
2.3.6.1	Origines et utilisation du logiciel <i>R</i>	49
2.3.6.2	L'ensemble des modèles <i>ETS</i>	50
2.4	Exemples d'applications analogues à celle de PdL	51
2.4.1	Le problème d'affectation multi-objectif à l'université Davidson : un cas d'étude	51

2.4.2	Combien de visiteurs à la Tour Eiffel ?	54
2.4.3	Domaine du textile	55
2.4.4	Centre d'appels	56
2.5	Conclusion	57
3	Optimisation bi-objectif et temps réel de l'adéquation offre/demande de prospects Internet dans un contexte mono-période	59
3.1	Offre et demande de prospects Internet	59
3.2	Un exemple illustratif	60
3.3	Formalisation du problème	61
3.3.1	Caractéristiques du problème	62
3.3.2	Variables de décision	62
3.3.3	Contraintes	63
3.3.4	Fonction objectif	66
3.4	Modèle d'optimisation mono-objectif de l'affectation offre/demande de prospects Internet	66
3.5	Algorithme pour l'affectation mono-objectif	68
3.6	Modèle d'optimisation bi-objectif de l'affectation offre/demande de prospects Internet	70
3.6.1	Fonctions de satisfaction des commandes	71
3.6.2	Programme mathématique bi-objectif	72
3.7	Algorithme pour l'affectation bi-objectif	74
3.8	Conclusion	76
4	Optimisation multi-période de l'adéquation offre/demande de prospects Internet	77
4.1	Quelques exemples illustratifs	78

4.1.1	Anticipation de l'évolution des flux de prospects Internet	78
4.1.2	Anticipation de la mise à jour des quotas	79
4.2	Modèle d'affectation multi-période de l'offre/demande de prospects Internet .	80
4.2.1	Variables de décision	81
4.2.2	Contraintes	82
4.2.3	Fonction objectif	84
4.2.4	Programme mathématique multi-période	84
4.3	Implication des futurs flux de prospects Internet et réinitialisations des quotas dans le modèle multi-période	85
4.3.1	Implication de l'évolution des flux de prospects Internet dans le mo- dèle d'optimisation multi-période	86
4.3.2	Implication des réinitialisations des quotas de commandes dans le mo- dèle d'optimisation multi-période	87
4.4	Algorithme pour l'affectation multi-période	88
4.5	Conclusion	90
5	Modèle de prévision pour l'optimisation multi-période	93
5.1	Introduction	93
5.2	Articulation prévision/optimisation	94
5.2.1	Pourquoi tenir compte des prévisions des flux de prospects Internet ? .	95
5.2.2	Données à prévoir	95
5.2.3	Pas de temps et horizon de la prévision	96
5.3	Système de prévision	97
5.3.1	Construction des séries chronologiques	98

5.3.2	Modèle de prévision et son intégration dans le modèle d'optimisation multi-période	100
5.3.3	Prévision des flux de prospects Internet pour une nouvelle commande	104
5.4	Conclusion	105
6	Résultats expérimentaux	107
6.1	Introduction	107
6.2	Évaluation des modèles d'optimisation	107
6.2.1	Schéma expérimental	108
6.2.2	Expériences et résultats	109
6.2.2.1	Capacité d'optimisation	109
6.2.2.2	Intervalle de distribution	113
6.2.2.3	Passage à l'échelle	114
6.2.2.4	Périodicité du modèle d'optimisation multi-période	116
6.2.2.5	Frontière efficace du modèle d'optimisation bi-objectif mono-période	118
6.3	Évaluation du modèle de prévision pour l'optimisation multi-période	120
6.3.1	Schéma expérimental	120
6.3.2	Expériences et résultats	122
6.3.2.1	Prévision exacte	122
6.3.2.2	Prévision réelle	125
6.3.2.3	Pas de temps de prévision	126
6.4	Conclusion	128
7	Généralisation des modèles	129
7.1	Introduction	129

7.2	Satisfaction des clients	130
7.3	Nouveaux flux de commandes et de clients	132
7.4	Nombre de ventes non exclusives des prospects Internet	135
7.5	Extension multi-période de la satisfaction des commandes et des clients . . .	137
7.6	Modèle d'optimisation bi-objectif multi-période	139
7.6.1	Variables de décision	139
7.6.2	Contraintes	140
7.6.3	Fonctions objectif	142
7.6.4	Programme mathématique bi-objectif multi-période	143
7.7	Conclusion	144
8	Conclusion et perspectives	145
8.1	Contributions	145
8.2	Perspectives de recherche	148
	Bibliographie	155
A	Implémentation des modèles d'optimisation	161
A.1	Lecture et structuration des données	161
A.2	Implémentation du modèle d'optimisation mono-objectif mono-période . . .	163
A.2.1	Création du projet <i>CPLEX</i> avec variables de décision binaires	163
A.2.2	Implémentation des contraintes et de la fonction objectif	164
A.2.3	Résolution du programme mathématique et affichage des résultats . .	166
A.3	Implémentation du modèle d'optimisation bi-objectif mono-période	167
A.4	Implémentation du modèle d'optimisation mono-objectif multi-période . . .	171
B	Implémentation du modèle de prévision	175

B.1	Structure du prospect Internet	175
B.2	Extraction des données	176
B.3	Ordonnancement des commandes	178
B.4	Construction des groupes de commandes	178
B.5	Ordonnancement des groupes de commandes	179
B.6	Affichage des groupes de commandes	181
B.7	Prévision sous le logiciel <i>R</i>	181
B.8	Calcul des erreurs de prévision	182

Liste des figures

1.1	Le site commercial de Place des Leads.	21
1.2	Un prospect Internet	22
1.3	Les commandes et leurs caractéristiques	23
1.4	Fonctionnement global de la place de marché.	25
6.1	Différences des chiffres d'affaires journaliers moyens cumulés.	113
6.2	Frontière efficace (instances 1 et 2).	119
6.3	Frontière efficace (instances 3 et 4).	119
6.4	Différence des chiffres d'affaires journaliers moyens cumulés, générés par les deux systèmes <i>Opt + Pré</i> et <i>Opt</i>	123
6.5	Différence des chiffres d'affaires hebdomadaires moyens cumulés, générés par les deux systèmes <i>Opt + Pré</i> et <i>Opt</i>	124
6.6	L'influence des erreurs de prévision sur le chiffre d'affaires hebdomadaire moyen, généré par le modèle d'optimisation multi-période.	126
A.1	Lecture des données	162
A.2	Structure des Commandes	162
A.3	Création du projet <i>CPLEX</i>	163

A.4	Implémentation de la contrainte liée aux volumes des commandes	164
A.5	Implémentation de la contrainte liée aux volumes des clients	164
A.6	Implémentation de la contrainte liée au nombre de ventes et l'exclusivité . . .	165
A.7	Implémentation de la contrainte liée à la duplication des prospectus Internet . .	165
A.8	Implémentation de la fonction objectif : maximisation du chiffre d'affaires . .	166
A.9	Résolution du programme mathématique mono-objectif	166
A.10	Affichage des résultats	167
A.11	Redéfinition du programme mathématique bi-objectif	168
A.12	Fonctions objectif	169
A.13	Condition d'arrêt pour la méthode ϵ constraint	170
A.14	Contrainte pour la méthode ϵ constraint	170
A.15	Calcul du chiffre d'affaires et de la satisfaction moyenne des clients	170
A.16	Résolution du programme mathématique bi-objectif	170
A.17	Matrices de données pour le modèle mono-objectif multi-période	172
A.18	Matrice unique pour le modèle mono-objectif multi-période	172
A.19	Résultat du modèle mono-objectif multi-période	173
B.1	Structure du prospect Internet	176
B.2	Extraction des données	177
B.3	Ordonnancement des commandes	178
B.4	Construction des groupes de commandes	179
B.5	Ordonnancement des groupes de commandes	180
B.6	Affichage des groupes de commandes	181
B.7	Prévision sous le logiciel <i>R</i>	182

B.8	Calcul des erreurs de prévision	183
B.9	Méthodes de calcul des $WAPE$, $WAPE^+$ et $WAPE^-$	184

Liste des tableaux

2.1	Domaines d'utilisation de la prévision.	45
3.1	Liste des commandes à l'instant t	60
3.2	Liste des prospects Internet à l'instant t	61
4.1	Commandes à l'instant t	78
4.2	Prospects Internet à l'instant t	78
4.3	Prospects Internet à l'instant $t + i$	79
4.4	Commandes à l'instant t	79
4.5	Prospects Internet à l'instant t	80
4.6	Commandes à l'instant $t + i$	80
4.7	Ensemble de toutes les affectations possibles à la période de temps 0.	86
4.8	Ensemble de toutes les affectations possibles à la période de temps 1.	87
4.9	Ensemble de toutes les affectations possibles à la période de temps 2.	87
5.1	Correspondances PI/C à la période historique h_1	99
5.2	Correspondances PI/C à la période historique h_2	99
5.3	Séries chronologiques	100
5.4	Ensemble de toutes les affectations possibles à la période de temps 0.	103

5.5	Prévision fournie par la méthode <i>ETS</i>	103
5.6	Ensemble de toutes les affectations possibles à la période de temps 1.	103
5.7	Ensemble de toutes les affectations possibles à la période de temps 2.	104
6.1	Paramètres du schéma expérimental.	109
6.2	Différences des chiffres d'affaires moyens à 10h00.	110
6.3	Différences des chiffres d'affaires moyens après 24 heures de simulation à partir de 10h00.	111
6.4	Différences des chiffres d'affaires journaliers moyens.	112
6.5	Chiffre d'affaires journalier moyen des modèles d'optimisation mono-obj mono-pér et mono-obj multi-pér, avec 10 et 60 minutes d'intervalle de distribution.	114
6.6	La différence entre le chiffre d'affaires hebdomadaire moyen généré par le système PdL et celui généré par le modèle d'optimisation mono-obj mono-pér, en fonction de la taille des données.	115
6.7	La différence entre le chiffre d'affaires hebdomadaire moyen généré par le système PdL et celui généré par le modèle d'optimisation mono-obj multi-pér, en fonction de la taille des données.	116
6.8	Temps de calcul.	116
6.9	Chiffre d'affaires journalier moyen du modèle d'optimisation multi-période en fonction des différentes périodicités.	117
6.10	Instances.	120
6.11	Chiffres d'affaires hebdomadaires moyens générés, en considérant les pas de temps de prévision 1h, 12h et 24h.	127
7.1	Ensemble de toutes les affectations possibles à la période de temps 0.	134

7.2	Ensemble de toutes les affectations possibles à la période de temps 1	135
7.3	Ensemble de toutes les affectations possibles à la période de temps 2	135
8.1	Table des commandes à l'instant t	151
8.2	Table des prospects Internet à l'instant t	151
8.3	Table des commandes à l'instant t	152
8.4	Table des prospects Internet à l'instant t	152

Liste des Algorithmes

3.1	Affectation mono-objectif()	69
3.2	Affectation bi-objectif()	75
4.1	Affectation multi-période()	89
5.1	Séries-Chronologiques()	101
5.2	Groupe-Commandes(j, h, O, Z)	102
A.1	ε - <i>constraint</i> ()	168

Introduction générale

Le problème d'affectation (PA) est un problème très répandu dans le domaine de la recherche opérationnelle. Dans sa version classique, l'objectif d'un PA est de résoudre de manière optimale l'attribution d'un ensemble donné de tâches à un groupe de personnes, tel que le coût minimum ou le profit maximum peut être obtenu. Le PA classique dispose d'un large éventail d'applications incluant la production, la planification du personnel, le transport, . . . etc. [Des-souky and Kijowski, 1997; Soumis et al., 1980; Campell and Langevin, 1995; Leblanc and Farhangian, 1981; Mckeown and Workman, 1976].

Du fait de cette diversité d'applications, plusieurs formulations du PA sont apparues prenant en compte diverses dimensions à satisfaire. En effet, un problème d'affectation standard ne tient compte que du coût ou du profit dans la formulation du problème, alors que différentes applications réelles nécessitent l'intégration d'autres aspects. À titre d'exemple, les décideurs peuvent avoir plusieurs objectifs à satisfaire et les moyens de les atteindre peuvent être conflictuels. C'est notamment le cas du problème d'affectation des travailleurs à formations multiples dans un environnement multiministériel. Afin de résoudre ce problème, [Campbell and Diaby, 2002] ont souligné que les niveaux de la demande dans les différents ministères et les capacités des travailleurs disponibles devrait être considérés comme entrées, et que les résultats d'affectation peuvent affecter la qualité du service et la satisfaction des employés. De même, [Suer and Bera, 1998] ont également affirmé que plusieurs facteurs pourraient influencer sur l'affectation de la main-d'oeuvre dans la cellule de fabrication. En général, les différentes unités d'évaluation peuvent être utilisées pour évaluer les mesures de performance des objectifs.

Le travail de cette thèse s'inscrit dans cette perspective, en d'autres termes, modéliser et formuler un problème d'affectation pour un nouveau contexte qui est celui de la place de marché Place des Leads (PdL) qui représente une place de marché de "prospects Internet". Un prospect Internet est un ensemble d'informations (département, profession, âge ... etc) ou caractéristiques concernant un internaute à la recherche d'un produit particulier (assurance santé, assurance automobiles, travaux de rénovation, achat/vente immobilier, ... etc). Ces prospects Internet sont automatiquement générés par la technologie PdL ou achetés auprès d'un fournisseur. En effet, PdL permet à des fournisseurs de prospects Internet (comparateurs d'assurance, de crédit, de prix, institutions financières compagnies d'assurance, ... etc) de les envoyer sur sa plateforme par Web service. PdL se charge ensuite de les revendre en temps réel à des clients ultimes, aussi bien de grands comptes (compagnies d'assurances, banques, ... etc) que de petits comptes (courtiers en assurance, en crédit, artisans, avocats, ... etc).

D'une part, PdL reçoit une quantité importante de prospects Internet en temps réel. D'autre part, les clients de PdL effectuent des commandes et peuvent filtrer les prospects Internet qu'ils souhaitent recevoir selon de multiples critères, à savoir : le nombre de prospects Internet par mois, par jour ou par heure, la nature de la transaction (prospect Internet vendu exclusivement à un client ou pas), l'âge du prospect Internet (moins d'une heure, de douze heures, ... etc), le profil de l'internaute (cadre ou pas, adulte ou pas, ... etc), ... etc. Par la suite, Place des Leads doit considérer l'ensemble de ces paramètres (à la fois du côté offre et du côté demande) pour établir une correspondance en temps réel entre les fournisseurs et les clients. Le défi de la place de marché PdL est de réussir non seulement à faire le plus de bénéfices possibles mais aussi de satisfaire au mieux ses clients.

Problématique

Cette thèse a pour ambition de solutionner le problème de la place de marché PdL à savoir rendre optimal l'affectation des prospects Internet à des commandes. L'optimalité devra répondre essentiellement à deux besoins. D'une part, nous devons garantir à la place de marché

de faire du bénéfice. D'autre part, nous devons garder en tête que PdL souhaite fidéliser ses clients et ne pas les perdre ¹. De ce fait, elle doit faire en sorte que la distribution des prospects soit la plus équitable possible entre eux. Plus précisément, ce travail a pour but de répondre au problème de l'adéquation de l'offre et la demande de sorte à optimiser deux objectifs, à savoir le chiffre d'affaire et le taux de satisfaction des clients.

À première vue, le problème d'affectation de Place des Leads semble être un problème classique d'affectation, à savoir, affecter un ensemble d'objets à un autre ensemble d'objets afin d'optimiser un objectif : maximiser le chiffre d'affaire de l'entreprise. Dans le cas le plus simple, le problème d'affectation linéaire possède un nombre de commandes égal au nombre de prospects Internet et un seul objectif [Burkard et al., 2012]. Dans notre cas, le nombre de commandes et de prospects Internet ne sont pas égaux. Cette question a créé le problème de l'affectation généralisée [Chen et al., 2010]. Bien que le problème de l'affectation standard et des variantes ont été bien étudiés et de bons algorithmes existent pour les résoudre [Barr et al., 1977; Hung and Rom, 1980; McGinnis, 1983], notre problème d'affectation diffère sur quelques points. En effet, à un instant donné, l'ensemble des prospects Internet disponibles (pouvant être vendus) ainsi que les commandes à satisfaire (clients souhaitant acheter des prospects Internet) sont connus. Il s'agit donc de répondre à la problématique de l'adéquation des prospects Internet et des commandes de sorte à optimiser simultanément deux objectifs. En sus, plusieurs caractéristiques sont à prendre en compte. En effet, une première caractéristique revient au fait que certaines commandes exigent des ventes de prospects Internet de manière exclusive ou pas. En effet, la vente du prospect se fait soit, au plus une fois et à une et une seule commande, ou bien plusieurs fois à plusieurs commandes.

Une autre caractéristique est celle de la temporalité du système. En effet, le système PdL fonctionne en continu et donc le modèle d'affectation doit prendre en compte cette exigence. En d'autres termes, l'affectation est effectuée en temps réel, dans un contexte caractérisé par des flux continus faisant évoluer en temps réel l'ensemble des prospects Internet disponibles

1. Il est clair que d'autres objectifs peuvent être pris en compte. Dans le cadre de ce travail nous nous sommes intéressés uniquement aux deux objectifs mentionnés. Nous reviendrons dans la conclusion sur la question de considérer ou pas d'autres objectifs et quelles sont les conséquences sur le problème d'affectation.

ainsi que les commandes à satisfaire.

De plus, le client de PdL possède la possibilité d'effectuer une ou plusieurs commandes avec des volumes différents. Le nombre de prospects Internet affecté à chaque commande est limité par un volume global, un quota par heure, un autre par jour et un dernier par mois. Le client peut aussi spécifier la quantité globale de prospects Internet correspondant à ses différentes commandes. Ainsi le nombre des prospects Internet affecté à l'ensemble des commandes d'un client ne doit pas dépasser le quota horaire, journalier et mensuel autorisé.

Ce travail s'intéresse donc à l'optimisation bi-objectif avec anticipation des flux de prospects Internet et des réactivations des quotas de commandes. L'objectif est double, d'une part, donner un cadre solide et formel pour répondre à la problématique de PdL. D'autre part, le cadre proposé devrait être aussi générique que possible afin de résoudre tout autre problème d'affectation partageant les mêmes caractéristiques que celui de PdL. Pour ce faire, nous nous sommes intéressés aux points suivants.

Formalisation du problème de l'adéquation offre/demande de prospects Internet. Le premier objectif de cette thèse est la formalisation de l'adéquation de l'offre et de la demande de prospects Internet. Cette adéquation est réalisée dans un contexte où des flux continus font évoluer en temps réel l'ensemble des offres disponibles et les demandes à satisfaire. En outre, le problème d'affectation auquel nous nous intéressons est caractérisé par la temporalité du système. En effet, le système PdL fonctionne en temps réel, ainsi à chaque instant nous devons décider quels prospects Internet seront affectés et à quelles commandes. Notre première contribution a été de formaliser le problème d'affectation afin de comprendre au mieux le fonctionnement et les enjeux liés à l'affectation dans la place de marché PdL.

Optimisation bi-objectif et temps réel de l'offre/demande de prospects Internet. Lors de la formalisation du problème de l'affectation de l'offre à la demande de prospects Internet, nous avons noté deux caractéristiques importantes du problème à prendre en compte.

1. La considération de deux objectifs à optimiser à travers le processus d'affectation, à sa-

voir : le chiffre d'affaires et le le niveau de satisfaction moyen des clients ; En effet, l'affectation a pour objectif de trouver un bon compromis conciliant ces deux objectifs contradictoires. En effet, d'une part, la maximisation du chiffre d'affaires nécessite d'affecter les prospects Internet aux clients proposant les prix les plus élevés. Ainsi, les clients proposant des prix moins intéressants seront marginalisés. D'autre part, la distribution équitable des prospects Internet à tous les clients générera un chiffre d'affaires moins importants. Cela est dû à la différence des prix proposés par les différents clients.

2. L'évolution des flux de prospects Internet et leur affectation se font en temps réel. En effet, un des enjeux majeurs de cette affectation consiste à permettre la prise en compte instantanée des nouvelles offres et demandes de prospects Internet. Cet aspect est d'autant plus crucial que la valeur des prospects Internet diminue à travers le temps. Pour réaliser une affectation optimale, PdL doit satisfaire deux objectifs à la fois, à savoir le chiffre d'affaires et la satisfaction des commandes, mais aussi tenir compte de l'aspect temps réel de cette affectation.

Prise en compte d'un modèle multi-période dans la phase d'optimisation. Comme nous l'avons souligné auparavant, une composante importante du problème d'affectation bi-objectif des prospects Internet est l'évolution continue des flux d'offres et de demandes de prospects Internet. La connaissance de ces flux s'avère donc un levier important pour réaliser des affectations optimales. En outre, la connaissance concernant les réactivations régulières des quotas de la demande permet également d'améliorer l'affectation. En effet, l'optimisation de l'affectation en considérant une seule période de temps conduit à une stratégie "myope", ne permettant pas d'anticiper les mises à jour futures du système. Pour répondre à ce besoin, nous considérons des affectations sur plusieurs périodes de temps. Une telle approche conduit à optimiser les résultats de l'affectation en tirant profit de la connaissance concernant les flux d'offre des prospects Internet et les réactivations des quotas de commandes.

En outre, le modèle d'optimisation que nous proposons permet d'effectuer des affectations optimales à un instant donné. Cependant, une caractéristique de la place de marché est son

fonctionnement en continu dans le temps. Ainsi, nous proposons un algorithme qui fonctionne de manière continue en optimisant à chaque instant l'adéquation offre/demande de prospects Internet tout en considérant instantanément les flux continus de prospects Internet et la mise à jour régulière des quotas de la demande.

Prise en compte d'un modèle de prévision dans l'optimisation multi-période. Nous travaillons dans un contexte caractérisé par des flux continus de prospects Internet faisant évoluer de façon continue et non déterministe l'ensemble de l'offre disponible. Afin d'améliorer le processus d'affectation de PdL, nous devons prendre en compte ces flux de prospects Internet dans le modèle d'optimisation multi-période. En d'autres termes, l'objectif est de permettre au modèle d'affectation multi-période d'optimiser l'affectation en tirant profit de la connaissance concernant l'évolution des flux de prospects Internet.

Nous proposons d'articuler un modèle de prévision et un modèle d'affectation multi-période afin de répondre au problème de PdL. Nous verrons dans la suite de ce travail que cette articulation pose des problèmes au niveau des choix des paramètres du modèle de prévision qui seront en adéquation avec le fonctionnement du modèle d'optimisation.

Généralisation des contributions. Un autre objectif visé par cette thèse est d'offrir un modèle réutilisable dans tout autre contexte ressemblant à celui de PdL. En effet, les contributions devraient être aussi génériques que possible afin de traiter tous les problèmes d'affectation partageant les mêmes caractéristiques que celles de PdL. Ainsi, nous présentons une extension de nos travaux dans le but d'avoir un modèle générique pouvant être appliqué à d'autres contextes analogues à celui de PdL.

Plan synthétique de la thèse

Ce manuscrit se décline en huit chapitres. Le chapitre 1 introduit le contexte général dans lequel s'inscrit notre travail. Le chapitre 2 réalise un état de l'art des différents outils utilisés

pour la suite de nos travaux. Les chapitres 3, 4, 5 et 7 présentent nos contributions. Le chapitre 6 présente les différentes expérimentations menées en vue de valider nos travaux. Enfin, le chapitre 8 introduit un bilan des travaux exposés dans cette thèse et présente les différentes perspectives ouvertes par ce travail.

Chapitre 1 : Place de marché de prospects Internet. Dans le premier chapitre nous décrivons le périmètre de travail dans le cadre de cette thèse, à savoir la place de marché Place des Leads. Nous exposons en premier le fonctionnement et les caractéristiques de cette place de marché. Ensuite, Nous discutons du problème rencontré par cette place et des enjeux derrière qui ont motivé ce travail de recherche. Enfin, nous détaillons le positionnement et les objectifs de cette thèse.

Chapitre 2 : Etat de l'art. En nous appuyant sur une étude bibliographique, nous introduisons les différents concepts et notions utilisés dans nos travaux. Dans un premier temps, nous rappelons quelques notions élémentaires de la recherche opérationnelle, à savoir les définitions de quelques classes de problèmes d'optimisation, d'affectation et les algorithmes de résolution en mettant l'accent sur quelques détails du solveur *CPLEX*. Dans un second temps, nous présentons les caractéristiques des problèmes de prévision, leurs typologies ainsi que l'intérêt des modèles de prévision. Aussi, nous abordons brièvement les différentes méthodes de prévision. Enfin, nous présentons quelques contextes analogues à celui de PdL afin de démontrer l'applicabilité de nos travaux dans d'autres contextes tels que l'affectation des équipages, l'ordonnancement des sessions parallèles dans les grandes conférences, les péages d'autoroutes, etc.

Chapitre 3 : Optimisation bi-objectif et temps réel de l'adéquation offre/demande de prospects Internet dans un contexte mono-période. Ce chapitre présente notre première contribution concernant le problème de l'optimisation de l'adéquation offre/demande posé par Place des Leads. Nous commençons par présenter une première formulation du problème posé. Nous développons ensuite un modèle d'optimisation bi-objectif de l'adéquation de l'offre et de

la demande de prospects Internet permettant de tenir compte des différents types de spécificités émanant de la place de marché. Le modèle devra permettre : (i) la considération instantannée des flux de l'offre et de la demande et (ii) le calcul en temps réel des affectations correspondant à de bonnes solutions de compromis portant sur deux objectifs à optimiser, à savoir le chiffre d'affaires et la satisfaction des commandes.

Chapitre 4 : Optimisation multi-période de l'adéquation offre/demande de prospects Internet. Une particularité de la place de marché est la temporalité de son système. De ce fait le modèle d'optimisation doit être en mesure de prendre en compte cette caractéristique pour rester fidèle au fonctionnement du système original. Dans ce chapitre nous décrivons notre proposition d'étendre le modèle d'optimisation introduit au chapitre précédent afin de prendre en compte plusieurs périodes de temps dans le processus de l'affectation. Cette extension devra conduire à optimiser l'affectation en considérant les éventuelles mises à jour futures du système, à savoir les nouveaux flux de l'offre ainsi que les réactualisations des quotas de la demande. Par ailleurs, une caractéristique de la place de marché est son fonctionnement en continu dans le temps. Ainsi, nous proposons un algorithme qui fonctionne de manière continue en optimisant à chaque instant l'adéquation offre/demande de prospects Internet tout en considérant instantanément les flux continus de prospects Internet et la mise à jour régulière des quotas de la demande.

Chapitre 5 : Modèle de prévision pour l'optimisation multi-période. Dans ce chapitre, nous proposons un modèle de prévision des flux de prospects Internet. Nous montrons comment construire un modèle de prévision des flux de prospects Internet tout en considérant les différentes caractéristiques du modèle d'affectation multi-période. Cette approche devra permettre d'améliorer l'optimisation en intégrant la connaissance sur l'évolution continue des flux de prospects Internet.

Chapitre 6 : Résultats expérimentaux. Ce chapitre s'attache à démontrer la validité de nos travaux proposés afin de répondre à la problématique posée par PdL. Nous nous intéressons

à mettre en évidence l'efficacité des modèles proposés ainsi que la qualité des résultats obtenus et ce en présence de différentes expériences concernant les modèles d'optimisation et de prévision. Pour ce faire, nous commençons par tester l'efficacité des modèles d'optimisation que nous avons proposés. Globalement, les expériences réalisées démontrent l'applicabilité de nos modèles dans des systèmes réels. Ils sont toujours capables d'améliorer considérablement les affectations en comparaison avec l'ancien système de PdL. Ils sont également capables de d'accroître le gain en passant outre des difficultés supplémentaires générées la croissance des données. Ensuite, nous penchons sur l'étude de l'efficacité de la combinaison d'un modèle d'optimisation multi-période avec un modèle de prévision. Les résultats obtenus sont très prometteurs et apportent une réelle valeur ajoutée à l'activité de la place de marché PdL. Ces expérimentations nous permettent alors de conclure à l'adéquation de nos modèles aux attentes initialement formulées et de valider l'utilité concrète des modèles proposés.

Chapitre 7 : Généralisation des modèles. Afin d'avoir un modèle générique pouvant être appliqué à d'autres contextes analogues à celui de la place de marché PdL, le chapitre 7 présente un aperçu de certaines extensions possibles des modèles proposés. Cette généralisation a pour ambition de définir un modèle aussi général que possible applicable à toute autre place de marché ayant les mêmes caractéristiques que celles de PdL.

Chapitre 8 : Conclusion et perspectives. Finalement, nous concluons en tirant le bilan des travaux exposés dans cette thèse et présentons les différentes perspectives ouvertes par ce travail.

Chapitre 1

Place de marché de prospects Internet

1.1 Introduction

Au cours de ces dernières années, beaucoup d'internautes ont recherché de l'information concernant des produits bancaires, financiers, d'assurances ou de crédits sur Internet. En effet, Internet demeure un canal prescripteur fort : les ventes sur Internet ont continué de progresser fortement en 2014. Au total, les français auront dépensé 57 milliards d'euros sur Internet, un chiffre en hausse de 11% sur un an. Le nombre de transactions augmente quant à lui de 15% (Source Fevad 28 janvier 2015). Selon les prévisions de la Fevad, les ventes sur Internet devraient connaître une progression de l'ordre de 10% en 2015. Ces estimations tablent sur une nouvelle poussée de l'offre, tirée par le nombre de sites marchands ainsi que sur une nouvelle progression du nombre de cyber-acheteurs, dont les effets conjugués devraient permettre au e-commerce français de franchir la barre des 60 milliards.

Pour attirer cette nouvelle clientèle, des moyens publicitaires importants sont déployés chaque année. À titre d'exemple, d'après le IAB (Interactive Advertising Bureau), les revenus publicitaires sur Internet pour 2013 ont atteint la somme folle de 42,8 milliards de dollars. Parmi ces publicités en ligne les plus rentables, ce sont celles associées aux moteurs de recherche qui dominent les dépenses du côté des annonceurs, mais aussi les publicités spécialement ciblées

sur les plateformes mobiles.

De ce fait, l'un des objectifs majeurs des programmes de communication qui sont déployés actuellement par les compagnies d'assurance ou les banques réside dans la génération ou l'acquisition de prospects Internet. En effet, les indicateurs de visiteurs uniques (VU), de pages vues ne sont plus les indicateurs les plus importants. Ils ne présentent que peu de valeur aux yeux des directions commerciales et marketing. Un site doit désormais prouver sa capacité à participer au Chiffre d'Affaires de la société et au développement commercial. De nouveaux indicateurs ont fait alors leur apparition : le coût par prospect Internet, le coût par contact qualifié, le cout par vente, ... etc

Dans ce contexte, Place des Leads (PdL) représente une place de marché (commerce électronique) de "leads". Un lead (un prospect Internet) est traditionnellement utilisé par les forces commerciales pour désigner un contact qualifié et en phase d'achat. PdL possède un site commercial www.placedesleads.com (voir la figure 1.1) offrant, sous la forme d'une plateforme technique, des mécanismes de transaction sécurisés permettant aux entreprises, particuliers, collectivités et associations de comparer et sélectionner différentes offres de fournisseurs pour des biens ou des services. De manière générale, l'objectif principal d'une place de marché est de satisfaire l'offre et la demande. Cela passe notamment par un taux de reventes minimum de prospects Internet. Aujourd'hui, ce taux de reventes est facilement atteignable lorsque les prospects Internet ne sont pas filtrés par les clients (donc plus faciles à revendre) : c'est ce qui est pratiqué aujourd'hui par les vendeurs de prospects Internet. Le problème lié à ce type de fonctionnement consiste en la perte induite par la non filtration des prospects Internet : les clients reçoivent des prospects Internet qui ne sont pas dans leur cible. Le progrès introduit par PdL consiste à maintenir ce taux de ventes même lorsque les clients possèdent des caractéristiques particulières afin de filtrer les prospects Internet qu'ils souhaitent recevoir. Dans son intégralité, le présent chapitre introduit le périmètre de travail dans le cadre de cette thèse, à savoir la place de marché Place des Leads.



Figure 1.1 — Le site commercial de Place des Leads.

1.2 Caractéristiques de la place de marché

1.2.1 Prospects Internet et commandes

Le système de la place de marché utilise deux entités principales, à savoir :

1. **Les prospects Internet** : un prospect Internet peut être défini dans un contexte web comme un contact qualifié et prêt à effectuer une action (achat, devis, demande de rendez-vous, demande de documentation, abonnement newsletter, ...). En d'autres termes, un prospect Internet représente un ensemble d'information (âge, département, profession, ... etc) ou de caractéristiques d'un internaute désirant un produit particulier (assurance santé, assurance automobile, ... etc). À titre d'exemple, la figure 1.2 représente un prospect Internet d'une dame retraitée, nommée "Chevallier Monique", née le 08/07/1943, qui habite au 8 route de Nangis 77000 VAUX LE PENIL FR qui est à la recherche d'une assurance santé. Ce prospect Internet a été créé le 13/10/2015 à 11 :13 :58. Les prospects Internet sont générés automatiquement par une technologie PdL ou bien achetés chez des

fournisseurs de prospects Internet. Pour générer ses propres prospects Internet, PdL diffuse sur le web des formulaires dépendant des commandes reçues ensuite elle collecte ces formulaires sous forme de prospects Internet (formulaires remplis et validés par des internautes). La topologie d'un prospect Internet est définie par : (i) sa date de création (date de validation du formulaire par l'internaute), (ii) son prix (coût de sa génération ou de son achat chez un fournisseur) ainsi que (iii) les informations concernant l'internaute (différents champs du formulaire : âge, département, profession, ... etc).

Détails du prospect			
Produit	Assurance Santé	Date de création	13/10/15 11:13:58
Apporteur	PIXEO SAS - Principale	ID Apporteur	103
Origine		ID Livraison	9434
ID Lead	38790900	Email	jacques405@laposte.net
ID Formulaire	-	Raison sociale	-
Statut	Transmis	SIRET	-
Situation familiale		Profession	Retraité
Civilité	Madame	Date naissance	08/07/1943
Nom Prénom	Chevallier Monique	Tél. mobile	06 07 65 31 76
Adresse	8 Route de Nangis	Tél. bureau	
CP / Ville / Pays	77000 VAUX LE PENIL FR	Tél. domicile	
IP		Statut email	

Figure 1.2 — Un prospect Internet

2. **Les commandes** : PdL interagit avec plusieurs clients pouvant effectuer une ou plusieurs commandes et filtrer les prospects Internet qu'ils souhaitent recevoir selon de multiples critères. Un filtre représente les caractéristiques de la commande, il peut s'effectuer selon de nombreux paramètres (voir la figure 1.3), à savoir :

- (a) Les champs du formulaire : un filtre sur les valeurs des différents champs du formulaire permet aux clients de PdL de viser par exemple un public d'un certain âge (âge de l'internaute) ou encore exerçant une profession particulière (profession de l'internaute) ... etc.
- (b) L'ancienneté : l'âge du prospect Internet est compté à partir de sa date de création. Ainsi, les clients de PdL possèdent la possibilité de faire un filtre sur l'ancienneté du

prospect Internet (moins d'une heure, de douze heures, ... etc). À titre d'exemple, la première commande sur la figure 1.3 exige des prospects Internet de moins de 24 heures alors que la seconde exige des prospects Internet de moins de 12 heures ... etc.

- (c) L'exclusivité : l'un des critères permettant aux clients de PdL de caractériser leurs commandes réside dans l'achat d'un prospect Internet en exclusivité ou pas. Autrement dit, certaines commandes exigent des achats exclusifs de prospects Internet (prospect Internet vendu au plus une fois) tandis que d'autres acceptent le fait que le prospect Internet soit revendu à plusieurs reprises (ventes multiples à plusieurs clients). À titre d'exemple, la première commande sur la figure 1.3 n'exige pas l'achat exclusif des prospects Internet contrairement à la seconde qui n'achète pas des prospects Internet déjà vendus à d'autres commandes.
- (d) Le volume : le nombre global de prospects Internet souhaité ainsi que le nombre par tranche de temps (par heure, jour et mois) représentent une autre caractéristique des commandes. À titre d'exemple, sur la figure 1.3, la première commande demande un quota global de 90000 prospects Internet mais pas plus de 1000 prospects Internet par mois, par jour et par heure.

Commande								Quota partenaire			Nb leads					Dispo.	PU HT	
Ref.	Prio.	Rej.	Type	Excl.	Trans.	Profil	Ancien.	Mois	Jour	Heure	Cmdés	Livrés						
												Total	Mois	Jour	Heure			
2602	9000	0%	Lead	Non	Auto	Agences...	< 24h	119	119	119	90 000	2 592	111 / 1 000	6 / 1 000	0 / 1 000	0	28,00 €	
1754	8000	5%	Lead	Oui	Auto	Affilia...	< 12h	3 500	50	300	100 000	19 517	444 / 5 000	15 / 200	0 / 20	0	13,00 €	
1586	5000	10%	Lead	Oui	Auto	Landing...	< 48h	10 000	77 / 75	1 000	100 000	24 215	328 / 5 000	27 / 300	1 / 12	0	0,00 €	
1587	5000	10%	Lead	Oui	Auto	Landing...	< 48h	10 000	77 / 75	1 000	100 000	27 459	471 / 5 000	34 / 300	4 / 15	0	0,00 €	
3007	2250	5%	Lead	Oui	Auto	Aff po...	< 5 j.	3 500	50	300	100 000	1 919	117 / 5 000	0 / 150	0 / 1 000	0	13,00 €	
1616	1030	20%	Lead	Non	Auto	geoloc	< 24h	1 750	38	30	15 000	14 983	162 / 1 000	15 / 18	0 / 50	109	5,00 €	
1614	1020	20%	Lead	Non	Auto	geoloc_...	< 24h	1 750	38	30	15 000	12 436	143 / 270	13 / 13	0 / 13	89	9,00 €	
1615	1010	20%	Lead	Non	Auto	geoloc_...	< 24h	1 750	38	30	10 000	2 115	55 / 1 000	5 / 5	0 / 50	19	9,00 €	
6624	100	30%	Lead	Non	Auto	Mutuell...	< 48h	90 000	1 500	1 500	20 000	18 788	997 / 20 000	75 / 75	0 / 75	125	7,50 €	
1547	50	30%	Lead	Non	Auto	Mutuell...	< 24h	90 000	1 500	1 500	30 000	22 415	375 / 5 000	25 / 25	0 / 25	93	5,02 €	

Figure 1.3 — Les commandes et leurs caractéristiques

1.2.2 Fonctionnement de la place de marché

Place des Leads s'intéresse en particulier à trois modules technologiques. Premièrement, la conversion des données hétérogènes provenant du web en un format normalisé et leur renvoi à d'autres systèmes ayant eux-même leurs propres formats. En effet, (1) PdL reçoit des prospects Internet (de la part des fournisseurs), (2) ensuite elle les intègre dans sa plateforme (3) et enfin elle les poste chez les clients quelque soit leurs systèmes d'information. L'objectif poursuivi par PdL est de concevoir un protocole d'échanges de flux pour un nombre indéterminé de fournisseurs et de clients qui possèdent, pour chacun d'entre eux, leur propre solution d'échange de flux de prospects Internet.

Deuxièmement, la génération de prospects Internet, autrement dit, produire ses propres prospects Internet via une technologie propriétaire permettant de déployer facilement et rapidement sur des sites Internet tiers, des formulaires de demande de devis concernant de nombreuses thématiques (banque, assurance, travaux, formation, ... etc).

Troisièmement, le traitement, l'analyse et la revente des prospects Internet. Ce module faisant l'objet de la thèse est au coeur du fonctionnement de la place de marché et par conséquent il est déterminant pour son efficience. En effet, dans le cadre de cette thèse, nous nous intéressons au troisième module technologique qui est illustré par la figure 1.4. D'une part (côté "offre"), PdL reçoit une quantité importante de prospects Internet en temps réel. D'autre part (côté "demande"), les clients de PdL passent des commandes et peuvent filtrer les prospects Internet qu'ils souhaitent recevoir selon de multiples critères. Ainsi, le travail que nous présentons dans cette thèse s'intéresse au système d'affectation des prospects Internet aux commandes.

Place des Leads interagit avec plusieurs clients (assurance, banque, ... etc) qui effectuent des commandes de prospects Internet. Pour répondre à cette demande, PdL procède en deux étapes :

1. Calculer un score pour chaque commande en se basant sur les informations techniques ainsi que les connaissances de PdL concernant les clients et le marché.
2. Utiliser ce score dans un algorithme glouton afin de décider de l'affectation des prospects

Internet aux commandes en scannant la liste de toutes les commandes du score le plus élevé au plus faible.

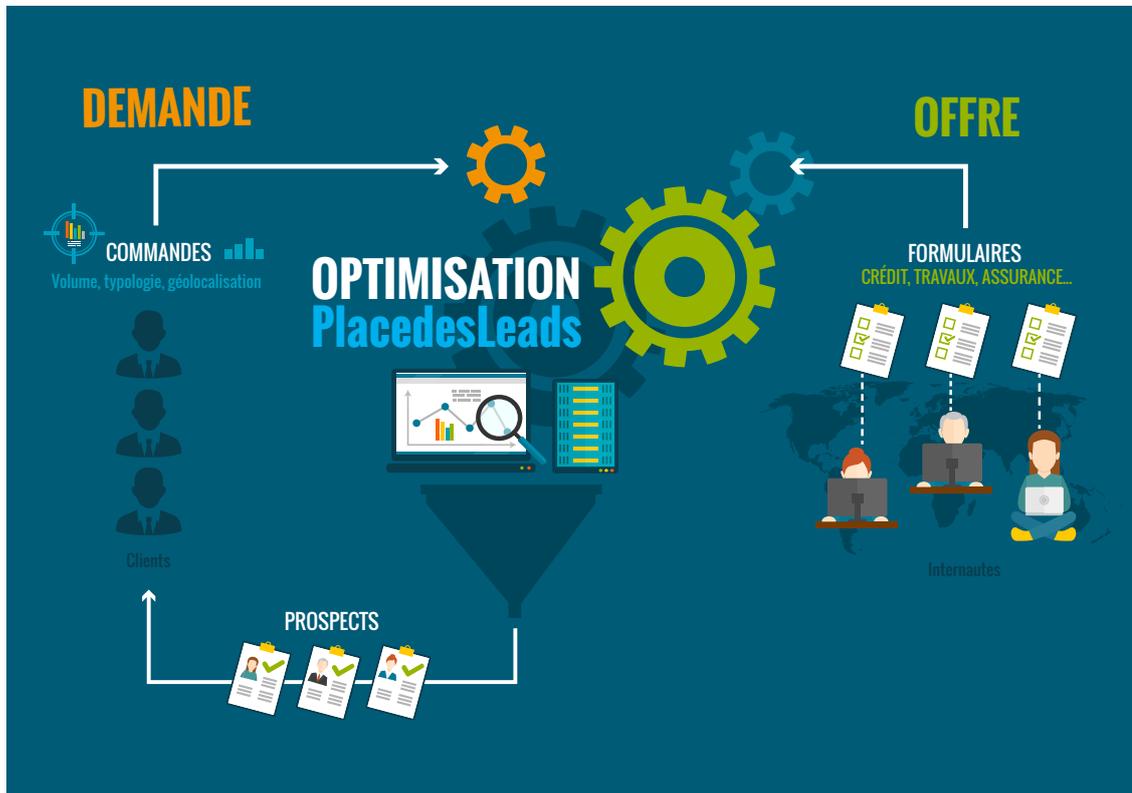


Figure 1.4 — Fonctionnement global de la place de marché.

Ce type de traitement reste inefficace. En effet, le score attribué aux commandes ne change pas au fil du temps, ainsi les commandes ayant les plus grands scores sont toujours prioritaires alors que les autres resteront mal servies. Cela génère le risque de perdre certains clients. En outre, toutes les caractéristiques des commandes ne sont pas considérées dans le calcul des scores de commandes. À titre d'exemple, une commande qui exige des prospects Internet jeunes devrait passer en priorité même si le prix qu'elle propose n'est pas élevé. Dans le cas contraire ces commandes ne pourront jamais être servies suite à l'expiration des prospects Internet par rapport à ces commandes. Enfin, les commandes exclusives ne sont pas forcément prioritaires malgré qu'elles proposent des prix plus élevés que ceux proposés par les commandes non exclusives. En effet, une revente jusqu'à 3 fois d'un prospect Internet à des commandes non

exclusives génèrent un gain plus important que celui généré par une revente exclusive à une seule commande. Partant de ces constats, le travail de cette thèse a pour objet d'offrir à PdL une nouvelle stratégie d'optimisation du système de vente des prospects Internet aux clients.

1.3 Problématique et enjeux

1.3.1 Problématique

Place des Leads fait face à une problématique technique complexe, à savoir "l'affectation des prospects Internet à ses clients souhaitant les transformer en ventes effectives". En effet, PdL établit une correspondance en temps réel entre l'offre et la demande de prospects Internet. À un instant donné, l'ensemble des prospects Internet disponibles (pouvant être vendus) et les commandes à satisfaire (clients souhaitant acheter des prospects Internet) sont connus. Une des caractéristiques du problème tient au fait que certaines commandes ciblent des prospects Internet bien particuliers qui exigent certains filtres sur ces derniers. De plus, la prise en compte de l'offre et de la demande ainsi que leur affectation, est effectuée en temps réel dans un contexte caractérisé par : (i) des flux continus faisant évoluer en temps réel l'ensemble des prospects Internet disponibles et (ii) des quotas mis à jour faisant varier régulièrement l'ensemble des commandes à satisfaire.

Pour avoir une affectation optimale, PdL doit satisfaire plusieurs objectifs à la fois (le chiffre d'affaires, la satisfaction des clients, ... etc) tout en respectant les caractéristiques des commandes mais aussi tenir compte de l'aspect temps réel de cette affectation.

1.3.2 Enjeux

Dans le cadre de Place des leads, une affectation vise à permettre la prise en compte de nombreux objectifs. Ces derniers ne sont pas considérés dans le fonctionnement du système d'affectation de PdL. En revanche, la considération de différents objectifs à optimiser au niveau du traitement des prospects Internet constitue un besoin réel pour la place de marché PdL, à

savoir :

1. Le chiffre d'affaires : l'aspect monétaire est indéniablement une question au coeur de toute entreprise souhaitant faire fleurir son activité. Donc, un des premiers enjeux de la place de marché à travers l'affectation est celui de maximiser son chiffre d'affaires.
2. La satisfaction des clients : PdL interagit avec différents clients. La favorisation de certains (suite aux prix intéressants qu'ils proposent) aboutit à la perte de ceux qui ont été marginalisés. D'où la nécessité de servir tous les clients de manière équitable. Ainsi, le deuxième objectif à prendre en compte par le processus d'affectation est celui de la satisfaction équitable des clients de PdL.
3. La satisfaction des fournisseurs de prospects Internet : de nombreux fournisseurs envoient une quantité importante de prospects Internet à PdL. Cette dernière s'occupe de leur revente à ses clients. Le fait de prendre en compte uniquement les prospects Internet provenant de certains fournisseurs (en raison du prix intéressant qu'ils proposent) conduit forcément à perdre ceux qui ne sont pas sollicités. Ainsi, satisfaire équitablement tous les fournisseurs s'avère un enjeu très important afin de continuer à alimenter PdL de prospects Internet. Par conséquent, un des objectifs à réaliser par le biais de l'affectation de l'offre des prospects Internet aux clients est la satisfaction équitable de tous les fournisseurs.
4. La fraîcheur des prospects Internet : l'un des critères définissant la topologie du prospect Internet est son âge, qui est calculé à partir de sa date de création. L'innovation apportée par PdL consiste à revendre des prospects Internet jeunes. En outre, certains clients effectuent un filtre sur l'âge des prospects Internet qu'ils souhaitent recevoir. Ainsi, plus le prospect Internet est jeune plus sa revente s'avère facile, d'où la nécessité de revendre en temps réel les prospects Internet arrivant sur la place de marché PdL.

Dans le cadre de cette thèse, nous nous intéressons uniquement au chiffre d'affaires et à la satisfaction des clients. Ainsi, nous cherchons à répondre à cette problématique de l'adéquation de l'offre et de la demande de prospects Internet de sorte à optimiser simultanément deux

objectifs à savoir maximiser le chiffre d'affaires et le taux de satisfaction des commandes. Cette optimisation s'effectue tout en prenant en compte les différentes caractéristiques de la place de marché PdL discutées ci-dessus.

1.4 Positionnement et objectifs de la thèse

Ce travail de thèse s'inscrit dans la problématique d'affectation de l'offre de prospects Internet aux clients de PdL. Pour répondre à cette problématique, cinq principaux points sont étudiés, à savoir :

1. La formalisation du problème de l'affectation de l'offre des prospects Internet à des commandes : le premier objectif de cette thèse est la formalisation de l'adéquation offre/demande de prospects Internet. Cette adéquation est réalisée dans un contexte où des flux continus font évoluer en temps réel l'ensemble de l'offre disponible et la demande à satisfaire. De plus, le problème d'affectation auquel nous nous intéressons est caractérisé par sa temporalité. En effet, le système PdL fonctionne en temps réel, ainsi à chaque instant nous devons décider quels prospects Internet seront affectés et à quelles commandes. Le volume des flux de l'offre et de la demande de prospects Internet ainsi que la fréquence de résolution (une affectation toutes les 10 minutes voire toutes les minutes) conduisent à un problème d'optimisation difficile à résoudre en temps réel.
2. L'optimisation bi-objectif de l'adéquation temps réel de l'offre et la demande : le second objectif de cette thèse concerne la prise en compte lors de l'affectation des prospects Internet à des commandes de deux principales caractéristiques. D'une part, les objectifs à optimiser par le processus d'affectation sont la maximisation du chiffre d'affaires total ainsi que le niveau de la satisfaction des clients. En effet, l'affectation a pour objectif de trouver de bons compromis conciliant ces deux objectifs contradictoires. En effet, afin de maximiser le chiffre d'affaires de PdL, il serait nécessaire d'affecter les prospects Internet aux clients proposant les prix les plus élevés. Par conséquent, ceux qui proposent des prix moins intéressants seront marginalisés. Cependant, la distribution équitable des prospects

Internet sur tous les clients génère un chiffre d'affaires moins important. Cela est dû à la différence des prix proposés par les différents clients.

D'autre part, l'un des enjeux forts de cette affectation est de permettre la prise en compte instantanée des nouvelles offres et demandes de prospects Internet et leur affectation en temps réel. Cet aspect est extrêmement important dans le sens où les prospects Internet vieillissent au fil du temps et deviennent donc difficiles à vendre. Notre idée ne consiste pas à mettre la notion d'âge comme critère à optimiser mais de construire un mécanisme qui permet de vendre au mieux les prospects Internet en temps réel.

Notre approche pour répondre à ce premier objectif consiste à formuler le problème par un programme mathématique bi-objectif dont la résolution conduit à identifier la surface de pareto des solutions pareto-optimales puis de sélectionner sur cette surface une bonne solution de compromis sur la base des arbitrages souhaités par le décideur.

3. L'extension multi-période de l'adéquation offre/demande de prospects Internet : une des caractéristiques importantes du problème d'affectation tient au fait que (i) les flux de prospects Internet évoluent continuellement et (ii) les quotas de commandes sont régulières mis à jour. La considération de ces deux connaissances dans le processus d'adéquation de l'offre et de la demande s'avère un levier important afin de réaliser des affectations pertinentes.

En effet, l'optimisation de l'affectation en considérant uniquement la période de temps courante ne permet pas l'anticipation des mises à jour futures du système. Pour répondre à ce besoin, nous considérons des affectations sur plusieurs périodes de temps futures. Cette extension prend en compte l'évolution continue des flux de prospects Internet ainsi que la mise à jour régulière des quotas de commandes. Une telle approche conduit à optimiser les résultats de l'affectation en tirant profit de ces deux informations.

4. L'intégration de la connaissance des flux de prospects Internet dans l'optimisation multi-période : notre contexte est caractérisé par des flux de prospects Internet évoluant de manière continue et non déterministe au fil du temps. Afin d'améliorer le processus d'affectation de PdL, nous devons intégrer la connaissance concernant l'évolution de ces flux

dans le modèle d'optimisation multi-période. En effet, nous cherchons à articuler un modèle de prévision avec un modèle d'optimisation multi-période. Le défi consiste donc à définir les paramètres du modèle de prévision qui conviendront au fonctionnement du modèle de l'optimisation multi-période.

5. La généralisation des modèles : l'un des objectifs de cette thèse est la définition des modèles aussi génériques que possible afin de traiter tous les problèmes d'affectation partageant les mêmes caractéristiques que celles de PdL. Parmi ces problèmes analogues nous pouvons citer : la programmation des sessions dans les grandes conférences, le péage dans les autoroutes, les centres d'appel . . . etc. Par conséquent, nous présentons une extension des modèles proposés pour PdL afin d'avoir un modèle plus générique pouvant être appliqué dans d'autres contextes analogues à celui de PdL.

1.5 Conclusion

Dans ce chapitre, nous avons présenté le contexte général dans lequel s'inscrit notre travail. Il est caractérisé par une place de marché dont l'objectif principal est de satisfaire l'offre et la demande de prospects Internet.

Ainsi, nous avons commencé par introduire les caractéristiques de PdL, à savoir les prospects Internet, les commandes et le fonctionnement global de la place de marché. Par la suite, nous avons présenté la problématique ainsi qu'un aperçu des différents enjeux. Enfin, nous avons exposé les objectifs de la thèse et les différentes approches proposées pour y parvenir.

Dans le prochain chapitre, nous allons introduire les différents concepts et notions utilisés dans nos travaux en s'appuyant sur un état de l'art balayant à la fois les problèmes d'affectation mais ceux de la prévision également.

Chapitre 2

Etat de l'art

2.1 Introduction

Dans ce chapitre nous présentons une vue d'ensemble des problèmes d'optimisation et de prévision. En effet, nous présentons plusieurs classes de problèmes d'optimisation. Le problème que nous traitons dans le cadre de cette thèse fait parti de la dernière famille de problèmes d'optimisation présentée. Par la suite, nous abordons les problèmes de prévision en mettant l'accent sur les différentes typologies de ces problèmes afin de situer le problème de prévision de prospects Internet de Place des Leads par rapport à la littérature. Enfin, nous présentons des exemples d'application illustrant l'utilité de tels modèles.

2.2 Modèles d'optimisation

L'optimisation joue un rôle primordial dans la recherche opérationnelle. Elle a pour objectif de modéliser, analyser et résoudre des problèmes inhérents à la détermination des solutions satisfaisant un objectif quantitatif tout en respectant d'éventuelles contraintes. Cependant, dans certains contextes décisionnels, la prise en compte d'un objectif unique est insuffisante. L'optimisation multi-objectif constitue donc une branche de l'optimisation classique dont la particu-

larité consiste à réaliser l'optimisation simultanée de plusieurs objectifs concernant un même problème.

Nous nous sommes appuyés principalement sur [Sakarovitch, 1984; Ahuja et al., 1993; Gondran and Minoux, 2009; Teghem, 2012] afin de réunir ces quelques éléments essentiels de la recherche opérationnelle. Loin d'être exhaustive, cette section a pour objectif de rappeler les notions classiques utilisées dans cette thèse.

2.2.1 Problèmes d'optimisation

Un problème d'optimisation, noté $P(X, f)$, est caractérisé par un ensemble réalisable ou admissible X non-vide et une fonction objectif f qui associe un scalaire dans R à chaque élément x de l'ensemble X . Les éléments de X sont dits solutions réalisables. Résoudre le problème $P(X, f)$ revient à trouver parmi les solutions réalisables, une qui minimise ou maximise f , c'est-à-dire dans le cas d'un problème de minimisation, trouver une solution $x^* \in X$ telle que $f(x) \geq f(x^*)$ pour tout élément x dans X . Une telle solution est dite optimale et sera notée $x(X, f)$.

Nous nous contenterons de présenter les problèmes de minimisation, les problèmes de maximisation obéissent aux mêmes règles, à quelques changements près. Nous ne considérons ici que les problèmes qui admettent au moins une solution optimale. Un problème d'optimisation peut avoir plusieurs fonctions objectif, il s'agit alors d'un problème d'optimisation multi-critère ou multi-objectif. Ces problèmes seront abordés dans la section 2.2.4.

L'ensemble réalisable X est habituellement défini comme partie de R^n où n est un entier positif désignant la taille du problème. Les solutions réalisables peuvent alors être représentées comme des vecteurs dont les n composantes sont les variables du problème. L'ensemble X est délimité couramment par un système d'inégalités appelées contraintes du problème. Les contraintes sont construites à l'aide de combinaisons des variables, et permettent de caractériser les propriétés communes aux solutions de X afin de les distinguer parmi toutes les solutions de R^n . La description de l'ensemble X est donc implicite. Les programmes linéaires sont sans

doute les problèmes d'optimisation les plus connus. La fonction objectif et les contraintes de ces problèmes sont linéaires.

2.2.2 Problèmes d'optimisation discrète

Les problèmes d'optimisation discrète, par opposition à l'optimisation continue, forment une classe de problèmes d'optimisation particulièrement étudiée. Tout ou partie des variables de ce type de problèmes appartiennent à l'ensemble des entiers, autrement dit $X \subseteq Z^m \times R^{n-m}$, avec $0 \leq m \leq n$. Lorsque $m = n$, $P(X, f)$ est dit problème en nombres entiers, sinon il fait partie des problèmes d'optimisation mixtes en nombres entiers. Bien que l'ensemble réalisable soit plus restreint dans le cas de l'optimisation discrète, ces problèmes sont souvent plus difficiles que leurs versions continues.

Les ensembles réalisables des problèmes d'optimisation et des problèmes mixtes peuvent être infinis, et ceux des problèmes en nombres entiers sont au plus dénombrables. Nous présentons dans la section 2.2.3 les problèmes auxquels nous nous intéressons dans cette thèse, à savoir les problèmes d'optimisation combinatoire.

2.2.3 Problèmes d'optimisation combinatoire

L'optimisation combinatoire définit un cadre formel pour de nombreux problèmes de l'industrie, de la finance ou de la vie quotidienne. Les problèmes d'optimisation combinatoire sont habituellement définis comme une problématique de choix d'une meilleure alternative dans un ensemble très grand mais fini d'alternatives. En raison du très grand nombre d'alternatives pour ces problèmes, l'ensemble des alternatives, dit aussi ensemble de solutions réalisables, est défini en compréhension. En d'autres termes les solutions réalisables se distinguent par un ensemble de propriétés ou de conditions, dites aussi contraintes, qu'elles doivent toutes remplir. Une évaluation est associée à toute solution réalisable à l'aide d'une fonction dite fonction objectif. Résoudre un tel problème consiste donc à trouver une solution optimale, c'est-à-dire trouver une solution réalisable qui minimise ou maximise, selon le contexte, la fonction objec-

tif.

Les problèmes d'optimisation combinatoire sont des problèmes d'optimisation dont les ensembles réalisables sont finis mais combinatoires. Aussi, le nombre de solutions réalisables des problèmes combinatoires augmente exponentiellement en fonction de la taille du problème, et c'est ce qui exclut des méthodes de résolution basées sur l'énumération de toutes les solutions réalisables.

Généralement l'ensemble admissible X d'un problème d'optimisation combinatoire est défini comme un sous-ensemble de l'ensemble des parties d'un ensemble fini d'éléments E , i.e. $E = \{e_1, \dots, e_n\}$ et $X \subseteq 2^E$. Un problème d'optimisation combinatoire peut aussi être formalisé comme un problème d'optimisation en variables binaires. En associant une variable x_i à tout élément e_i de E , tout sous-ensemble de E peut être représenté par un vecteur $x = (x_1, \dots, x_n)$, avec $x_i = 1$ si e_i appartient au sous-ensemble et $x_i = 0$ sinon.

Ce qui fait l'élégance des problèmes d'optimisation combinatoire, c'est qu'en plus des formulations en programmes mathématiques sur lesquelles nous nous basons principalement dans le travail présenté dans cette thèse, il est souvent possible de les formuler comme des problèmes de la théorie des graphes [Burkard et al., 2012].

2.2.4 Problèmes d'optimisation combinatoire multi-objectif

Nous définissons dans cette section le problème de recherche de solutions de bon compromis pour un problème d'optimisation combinatoire multi-objectif. Dans un premier temps nous rappelons les notions préliminaires de l'optimisation multi-objectif, ensuite nous précisons la notion de compromis et nous présentons la formalisation de ces problèmes.

Un problème d'optimisation combinatoire multi-objectif peut s'écrire sous la forme du problème (2.1), où p représente le nombre d'objectifs et X un ensemble combinatoire.

$$\min_{x \in X} Z(x) = (z_1(x), \dots, z_p(x)) \quad (2.1)$$

Afin de ne pas confondre les solutions réalisables et leurs images dans R^p , il est d'usage en

optimisation multi-objectif de distinguer l'espace des décisions X , qui contient l'ensemble admissible X , et l'espace des objectifs $Z \subseteq R^p$ qui lui contient l'image de l'ensemble admissible X dans R^p . A toute solution réalisable x est associé un vecteur critère $z(x) = (z_1(x), \dots, z_p(x))$ qui correspond à son image dans Z , telle que z_k est une fonction objectif pour tout k dans $\{1, \dots, p\}$. L'ensemble des points correspondant aux images des solutions admissibles dans Z est noté $Z = \{z(x) : x \in X\}$.

En raison de la nature conflictuelle des objectifs, il n'y a généralement pas une solution réalisable qui minimise simultanément tous les objectifs. L'optimalité dans un contexte multi-objectif est basée sur la notion de dominance et d'efficacité au sens de pareto dont nous rappelons le principe dans les définitions 2.1 et 2.2.

Définition 2.1 Soient $z, z' \in R^p$ deux vecteurs.

1. z domine largement z' , noté $z \leq z'$, si $z_k \leq z'_k$ pour tout k dans $\{1, \dots, p\}$.
2. z domine z' , noté $z \leq z'$, si $z_k \leq z'_k$ pour tout k dans $\{1, \dots, p\}$ et $z \neq z'$.
3. z domine strictement z' , noté $z < z'$, si $z_k < z'_k$ pour tout k dans $\{1, \dots, p\}$.

Définition 2.2 Soient z, z' deux points.

1. Un point z dans Z est (faiblement) non-dominé s'il n'existe pas un autre point z' dans Z tel que z' domine (strictement) z .
2. Une solution x dont l'image $z(x)$ dans Z est (faiblement) non-dominé est dite (faiblement) efficace.
3. On notera Z_N et Z_{WN} les ensembles de points non-dominés et faiblement non-dominés. Il est facile de voir que $Z_N \subset Z_{WN}$. Les ensembles des solutions efficaces et faiblement-efficaces seront notés respectivement X_E et X_{XE} .

2.2.4.1 Solutions de bon compromis

Un point non-dominé est l'image d'une ou de plusieurs solutions réalisables, tel que la performance d'un objectif ne peut être améliorée sans détériorer au moins un autre objectif. Les solutions efficaces ne sont donc pas comparables entre elles et représentent de bons compromis potentiels. Cependant, un décideur peut vouloir exprimer des préférences qui lui sont propres. Ces préférences seraient alors de nature à privilégier un certain type de solutions, par exemple

il est d'usage de préférer des solutions équilibrées, c'est-à-dire des solutions avec des valeurs moyennes sur les objectifs.

La solution des problèmes de la programmation mathématique (PM) avec une seule fonction objectif est une tâche difficile. La sortie est la solution optimale et toutes les informations pertinentes (prix, coûts, ... etc). D'autre part, la solution des problèmes de la programmation mathématique multi-objectif (PMMO) est une tâche beaucoup plus complexe que la précédente. En effet, dans la PMMO, il y'a plus d'une fonction objectif et, en général, il n'y a pas de solution optimal unique qui optimise simultanément toutes les fonctions objectif. Dans ces cas, les décideurs sont à la recherche de la solution "la plus préférée", contrairement à la solution optimale. Dans un PMMO, le concept d'optimalité est remplacé par celui de l'optimalité de Pareto ou l'efficacité. Les solutions Pareto optimales (ou efficaces, non-dominées) sont les solutions qui ne peuvent pas être améliorées dans une fonction objectif sans altérer leur performance dans au moins une partie du reste. La définition mathématique de la solution efficace est la suivante (sans perte de généralité supposer que toutes les fonctions objectif $f_i, i = 1, \dots, p$ sont pour la maximisation) : une solution réalisable x d'un problème de PMMO est efficace s'il n'y a pas une autre solution possible x' tel que $f_i(x') \geq f_i(x)$ pour tout $i = 1, 2, \dots, p$ avec au moins une inégalité stricte. Chaque solution efficace correspond à un vecteur non-dominé ou non améliorable dans l'espace de critères. Si nous remplaçons la condition $f_i(x') \geq f_i(x)$ avec $f_i(x') > f_i(x)$, nous obtenons les solutions faiblement efficaces. Les solutions faiblement efficaces ne sont généralement pas menées dans un PMMO, parce qu'elles peuvent être dominées par d'autres solutions efficaces. Le décideur rationnel est à la recherche de la solution la plus préférée parmi les solutions optimales de Pareto du PMMO.

2.2.4.2 Approche de résolution.

Selon [Hwang and Masud, 1979], les méthodes pour résoudre les problèmes de PMMO peuvent être classées en trois catégories, en fonction de la phase dans laquelle le décideur intervient dans le processus de prise de décision en exprimant ses préférences : les méthodes a priori, les méthodes interactives et les méthodes a posteriori ou de génération. Dans des mé-

thodes a priori, le décideur exprime ses préférences avant le processus de solution (par exemple fixer des poids aux fonctions objectif). La critique envers les méthodes a priori est qu'il est très difficile pour le décideur de savoir à l'avance et d'être en mesure de quantifier avec précision ses préférences. Dans les méthodes interactives, les phases du dialogue avec le décideur sont échangées avec les phases de calcul. Le processus converge habituellement à la solution la plus préférée, après quelques itérations. L'inconvénient est que le décideur ne voit jamais l'image entière (l'ensemble de Pareto) ou un rapprochement de celui-ci. Par conséquent, la solution la plus préférée est "la plus préférée" par rapport à ce qu'il a vu. Dans les méthodes a posteriori, les solutions efficaces du problème (toutes ou une représentation suffisante) sont générées et ensuite le décideur intervient afin de sélectionner parmi elles, celle la plus préférée.

Les méthodes de génération sont les moins populaires en raison de leur effort de calcul (le calcul des solutions efficaces est généralement un processus qui prend du temps) et le manque de logiciels largement disponibles [Hwang and Masud, 1979]. Cependant, elles présentent certains avantages importants. Le processus de solution est divisé en deux phases indépendantes : d'abord, la génération des solutions efficaces et par la suite la participation du décideur lorsque toutes les informations sont sur la table. Par conséquent, les méthodes de génération sont favorables dans le cas où le décideur, est peu disponible et l'interaction avec lui est difficile, car il n'est impliqué que dans la deuxième phase, ayant à portée de main toutes les alternatives possibles (l'ensemble de Pareto ou une représentation adéquate). D'ailleurs, le fait qu'aucune des solutions potentielles n'a été laissée à découvrir, renforce la confiance du décideur sur la décision finale.

Pour un type particulier de problèmes de PMMO (surtout les problèmes linéaires) de petite et moyenne taille, il y'a des méthodes qui calculent l'ensemble de solutions efficaces [Steuer, 1986; Miettinen, 1999; Mavrotas and Diakoulaki, 2005]. Cependant, dans le cadre de la thèse, nous n'allons pas nous intéresser à uniquement un type particulier de de problèmes de PMMO mais au cas général, où des problèmes importants de PMMO peuvent être abordés. En général, la méthode de génération la plus utilisée est la méthode de ϵ - *constraint*. Cette méthode peut fournir un échantillon représentatif de l'ensemble de Pareto, qui dans la plupart des cas est

suffisant. Dans ce contexte, nous proposons l'utilisation de la méthode $\varepsilon - constraint$ afin de répondre au problème d'optimisation bi-objectif posé par PdL que nous avons abordé dans la section 3.6.

La méthode $\varepsilon - constraint$. Considérons le problème PMMO suivant :

$$\left\{ \begin{array}{l} \max \quad (f_1(x), f_2(x), \dots, f_p(x)) \\ s.t. \\ \quad \quad \quad x \in S \end{array} \right. \quad (2.2)$$

Où x est le vecteur des variables de décision, $f_1(x), \dots, f_p(x)$ sont les p fonctions objectif et S est la région réalisable.

Dans la méthode $\varepsilon - constraint$, nous optimisons l'une des fonctions objectif en utilisant les autres fonctions objectif comme contraintes, en les intégrant dans la partie contrainte du modèle comme ci-dessous [Chankong and Haimes, 1983; Cohon, 1978].

$$\left\{ \begin{array}{l} \max f_1(x) \\ s.t. \\ \quad \quad \quad f_2(x) \geq e_2 \\ \quad \quad \quad f_3(x) \geq e_3 \\ \quad \quad \quad \dots \\ \quad \quad \quad f_p(x) \geq e_p \\ \quad \quad \quad x \in S \end{array} \right. \quad (2.3)$$

Les solutions efficaces du problème sont obtenues par une variation paramétrique dans les fonctions objectif (e_i).

D'une part, plusieurs versions de la méthode $\varepsilon - constraint$ sont apparues dans la littérature et tentent d'améliorer ses performances ou de l'adapter à un type spécifique de problèmes [Ehrgott and Ryan, 2002; Hamacher et al., 2007; Laumanns et al., 2006]. D'autre part, une étape essentielle vers une plus grande pénétration des méthodes de génération dans les applications est

de fournir les codes appropriés pour résoudre les problèmes bi-objectif qui sont largement utilisés par les personnes dans l'ingénierie, l'économie, l'agriculture, ... A titre d'exemple, [Mavrotas, 2009] présente un effort pour mettre en oeuvre efficacement la méthode $\epsilon - constraint$ pour produire les solutions optimales de pareto pour la programmation mathématique multi-objectif. Il propose une nouvelle version de la méthode (la méthode $\epsilon - constraint$ augmentée - AUGMECON) qui évite la production de solutions faiblement efficaces sur la surface de pareto et accélère l'ensemble du processus, en évitant les itérations redondantes. Dans le cadre de cette thèse, nous implémentons la méthode $\epsilon - constraint$ afin de résoudre le modèle d'optimisation bi-objectif que nous proposons dans la section 3.6.2.

2.2.5 Problèmes d'affectation

Le problème d'affectation joue un rôle important dans cette thèse. En effet, la problématique principale traitée dans le cadre de cette thèse est l'optimisation de l'affectation des prospects Internet aux clients. C'est la raison pour laquelle nous dédions cette section au rappel de quelques définitions autour de ce problème. Nous rappelons également différentes formulations du problème d'affectation.

Nous nous appuyons dans cette section essentiellement sur l'ouvrage spécialisé dans le problème d'affectation [Burkard et al., 2012].

2.2.5.1 Formulations

Le problème d'affectation est un problème classique de recherche opérationnelle et d'optimisation combinatoire. Informellement, nous entendons par problème d'affectation le problème qui consiste à associer chaque élément d'un ensemble de n items à un seul élément d'un autre ensemble de n items avec un coût minimal [Mckeown and Workman, 1976; Soumis et al., 1980; Campell and Langevin, 1995; Dessouky and Kijowski, 1997; Pentico, 2007]. Un problème d'affectation standard ne tient compte que du coût ou du profit dans la formulation du problème. Cependant, les différentes applications réelles ont exprimé la nécessité d'intégrer d'autres aspects [Campbell and Diaby, 2002; Suer and Bera, 1998]. En outre, une générali-

sation du problème d'affectation bi-objectif avec relaxation linéaire a été étudiée dans [Karsu and Azizoglu, 2014].

Plus formellement, il existe plusieurs formulations du problème d'affectation, à savoir, la formulation par la programmation linéaire, la formulation en théorie des graphes, la formulation en problème de flot maximum à coût minimum. Parmi ces trois formulations nous en retenons celle par la programmation linéaire que nous utilisons dans le cadre de nos travaux.

En effet, la première est la formulation par la programmation linéaire (2.4).

$$\left\{ \begin{array}{l} \min \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \\ s.t. \\ \sum_{j=1}^n x_{ij} = 1 \quad i = 1, \dots, n \\ \sum_{i=1}^n x_{ij} = 1 \quad j = 1, \dots, n \\ x_{ij} \geq 0 \quad i, j = 1, \dots, n \end{array} \right. \quad (2.4)$$

Une solution réalisable pour le problème d'affectation dans ce cas est un vecteur ou une matrice carrée de n^2 éléments, où :

$$x_{ij} = \begin{cases} 1 & \text{si } i \text{ est associé à } j \\ 0 & \text{sinon} \end{cases}$$

2.2.5.2 Quelques algorithmes pour le problème d'affectation

Différents algorithmes comprenant la programmation linéaire standard [Balinski, 1986; Barr et al., 1977; Hung and Rom, 1980; McGinnis, 1983], l'algorithme hongrois [Kuhn, 1955], les réseaux de neurones [Eberhardt et al., 1991], et l'algorithme génétique [Avis and Devroye, 1985] ont été développés pour trouver des solutions de manière efficace. Dans ce contexte, [Burkard et al., 2012] présentent une comparaison des différentes méthodes de résolution du problème d'affectation, allant du premier algorithme exponentiel proposé pour ce problème

jusqu'aux dernières implémentations en programmation parallèle. S'il fallait faire une classification de la plus grande partie de ces méthodes, nous pouvons distinguer deux types d'approches. Une première classe de méthodes vise à construire une solution réalisable puis à l'améliorer progressivement jusqu'à l'obtention d'une preuve d'optimalité. La seconde classe de méthodes est constituée de celles qui construisent une solution partielle et qui l'augmentent pas à pas, jusqu'à l'affectation de tous les éléments de l'ensemble U à ceux de l'ensemble V .

Tous les algorithmes de recherche des k meilleures solutions ($k - best$) pour le problème d'affectation sont basés sur des méthodes arborescentes, où à chaque noeud un problème d'affectation avec des contraintes supplémentaires est résolu (voir [Murty, 1968; Chegireddy and Hamacher, 1987; Miller et al., 1997; Pascoal et al., 2003; Pedersen et al., 2008]).

Dans le cadre de la thèse, nous avons utilisé un solveur commercial afin de résoudre le problème de l'affectation de l'offre et la demande chez Place des Leads. En effet, notre objectif était de proposer un cadre formel modélisant le problème de l'adéquation offre/demande de prospects Internet. Ainsi, nous nous sommes focalisés sur la partie modélisation en considérant :

1. Les différentes caractéristiques du problème d'affectation de PdL, à savoir : les quotas horaires, journaliers et mensuels des commandes et des clients, le nombre de reventes des prospects Internet (vente exclusive ou pas), ... etc.
2. l'aspect bi-objectif du problème d'affectation posé.
3. l'aspect temps réel des affectations s'effectuant toutes les tranches horaires dans un contexte caractérisé par des flux continus de prospects Internet.

Afin de résoudre les modèles d'affectation proposés, nous avons utilisé le solveur *CPLEX* qui est un solveur de programmation linéaire permettant de résoudre des problèmes d'optimisation linéaire. Nous avons utilisé *CPLEX* via une API, permettant de l'interfacer à l'aide du langage de programmation Java.

2.3 Modèles de prévision

Le contexte de la place de marché PdL est caractérisé par des flux continus de prospects Internet faisant évoluer de façon continue et non déterministe l'ensemble de l'offre disponible. La connaissance de ces flux permet d'améliorer le processus d'affectation de PdL. Notre objectif est de proposer un modèle de prévision de ces flux de prospects Internet. Nous verrons dans la suite de ce travail qu'une articulation d'un modèle de prévision et un modèle d'affectation est nécessaire afin de répondre au problème de PdL. Nous verrons également que cette articulation posera des problèmes au niveau des choix des paramètres du modèle de prévision. Ces paramètres seront ajustés en adéquation avec le fonctionnement du modèle d'optimisation.

La prévision recouvre un ensemble de méthodes très diverses qui ont en commun de chercher à réduire l'incertitude liée à la non-connaissance du futur. Dans le cadre de cette thèse, nous nous intéressons à prévoir l'évolution des flux de prospects Internet chez Place des Leads. L'enjeu est important car la prise de décision risque d'être coûteuse si les prévisions sur lesquelles elle est fondée sont infirmées.

Tout comme les besoins, les méthodes sont multiples et d'inspiration très différentes. Deux caractères permettent de distinguer les types de prévision :

1. La période : le court, le moyen ou le long terme.
2. La technique : qualitative (appel au jugement humain) ou quantitative (formalisation et estimation statistique).

Les critères sont indépendants les uns des autres : il est possible d'élaborer une prévision à court terme à l'aide d'une technique qualitative ou bien une prévision à long terme par une technique quantitative ... etc. Tous ces éléments différencient nettement les méthodes de prévision, qui seront retenues en fonction des besoins d'information prévisionnelle. Dans le contexte de PdL, nous faisons face à une prévision à court terme par une technique quantitative. En effet, PdL reçoit une quantité importante de prospects Internet en temps réel qui seront affectés toutes les minutes à ses clients. Ainsi, nous avons besoin de prévoir les arrivées de

ces prospects Internet sur des courtes durées afin de pouvoir intégrer cette prévision dans le système d'affectation de l'offre à la demande de prospects Internet.

Pourquoi des modèles ?

Nombreux sont ceux qui élaborent des prévisions de toutes sortes. Peu de personnes cependant reconnaissent qu'une certaine structure logique, ou modèle, est implicite dans toute prévision, ou bien dans toute analyse d'un système social ou physique. Considérez, par exemple, un agent de change qui vous dit que l'indice de la bourse sera à la hausse l'année prochaine. Il peut faire cette prévision parce qu'il a vu l'indice augmenter durant les dernières années, et pense que, quelles que soient les raisons qui le faisaient croître dans le passé, celles-ci continueront à agir dans le futur.

Par ailleurs, il peut penser que l'indice augmentera l'année prochaine parce qu'il est en relation avec un ensemble de variables économiques et politiques à travers des relations complexes. Par exemple, l'agent de change imagine que l'indice est lié d'une certaine façon au produit national brut, au résultat de l'entreprise et aux taux d'intérêt, si bien qu'à partir d'hypothèses très probables concernant l'évolution future de ces variables, il est amené à envisager comme vraisemblable une hausse de l'indice.

Si un mot doit décrire cette méthode de prévision, nous dirons qu'il s'agit d'intuition, bien que les modalités de raisonnement diffèrent dans les deux cas cités plus haut. L'agent de change ne dira sûrement pas que la prévision a été établie à partir d'une modélisation des cours de bourse : aucune équation n'a été écrite. Néanmoins, il a établi une sorte de modèle implicite. À titre d'exemple, si l'agent de change a établi sa prévision optimiste sur l'indice à partir des taux de croissance du passé, il a bâti un modèle d'extrapolation de séries chronologiques, que nous présentons dans ce qui suit.

Même inconsciemment, le prévisionniste intuitif construit implicitement des modèles. Une question se pose alors : pourquoi les construire explicitement, les estimer, et les tester statistiquement ? Plusieurs raisons incitent à cette démarche de modélisation.

Tout d'abord, cela force l'individu à établir clairement et à estimer les inter-relation sous-jacentes. Ensuite, la confiance aveugle dans l'intuition peut amener à l'ignorance de liaisons importantes ou à leur mauvaise utilisation. Enfin, il est nécessaire de fournir en même temps que la prévision une mesure de la confiance que l'utilisateur peut avoir en celle-ci, c'est-à-dire la précision que l'on peut en attendre. L'utilisation de méthodes purement intuitives exclut toute mesure quantitative de fiabilité d'une prévision.

Dans le but de réunir ces éléments essentiels de la prévision, nous nous sommes appuyés principalement sur [Box et al., 2008; Bourbonnais and Usunier, 2007; Makridakis et al., 1997].

2.3.1 Pourquoi élaborer des prévisions pour le processus d'affectation ?

Cette question peut sembler triviale, cependant elle mérite réflexion et s'avère plus complexe qu'il n'y paraît dans sa réponse. En effet, la prévision conditionne l'optimisation dans le sens où la prévision de l'offre est à la base de l'optimisation du système de vente de l'entreprise. Quel que soit le type d'organisation d'une entreprise et son degré d'intégration dans le pilotage des flux, la prévision de l'offre et la fiabilité attendue de cette prévision présentent deux éléments indispensables afin d'optimiser au mieux le système de vente de l'entreprise. La fiabilité attendue de cette prévision, ou la crédibilité de la prévision, permet de dimensionner de manière optimale les affectations de l'offre dans le système de vente de l'entreprise. Dans ce contexte, PdL a besoin d'établir de bonnes prévisions des flux d'offres de prospects Internet afin de parvenir à effectuer des affectations optimales. En effet, le système de prévision des arrivées de prospects Internet sera intégré dans celui de l'affectation de l'offre à la demande afin de tirer profit de l'évolution des flux de prospects Internet. Ainsi, l'optimisation des affectations de prospects Internet va dépendre de la qualité de la prévision de ces derniers.

2.3.2 Ce qui caractérise un problème de prévision

La première approche de la prévision consiste à en mesurer les spécificités. Pour envisager une typologie des problèmes de prévision en terme d'horizon ou de secteur, il est utile de dégager

quelques traits qui différencient fondamentalement le secteur d'activité et l'horizon.

L'approche est très dépendante du secteur d'activité : on ne prévoit pas des livraisons de ciment par les mêmes méthodes que des ventes de savons. Les causalités économiques sous-jacentes sont différentes selon : (i) le secteur est plus ou moins en amont dans le circuit industriel donc plus ou moins proche de la demande finale, (ii) le produit est stockable ou pas, qu'il donne lieu à un marché de renouvellement (télévision) ou pas (acier) et (iii) le produit est standardisé ou pas. Tous ces éléments jouent un rôle dans la définition d'une typologie de trois grands secteurs, à savoir : la prévision en milieu industriel, le secteur des biens de consommation durables et les produits de grande consommation, qui sont présentés dans la section 2.3.3.

Le tableau 2.1 indique les principales utilisations, indépendamment du secteur d'activité de l'entreprise, c'est à dire suivant l'horizon de prévision.

Application	Horizon
Prévision de commandes et de livraisons	3 à 6 mois
Prévision des ventes	6 à 12 mois
Prévision budgétaire	15 à 18 mois

Tableau 2.1 — Domaines d'utilisation de la prévision.

Ce tableau ne donne cependant qu'un aperçu très simplifié des horizons. Ceux-ci peuvent également être très courts[Mélard and Carbon, 2008] :

1. Prévision par tranche horaire à la journée ou à quelques journées (centre d'appels, fréquentation des péages d'autoroutes, passages aux caisses d'hypermarché ... etc).
2. Prévisions journalières pour les activités de service (restauration, hôtellerie, cinéma ... etc).

2.3.3 Typologie des problèmes de prévision

L'approche de prévision est très dépendante du secteur d'activité de l'entreprise, nous pouvons distinguer trois grands secteurs, à savoir :

1. Le domaine des biens industriels intermédiaires (acier, produits chimiques, matériaux de construction, ... etc) : qui dépendent étroitement de l'environnement économique général, et de l'évolution de la demande dans de grandes branches (bâtiment, automobile ... etc).
2. Le domaine des produits de grande consommation : est à l'extrême opposé, l'influence de la conjoncture générale est relativement faible, ceci est largement compensé par une forte sensibilité à l'environnement concurrentiel.
3. Les biens de consommation durables (automobile, électro-ménager, télévision, ... etc) : ils se situent à mi-chemin. Sensibles à la fois à la conjoncture économique générale et à des variables marketing (efforts promotionnels, effet de mode, ... etc).

Par ailleurs, d'autres problèmes de prévision sont liés à des horizons courts et des activités particulières, à savoir :

1. La prévision d'activités quotidiennes dans le domaine des services et des produits à date de péremption courte.
2. La prévision pour les articles à durée de vie limitée, particulièrement les articles de collection dans le domaine du textile.
3. La prévision de consommation pour les articles à très faible rotation ou à demande intermittente.

C'est dans ce contexte à horizon court et activité particulière où se place le problème de prévision des flux de prospects Internet chez PdL. Le rythme et le type d'activité de PdL justifient une prévision à court terme. En effet, PdL exerce une activité quotidienne dans sur des produits à date de péremption courte. En d'autres termes, PdL reçoit une quantité importante de prospects Internet en temps réel. Ces derniers sont ensuite intégrés dans le système de vente de PdL qui effectue, toutes les dix minutes, des affectations de l'offre à la demande des prospects Internet. Ainsi, nous devons prévoir par tranche horaire. En outre, la valeur des prospects Internet est définie principalement par leurs âges (plus le prospect Internet est jeune plus sa vente est facile). Par conséquent, l'exactitude de la prévision est très importante car les prospects Internet possèdent une durée de vie limitée. En effet, ils peuvent expirer par rapport à certaines

commandes exigeant des prospects Internet très jeunes ou encore ils ne sont plus d'actualité car les internautes générant ces prospects Internet suite à un besoin d'un produit particulier l'ont bien retrouvé ailleurs (voir la section 1.2.1).

2.3.4 Ce qui caractérise une solution de prévision

La démarche qu'implique la construction d'un système de prévision enchaîne un certain nombre d'opérations. Lorsque se pose un problème de prévision dans une entreprise, il convient de s'interroger sur un certain nombre d'éléments qui vont conditionner le choix de la solution à savoir les données à prévoir et la méthodologie prévisionnelle que nous détaillerons dans ce qui suit.

2.3.4.1 Les données à prévoir

Le choix des données historiques servant de base à la prévision est important. La profondeur de l'historique (le nombre d'observations utiles pour calculer la prévision) dépend de la périodicité des données (annuelle, trimestrielle, mensuelle, ... etc) et ne doit être :

1. Ni trop courte, par exemple, pour calculer des coefficients saisonniers robustes, il convient d'avoir au minimum trois ans.
2. Ni trop importante (cinq ans est un maximum) au risque de créer des rigidités ou des comportements d'inertie.

2.3.4.2 Horizon et périodicité de la prévision

Le besoin de réactiver de l'entreprise détermine la périodicité des données historiques et de la prévision :

1. Pour des produits industriels ou de consommation courante, la périodicité mensuelle s'avère suffisante et l'horizon de prévision peut être compris entre six et dix-huit mois. C'est le cas le plus classique.

2. Dans le domaine des produits frais, la prévision doit être calculée de manière hebdomadaire afin d'être plus réactive, la maille mensuelle n'est alors plus suffisante. L'horizon de prévision peut aller jusqu'à quinze semaines.
3. Enfin, pour certains secteurs de services (centres d'appels, péages d'autoroutes, ...) ou pour des produits ultra-frais, la prévision journalière (voire horaire, qui est le cas pour la prévision des prospects Internet chez PdL) s'impose. Les produits sont stockables peu de temps et la demande journalière est variable. La détermination la plus exacte possible du rythme d'activité quotidien est essentielle, l'horizon de prévision est limité à quelques jours (une trentaine en pratique).

2.3.4.3 Quelle(s) méthode(s) choisir ?

La deuxième étape est la recherche d'une méthodologie prévisionnelle. Deux principaux types de techniques sont possibles :

1. Des méthodes extrapolatives, appelées endogènes, consistant à dégager dans la série elle-même un certain nombre de composantes que l'on peut prolonger dans le futur.
2. Des méthodes explicatives, appelées exogènes, recherchant des liaisons entre la série à prévoir et des séries explicatives, à travers les calculs de corrélation.

2.3.5 Principales méthodes de prévision

Cette section a pour finalité de passer en revue les principales méthodes de prévision. Nous fournissons pour chacune d'entre elles des définitions minimales nécessaires à sa compréhension. Le lecteur intéressé par plus de détails se reportera à l'ouvrage [Bourbonnais and Usunier, 2007].

2.3.5.1 Les méthodes fondées sur le lissage exponentiel

La première catégorie des modèles de prévision représente les méthodes extrapolatives fondées sur le principe de lissage exponentiel. Les méthodes de lissage présentent certains principes

de base, à savoir : la dévalorisation croissante de l'information avec l'âge, la synthétisation des informations et la réactualisation permanente des paramètres moyennant des calculs relativement simples. À titre d'exemple, le lissage de Holt-Winters présente l'avantage d'intégrer une composante saisonnière et donc de réaliser le calcul de la prévision en un seul traitement. De ce fait, le lissage de Holt-Winters est le plus couramment employé dans les progiciels des prévisions de ventes.

2.3.5.2 Box-Jenkins et modèle causal

La méthode de Box et Jenkins est une technique d'extrapolation par référence à une loi de reproduction du phénomène étudié. L'information permettant de mettre en évidence le processus est contenue dans la série chronologique elle-même, sans apport externe, d'où le nom de prévision endogène.

Les modèles causaux ou modèles explicatifs cherchent à expliquer les fluctuations de la série des ventes par référence à des facteurs explicatifs externes représentatifs soit de la politique marketing de la firme, soit de la demande. D'où le nom de prévision exogènes.

Dans le cadre de nos travaux et afin de mettre en oeuvre le système de prévision que nous allons proposer dans la section 5.3, nous utilisons l'ensemble des modèles *ETS* implémenté dans le logiciel *R* que nous détaillons dans ce qui suit.

2.3.6 L'ensemble des modèles *ETS* au sein du logiciel *R*

2.3.6.1 Origines et utilisation du logiciel *R*

Le logiciel *R* [Micheaux et al., 2014] est un logiciel de statistique créé par Ross Ihaka et Robert Gentleman. Il est à la fois un langage informatique et un environnement de travail.

C'est aussi un outil très puissant et très complet, particulièrement bien adapté pour la mise en oeuvre informatique de méthodes statistiques. Le logiciel *R* est particulièrement performant pour la manipulation de données, le calcul et l'affichage de graphiques. Il possède :

1. Des procédures efficaces de traitement des données et des capacités de stockage de ces données.
2. Une suite d'opérateurs pour des calculs sur des tableaux et en particulier sur des matrices.
3. Une vaste et cohérente collection de procédures statistiques pour l'analyse de données.
4. Un langage de programmation simple et efficace intégrant les conditions, les boucles, la récursivité, et des possibilités d'entrée-sortie.

R est un logiciel dans lequel de nombreuses techniques statistiques modernes et classiques ont été implémentées [Shumway and Stoffer, 2010]. Les méthodes les plus courantes permettant de réaliser une analyse statistique sont enchâssées directement dans le coeur du système. À titre d'exemple :

1. Analyse de la variance.
2. Méthodes de régression linéaire (simple et multiple).

Plus de 80 livres fournissent des documentations sur le logiciel *R*. Entre autres [Dalgaard, 2008; Albert, 2009; Foulkes, 2009; Wood, 2006; Faraway, 2005].

2.3.6.2 L'ensemble des modèles *ETS*

ETS (ExponenTial Smoothing) est un ensemble de modèles de prévision. Afin d'appliquer l'un de ces modèles de prévision regroupés dans *ETS*, nous faisons appel à la méthode *ets(forecast)*. Cette dernière retourne le modèle de prévision le plus adéquat à notre série chronologique. En effet, *ets(y)* où *y* est un argument représentant un vecteur numérique ou une série chronologique, est basée sur la classification des méthodes comme décrit dans [Hyndman et al., 2008]. La méthodologie est complètement automatisée. Le seul agrument demandé pour la méthode *ets(y)* est la série chronologique. Ensuite, le modèle est choisi automatiquement s'il n'est pas spécifié par l'utilisateur. En effet, l'environnement *ETS* fournit un moyen automatique afin de choisir le modèle de prévision le plus adapté à la série chronologique fournie. La méthode *ets(y)* fonctionne extrêmement bien sur les données M3-compétition [Hyndman et al., 2002].

2.4 Exemples d'applications analogues à celle de PdL

Bien que dans le cadre de cette thèse nous nous intéressons spécifiquement au problème d'affectation chez PdL, les modèles que nous proposons peuvent être adaptés à plusieurs autres contextes analogues à celui de PdL. Le contexte de PdL est principalement caractérisé par : (i) l'établissement d'une correspondance en temps réel entre l'offre et la demande afin d'optimiser deux objectifs, (ii) cette correspondance s'effectue en temps réel et se répète plusieurs fois par jour, ce qui permet d'envisager un système d'optimisation multi-période et (iii) l'intégration de la prévision des flux de prospects Internet dans le système d'optimisation multi-période.

Dans cette section, nous allons aborder quelques problèmes existants dans la littérature, en comparant leurs caractéristiques d'un point de vue contexte et positionnement à celles du problème de PdL. Cette comparaison a pour objectif de démontrer l'applicabilité de nos travaux dans plusieurs autres contextes analogues à celui de PdL.

2.4.1 Le problème d'affectation multi-objectif à l'université Davidson : un cas d'étude

[Chartier et al., 2014] présente un programme linéaire multi-objectif en nombres entiers qui attribue des étudiants volontaires pour présenter des cours aux classes participantes dans les écoles locales. L'affectation d'une classe à un étudiant est basée sur sa disponibilité pour l'heure ainsi que plusieurs autres facteurs, y compris les préférences des étudiants concernant les déplacements et les partenaires ainsi que l'objectif de l'institution de créer des groupes d'étudiants divers.

Pour commencer le projet, le bureau de communication a fourni les détails suivants concernant le programme. Chaque étudiant rejoint une équipe de 3-5 volontaires. Chaque équipe visite la même salle de classe (même jour et même heure) pendant trois semaines et présente une leçon différente à chaque visite. Par exemple, une équipe peut visiter sa classe affectée le Mercredi à 09h00 la première et la troisième semaine d'Avril et la première semaine de Mai. Les salles de classe sont en dehors du campus et les cours durent environ 50-60 minutes. En outre, le

temps de trajet doit être considéré lors de l'affectation des étudiants aux salles de classe. En règle générale, le bureau de communication a recommandé l'attribution de 25 minutes avant et après un intervalle d'affectation puisque les écoles varient en distance de 5 minutes à pied à 20 minutes en voiture à partir du campus universitaire.

En particulier, l'université de Davidson a collecté les données additionnelles suivantes au sujet de leurs étudiants volontaires, qui ont été intégrées dans les objectifs subsidiaires. Lors de l'inscription, un étudiant remplit un formulaire écrit à la main, qui demande les informations suivantes :

1. Quelles sont les heures auxquelles vous êtes disponibles pour enseigner et se déplacer vers et depuis l'école locale ?
2. Quelles heures préféreriez-vous pour enseigner et se déplacer vers et depuis l'école locale ?
3. Avez-vous un véhicule ?
4. Avez-vous un partenaire privilégié avec lequel vous souhaitez travailler ?
5. Quel est votre sexe ?
6. Quel est votre cours ?

Le programme de service communautaire a cherché à maximiser les quatre objectifs suivants, classés du plus au moins importants :

1. Le nombre d'étudiants volontaires pouvant participer.
2. Le nombre d'étudiants volontaires pouvant participer à leurs heures spécifiées préférées ?
3. Le nombre d'étudiants placés dans le même groupe que leurs partenaires spécifiés.
4. La diversité de chaque groupe en termes de sexe et cours.

En examinant le cadre d'étude du problème d'affectation de l'université Davidson, nous constatons qu'il est analogue à celui de PdL. En effet, le nombre d'étudiants et de classes ne sont pas égaux, c'est le cas pour PdL : nous n'avons pas le même nombre de commandes et

de prospects Internet. Une autre caractéristique du problème d'affectation de l'université Davidson est qu'il est possible et probable que l'affectation optimale ignore certains étudiants et certaines classes, c'est encore le cas pour PdL : toutes les commandes et tous les prospects Internet ne sont pas pris en compte dans l'affectation optimale. De plus, dans les deux contextes, nous ne considérons pas seulement un seul objectif comme c'est le cas dans le problème d'affectation standard, mais de multiples objectifs à optimiser. Enfin, l'université Davidson met l'accent sur certaines spécificités telles que chaque groupe possède un membre avec un véhicule pour le transporter à sa classe attribuée, de même, les clients de PdL précisent des quots horaires, journaliers et mensuels à respecter, ils peuvent aussi caractériser leurs commandes en effectuant un filtre sur la typologie des prospects Internet comme l'âge, la catégorie professionnelle ainsi que la vente exclusive ou pas.

Bien évidemment, il existe de nombreux autres contextes semblables au notre, à savoir, l'affectation des équipages. Il s'agit de regrouper les gens afin de les affecter à des missions spécifiques. A titre d'exemple, un service de nettoyage ou d'aménagement regroupe ses employés en équipages qui sont ensuite affectés à des tâches spécifiques. Les compagnies aériennes sont également concernées par ce type de problème [Yu et al., 2003]. En outre, le domaine de la programmation des sessions parallèles dans de grandes réunions annuelles telles que les conférences [Emrah and Edis, 2013; Potthoff and Brams, 2007; Sampson, 2004]. Dans ce cas, l'objectif est de regrouper sept présentateurs dans un intervalle de temps, créant ainsi sept sessions parallèles tout en respectant les disponibilités et les préférences des présentateurs. À titre d'exemple, deux présentateurs ne voudraient pas présenter simultanément afin que chacun puisse assister à la présentation de son collègue.

Cependant, dans le contexte de PdL, nous nous intéressons également au problème de prévision des flux de prospects Internet. Cette prévision est ensuite intégrée dans le modèle d'affectation afin de réaliser des affectations pertinentes. Nous citons, dans ce qui suit, quelques exemples de contextes traitant le problème de prévision. Nous nous basons principalement sur [Bourbonnais and Usunier, 2007; Koole, 2013].

2.4.2 Combien de visiteurs à la Tour Eiffel ?

La Tour Eiffel est le monument payant le plus visité de France. Les visites se font tous les jours avec un horaire d'hiver et un horaire d'été. Quatre ascenseurs sont disponibles (un par palier), mais pour des raisons de sécurité trois ascenseurs au maximum sont en activité, le quatrième étant un ascenseur de secours.

La société d'exploitation de la Tour Eiffel souhaite avoir une prévision concernant le nombre de visiteurs par jour. Cette information aidera la société à déterminer le nombre d'ascenseurs et de caisses à ouvrir chaque jour, afin de minimiser le temps d'attente des visiteurs.

L'objectif est de pouvoir déterminer la fréquentation quotidienne, à un horizon de 35 jours glissants, avec suffisamment de précision pour déterminer le nombre d'ascenseurs à mettre en service. Dans ce contexte, ce n'est pas la valeur de la prévision elle-même qui est intéressante mais son intervalle de confiance, car l'ouverture d'un ascenseur supplémentaire est conditionnée par l'atteinte de seuils de fréquentation. Les difficultés majeures d'élaboration de la prévision sont surtout ressenties pour les jours de fin de semaine, de pont et fériés.

Le système de prévision de fréquentation de la Tour Eiffel représente une prévision d'activité quotidienne dans le domaine des services. Nous pouvons également citer dans ce contexte, la fréquentation des péages d'autoroutes, les passages aux caisses d'hypermarché, la restauration, l'hôtellerie, le cinéma . . .

La société d'exploitation de la Tour Eiffel désire améliorer le service offert aux visiteurs en terme de minimisation du temps d'attente. Pour ce faire, il convient de dimensionner au mieux le besoin en infrastructure (nombre d'ascenseurs en service et de caisses ouvertes). La qualité et la précision des prévisions de fréquentation conditionnent cette optimisation. Dans ce contexte, le système de prévision des flux de prospects Internet chez PdL représente aussi une prévision à court terme. En effet, il s'agit de répondre à la problématique de l'adéquation de l'offre et de la demande de sorte à maximiser simultanément deux objectifs, dans un contexte caractérisé par des flux continus faisant évoluer en temps réel l'ensemble des prospects Internet disponibles. Pour ce faire, il convient de déterminer au mieux le nombre de prospects Internet arrivant sur

la plate forme PdL par tranche horaire à la journée ou à quelques journées.

2.4.3 Domaine du textile

Un système de prévision pour le domaine du textile représente une prévision pour les articles à durée de vie limitée. Certains produits ont une durée de vie très courte, particulièrement dans le secteur textile où des articles sont créés et distribués le temps d'une saison (saison printemps-été, saison automne-hiver). Ces articles sont appelés "article de collection". Dans le cas des produits textiles, les critères liés au produit susceptibles d'expliquer les ventes sont nombreux. Il s'agit notamment du dessin, de la texture, du prix ... etc.

Le problème de prévision pour ces articles de collection est multiple :

1. L'absence de l'historique.
2. La multiplicité des préférences : article \times taille \times couleur.
3. Les articles à prévoir possèdent une durée de vie limitée.

La prévision pour les flux de prospects Internet chez PdL fait également face à cette même multiplicité de problème. Dans un premier temps, une nouvelle commande qui arrive sur la plate forme PdL ne possède pas un historique pour les arrivées des flux de prospects Internet, ce point est discuté dans le détail dans la section 5.3.3. Dans un second temps, la multiplicité des variables à prévoir représente une autre complexité pour la prévision des flux de prospects Internet. Nous rappelons qu'un prospect Internet représente un ensemble d'information (âge, département, profession, ... etc) d'un internaute dépendant du produit concerné (assurance santé, travaux, ... etc). Par conséquent, prévoir un prospect Internet revient à prévoir une multiplicité de variable âge \times département \times profession \times ... etc. Enfin, la durée de vie limitée des prospects Internet rend la prévision particulièrement difficile. En effet, un prospect Internet possède une durée de vie très courte, car il représente un contact qualifié et prêt à effectuer une action : achat, devis, demande de rendez-vous, demande de documentation, ... etc. Ainsi, nous devons donner suite très rapidement au besoin de ce contact avant qu'il ne soit plus d'actualité. En outre, les commandes possèdent la possibilité de filtrer les prospects Internet

qu'ils souhaitent recevoir selon de multiples critères, entre autres l'âge du prospects Internet. Par conséquent, un prospect Internet peut expirer par rapport à certaines commandes recherchant des prospects Internet jeunes (prospects Internet ayant un âge inférieur à une heure, deux heures, ... etc). Par conséquent, un des enjeux forts de l'affectation chez PdL est de permettre la prise en compte instantanée des nouvelles offres de prospects Internet. Cet aspect est d'autant plus crucial que la valeur des prospects Internet diminue avec le temps.

2.4.4 Centre d'appels

L'estimation des charges de travail est la tâche la plus difficile dans un centre d'appel. Afin de comprendre la prévision dans les centres d'appels nous devons comprendre la nature du processus des arrivées des appels. L'activité où les techniques analytiques sont utilisées le plus souvent est la gestion de la main-d'oeuvre. cette dernière est le nom commun du cycle de planification qui se traduit par des horaires des agents du centre d'appels, généralement quelques semaines avant la période (souvent une semaine) pour laquelle le calendrier est fait.

Afin d'optimiser la qualité de la production et/ou réduire les coûts, nous devons être capable de prévoir les arrivées des flux d'appels. Comme entrée, les prévisionnistes utilisent les données historiques du centre d'appels concernant les charges de trafic et les informations sur la disponibilité de l'agent, la sortie représente les horaires de l'agent.

Un système de prévision pour un centre d'appels représente une prévision par tranche horaire à la journée ou à quelques journées, qui est également le cas pour place des leads. En effet, une des caractéristiques du problème d'affectation de PdL tient à son aspect temps réel et donc au fait que l'affectation s'effectue à chaque tranche horaire dans un contexte où les flux de prospects Internet évoluent continuellement. La prévision de ces flux à chaque tranche horaire permet d'envisager l'affectation des prospects Internet sur plusieurs tranches horaire. Une telle approche conduit à optimiser les résultats de l'affectation en tirant profit de la connaissance des flux de prospects Internet sur les futures tranches horaires. Cette approche est étudié dans le chapitre 5 de ce manuscrit.

2.5 Conclusion

Dans ce chapitre, nous avons présenté le cadre général dans lequel s'inscrivent nos travaux. Ainsi nous avons commencé par introduire quelques notions élémentaires de la recherche opérationnelle. Ensuite nous avons rappelé les définitions de quelques classes de problèmes d'optimisation et de la théorie des graphes.

L'application que nous proposons dans cette thèse est basée sur le problème d'affectation, aussi nous avons rappelé différentes formulations du problème d'affectation. Ensuite nous avons abordé brièvement les algorithmes permettant de les résoudre en mettant l'accent sur quelques détails du solveur *CPLEX*.

Dans la section 2.3, nous avons présenté les caractéristiques des problèmes de prévision, leurs typologies ainsi que l'intérêt des modèles de prévision. Nous avons aussi présenté brièvement les différentes méthodes de prévision.

Enfin, nous avons présenté quelques contextes analogues à celui de PdL afin de démontrer l'applicabilité de nos travaux dans plusieurs autres contextes tels que l'affectation des équipes, l'ordonnancement des sessions parallèles dans les grandes conférences, le domaine des services : péages d'autoroutes, passages aux caisses, centre d'appels ... etc.

Chapitre 3

Optimisation bi-objectif et temps réel de l'adéquation offre/demande de prospects Internet dans un contexte mono-période

Ce chapitre a pour ambition de répondre au premier objectif de la thèse qui est la formalisation de l'adéquation de l'offre et de la demande en temps réel de sorte à maximiser simultanément et à chaque instant deux objectifs. Cette adéquation est réalisée dans un contexte caractérisé par des flux continus qui font évoluer en temps réel l'ensemble de offre disponible et la demande à satisfaire. Notre proposition consiste à : (1) formuler ce problème d'affectation de l'offre à la demande par un programme mathématique bi-objectif dont la résolution conduit à identifier la surface de pareto des solutions pareto-optimales, puis (2) sélectionner sur cette surface une bonne solution de compromis sur la base des arbitrages souhaités par le décideur.

3.1 Offre et demande de prospects Internet

Nous utilisons deux données principales à savoir les prospects Internet et les commandes. PdL reçoit continuellement des flux d'offres et de demandes de prospects Internet. Nous rappelons qu'un prospect internet représente un ensemble de caractéristiques (âge, département,

profession, ... etc) d'un internaute désirant un produit particulier (assurance santé, assurance automobile, travaux, ... etc). Sa topologie est définie par cet ensemble de caractéristiques ainsi que son âge et son prix d'achat. Par ailleurs, les clients de PdL effectuent des commandes de prospects Internet personnalisées en fonction de plusieurs critères à savoir le nombre de prospects Internet souhaité (il peut s'agir d'un volume global ou d'une quantité par mois, par jour ou par heure), l'âge des prospects Internet (moins d'une heure, douze heures, ou plus), la nature de la transaction (prospect Internet vendu en exclusivité à un seul client ou pas), ... etc. La finalité de la prise en compte de tous ces paramètres et contraintes est d'essayer d'établir la meilleure correspondance en temps réel entre l'offre et la demande de prospects Internet que nous détaillons dans ce chapitre.. Mais avant cela, nous présentons dans la section 3.2 un exemple qui a pour objectif de décrire notre contexte.

3.2 Un exemple illustratif

Considérons trois commandes et un prospect Internet à un instant donné t , décrits dans les tables 3.1 et 3.2 respectivement. Chaque commande est définie par deux types de caractéristiques. Le premier est lié à la topologie du prospect Internet (coordonnées de l'internaute : âge, profession (Prof), département (Dép). Le second concerne la commande elle-même, montrant si le client exige l'exclusivité (Exclu) des prospects Internet ou pas, le volume restant à délivrer, le prix de vente et le score de la commande. Nous rappelons que ce score représente le degré de la priorité de la commande. Il est calculé par PdL en se basant principalement sur l'expertise de l'entreprise. Le score est l'élément essentiel utilisé par PdL pour classer ses commandes et faire ses affectations.

Commande	Âge	Prof	Dép	Exclu	Volume	Prix	Score
1	[18,28]	étudiant	{75,92}	oui	15	8€	10
2	[18,26]	étudiant	{75}	non	40	6€	9
3	[18,25]	étudiant	{75}	non	45	6€	8

Tableau 3.1 — Liste des commandes à l'instant t .

En observant le prospect Internet dans la table 3.2, nous constatons qu'il n'a jamais été vendu

et que les informations concernant l'internaute correspondent à celles indiquées pour chaque commande. Par conséquent, ce prospect Internet peut être délivré exclusivement ou pas, soit à la commande 1, 2 ou 3. Si nous utilisons le processus de PdL décrit dans le chapitre 1, le prospect Internet 1 sera affecté en priorité et exclusivement à la commande 1 parce qu'elle possède le score le plus élevé. Fournissant ainsi un gain de 8€. Cependant, si nous considérons la règle de PdL qui permet de vendre les prospects Internet plusieurs fois à des commandes différentes non exclusives, alors la stratégie devrait être de vendre le prospect Internet 1 aux commandes 2 et 3 simultanément. Une telle stratégie génère un gain de 16€ et permet de satisfaire plusieurs clients à la fois. Dans le premier cas, le processus d'affectation utilisant les scores des commandes n'aboutit pas à une affectation optimale si nous considérons à la fois le chiffre d'affaires et la satisfaction des clients. Dans ce chapitre, nous proposons une nouvelle stratégie d'affectation basée sur ces deux critères, à savoir l'optimisation du chiffre d'affaires et le taux de la satisfaction des clients de façon simultanée.

	Âge	Prof	Dép	Ventes antérieures
Prospect Internet 1	19	étudiant	75	0

Tableau 3.2 — Liste des prospects Internet à l'instant t .

3.3 Formalisation du problème

Dans cette section, nous formulons le problème d'affectation des prospects Internet aux commandes en prenant en compte leurs différentes spécificités. L'enjeu est double : (i) optimiser le chiffre d'affaires et (ii) optimiser la satisfaction des clients, tout en considérant instantanément les nouvelles offres et demandes de prospects Internet et leur affectation en temps réel. En effet, notre contexte est caractérisé par des flux continus de prospects Internet faisant évoluer en temps réel l'offre disponible et la demande à satisfaire. L'un des enjeux forts de cette affectation est de permettre la prise en compte instantanée de ces flux. Cet aspect est d'autant plus crucial que la valeur des prospects Internet diminue en fonction du temps.

3.3.1 Caractéristiques du problème

Dans ce chapitre, nous étudions l'optimisation de l'affectation dans un contexte mono-période. Ainsi, nous nous limitons à une seule période de temps. En effet, à un instant t donné, l'ensemble des prospects Internet disponibles et les commandes à satisfaire sont supposés connus. Il s'agit donc de répondre à la problématique de l'adéquation de l'offre et de la demande à un instant t donné et en considérant une seule période de temps. En effet, l'optimisation s'effectue en prenant en compte uniquement la période de temps courante indépendamment des autres périodes de temps futures. Dans le prochain chapitre, nous montrerons comment étendre le modèle d'optimisation mono-période que nous allons proposer dans ce chapitre pour prendre en compte à un instant t donné plusieurs périodes de temps futures.

Afin de formuler le problème d'affectation des prospects Internet aux commandes, nous considérons à un instant donné : m prospects Internet et n commandes provenant de r clients pouvant effectuer une ou plusieurs commandes, représentés respectivement par les ensembles suivants :

- Ensemble de prospects Internet : $P = \{j : j \in \{1, \dots, m\}\}$.
- Ensemble de commandes : $O = \{i : i \in \{1, \dots, n\}\}$.
- Ensemble de clients : $C = \{k : k \in \{1, \dots, r\}\}$.

Le modèle est représenté par un programme mathématique linéaire, déterministe et mono-période dont les différents composants sont définis dans ce qui suit.

3.3.2 Variables de décision

Le problème de correspondance entre les prospects Internet et les commandes sous entend que nous devons décider quels prospects internet seront affectés à quelles commandes. Ainsi, pour représenter cette correspondance, nous proposons d'utiliser une variable binaire x_{ij} définie comme suit :

$$x_{ij} = \begin{cases} 1 & \text{si le prospect Internet } j \text{ est affecté à la commande } i \\ 0 & \text{sinon} \end{cases}$$

Par ailleurs, l'ensemble de toutes les affectations "commande/prospect Internet" existantes est décrit par l'ensemble $X = \{x_{ij} : i \in O, j \in P\}$. Néanmoins, les commandes acceptent seulement les prospects Internet avec certaines caractéristiques. En effet, les clients peuvent personnaliser leurs commandes en imposant un filtre sur la topologie des prospects Internet. Par exemple, des prospects Internet qui correspondent à un département donné, profession ou une combinaison particulière de ces caractéristiques. Par conséquent, certaines paires x_{ij} ne sont pas possibles. Ainsi, nous allons considérer que $x_{ij} \in X$ seulement si un prospect Internet j correspond à une commande i en fonction des exigences de cette dernière. Cela permet de réduire considérablement l'espace de recherche des solutions optimales.

En outre, certaines commandes demandent des prospects Internet en exclusivité. Cela signifie que le prospect Internet est vendu au plus une fois. Cette vente est appelée une vente exclusive. Dans le cas contraire, nous pouvons vendre le prospect Internet plusieurs fois à des commandes différentes. Cette vente est appelée une vente non exclusive. En conséquence, nous avons $X = E \cup \bar{E}$, où $E = \{x_{ij} : i \in O, j \in P, i \text{ est une commande exclusive}\}$ et $\bar{E} = \{x_{ij} : i \in O, j \in P, i \text{ est une commande non exclusive}\}$.

3.3.3 Contraintes

Afin de pouvoir construire un modèle mathématique qui répond aux besoins de PdL, nous définissons les différentes contraintes de transactions entre PdL et ses clients.

En effet, l'affectation des prospects Internet aux commandes est soumise à certaines contraintes que nous détaillons dans ce qui suit.

Contraintes liées aux volumes : le client de PdL possède la possibilité d'effectuer une ou plusieurs commandes avec des volumes différents. Le nombre de prospects Internet affecté à chaque commande est limité par un volume global, un quota par mois, un autre par jour et un dernier par heure. Ainsi, la contrainte (3.1) assure que le nombre de prospects Internet affecté à chaque commande respecte son volume global ainsi que son quota horaire, journalier

et mensuel.

$$\sum_{j:x_{ij} \in X} x_{ij} \leq Ro_i, \quad \forall i \in O \quad (3.1)$$

Tel que, Ro_i est le nombre de prospects Internet restant à délivrer à la commande i . Il est décrit comme suit :

$$Ro_i = \min\{Go_i, Ho_i, Do_i, Mo_i\} \quad (3.2)$$

Où, Go_i est le volume global de la commande i , Ho_i est le quota horaire de la commande i , Do_i est le quota journalier de la commande i et Mo_i est le quota mensuel de la commande i .

De plus, le client peut spécifier la quantité globale de prospects Internet correspondant à ses différentes commandes. Par conséquent, la contrainte (3.3) exprime le fait que le nombre des prospects Internet affecté à l'ensemble des commandes d'un client ne doit pas dépasser le quota horaire, journalier et mensuel autorisé.

$$\sum_{i:x_{ij} \in X} a_{ik}x_{ij} \leq Rc_k, \quad \forall k \in C \quad (3.3)$$

Tel que, Rc_k est le nombre de prospects Internet restant à délivrer au client k , qui est défini comme suit :

$$Rc_k = \min\{Hc_k, Dc_k, Mc_k\} \quad (3.4)$$

Où, Hc_k est le quota horaire du client k , Dc_k est le quota journalier du client k , Mc_k est le quota mensuel du client k et a_{ik} est une variable binaire décrite de la manière suivante :

$$a_{ik} = \begin{cases} 1 & \text{si la commande } i \text{ provient du client } k \\ 0 & \text{sinon} \end{cases}$$

Contrainte liée au nombre de ventes et à l'exclusivité : comme décrit précédemment, il existe deux types de ventes à savoir une vente exclusive, où un prospect Internet est vendu au plus une fois et une vente non exclusive, où un prospect Internet est vendu au plus trois fois. Le fait que le nombre de ventes autorisées est limité à trois fois est exigé par PdL. Toutefois, il est facile de considérer un nombre quelconque de vente de la façon suivante :

$$S_j + b \sum_{i:x_{ij} \in E} x_{ij} + \sum_{i:x_{ij} \in \bar{E}} x_{ij} \leq b, \quad \forall j \in P \quad (3.5)$$

où,

- S_j est le nombre de ventes antérieures du prospect Internet j .
- b est le nombre de ventes non exclusives autorisées d'un prospect Internet.

La contrainte (3.5) nous permet de vérifier qu'un prospect Internet exclusif n'est pas vendu plus qu'une fois, et qu'un prospect Internet non exclusif n'est pas vendu plus que b fois.

Etant donné que PdL fixe le nombre de ventes non exclusives à 3, alors la contrainte 3.5 devient :

$$S_j + 3 \sum_{i:x_{ij} \in E} x_{ij} + \sum_{i:x_{ij} \in \bar{E}} x_{ij} \leq 3, \quad \forall j \in P \quad (3.6)$$

La contrainte (3.6) nous permet de vérifier qu'un prospect Internet exclusif est vendu au plus une fois, et qu'un prospect Internet non exclusif est vendu au plus 3 fois.

Le fait que le nombre de ventes non exclusives des prospects Internet soit le même pour toutes les commandes est spécifique à la place de marché PdL. Dans le sens où, si un prospect Internet est vendu en non exclusivité à une commande, il pourra être encore revendu jusqu'à 3 fois à d'autres commandes différentes. Toutefois, nous pouvons personnaliser le nombre de ventes non exclusives des prospects Internet. En effet, il est possible d'envisager un nombre de ventes non exclusives différent pour chaque commande n'exigeant pas l'exclusivité des prospects Internet. Nous détaillerons cela dans le chapitre 7.

Contrainte liée à la duplication des prospects Internet : le client de PdL possède la possibilité d'effectuer une ou plusieurs commandes. Cependant, nous ne pouvons pas affecter le même prospect Internet au même client plusieurs fois.

La contrainte (3.6) permet de vendre un prospect Internet non exclusif à plusieurs reprises à des commandes différentes. Ainsi, nous devons vérifier que ces commandes correspondent à des clients différents. Par conséquent, nous définissons la contrainte (3.7) qui garantit que chaque prospect Internet est affecté au plus une fois au même client.

$$\sum_{i:x_{ij} \in X} a_{ik} x_{ij} \leq 1, \quad \forall k \in C \quad (3.7)$$

3.3.4 Fonction objectif

Comme toute place de marché, l'objectif principal de PdL est de maximiser ses profits. Il s'agit donc de répondre à la problématique de l'adéquation de l'offre et de la demande de prospects Internet à un instant donné de sorte à maximiser le chiffre d'affaires.

$$\max \sum_{i,j:x_{ij} \in X} Pr_i x_{ij} \quad (3.8)$$

Où Pr_i est le prix de vente de la commande i .

Ainsi, la fonction objectif décrite par l'équation (3.8) cherche l'adéquation offre/demande de prospects Internet maximisant au mieux le chiffre d'affaires de PdL.

3.4 Modèle d'optimisation mono-objectif de l'affectation offre/demande de prospects Internet

Dans cette section, nous proposons un modèle mono-objectif à l'aide de la programmation mathématique. Ce modèle prend en compte les différentes spécificités des commandes et des prospects Internet discutées dans les sections précédentes. Il permet ainsi d'optimiser, à un

instant donné et en considérant une seule période de temps, le chiffre d'affaires de PdL. De plus, ce modèle permet de prendre en compte instantanément les nouvelles offres et demandes de prospects Internet et leur adéquation en temps réel.

$$P_1 \left\{ \begin{array}{l} \max \quad f_1(X) = \sum_{i,j:x_{ij} \in X} Pr_i x_{ij} \\ s.t. \\ \sum_{j:x_{ij} \in X} x_{ij} \leq Ro_i, \quad \forall i \in O \\ \sum_{i,j:x_{ij} \in X} a_{ik} x_{ij} \leq Rc_k, \quad \forall k \in C \\ S_j + b \sum_{i:x_{ij} \in E} x_{ij} + \sum_{i:x_{ij} \in \bar{E}} x_{ij} \leq b, \quad \forall j \in P \\ \sum_{i:x_{ij} \in X} a_{ik} x_{ij} \leq 1, \quad \forall j \in P, \forall k \in C \\ x_{ij} \in \{0, 1\} \end{array} \right. \quad (3.9)$$

P_1 représente un programme mathématique linéaire, déterministe et mono-période qui optimise à un instant t donné l'affectation des prospects Internet aux commandes en guise de maximisation du chiffre d'affaires. La fonction objectif de ce modèle mono-objectif maximisant le chiffre d'affaires est donnée par l'équation (3.8). Quant aux contraintes à savoir la contrainte des volumes des commandes et des clients, la contrainte liée au nombre de ventes et l'exclusivité ainsi que la contrainte de la duplication des prospects Internet, sont décrites par les équations (3.1), (3.3), (3.5) et (3.7) respectivement.

Dans cette section, nous avons formalisé le problème de la place de marché à travers la programmation mathématique. Ainsi, nous avons défini les variables de décision, les contraintes et la fonction objectif. Une deuxième possibilité est de proposer la même formalisation mais en utilisant la théorie des graphes. Plus formellement, l'objectif pour la formulation en théorie des graphes sera de déterminer un couplage parfait de coût minimum dans un graphe biparti complet.

3.5 Algorithme pour l'affectation mono-objectif

Le problème d'affectation auquel nous nous intéressons est caractérisé par la notion de temps. En effet, la place de marché PdL est caractérisée par la temporalité de son système. Le processus PdL fonctionne en temps réel, ainsi à chaque instant nous devons décider quels prospects Internet seront affectés et à quelles commandes.

L'optimisation de l'affectation à un instant t donné est assurée par le modèle mathématique P_1 présenté dans la section 3.4. Cependant, le système PdL fonctionne en continu. En effet, une des caractéristiques du processus d'affectation de PdL tient au fait que l'adéquation offre/demande de prospects Internet s'effectue à chaque instant t à longueur de journée dans un contexte caractérisé par des flux continus de prospects Internet. Ainsi, l'enjeu principal de ce travail est d'établir à chaque instant t une correspondance en temps réel entre les prospects Internet et les commandes. Ainsi, dans cette section, nous présentons un algorithme assurant l'aspect temps réel du système PdL en effectuant des affectations plusieurs fois par jour (toutes les heures voire toutes les minutes) et en considérant à chaque fois une seule période de temps. En outre, cet algorithme prend en compte de manière instantanée les nouveaux flux, leur affectation en temps réel ainsi que la mise à jour des quotas des commandes et des nombres de ventes des prospects Internet après chaque affectation.

L'algorithme 3.1 fonctionne de manière continue tout au long de la journée. En effet, à chaque instant t de la journée, l'algorithme 3.1 considère un ensemble de prospects Internet ainsi qu'un ensemble de commandes provenant des différents clients de PdL, représentés par P, O et C respectivement.

Nous rappelons que les commandes acceptent seulement les prospects Internet avec certaines caractéristiques. En effet, les clients peuvent personnaliser leurs commandes en imposant un filtre sur la topologie des prospects Internet qu'ils souhaitent recevoir. Par conséquent, l'algorithme 3.1 considère que $x_{ij} \in X$ seulement si un prospect Internet j correspond à une commande i en fonction des exigences de cette dernière. Cela permet de réduire considérablement l'espace de recherche des solutions optimales.

Algorithme 3.1 — Affectation mono-objectif()

Data : $T = \{t : t \text{ est l'heure pile} \}$. $P = \{j : j \in \{1, \dots, m\}\}$ est l'ensemble de prospects Internet à l'instant t . $O = \{i : i \in \{1, \dots, n\}\}$ est l'ensemble de commandes à l'instant t . $C = \{k : k \in \{1, \dots, r\}\}$ est l'ensemble de clients à l'instant t . $X = \{x_{ij} : i \in O, j \in P, x_{ij} \text{ est une variable binaire} \}$ est l'ensemble des variables de décision. $X' = \{x'_{ij} : i \in O, j \in P, x'_{ij} \text{ est une variable binaire} \}$ est le résultat fourni par le programme mathématique P_1 présenté dans la section 3.4.

```

1 while  $t \in T$  do
2    $X \leftarrow \emptyset$ 
3   while  $i \in O$  do
4     while  $j \in P$  do
5       if le prospect Internet  $j$  correspond aux caractéristiques de la
6         commande  $i$  then
7          $X \leftarrow X \cup \{x_{ij}\}$ 
8       end
9     end
10     $X' \leftarrow P_1(X)$ 
11    while  $x'_{ij} \in X'$  do
12      if  $x'_{ij} = 1$  then
13        Affecter le prospect Internet  $j$  à la commande  $i$ 
14      end
15    end
16    Mettre à jour les quotas des commandes ainsi que le nombre de reventes
17    des prospects Internet.
18    Intégrer les nouveaux flux de l'offre et de la demande.
19  end

```

Ensuite, il fait appel au modèle d'optimisation P_1 présenté dans la section 3.4 en lui fournissant l'ensemble des variables de décision X . Suite au résultat X' retourné par le modèle d'optimisation P_1 , l'algorithme 3.1 effectue les affectations des prospects Internet aux commandes. En effet, si $x'_{ij} = 1$ alors le prospect Internet j sera affecté à la commande i , sinon ($x'_{ij} = 0$) l'affectation n'aura pas lieu.

Enfin, l'algorithme 3.1 établit les mises à jour nécessaires des quotas de commandes et des nombres de reventes des prospects Internet suite aux affectations effectuées.

L'intégration des nouveaux flux de l'offre et de la demande des prospects Internet est effectuée avant d'entamer la prochaine itération traitant l'instant suivant de la journée.

3.6 Modèle d'optimisation bi-objectif de l'affectation offre/demande de prospects Internet

Dans le modèle décrit dans la section précédente, seul le chiffre d'affaires est maximisé, alors que d'autres objectifs pourraient être également intéressants. En effet, un problème d'affectation classique considère uniquement le coût ou le profit dans la formulation du problème. Cependant, les différentes applications réelles ont montré la nécessité d'intégrer d'autres objectifs [Keeney and Raiffa, 1993; Roy, 1996; Adeyefa and Luhandjula, 2011]. Dans ce contexte, un aspect important pour PdL est la satisfaction de ses clients, dans la mesure où la favorisation de certains clients se traduit par la perte des clients marginalisés. Ainsi, satisfaire équitablement tous les clients s'avère un enjeu très important afin de continuer à alimenter PdL de commandes.

Ainsi, l'objectif de la place de marché est de maximiser ses profits tout en conservant ses clients. Pour ce faire, nous proposons de considérer dans la formalisation du problème deux objectifs simultanément à savoir le chiffre d'affaires et la satisfactions des clients. Le premier est un objectif classique pour toute place de marché voulant faire prospérer son activité. Le second objectif est une conséquence du fait qu'il n'existe aucune façon de satisfaire tous les clients en même temps. En effet, il n'est pas toujours possible de servir toutes les com-

mandes parce que nous n'avons pas nécessairement des prospects Internet qui répondent à leurs caractéristiques. L'objectif recherché à travers la maximisation du taux de satisfaction des commandes est celui d'être juste envers tous les clients afin d'éviter une situation où les clients négligés se sentent mal servis et peuvent donc quitter le système de PdL.

3.6.1 Fonctions de satisfaction des commandes

Nous pouvons interpréter le problème de la satisfaction des commandes comme un problème de distribution équitable de prospects Internet. La satisfaction des commandes peut être considérée sur plusieurs horizons de temps : satisfaction sur la semaine, le mois ou l'année . . . Dans le cadre de ce travail nous nous intéressons à la satisfaction journalière des commandes. En effet, à la fin de chaque journée, PdL fait le bilan sur l'ensemble de ses transactions. Ainsi, le but est d'avoir une distribution journalière juste et équitable des prospects Internet entre les différentes commandes sur la journée. Cette notion d'équité peut être décrite selon différents critères (voir [Pettypool and Karathanos, 2004; Larusic and Punnen, 2011; Turner et al., 2011; Duin and Volgenant, 2012]).

La politique la plus populaire est l'équité max-min (l'équité min-max), qui maximise (minimise respectivement) la distribution minimum (maximum respectivement) reçue par l'une des commandes dans le système [Nace et al., 2008]. Ces deux politiques peuvent être facilement appliquées à notre contexte. Cependant, elles ne sont pas satisfaisantes pour notre problème, dans la mesure où MaxMin et MinMax sont des méthodes conservatrices. Autrement dit, elles fixent une limite à la valeur maximale afin de servir la deuxième valeur la plus élevée. Néanmoins, dans notre contexte, il n'est pas toujours possible de servir les commandes les moins bien satisfaites parce que nous n'avons pas nécessairement des prospects Internet qui répondent à leurs caractéristiques. Par conséquent, nous privons la valeur maximale (commandes bien servies) de recevoir de nouveaux prospects Internet sans être en mesure de servir la deuxième valeur la plus élevée (commandes mal servies). Nous pouvons interpréter ce cas par un point dominant sur une frontière efficace : bien que le premier critère (le chiffre d'affaires) est réduit, nous ne pouvons pas améliorer le deuxième critère (la satisfaction des commandes). Nous

proposons donc deux formalisations différentes de la satisfaction des commandes.

Premièrement, nous considérons la satisfaction moyenne des commandes sur la journée. Formellement,

$$\max \frac{1}{n} \left[\sum_{i:x_{ij} \in X} (V_i + \sum_{j:x_{ij} \in X} x_{ij}) / Do_i \right] \quad (3.10)$$

Où, V_i est le nombre de prospects Internet délivré à la commande i à partir du début de la journée.

Deuxièmement, nous pouvons essayer d'avoir un équilibre entre la minimisation de la valeur maximale et la maximisation de la valeur minimale. Ceci est réalisé en minimisant la différence entre les deux valeurs maximale et minimale. Formellement,

$$\min \left[\max \left[(V_i + \sum_{j:x_{ij} \in X} x_{ij}) / Do_i \right] - \min \left[(V_i + \sum_{j:x_{ij} \in X} x_{ij}) / Do_i \right] \right], \quad \forall i \in O \quad (3.11)$$

La fonction objectif (3.11) minimise la différence de la satisfaction entre les commandes bien servies (commandes ayant un taux de satisfaction élevé) et celles mal servies (commandes ayant un taux de satisfaction moins élevé). En d'autres termes, le but est d'homogénéiser la satisfaction des commandes dans son ensemble : minimisation de l'écart entre les commandes bien satisfaites et celles moins bien satisfaites.

3.6.2 Programme mathématique bi-objectif

L'affectation a pour objectif de trouver le bon compromis qui concilie deux objectifs contradictoires, à savoir, le chiffre d'affaire et le taux de satisfaction des commandes. En effet, d'une part, la maximisation du chiffre d'affaires nécessite d'affecter les prospects Internet aux clients proposant les prix les plus élevés. Ainsi, ceux qui proposent des prix moins intéressants seront marginalisés. D'autre part, la distribution équitable des prospects Internet sur tous les clients génère un chiffre d'affaires moins important. Cela est dû à la différence des prix proposés par les différents clients.

Ainsi, nous formulons un modèle d'optimisation bi-objectif en utilisant la programmation mathématique. Il s'agit de permettre le calcul en temps réel des affectations correspondant à de bonnes solutions de compromis sur les deux objectifs à optimiser.

P_2 représente un programme mathématique linéaire, déterministe et mono-période. Le programme mathématique P_2 permet de répondre à la problématique de l'affectation temps réel de l'offre et de la demande, à un instant donné et en considérant une seule période de temps, de sorte à maximiser deux objectifs simultanément. Les deux fonctions objectif de ce modèle sont données par les équations (3.8) et (3.10) permettant de maximiser simultanément le chiffre d'affaires et la satisfaction moyenne des commandes sur la journée. Les contraintes des deux modèles P_1 et P_2 sont identiques.

$$P_2 \left\{ \begin{array}{l} \max \quad f_1(X) = \sum_{i,j:x_{ij} \in X} Pr_i x_{ij} \\ \max \quad f_2(X) = \frac{1}{n} \left[\sum_{i:x_{ij} \in X} (V_i + \sum_{j:x_{ij} \in X} x_{ij}) / Do_i \right] \\ s.t. \\ \sum_{j:x_{ij} \in X} x_{ij} \leq Ro_i, \quad \forall i \in O \\ \sum_{i,j:x_{ij} \in X} a_{ik} x_{ij} \leq Rc_k, \quad \forall k \in C \\ S_j + b \sum_{i:x_{ij} \in E} x_{ij} + \sum_{i:x_{ij} \in \bar{E}} x_{ij} \leq b, \quad \forall j \in P \\ \sum_{i:x_{ij} \in X} a_{ik} x_{ij} \leq 1, \quad \forall j \in P, \forall k \in C \\ x_{ij} \in \{0,1\} \end{array} \right. \quad (3.12)$$

Un deuxième programme mathématique P'_2 peut être défini à partir de P_2 en remplaçant la fonction objectif (3.10) par la fonction objectif (3.11). Ainsi, P'_2 optimise, à un instant donné et en considérant une seule période de temps, l'affectation des prospects Internet aux commandes. Cette optimisation s'effectue en prenant en compte : (i) la maximisation du chiffre d'affaires et (ii) la minimisation de l'écart entre les satisfactions des commandes sur la journée.

3.7 Algorithme pour l'affectation bi-objectif

L'algorithme 3.2 effectue des affectations de manière continue à longueur de journée tout en considérant deux objectifs à chaque affectation. Ce fonctionnement est similaire à celui de l'algorithme 3.1 présenté dans la section 3.5, à l'exception de l'instruction ($X' \leftarrow P_1(X)$) qui sera remplacée par ($X' \leftarrow P_2(X)$).

Dans un contexte bi-objectif, nous faisons appel au modèle d'optimisation bi-objectif P_2 présenté dans la section 3.6.2. Ainsi, plusieurs solutions optimales sont calculées par le biais de la méthode $\varepsilon - constraint$ (voir la section 2.2.4.2). Ces solutions constituent un échantillon représentatif de l'ensemble de pareto.

Afin de pouvoir sélectionner une solution optimale parmi l'ensemble de pareto, nous faisons appel à un décideur du domaine. Ainsi, le processus de solution est divisé en deux phases indépendantes : d'abord, la génération des solutions efficaces et par la suite la participation du décideur lorsque toutes les solutions sont calculées.

Le décideur choisira une solution optimale parmi toutes les solutions de la frontière efficace. Ainsi, il peut exprimer des préférences qui lui sont propres. Ces préférences seraient alors de nature à privilégier un certain type de solutions, par exemple il est d'usage de préférer des solutions équilibrées, c'est-à-dire des solutions avec des valeurs moyennes sur les objectifs.

Une fois la bonne solution de compromis est choisie par le décideur, le processus restera le même que celui de l'algorithme 3.1. En effet, l'algorithme 3.2 effectue l'affectation du prospect Internet j à la commande i si $x'_{ij} = 1$, sinon ($x'_{ij} = 0$) l'affectation n'aura pas lieu. Ensuite, il passe à l'instant t suivant après avoir établi les mises à jour nécessaires des quotas de commandes et du nombre de reventes des prospects Internet.

Les algorithmes 3.1 et 3.2 présentés dans les sections 3.5 et 3.7 respectivement sont caractérisés par un intervalle de distribution : il s'agit de l'instant t qui représente l'intervalle de temps entre deux affectations.

Ainsi, nous pouvons envisager des affectations toutes les minutes, les 10 minutes, ou plus.

Dans la section 6.2.2.2, nous traiterons dans le détail l'impact du choix de l'intervalle de distribution sur l'optimisation.

Algorithme 3.2 — Affectation bi-objectif()

Data :

$T = \{t : t \text{ est l'heure pile}\}$.

$P = \{j : j \in \{1, \dots, m\}\}$ est l'ensemble de prospects Internet à l'instant t .

$O = \{i : i \in \{1, \dots, n\}\}$ est l'ensemble de commandes à l'instant t .

$C = \{k : k \in \{1, \dots, r\}\}$ est l'ensemble de clients à l'instant t .

$X = \{x_{ij} : i \in O, j \in P, x_{ij} \text{ est une variable binaire}\}$ est l'ensemble des variables de décision.

$X' = \{x'_{ij} : i \in O, j \in P, x'_{ij} \text{ est une variable binaire}\}$ est le résultat fourni par le programme mathématique P_2 présenté dans la section 3.6.2.

```

1 while  $t \in T$  do
2    $X \leftarrow \emptyset$ 
3   while  $i \in O$  do
4     while  $j \in P$  do
5       if le prospect Internet  $j$  correspond aux caractéristiques de la
6         commande  $i$  then
7          $X \leftarrow X \cup \{x_{ij}\}$ 
8       end
9     end
10    Calculer la frontière efficace du programme mathématique bi-objectif
11     $P_2(X)$  en utilisant la méthode  $\varepsilon$  – constraint.
12     $X' \leftarrow$  la solution optimale choisie par le décideur.
13    while  $x'_{ij} \in X'$  do
14      if  $x'_{ij} = 1$  then
15        Affecter le prospect Internet  $j$  à la commande  $i$ 
16      end
17    end
18    Mettre à jour les quotas des commandes ainsi que le nombre de reventes
19    des prospects Internet.
20    Intégrer les nouveaux flux de l'offre et de la demande.
21  end

```

3.8 Conclusion

Dans ce chapitre, nous avons proposé un modèle d'optimisation bi-objectif qui maximise, à un instant donné et en considérant une seule période de temps, deux objectifs simultanément à savoir le chiffre d'affaires et la satisfaction des clients. Cette optimisation considère instantanément les nouvelles offres et demandes de prospects Internet et leur adéquation en temps réel.

Par ailleurs, la place de marché PdL est caractérisée par la temporalité de son système. Ainsi, nous avons proposé un algorithme assurant l'aspect temps réel du processus PdL en effectuant des affectations continues. En effet, l'algorithme proposé permet d'assurer l'aspect temps réel du système PdL en effectuant des affectations plusieurs fois par jour et en considérant à chaque fois une seule période de temps. Il permet également de considérer de manière instantanée les nouveaux flux et de mettre à jour les quotas des commandes ainsi que le nombre de reventes des prospects Internet après chaque affectation.

Cependant, l'optimisation en considérant une seule période de temps ne permet pas de réaliser des affectations pertinentes. En effet, l'optimisation mono-période s'effectue en prenant en compte uniquement la période de temps courante indépendamment des autres périodes de temps futures. Ce type de fonctionnement conduit à une stratégie myope, ne permettant pas l'anticipation des mises à jour du système. Par conséquent, nous proposons dans le chapitre suivant, un modèle d'optimisation multi-période qui étend le problème sur plusieurs périodes de temps futures. Cette extension conduit à optimiser les résultats de l'affectation en tirant profit de la connaissance concernant l'évolution du système : nouveaux flux d'offres et réactivation des quotas de la demande.

Chapitre 4

Optimisation multi-période de l'adéquation offre/demande de prospects Internet

Les modèles d'optimisation proposés dans le chapitre précédent optimisent l'affectation des prospects Internet à un instant donné et en considérant une seule période de temps. En effet, cette première formalisation est mono-période, ce qui signifie que l'optimisation considère uniquement la période de temps courante sans prendre en compte l'évolution du système sur les périodes de temps futures. Le fait de considérer une seule période de temps ne permet pas d'anticiper les mises à jour du système. Cependant, la considération d'une période de temps continue conduit à optimiser les résultats de l'affectation en tirant profit de la connaissance de l'évolution du système sur les périodes de temps futures.

Dans ce chapitre, nous présentons une extension du modèle mono-période afin de surmonter les limitations qu'il présente. Pour ce faire, nous proposons un modèle d'optimisation multi-période qui étend le problème sur plusieurs périodes de temps futures [Maxon and Bhadury, 2001; Zhang and Barad, 2006]. En effet, le modèle d'optimisation multi-période permet d'établir des affectations à un instant donné et en considérant plusieurs périodes de temps futures. Cette extension conduit à optimiser l'affectation en considérant les éventuelles futures mises

à jour du système, à savoir, nouveaux flux d'offres et réactualisation des quotas de la demande.

4.1 Quelques exemples illustratifs

Notre objectif est d'établir une adéquation entre l'offre et la demande de prospects Internet, à un instant donné et en considérant plusieurs périodes de temps futures. Ainsi, cette extension multi-période permet d'optimiser les résultats de l'affectation en tirant parti de la connaissance de l'évolution des : (i) flux de prospects Internet et (ii) mises à jour des quotas des commandes, sur toutes les périodes de temps futures incluses dans l'optimisation. Deux exemples mettant en évidence l'avantage de la prise en compte de ces informations dans le processus d'affectation, sont détaillés dans ce qui suit.

4.1.1 Anticipation de l'évolution des flux de prospects Internet

Supposons que nous possédons, à un instant t donné, les commandes et les prospects Internet présentés dans les tables 4.1 et 4.2 respectivement. En utilisant le programme mathématique P_1 défini dans la section 3.4, le prospect Internet 1 sera affecté prioritairement et exclusivement à la commande 1. Une telle affectation épuise le quota de la commande 1.

t	Exclusivité	Prix	Volume
Commande 1	oui	7€	1
Commande 2	non	2€	10

Tableau 4.1 — Commandes à l'instant t .

t	Commandes compatibles
Prospect Internet 1	commande 1, commande 2

Tableau 4.2 — Prospects Internet à l'instant t .

Supposons qu'à l'instant t , nous avons la possibilité de prédire l'arrivée des prospects Internet sur toutes les périodes de temps futures prises en compte dans l'optimisation en utilisant, par exemple, des techniques de prévision [Makridakis et al., 1997; Bourbonnais and Usunier,

2007]. Le but de ces techniques est de réduire l'incertitude liée à la méconnaissance de l'avenir. À titre d'exemple, cette prévision montrera que nous pourrions disposer d'un prospect Internet 2 à l'instant $t + i$ (voir la table 4.3). Ce prospect Internet prévu correspond seulement aux caractéristiques de la commande 1 qui n'est plus disponible, car son quota est épuisé. Ainsi, le prospect Internet 2 ne peut pas être distribué. Le gain généré sera de 7€.

$t + i$	Commandes compatibles
Prospect Internet 1	commande 1, commande 2
Prospect Internet 2	commande 1

Tableau 4.3 — Prospects Internet à l'instant $t + i$.

Cependant, si nous avons la possibilité d'avoir cette information à l'instant t , nous serions en mesure d'anticiper et de décider de l'affectation du prospect Internet 1 à la commande 2 à l'instant t et d'attendre l'instant $t + i$ pour affecter le prospect Internet 2 à la commande 1. Ainsi, le gain sera de 9€.

4.1.2 Anticipation de la mise à jour des quotas

Prenons l'exemple présenté dans la table 4.4 où nous possédons deux commandes à l'instant t , de sorte à ce que le quota de la première commande soit épuisé alors qu'il reste encore 4 prospects Internet à fournir à la seconde commande. Le prospects Internet disponible à cet instant est représenté dans la table 4.5. Par conséquent, l'optimisation de l'affectation sur un seul instant conduit à affecter de manière exclusive le prospect Internet 1 à la commande 2 par suite de l'épuisement du quota de la première commande.

t	Exclusivité	Prix	Volume
Commande 1	oui	8€	0
Commande 2	oui	5€	4

Tableau 4.4 — Commandes à l'instant t .

Considérons, maintenant, que les quotas des deux commandes à l'instant $t + i$ sont représentés dans la table 4.6. À ce moment, le volume de la commande 1 est mis à jour, mais il n'est plus

t	Commandes compatibles
Prospect Internet 1	commande 1, commande 2

Tableau 4.5 — Prospects Internet à l'instant t .

possible d'affecter le prospect Internet 1 à la commande 1 parce qu'il a été vendu exclusivement à la commande 2. Le gain total réalisé sera de 5€. Cependant, une stratégie multi-période permet d'attendre l'instant $t + i$ pour affecter le prospect Internet 1 à la commande 1. Une telle approche génèrera un gain de 8€.

$t + i$	Exclusivité	Prix	Volume
Commande 1	oui	8€	4 (réinitialisation)
Commande 2	oui	5€	2

Tableau 4.6 — Commandes à l'instant $t + i$.

À travers ces deux exemples, nous observons la nécessité d'étendre le modèle mono-période pour considérer différentes périodes de temps. Dans ce qui suit nous détaillons les caractéristiques du modèle d'optimisation multi-période. Ce dernier permet de considérer l'évolution des flux de prospects Internet ainsi que les mises à jour régulières des quotas des commandes dans le processus d'affectation.

4.2 Modèle d'affectation multi-période de l'offre/demande de prospects Internet

Dans cette section, nous introduisons un modèle d'affectation multi-période afin d'optimiser l'affectation à un instant donné en tirant profit de la connaissance de l'évolution des flux de l'offre de prospects Internet ainsi que des mises à jour des quotas de commandes sur des périodes de temps futures.

La problématique de l'affectation multi-période est caractérisée par un horizon glissant, ce qui signifie que le temps est modélisé par plusieurs instants. À un instant t donné, nous considérons un ensemble T' de périodes de temps de sorte que $T' = \{t' : t' \in \{1, \dots, g\}\}$, où $t' = 1$ est la période de temps courante et $t' \in \{2, \dots, g\}$ sont les périodes de temps futures. Ainsi, nous

devons décider, à l'instant t , quels prospects Internet seront affectés à quelles commandes tout en considérant toutes les périodes de temps de l'ensemble T' . Par conséquent, l'évolution des flux de prospects Internet et la mise à jour des quotas des commandes sur toutes les périodes de temps de l'ensemble T' , sont prises en compte dans l'affectation effectuée à l'instant t courant.

Nous supposons que les commandes et les clients restent inchangés pour toutes les périodes de temps t' (cette hypothèse est levée dans le chapitre 7 où nous proposons une nouvelle formulation sans cette hypothèse). Par conséquent, les ensembles de commandes et de clients (O et C respectivement) gardent la même définition que dans la section 3.3.1. Cependant, il est possible d'avoir des nouveaux flux d'offres de prospects Internet arrivant de façon non déterministe sur l'ensemble des périodes de temps t' . Ainsi, nous redéfinissons l'ensemble des prospects Internet de sorte qu'à chaque période de temps t' , nous avons $P_{t'} = \{j : j \in \{1, \dots, m_{t'}\}\}$, l'ensemble des prospects Internet à la période de temps t' et $m_{t'}$ est le nombre de prospects Internet à cette même période de temps.

Nous notons que dans ce chapitre nous supposons que les flux de prospects Internet sont des données connues sur toutes les périodes de temps futures prises en compte dans l'optimisation. Par ailleurs, nous montrerons dans le chapitre suivant la manière dont ces flux futurs de prospects Internet sont calculés.

4.2.1 Variables de décision

Pour chacune des périodes de temps t' de l'ensemble T' , l'affectation d'un prospect Internet à une commande est représentée par la variable binaire suivante :

$$x_{ijt'} = \begin{cases} 1 & \text{si le prospect Internet } j \text{ est affecté à la commande } i \text{ à la période de temps } t' \\ 0 & \text{sinon} \end{cases}$$

Nous définissons ainsi $X_{t'} = \{x_{ijt'} : i \in O, j \in P_{t'}, t' \in T'\}$, qui est l'ensemble de toutes les affectations "commande/prospect Internet" existantes à toutes les périodes de temps de l'ensemble T' . Nous considérerons que $x_{ijt'} \in X_{t'}$ seulement si le prospect Internet j corres-

pond à la commande i , à la période de temps t' . En outre, pour être en mesure de séparer les prospects Internet exclusifs de ceux qui ne le sont pas, nous définissons à chaque période de temps $t' : E_{t'} = \{x_{ijt'} : i \in O, j \in P_{t'}, t' \in T', i \text{ est une commande exclusive} \}$ et $\overline{E}_{t'} = \{x_{ij} : i \in O, j \in P_{t'}, t' \in T', i \text{ est une commande non exclusive} \}$. Ainsi, $X_{t'} = E_{t'} \cup \overline{E}_{t'}$ tel que $t' \in T'$.

4.2.2 Contraintes

Nous rappelons qu'afin de pouvoir construire un modèle mathématique multi-période répondant aux besoins de PdL, nous devons définir les contraintes de transactions entre PdL et ses clients dans un contexte multi-période.

En effet, le problème d'affectation de prospects Internet en considérant, à chaque instant t , l'ensemble T' des périodes de temps futures est soumis aux contraintes suivantes.

Contraintes liées aux volumes : la contrainte (4.1) vérifie que le nombre de prospects Internet affecté à chaque commande i , à chaque période de temps de l'ensemble T' , respecte le volume de la commande. Ce volume représente la différence entre le volume restant à délivrer pour la commande i à la période de temps 1, et le nombre de prospects Internet délivré pour la commande i depuis la période de temps 1 à la période de temps courante t' .

En outre, la contrainte (4.2) vérifie que le nombre de prospects Internet affecté à chaque client k , à chaque période de temps de l'ensemble T' , respecte le volume du client. Ce volume représente la différence entre le volume restant à délivrer pour le client k à la période de temps 1, et le nombre de prospects Internet délivré pour le client k depuis la période de temps 1 à la période de temps courante t' .

Formellement,

$$\sum_{j:x_{ijt'} \in X_{t'}} x_{ijt'} \leq Ro_i - \sum_{t''=1}^{t'-1} \sum_{j:x_{ijt''} \in X_{t''}} x_{ijt''}, \quad \forall i \in O, \forall t' \in T' \quad (4.1)$$

Tel que, Ro_i est le nombre de prospects Internet restant à délivrer à la commande i , à la période

de temps 1, qui est défini dans l'équation (3.2) dans la section 3.3.3.

$$\sum_{i,j:x_{ij}^{t'} \in X_{t'}} a_{ik} x_{ij}^{t'} \leq Rc_k - \sum_{t''=1}^{t'-1} \sum_{i,j:x_{ij}^{t''} \in X_{t''}} a_{ik} x_{ij}^{t''}, \quad \forall k \in C, \forall t' \in T' \quad (4.2)$$

Où,

– a_{ik} est une variable binaire décrite de la manière suivante :

$$a_{ik} = \begin{cases} 1 & \text{si la commande } i \text{ provient du client } k \\ 0 & \text{sinon} \end{cases}$$

– Rc_k est le nombre de prospects Internet restant à délivrer au client k , à la période de temps 1, qui est défini dans l'équation (3.4) dans la section 3.3.3.

Contrainte liée au nombre de ventes et à l'exclusivité : la contrainte (4.3) exprime le fait que, sur toutes les g périodes de temps prises en compte dans l'optimisation, tout prospect Internet exclusif (tout prospect Internet non exclusif) peut être vendu au plus une fois (au plus b fois respectivement). Formellement,

$$S_j + \sum_{t''=1}^{t'-1} \sum_{i:x_{ij}^{t''} \in X_{t''}} x_{ij}^{t''} + b \sum_{i:x_{ij}^{t'} \in E_{t'}} x_{ij}^{t'} + \sum_{i:x_{ij}^{t'} \in \bar{E}_{t'}} x_{ij}^{t'} \leq b, \quad \forall j \in P_{t'}, \forall t' \in T' \quad (4.3)$$

Où, S_j est le nombre de ventes antérieures (avant la période de temps 1) du prospect Internet j et b est le nombre de reventes autorisées des prospects Internet non exclusifs.

Nous notons qu'à la période de temps 1, les termes $\sum_{t''=1}^{t'-1} \sum_{j:x_{ij}^{t''} \in X_{t''}} x_{ij}^{t''}$, $\sum_{t''=1}^{t'-1} \sum_{i,j:(ij)^{t''} \in X_{t''}} a_{ik} x_{ij}^{t''}$

et $\sum_{t''=1}^{t'-1} \sum_{i:x_{ij}^{t''} \in X_{t''}} x_{ij}^{t''}$ des contraintes (4.1), (4.2) et (4.3) respectivement, sont nuls. Par conséquent, les contraintes (4.1), (4.2) et (4.3) sont similaires à leurs analogues dans le programme mathématique mono-période à savoir les contraintes (3.1), (3.3) et (3.6) respectivement. Plus précisément, dans les contraintes (4.1) et (4.2), nous avons juste à vérifier que le nombre de prospects Internet affecté à chaque commande i (chaque client k) respecte le volume restant de la commande i (client k respectivement) à la période de temps 1, qui est Ro_k (Rc_k respective-

ment). Dans la contrainte (4.3), le nombre de ventes antérieures de chaque prospect Internet j est seulement le nombre de ventes avant la période de temps 1, qui est S_j .

Contrainte liée à la duplication des prospects Internet : le client de PdL possède la possibilité d'effectuer une ou plusieurs commandes. Cependant, nous ne pouvons pas affecter le même prospect Internet au même client plusieurs fois.

La contrainte (4.4) garantit qu'un client k ne reçoit pas un prospect Internet j plus qu'une fois, sur toutes les g périodes de temps prises en compte dans l'optimisation. Formellement,

$$\sum_{t'=1}^g \sum_{i:x_{ijt'} \in X_{t'}} a_{ik} x_{ijt'} \leq 1, \quad \forall k \in C, \quad \forall j \in P_{t'} \quad (4.4)$$

4.2.3 Fonction objectif

L'équation (4.5) représente la fonction objectif du modèle d'affectation multi-période qui cherche à maximiser le chiffre d'affaires, sur toutes les g périodes de temps considérées dans l'optimisation.

$$\max \sum_{t'=1}^g \sum_{i,j:x_{ijt'} \in X_{t'}} Pr_i x_{ijt'} \quad (4.5)$$

4.2.4 Programme mathématique multi-période

Dans cette section, nous présentons un modèle d'optimisation déterministe et multi-période P_3 en utilisant la programmation mathématique linéaire avec des variables binaires. P_3 permet d'optimiser à un instant t donné l'affectation des prospects Internet aux commandes en termes de chiffre d'affaires. Cette optimisation est effectuée en considérant plusieurs périodes de temps futures. Par conséquent, la connaissance concernant les mises à jour des quotas ainsi que les arrivées des nouveaux flux de prospects Internet sur les périodes de temps futures, est

prise en compte dans le processus d'affectation afin d'aboutir à une meilleure optimisation. Ceci est détaillé dans la prochaine section.

La fonction objectif de ce modèle est donnée par l'équation (4.5). Concernant les contraintes, elles sont décrites par les équations (4.1), (4.2), (4.3) et (4.4).

$$\begin{array}{l}
 \left. \begin{array}{l}
 \max \quad f(X_{T'}) = \sum_{t'=1}^g \sum_{i,j:x_{ijt'} \in X_{t'}} Pr_i x_{ijt'} \\
 s.t. \\
 \sum_{j:x_{ijt'} \in X_{t'}} x_{ijt'} \leq Ro_i - \sum_{t''=1}^{t'-1} \sum_{j:x_{ijt''} \in X_{t''}} x_{ijt''}, \quad \forall i \in O, \forall t' \in T' \\
 \sum_{i,j:x_{ijt'} \in X_{t'}} a_{ik} x_{ijt'} \leq Rc_k - \sum_{t''=1}^{t'-1} \sum_{i,j:x_{ijt''} \in X_{t''}} a_{ik} x_{ijt''}, \quad \forall k \in C, \forall t' \in T' \\
 S_j + \sum_{t''=1}^{t'-1} \sum_{i:x_{ijt''} \in X_{t''}} x_{ijt''} + b \sum_{i:x_{ijt'} \in E_{t'}} x_{ijt'} + \sum_{i:x_{ijt'} \in \bar{E}_{t'}} x_{ijt'} \leq b, \quad \forall j \in P_{t'}, \forall t' \in T' \\
 \sum_{t'=1}^g \sum_{i:x_{ijt'} \in X_{t'}} a_{ik} x_{ijt'} \leq 1, \quad \forall j \in P_{t'}, \forall k \in C \\
 x_{ijt'} \in \{0,1\} \quad \forall i \in O, \forall j \in P_{t'}, \forall t' \in T'
 \end{array} \right\} P_3 \quad (4.6)
 \end{array}$$

4.3 Implication des futurs flux de prospects Internet et ré-initialisations des quotas dans le modèle multi-période

Afin d'obtenir une meilleure optimisation, nous devons tenir compte à la fois de l'évolution continue des flux de prospects Internet (PI) ainsi que des mises à jour régulières des quotas de commandes dans le processus d'affectation. Dans cette section, nous montrons comment ces deux exigences sont prises en compte dans le programme mathématique P_3 .

4.3.1 Implication de l'évolution des flux de prospects Internet dans le modèle d'optimisation multi-période

À présent, nous montrons la manière dont la connaissance concernant les nouveaux flux de prospects Internet arrivant de façon continue et non déterministe sur les futures périodes de temps, est incluse dans le modèle d'optimisation multi-période. Tel qu'il a été précisé précédemment, nous supposons que la prévision des flux de prospects Internet, à chaque période de temps t' , est une donnée connue à l'avance. Ainsi, nous pouvons construire, à la période de temps 0 (période de temps courante), l'ensemble de toutes les affectations possibles pour chacune des périodes de temps t' prises en compte dans le processus d'optimisation. Ces ensembles sont utilisés dans notre modèle d'optimisation multi-période afin d'affecter les prospects Internet aux commandes sur toutes les périodes de temps de l'ensemble T . Par exemple, considérons que :

1. Nous effectuons l'affectation sur 3 périodes de temps à savoir les périodes de temps 0, 1 et 2.
2. Nous disposons de 2 commandes et 2 prospects Internet à la période de temps 0.
3. La prévision pour les périodes de temps 1 et 2 présente l'arrivée de 3 prospects Internet à savoir le prospect Internet 3 et 4 à la période de temps 1 et le prospect Internet 5 à la période de temps 2.

Par conséquent, l'ensemble de toutes les affectations possibles de chacune des périodes de temps 0, 1 et 2 est présenté dans les tables 4.7, 4.8 et 4.9 respectivement.

t_0	PI 1	PI 2
Commande 1	x_{110}	x_{120}
Commande 2	x_{210}	/

Tableau 4.7 — Ensemble de toutes les affectations possibles à la période de temps 0.

Ces ensembles sont construits afin d'être fournis au modèle d'optimisation multi-période à la période de temps 0. Ainsi, la connaissance concernant les flux de prospects Internet sur les 3

4.3. Implication des futurs flux de prospects Internet et réinitialisations des quotas dans le modèle multi-période

t_1	PI 1	PI 2	PI 3	PI 4
Commande 1	x_{110}	x_{120}	prévision	prévision
Commande 2	x_{210}	/	prévision	prévision

Tableau 4.8 — Ensemble de toutes les affectations possibles à la période de temps 1.

t_2	PI 1	PI 2	PI 3	PI 4	PI 5
Commande 1	x_{110}	x_{120}	prévision	prévision	prévision
Commande 2	x_{210}	/	prévision	prévision	prévision

Tableau 4.9 — Ensemble de toutes les affectations possibles à la période de temps 2.

périodes de temps, est prise en compte dans le processus d'affectation afin de procéder à une meilleure optimisation.

4.3.2 Implication des réinitialisations des quotas de commandes dans le modèle d'optimisation multi-période

Dans cette section, nous montrons comment les réinitialisations des quotas de commandes sont prises en compte dans le modèle d'optimisation multi-période. Tel qu'il a été décrit précédemment, les clients ainsi que les commandes possèdent des quotas horaires, journaliers et mensuels, qui sont remis à zéro à chaque heure, jour et mois, respectivement. Pour intégrer cette connaissance nous devons inclure dans le modèle d'optimisation multi-période les périodes de temps dans lesquelles nous effectuons les réinitialisations. À titre d'exemple, afin d'intégrer à la fois la réinitialisation horaire et journalière des quotas de commandes (quotas de clients), nous pouvons envisager 24 périodes d'une heure ou 48 périodes de 30 minutes ... etc. Cependant, si nous utilisons des périodes de 2 heures, il suffit de considérer 12 périodes afin d'inclure la réinitialisation journalière, cela dit il n'est plus possible d'envisager toutes les réinitialisations horaires. En d'autres termes, la durée de la période ne doit pas dépasser une heure afin d'inclure toutes les réinitialisations horaires. En outre, les périodes de temps doivent ensuite être cumulées jusqu'à l'obtention d'un horizon de 24 heures afin d'inclure la réinitialisation journalière. En effet, si nous considérons, par exemple, 12 périodes de 2 heures ou 6 périodes de 4 heures, la réinitialisation quotidienne est prise en compte dans les deux cas,

mais seulement 12 réinitialisations horaires parmi 24 sont prises en compte pour le premier cas et 6 réinitialisations horaires parmi 24 pour le second. Par ailleurs, les périodes t' de 4 heures ne permettent de prendre en compte que les mises à jour à : 00h00, 4h00, 8h00 ... etc, alors que nous avons une mise à jour toutes les heures.

Afin de définir la configuration la plus pertinente nous avons menés des expérimentations qui sont détaillées dans la section 6.2.2.4.

4.4 Algorithme pour l'affectation multi-période

Le modèle mathématique P_3 présenté dans la section 4.2.4 permet d'optimiser, à un instant t donné, l'affectation des prospects Internet aux commandes en considérant plusieurs périodes de temps futures. Cependant, l'une des caractéristiques principales de PdL réside dans le fonctionnement continu de son système. Ainsi, nous devons optimiser l'affectation des prospects Internet plusieurs fois par jour. Cette optimisation est assurée par l'algorithme 4.1 présenté ci-dessous.

La problématique de l'affectation multi-période est caractérisée par un horizon glissant, ce qui signifie que le temps est modélisé par plusieurs instants. Ainsi, l'algorithme 4.1 établit des affectations de façon successives tout au long de la journée. En effet, l'algorithme 4.1 optimise l'affectation des prospects Internet aux commandes à chaque instant t en considérant plusieurs périodes de temps.

Dans un premier temps, l'algorithme 4.1 effectue un filtre sur l'espace de recherche en ne gardant dans l'ensemble $X_{t'}$ que les $x_{ij_{t'}}$ dont le prospect Internet j correspond aux caractéristiques de la commande i , à la période de temps t' .

Dans un second temps et afin de décider à chaque instant t de l'ensemble T , quels prospects Internet affecter à quelles commandes et à quelles périodes de temps, l'algorithme 4.1 fait appel au modèle d'optimisation multi-période P_3 présenté dans la section 4.2.4 en lui fournissant l'ensemble des variables de décision $X_{t'}$. Suite au résultat $X'_{t'}$ retourné par le modèle d'optimisation P_3 , l'algorithme 4.1 effectue les affectations des prospects Internet aux commandes aux

Algorithme 4.1 — Affectation multi-période()

Data :

$T = \{t : t \text{ est l'heure pile} \}$.

$T' = \{t' : t' \in \{1, \dots, g\}\}$, où $t' = 1$ est la période de temps courante et $t' \in \{2, \dots, g\}$ sont les périodes de temps futures.

$P_{t'} = \{j : j \in \{1, \dots, m_{t'}\}\}$ est l'ensemble de prospects Internet à l'instant t' et $m_{t'}$ est le nombre de prospects Internet à la période de temps t' .

$O = \{i : i \in \{1, \dots, n\}\}$ est l'ensemble de commandes. Cet ensemble reste constant sur toutes les périodes de temps t' .

$C = \{k : k \in \{1, \dots, r\}\}$ est l'ensemble de clients. Cet ensemble reste constant sur toutes les périodes de temps t' .

$X_{t'} = \{x_{ij_{t'}} : i \in O, j \in P_{t'}, t' \in T', x_{ij_{t'}}$ est une variable binaire $\}$ est l'ensemble des variables de décision de chacune des périodes de temps de l'ensemble T' .

$X'_{t'} = \{x'_{ij_{t'}} : i \in O, j \in P_{t'}, t' \in T', x'_{ij_{t'}}$ est une variable binaire $\}$ est le résultat fourni par le programme mathématique P_3 présenté dans la section 4.2.4.

```

1 while  $t \in T$  do
2    $X_{t'} \leftarrow \emptyset$ 
3   while  $t' \in T'$  do
4     while  $i \in O$  do
5       while  $j \in P_{t'}$  do
6         if le prospect Internet  $j$  correspond aux caractéristiques de
           la commande  $i$  à la période de temps  $t'$  then
7            $X_{t'} \leftarrow X_{t'} \cup \{x_{ij_{t'}}\}$ 
8         end
9       end
10    end
11  end
12   $X'_{t'} \leftarrow P_3(X_{t'})$ 
13  while  $x'_{ij_{t'}} \in X'_{t'}$  do
14    if  $x'_{ij_{t'}} = 1$  then
15      Affecter le prospect Internet  $j$  à la commande  $i$  à la période de
        temps  $t'$ 
16    end
17  end
18  Mettre à jour les quotas des commandes ainsi que le nombre de reventes
    des prospects Internet.
19  Intégrer les nouveaux flux de l'offre et de la demande.
20 end

```

périodes de temps adéquates. En effet, si $x'_{ijt'} = 1$ alors le prospect Internet j sera affecté à la commande i à la période de temps t' , sinon ($x'_{ijt'} = 0$) l'affectation n'aura pas lieu.

Enfin, l'algorithme 4.1 intègre les nouveaux flux de l'offre et de la demande des prospects Internet et passe à l'instant t suivant après avoir établi les mises à jour des quotas des commandes et des nombres de reventes de prospects Internet.

Ainsi, l'algorithme 4.1 est caractérisé par deux principaux paramètres, à savoir :

1. L'intervalle de distribution : il s'agit de l'instant t qui représente l'intervalle de temps entre deux affectations.
2. La périodicité : elle représente le nombre de périodes de temps t' ainsi que leur durée, considérés à chaque instant t .

Nous traiterons dans le détail l'impact du choix de l'intervalle de distribution ainsi que la périodicité sur l'optimisation, dans les sections 6.2.2.2 et 6.2.2.4 respectivement.

4.5 Conclusion

L'objectif de ce chapitre était de présenter une extension multi-période du modèle d'affectation. Nous avons commencé par introduire des exemples explicatifs dans le but de montrer l'intérêt d'une telle extension. Par la suite, nous avons présenté les variables de décision, les contraintes ainsi que la fonction objectif du modèle d'optimisation multi-période. Enfin, nous avons montré comment le modèle d'affectation multi-période considère à la fois l'évolution continue des flux de prospects Internet et les mises à jour régulières des quotas de commandes, afin d'obtenir une meilleure optimisation.

En outre, suite au fonctionnement continu du système de la place de marché PdL, nous avons proposé un algorithme qui permet d'établir des affectations plusieurs fois par jour et en considérant à chaque fois plusieurs périodes de temps futures. Il permet également de considérer de manière instantanée les nouveaux flux et de mettre à jour les quotas des commandes ainsi que le nombre de reventes des prospects Internet après chaque affectation.

Une hypothèse forte considérée dans ce chapitre est la connaissance à l'avance des futurs flux de prospects Internet pris en compte dans le modèle d'optimisation multi-période, sont des données connues à l'avance par le biais d'un modèle de prévision. Le défi consiste donc à définir les paramètres de ce modèle de prévision qui conviendront au fonctionnement du modèle d'optimisation multi-période, afin de pouvoir combiner les deux modèles. Dans le prochain chapitre, nous montrerons comment nous construisons un modèle de prévision des flux de prospects Internet tout en tenant compte des caractéristiques de notre modèle d'optimisation multi-période.

Chapitre 5

Modèle de prévision pour l'optimisation multi-période

5.1 Introduction

Nous rappelons que nous avons commencé par proposer un premier modèle prenant en compte les contraintes du problème, ainsi que les deux objectifs à savoir la maximisation du chiffre d'affaires et la satisfaction des clients. Après avoir testé la robustesse du modèle, nous avons constaté que ce modèle ne considère pas l'évolution du système sur les périodes de temps à venir. Par conséquent, nous avons pensé à un modèle permettant l'affectation des prospects Internet sur plusieurs périodes de temps. De ce fait, le modèle multi-période a été proposé. Cependant, nous pourrions avoir une meilleure optimisation en prévoyant l'arrivée des nouveaux prospects Internet sur les périodes de temps prises en compte dans le processus d'affectation. Ainsi, définir un modèle de prévision s'avère donc nécessaire pour améliorer les résultats fournis par le modèle d'optimisation multi-période.

Dans ce chapitre, nous cherchons à combiner un modèle de prévision avec notre modèle d'optimisation multi-période afin d'améliorer les résultats du système d'affectation chez PdL.

La majorité des travaux de recherche traite soit des modèles de prévision dont l'objectif est de

réduire l'incertitude liée à la méconnaissance de l'avenir, soit ceux de l'affectation standard dont le but est de résoudre de manière optimale l'affectation de n tâches à n machines de sorte à ce que le coût minimum ou maximum de profit peut être obtenu (voir le chapitre 2).

Notre proposition consiste à articuler un modèle de prévision et un modèle d'affectation multi-période. Cette approche comme nous le verrons par la suite va permettre d'améliorer l'optimisation, en intégrant la connaissance concernant l'évolution continue des flux de prospects Internet dans le modèle d'optimisation multi-période.

5.2 Articulation prévision/optimisation

Dans le but de pouvoir cerner tous les aspects du problème, nous avons constitué un contexte très particulier dans lequel il faudra articuler un modèle de prévision avec un autre modèle d'optimisation multi-période.

Ainsi, dans cette section, nous proposons un système de prévision pour notre modèle d'affectation multi-période dans un contexte caractérisé par des flux de prospects Internet évoluant constamment au fil du temps.

En particulier, nous montrons comment nous construisons un modèle de prévision des flux de prospects Internet tout en tenant compte des caractéristiques de notre modèle d'optimisation multi-période. Cette combinaison a pour objectif d'optimiser le système d'affectation de la place de marché PdL.

Par ailleurs, construire un modèle de prévision nécessite de définir les données à prévoir avant d'envisager toute méthode de prévision. En d'autres termes, nous devons choisir les paramètres du modèle de prévision à savoir les données historiques appropriées, le pas de temps de la prévision ainsi que l'horizon de la prévision (voir la section 2.3.4).

Le défi consiste donc à définir les paramètres du modèle de prévision qui conviendront au fonctionnement de notre modèle d'optimisation multi-période.

5.2.1 Pourquoi tenir compte des prévisions des flux de prospects Internet ?

Nous travaillons dans un contexte caractérisé par des flux de prospects Internet évoluant en permanence de manière non déterministe au fil du temps. Afin d'optimiser le chiffre d'affaires de PdL, nous devons prendre en compte ces flux de prospects Internet dans le processus d'affectation. En d'autres termes, l'objectif est de permettre au modèle d'affectation multi-période défini dans le chapitre précédent d'optimiser l'affectation en tirant profit de la connaissance concernant l'évolution des flux de prospects Internet. En effet, nous fournissons à notre modèle d'optimisation multi-période l'ensemble de toutes les affectations possibles à chaque période de temps t' prise en compte dans l'optimisation multi-période. Par conséquent, nous devons calculer les flux de prospects Internet arrivant sur les périodes de temps futures.

La définition d'un modèle de prévision des flux de prospects Internet est donc nécessaire afin d'améliorer le modèle d'optimisation multi-période. Par conséquent, nous proposons dans ce qui suit, d'articuler un modèle de prévision avec notre modèle d'optimisation multi-période.

5.2.2 Données à prévoir

Afin d'intégrer les flux de prospects Internet dans notre modèle d'optimisation multi-période, nous devons prévoir pour chaque période de temps t' prise en compte dans le modèle d'optimisation multi-période : (i) le nombre de prospects Internet arrivant sur cette période de temps ainsi que (ii) les commandes pour lesquelles ces prospects Internet peuvent être affectés.

Une première approche pourrait consister à prévoir le nombre de prospects Internet pour chaque commande. Cependant, cette prévision reste insatisfaisante car elle génère un problème de doublons. En d'autres termes, une grande partie des prospects Internet est partagée par plusieurs commandes. Par conséquent, si nous prévoyons le nombre de prospects Internet pour chaque commande, les prospects Internet partagés par plusieurs commandes apparaîtront dans les données historiques de chacune de ces commandes et seront donc comptabilisés plusieurs fois, d'où l'appellation "problème des doublons". En outre, nous avons besoin de savoir pour

chaque prospect Internet prévu, l'ensemble des commandes dont les caractéristiques ont été satisfaites. Cependant, la prévision du nombre de prospects Internet pour chaque commande ne fournit pas l'information concernant la correspondance ou pas entre les prospects Internet et les commandes.

Afin de surmonter ces problèmes, nous proposons de prévoir le nombre de prospects Internet pour chaque groupe de commandes partageant les mêmes prospects Internet. Cette approche permet de prévoir le nombre de prospects Internet ainsi que les commandes pour lesquelles chaque prospect Internet prévu peut être affecté. Par conséquent, chaque groupe de commandes partageant les mêmes prospects Internet représente une série chronologique pour laquelle une étude de prévision sera effectuée. L'algorithme de construction de telles séries chronologiques est discuté dans la section 5.3.1.

5.2.3 Pas de temps et horizon de la prévision

Nous sommes confrontés à un problème de prévision complexe, où le défi consiste à intégrer les résultats du modèle de prévision dans notre modèle d'optimisation multi-période.

La première difficulté vient du fait que la prévision des prospects Internet devrait être compatible avec le problème d'optimisation multi-période décrit dans le chapitre précédent. En effet, nous devons prévoir le nombre de prospects Internet ainsi que les commandes pour lesquelles chaque prospect Internet prévu peut être affecté. La deuxième difficulté est liée au fait que l'affectation des prospects Internet aux commandes est fondée, à chaque instant t , sur un horizon glissant des périodes de temps futures. Par conséquent, nous devons calculer les flux de prospects Internet sur ces périodes de temps futures de sorte que, d'une part, le pas de temps de la prévision doit être égale à la durée de la période du modèle d'optimisation multi-période, d'autre part, l'horizon doit représenter le nombre des futures périodes de temps considérées dans le modèle d'optimisation multi-période.

Le fait que le pas de temps de prévision et la durée des périodes de temps du modèle d'optimisation multi-période doivent être les mêmes génère une autre difficulté. En effet, les quotas

des commandes sont mis à jour à chaque heure. Par conséquent, l'intégration de cette information dans le modèle d'optimisation multi-période nécessite la considération d'une durée de périodes de temps d'une heure. Cependant, un pas de temps de prévision d'une heure ne garantit pas nécessairement une bonne prévision en raison de la faible densité des flux de prospects Internet sur une telle durée de période de temps. Par conséquent, nous devrions envisager une durée de périodes de temps qui dépasse l'heure afin d'améliorer les prévisions. Avec une telle approche, toutes les mises à jour horaires des quotas de commandes ne sont pas intégrées dans le modèle d'optimisation multi-période.

À titre d'exemple, la considération d'une durée de périodes de 4 heures conduit à prendre en compte les mises à jour des quotas à : 00h00, 04h00, 08h00 ... etc, alors que les quotas des commandes sont mises toutes les heures. Cependant, un pas de temps de prévision de 4 heures fournira des résultats meilleurs qu'un pas de temps de prévision d'une heure. Cela est dû au fait que la densité des flux de prospects Internet est plus importante sur 4 heures que sur une heure. En conséquence, nous devons choisir une bonne solution de compromis. Il s'agit de considérer une certaine durée qui permet d'inclure un maximum de réinitialisations de quotas de commandes dans le de modèle d'optimisation multi-période, mais aussi qui soit suffisante pour avoir un flux important de prospects Internet permettant d'avoir une bonne prévision. Nous étudierons cela dans le détail dans la section 6.3.2.3.

5.3 Système de prévision

Dans les sections précédentes, nous avons souligné les avantages ainsi que les difficultés liées à l'intégration de la prévision dans le modèle d'optimisation multi-période. Dans cette section, nous présentons, dans un premier temps, un algorithme calculant les séries chronologiques discutées dans la section 5.2.2. Dans un second temps, nous définissons un modèle de prévision calculant les arrivées des flux de prospects Internet en se basant sur ces séries chronologiques. Ce modèle de prévision est ensuite intégré dans le modèle d'optimisation multi-période.

5.3.1 Construction des séries chronologiques

Nous proposons un algorithme qui calcule l'ensemble des séries chronologiques. Plus précisément, chaque série chronologique représente un groupe de commandes partageant les mêmes prospects Internet. Ainsi, l'algorithme que nous proposons calcule à la fois : (i) l'historique des groupes de commande partageant les mêmes prospects Internet et (ii) pour chaque groupe de commandes, l'historique du nombre de prospects Internet satisfaisant simultanément toutes les commandes de ce groupe.

Pour ce faire, l'algorithme considère l'ensemble des périodes historiques. Supposons que nous possédons deux périodes historiques h_1 et h_2 . Les correspondances prospect Internet(PI)/commande(C) aux périodes historiques h_1 et h_2 sont définies par les tables 5.1 et 5.2 respectivement. Autrement dit, chacune de ces tables définit l'ensemble des prospects Internet reçu à une période historique donnée ainsi que l'information concernant la correspondance ou pas de chacun de ces prospects Internet par rapport à chaque commande (1 si le prospect Internet vérifie les caractéristiques de la commande, 0 sinon). L'algorithme que nous proposons traite successivement l'ensemble des périodes historiques. Dans notre exemple, il s'agit de :

1. Parcourir l'ensemble des prospects Internet de la période historique h_1 .
2. Pour chaque prospect Internet reçu à la période historique h_1 , construire le groupe de commandes partageant ce prospect Internet :
 - (a) si ce groupe de commandes a déjà été créé alors incrémenter de un le nombre de prospects Internet satisfaisant simultanément ce groupe de commandes à la période historique h_1 .
 - (b) sinon : (i) intégrer ce groupe de commande à l'ensemble des groupes de commandes, (ii) initialiser à 1 le nombre de prospects Internet satisfaisant simultanément ce groupe de commande à la période historique h_1 et (iii) initialiser à 0 le nombre de prospects Internet satisfaisant simultanément ce groupe de commandes aux autres périodes historiques (la période historique h_2 pour notre exemple).

3. Refaire les étapes 1 et 2 pour le reste des périodes historiques (la période historique h_2 pour notre exemple).

PI/C	C1	C2	C3
PI 1	1	0	0
PI 2	0	1	0
PI 3	1	0	0
PI 4	1	1	0
PI 5	1	0	1
PI 6	0	1	0
PI 7	0	0	1
PI 8	1	1	0
PI 9	1	1	0
PI 10	1	0	1

Tableau 5.1 — Correspondances PI/C à la période historique h_1

PI/C	C1	C4	C5
PI 11	1	0	0
PI 12	1	1	1
PI 13	1	1	0

Tableau 5.2 — Correspondances PI/C à la période historique h_2

L'application de cet algorithme sur les deux périodes historiques représentées par les deux tables 5.1 et 5.2 fournira les séries chronologiques définies dans la table 5.3. En effet, chaque ligne de la table 5.3 représente une série chronologique définissant pour un groupe de commandes donné partageant les mêmes prospects Internet, le nombre de prospects Internet satisfaisant simultanément ce groupe de commandes à chaque période historique. La définition formelle de cette méthode est présentée par l'algorithme 5.1.

L'algorithme 5.1 génère l'ensemble de toutes les séries chronologiques. En effet, pour chaque groupe de commandes l appartenant à l'ensemble des groupes de commandes ∂ , nous avons la série chronologique $(y_{lh})_{h \in H}$ représentant le nombre de prospects Internet satisfaisant simultanément le groupe de commandes l à chaque période historique h . En outre, pour calculer un groupe de commandes l partageant un prospect Internet donné j à une période histo-

Groupe de commandes / Période historique	h_1	h_2
{C1}	2	1
{C2}	2	0
{C1,C2}	3	0
{C1,C3}	2	0
{C3}	1	0
{C1,C4,C5}	0	1
{C1,C4}	0	1

Tableau 5.3 — Séries chronologiques

rique donnée h , l'algorithme Séries-Chronologiques() fait appel à l'algorithme 5.2 : Groupe-Commandes(j, h, O, Z).

Nous montrerons dans la section 5.3.2 comment le résultat de l'algorithme 5.1 à savoir la table des séries chronologiques, servira comme données historiques afin de calculer les prévisions pour notre modèle d'optimisation multi-période.

5.3.2 Modèle de prévision et son intégration dans le modèle d'optimisation multi-période

Afin de calculer la prévision de chaque série chronologique $(y_{lh})_{h \in H}$ tel que $l \in \partial$ définie dans la section 5.3.1, nous utilisons l'ensemble des modèles *ETS* implémenté dans le logiciel *R* (voir la section 2.3.6). L'ensemble des modèles *ETS* prend en entrée une série chronologique et retourne la méthode de lissage appropriée à la série chronologique. Dans notre cas, la méthode de lissage retourne pour chaque groupe de commandes partageant les mêmes prospects Internet, le nombre de prospects Internet satisfaisant simultanément toutes les commandes de ce groupe à chaque période de l'horizon de la prévision.

Nous rappelons que dans le modèle d'optimisation multi-période défini dans le chapitre précédent, nous considérons, à chaque instant t , un ensemble T' de périodes de temps de sorte que $T' = \{t' : t' \in \{1, \dots, g\}\}$, où $t' = 1$ est la période de temps courante et $t' \in \{2, \dots, g\}$ sont les périodes de temps futures. Pour pouvoir intégrer la connaissance concernant les arrivées de nouveaux flux de prospects Internet sur les g périodes de temps considérées dans le modèle

Algorithme 5.1 — Séries-Chronologiques()

Output : Un ensemble A représentant un historique des arrivées de prospects Internet pour chaque groupe de commandes l partageant les mêmes prospects Internet ;

Data :

$H = \{h : h \in \{1, \dots, d\}\}$ est l'ensemble des périodes historiques.

$P_h = \{j : j \in \{1, \dots, m_h\}\}$ est l'ensemble des prospects Internet.

$O = \{i : i \in \{1, \dots, n\}\}$ est l'ensemble des commandes.

$\partial = \{l : l \in \{1, \dots, e\}\}$ est l'ensemble des groupes de commandes.

$A = \{y_{lh} : l \in \partial, h \in H\}$ tel que y_{lh} est le nombre de prospects Internet partagés par le groupe de commandes l à la période de temps historique h .

$Z = \{z_{ijh} : i \in O, j \in P_h, h \in H\}$ tel que $z_{ijh} = 1$ si le prospect Internet j correspond aux caractéristiques de la commande i à la période de temps historique h , 0 sinon.

```

1  $A \leftarrow \emptyset$ 
2  $\partial \leftarrow \emptyset$ 
3 while ( $h \in H$ ) do
4   while ( $j \in P_h$ ) do
5      $l \leftarrow \text{Groupe} - \text{Commande}(j, h, O, Z)$ 
6     if  $l \in \partial$  then
7        $y_{lh} \leftarrow y_{lh} + 1$ 
8     end
9     else
10       $\partial \leftarrow \partial \cup \{l\}$ 
11      while  $h' \in H$  do
12        if  $h' = h$  then
13           $y_{lh} \leftarrow 1$ 
14           $A \leftarrow A \cup \{y_{lh}\}$ 
15        end
16        else
17           $y_{lh} \leftarrow 0$ 
18           $A \leftarrow A \cup \{y_{lh}\}$ 
19        end
20      end
21    end
22  end
23 end
24 return  $A$ 

```

Algorithme 5.2 — Groupe-Commandes(j, h, O, Z)

Output : Un groupe l de commandes partageant les mêmes prospects Internet ;

```

1  $l \leftarrow \emptyset$ 
2 while ( $i \in O$ ) do
3   | if ( $z_{ijh} = 1$ ) then
4   |   |  $l \leftarrow l \cup \{i\}$ 
5   | end
6 end
7 return  $l$ 

```

d'optimisation multi-période, nous avons besoin de savoir à la fois le nombre de prospects Internet, à chaque période de temps t' , ainsi que les commandes pour lesquelles ces prospects Internet peuvent être affectés. Ces deux informations sont fournies directement par la méthode de prévision *ETS* si nous remplissons les trois conditions suivantes : (i) la construction des séries chronologiques en utilisant l'algorithme 5.1 défini dans la section 5.3.1, (ii) le pas de temps de la prévision doit être égal à la durée de la période du modèle d'optimisation multi-période et (iii) l'horizon de la prévision doit correspondre au nombre des périodes de temps t' considérées dans le modèle d'optimisation multi-période. Ainsi, la méthode *ETS* retourne pour chaque période de l'horizon t' et pour chaque groupe de commandes partageant les mêmes prospects Internet, le nombre de prospects Internet satisfaisant simultanément toutes les commandes de ce groupe.

Afin d'illustrer le fonctionnement du modèle de prévision ainsi que son intégration dans le modèle d'optimisation multi-période, nous proposons cet exemple. Considérons que :

1. L'affectation multi-période s'effectue sur trois périodes de temps t' à savoir les périodes 0, 1 et 2.
2. Trois commandes et deux prospects Internet sont disponibles à la période de temps courante 0, de sorte à ce que l'ensemble des affectations possibles soit décrit dans la table 5.4.
3. Les résultats de la prévision calculés avec la méthode *ETS* pour les périodes de temps

futures (les périodes 1 et 2) sont représentés dans la table 5.5.

Compte tenu de la table des prévisions 5.5, nous avons, à la période de temps courante 0, l'information concernant : (i) le nombre de prospects Internet à chacune des périodes de temps futures 1 et 2 et (ii) les commandes pour lesquelles ces prospects Internet peuvent être affectés. Par conséquent, nous pouvons construire, à la période de temps courante 0, l'ensemble de toutes les affectations possibles concernant les futures périodes de temps 1 et 2, nécessaires au modèle d'optimisation multi-période. En d'autres termes, à la période de temps 1 nous possédons un PI pour la commande 1 (voir la table 5.6) et, à la période de temps 2, nous possédons deux PIs pour les commandes 2 et 3 (voir la table 5.7). Ainsi, les variables de décision du modèle d'optimisation multi-période sont : x_{111} , x_{121} , x_{131} , x_{211} , x_{242} , x_{252} , x_{321} , x_{342} et x_{352} . À titre d'exemple, la variable de décision x_{342} représente l'affectation ou pas du prospect Internet 4 à la commande 3 à la période de temps 2.

C / PI	PI 1	PI 2
C 1	x_{110}	x_{120}
C 2	x_{210}	/
C 3	/	x_{320}

Tableau 5.4 — Ensemble de toutes les affectations possibles à la période de temps 0.

Groupe de commandes / Horizon	Période 1	Période 2
{1}	1	0
{2,3}	0	2

Tableau 5.5 — Prévision fournie par la méthode *ETS*.

C / PI	PI 1	PI 2	PI 3
C 1	x_{110}	x_{120}	x_{131}
C 2	x_{210}	/	/
C 3	/	x_{320}	/

Tableau 5.6 — Ensemble de toutes les affectations possibles à la période de temps 1.

C / PI	PI 1	PI 2	PI 3	PI 4	PI 5
C 1	x_{111}	x_{121}	x_{131}	/	/
C 2	x_{211}	/	/	x_{242}	x_{252}
C 3	/	x_{321}	/	x_{342}	x_{352}

Tableau 5.7 — Ensemble de toutes les affectations possibles à la période de temps 2.

5.3.3 Prévision des flux de prospects Internet pour une nouvelle commande

Côté "offre", PdL reçoit une quantité importante de prospects Internet en temps réel, ces flux continus font évoluer en temps réel l'ensemble des prospects Internet disponibles. D'où la nécessité d'effectuer une étude prévisionnelle de ces flux. Cependant, l'étude prévisionnelle n'est pas utile côté "demande". En effet, PdL interagit avec de gros clients effectuant des commandes sur une longue durée appelée "période de contrat", en précisant les quotas souhaités par heure, jour et mois. Ainsi, PdL sait exactement les quotas demandés ainsi que l'heure de leurs réinitialisations et cela durant toute la période du contrat. En outre, PdL interagit avec un ensemble constant de clients, dans le sens où l'établissement de nouveaux contrats se fait de manière fortuite.

L'établissement d'un nouveau contrat enrichit l'ensemble des commandes disponibles. Cependant, ces nouvelles commandes manquent d'historique de flux des prospects Internet nécessaire à la construction des groupes de commandes et le calcul de la prévision de leurs flux de prospects Internet. En effet, nous devons fournir à l'algorithme 5.1 présenté dans la section 5.3.1 l'historique des flux de prospects Internet de toutes les commandes disponibles chez PdL. Par la suite, l'algorithme 5.1 retourne : (i) l'historique des groupes de commande partageant les mêmes prospects Internet et (ii) pour chaque groupe de commandes, l'historique du nombre de prospects Internet satisfaisant simultanément toutes les commandes de ce groupe.

Dans cette section nous proposons deux solutions pour remédier à l'absence de l'historique pour les nouvelles commandes. En effet, soit les caractéristiques de la nouvelle commande sont similaires à une ancienne commande déjà intégrée dans le système de la place de marché PdL, dans ce cas l'historique des flux de prospects Internet de la nouvelle commande est le même

que celui de l'ancienne commande. Sinon, nous récupérons de la base de données PdL tous les prospects Internet qui sont arrivés sur la plate forme PdL. Ceux vérifiant les caractéristiques de la nouvelle commande représenteront son historique. Ainsi, nous pouvons construire un historique des flux de prospects Internet pour toute nouvelle commande.

5.4 Conclusion

Dans ce chapitre, nous avons proposé un système de prévision pour notre modèle d'optimisation multi-période, dans un contexte caractérisé par des flux de prospects Internet évoluant continuellement au fil du temps. En particulier, nous avons montré comment nous pouvons construire un modèle de prévision de flux de prospects Internet tout en considérant les différentes caractéristiques de notre modèle d'optimisation multi-période.

Par la suite, nous avons expliqué dans la section 5.2.1 l'intérêt d'une telle prévision dans le modèle d'optimisation multi-période. Puis, nous avons discuté dans la section 5.2 la relation entre les deux modèles : prévision et optimisation multi-période et la possibilité de leur articulation. Dans la section 5.3, nous avons proposé un système de prévision considérant toutes les caractéristiques de notre modèle d'optimisation multi-période ainsi que la façon dont ce système de prévision est intégré dans le modèle d'optimisation multi-période.

Dans les chapitres précédents, nous avons proposé des modèles et des algorithmes d'optimisation et de prévision afin de répondre à la problématique posée par la place de marché PdL. Dans le chapitre suivant nous allons présenter une généralisation de ces modèles. Cette généralisation a pour ambition de définir un modèle aussi général que possible pouvant être appliqué à toute autre place de marché ayant les mêmes caractéristiques que celles de PdL.

Chapitre 6

Résultats expérimentaux

6.1 Introduction

Nous allons développer, dans ce chapitre, différentes expérimentations permettant d'étudier et de tester l'efficacité de nos modèles et algorithmes précédemment présentés ainsi que la qualité des résultats obtenus. Pour ce faire, ce chapitre est divisé en deux parties. La première évaluera les différents modèles et algorithmes d'optimisation. Quant à la seconde, elle appréciera le modèle de prévision. Ces expérimentations nous permettront alors de : (i) conclure sur l'adéquation de nos modèles et algorithmes aux attentes initialement formulées, (ii) valider l'utilité concrète des modèles et algorithmes proposés et (iii) fournir des orientations aux analystes qui voudront utiliser l'approche proposée.

6.2 Évaluation des modèles d'optimisation

Dans cette section, nous analysons le comportement des modèles d'optimisation proposés dans les sections 3.4, 3.6.2 et 4.2.4, en utilisant des exemples numériques. Le cadre d'application nécessaire aux expérimentations est réalisé en effectuant une étude numérique par l'utilisation des données réelles récupérées de la base de données PdL. Les modèles d'optimisation ont

été : (1) implémentés avec le logiciel *CPLEX* et le langage *JAVA* ensuite (2) testés directement sur le serveur de la place de marché PdL.

6.2.1 Schéma expérimental

Les expériences sont conçues pour répondre à cinq enjeux, à savoir :

1. La capacité d'optimisation des modèles proposés :
 - (a) Le modèle d'optimisation mono-objectif mono-période (mono-obj mono-pér) présenté dans la section 3.4.
 - (b) Le modèle d'optimisation mono-objectif multi-période (mono-obj multi-pér) présenté dans la section 4.2.4.
 - (c) Le modèle d'optimisation bi-objectif mono-période (bi-obj mono-pér) présenté dans la section 3.6.2.
2. Le choix d'un intervalle de distribution adéquat. En effet, certains paramètres peuvent influencer la performance des modèles d'optimisation proposés, comme l'intervalle de temps entre deux affectations.
3. Le passage à l'échelle des modèles d'optimisation en fonction de la taille des données.
4. La périodicité appropriée concernant le modèle d'optimisation multi-période (le nombre de périodes de temps ainsi que leur durée).
5. La densité des points sur la frontière efficace concernant le modèle d'optimisation bi-objectif.

Pour chacune de ces expériences, nous considérons une combinaison d'un certain nombre de paramètres, décrits dans la table 6.1. Plus spécifiquement :

1. Nous commençons par tester chacune des expériences à un instant t donné et pour généraliser les résultats, nous répétons l'expérience sur 30 instants différents.

Paramètre	Valeur considérée
Modèle	PdL, mono-obj mono-pér, mono-obj multi-pér, bi-obj mono-pér
Taille des données	$\times 1, \times 5, \times 10$
Intervalle de distribution	10 minutes, une heure
Nombre de périodes de temps et leur durée	24 périodes d'une heure, 12 périodes de 2 heures, 6 périodes de 4 heures

Tableau 6.1 — Paramètres du schéma expérimental.

2. Ensuite, nous testons chacune des expériences sur un jour. Pour ce faire, nous simulons un jour en effectuant les expériences à chaque instant t et pour généraliser les résultats, nous simulons 30 jours différents.
3. Enfin, nous testons chacune des expériences sur une période d'une semaine. Pour ce faire, nous simulons la semaine en effectuant les expériences à chaque instant t et pour généraliser les résultats, nous simulons 5 semaines différentes.

6.2.2 Expériences et résultats

6.2.2.1 Capacité d'optimisation

Expérience : l'objectif est d'étudier la capacité des modèles d'optimisation proposés de réaliser un meilleur chiffre d'affaires que celui de PdL.

Pour ce faire, nous récupérons d'abord, de la base de données PdL, les commandes et les prospects Internet de 30 instants t différents (tous à 10h00). En second lieu, nous réalisons l'affectation à chacun de ces instants, en utilisant deux modèles d'optimisation. Le premier constitue le modèle d'optimisation mono-obj mono-pér. Quant au second, il représente le modèle d'optimisation mono-obj multi-pér. Les résultats de l'affectation à savoir le chiffre d'affaires moyen généré par le modèle d'optimisation mono-obj mono-pér ($CA_{mono-obj-mono-pér}$) et celui généré par le modèle d'optimisation mono-obj multi-pér ($CA_{mono-obj-multi-pér}$), sont ensuite comparés au chiffre d'affaires généré par le système PdL (CA_{PdL}) et la proportion du gain est calculée.

	Différence moyenne	Proportion moyenne
$CA_{mono-obj-mono-pér} - CA_{PdL}$	3,63€	0,46%
$CA_{mono-obj-multi-pér} - CA_{PdL}$	-35,58€	-4,53%

Tableau 6.2 — Différences des chiffres d'affaires moyens à 10h00.

Résultats, description et interprétation : la table 6.2 présente la différence du chiffre d'affaires moyen, à un instant t donné (10h00 pour notre expérience), en comparant le chiffre d'affaires moyen généré par le système PdL aux chiffres d'affaires moyens générés par les deux modèles d'optimisation mono-obj mono-pér et mono-obj multi-pér.

Dans la table 6.2, nous observons qu'à un instant t donné, le chiffre d'affaires moyen généré par le modèle d'optimisation mono-obj mono-pér est plus important, que celui généré par le système PdL, de 0,46%. Cependant, le chiffre d'affaires moyen généré par le modèle d'optimisation mono-obj multi-pér est moins important, que celui généré par le système PdL, de 4,53%. Ainsi, le chiffre d'affaires moyen généré par le modèle d'optimisation mono-obj multi-pér est également moins important que celui généré par le modèle d'optimisation mono-obj mono-pér. Cette observation n'est pas surprenante car le modèle d'optimisation mono-obj mono-pér période effectue toutes les affectations possibles à l'instant courant (10h00 pour notre expérience). En revanche, le modèle d'optimisation mono-obj multi-pér ne vend pas certains prospects Internet disponibles à l'instant courant afin de les vendre ultérieurement à des prix plus élevés.

Pour confirmer cela, nous procédons au calcul des chiffres d'affaires moyens en simulant les deux modèles d'optimisation mono-obj mono-pér et mono-obj multi-pér, sur une période de 24 heures à partir de 10h00. Cette simulation s'effectue en mode hors ligne, ce qui signifie que nous ne considérons pas les nouveaux flux de prospects Internet arrivant au cours de ces 24 heures de simulation. Les résultats sont présentés dans la table 6.3 et indiquent qu'à la fin des 24 heures, le gain moyen du modèle d'optimisation mono-obj mono-pér s'est légèrement amélioré. En effet, il était de 0,46% à 10h00 (voir la table 6.2) et passe à 0,48% à la fin des 24 heures (voir la table 6.3). Cependant, le gain moyen du modèle d'optimisation mono-obj multi-pér était de -4,53% à 10h00 (voir la table 6.2) et passe à 4,14% à la fin des 24 heures

(voir la table 6.3).

Cette observation est cohérente avec les résultats des expériences précédentes. D'une part, le modèle d'optimisation mono-obj mono-pér n'arrive pas à améliorer son gain durant les 24 heures de simulation parce qu'il a vendu tous les prospects internet à 10h00. La légère amélioration du gain réalisé est expliquée par les mises à jour des quotas de commandes durant les 24 heures, ce qui rend certains prospects Internet disponibles (selon notre hypothèse, nous n'avons pas de nouveaux flux de prospects Internet qui arrivent durant les 24 heures de simulation). D'autre part, le modèle d'optimisation mono-obj multi-pér effectue progressivement les affectations au cours des 24 heures en tenant compte de l'évolution du système sur cette période. Par conséquent, à la fin des 24 heures, le modèle d'optimisation mono-obj multi-pér génère un chiffre d'affaires de 4,14% supérieur à celui du système PdL. Cependant, le modèle d'optimisation mono-obj mono-pér génère un chiffre d'affaires de seulement 0,48% supérieur à celui du système PdL. Ainsi, nous pouvons conclure que le modèle d'optimisation mono-obj mono-pér est meilleur que le système PdL et que le modèle d'optimisation mono-obj multi-pér est meilleur que le modèle d'optimisation mono-obj mono-pér.

	Différence moyenne	Proportion moyenne
$CA_{mono-obj-mono-pér} - CA_{PdL}$	4,70€	0,48%
$CA_{mono-obj-multi-pér} - CA_{PdL}$	40,17€	4,14%

Tableau 6.3 — Différences des chiffres d'affaires moyens après 24 heures de simulation à partir de 10h00.

La table 6.4 indique la différence entre les chiffres d'affaires moyens sur une journée générés par le système PdL, le modèle d'optimisation mono-obj mono-pér et le modèle d'optimisation mono-obj multi-pér. Nous observons que le modèle d'optimisation mono-obj multi-pér génère un chiffre d'affaires de 2% supérieur à celui du système PdL. Cependant, le modèle d'optimisation mono-obj mono-pér génère un chiffre d'affaires de seulement 0,5% supérieur à celui du système PdL. Nous expliquons cela par le fait que le modèle d'optimisation mono-obj mono-pér ne permet pas d'anticiper l'évolution du système sur les périodes de temps futures. En effet, il effectue des affectations à chaque instant t sans tenir compte de l'influence des résultats ac-

tuels sur les périodes de temps futures. Cependant, l'optimisation multi-période surmonte cette limite en optimisant les affectations à chaque instant t en tirant profit de la connaissance de l'évolution des flux de prospects Internet et des mises à jour des quotas de commandes, sur toutes les périodes de temps considérées dans l'optimisation.

	Différence moyenne	Proportion moyenne
$CA_{mono-obj-mono-pér} - CA_{PdL}$	14,27€	0,50%
$CA_{mono-obj-multi-pér} - CA_{PdL}$	56,06€	2,00%

Tableau 6.4 — Différences des chiffres d'affaires journaliers moyens.

L'évolution moyenne de la vente des prospects Internet sur une journée est détaillée dans la figure 6.1. D'une part, le chiffre d'affaires moyen généré par le modèle d'optimisation mono-obj mono-pér demeure plus important que celui généré par le système PdL durant la journée entière. D'autre part, nous pouvons voir qu'aux premières heures le modèle d'optimisation mono-obj multi-pér génère un chiffre d'affaires moyen inférieur à celui généré par le système PdL. Cette situation est maintenue jusqu'à 15h00. À partir de ce moment, le modèle d'optimisation mono-obj multi-pér entame une amélioration par rapport au système PdL. À 16h00, le modèle d'optimisation mono-obj multi-pér récupère tout le chiffre d'affaires perdu par rapport au système PdL et continue d'être meilleur jusqu'à 18h00. Néanmoins, à partir de 18h00, le modèle d'optimisation mono-obj multi-pér commence à perdre par rapport au système PdL. En fin de journée, le modèle d'optimisation mono-obj multi-pér génère un chiffre d'affaires plus important que celui généré par le système PdL. En outre, quelques minutes après 16h00, le chiffre d'affaires moyen généré par le modèle d'optimisation mono-obj multi-pér se maintient à un niveau important que celui généré par le modèle d'optimisation mono-obj mono-pér et ce jusqu'à la fin de la journée.

Ces observations confirment les résultats des expériences précédentes, dans le sens où le modèle d'optimisation mono-obj multi-pér tire profit de sa connaissance sur l'évolution des flux de prospects Internet ainsi que des mises à jour des quotas de commandes au cours des périodes de temps futures. Il préserve ainsi certains prospects Internet pour les vendre ultérieurement à des prix plus élevés. Nous observons ce phénomène en début ainsi qu'en fin de journée.

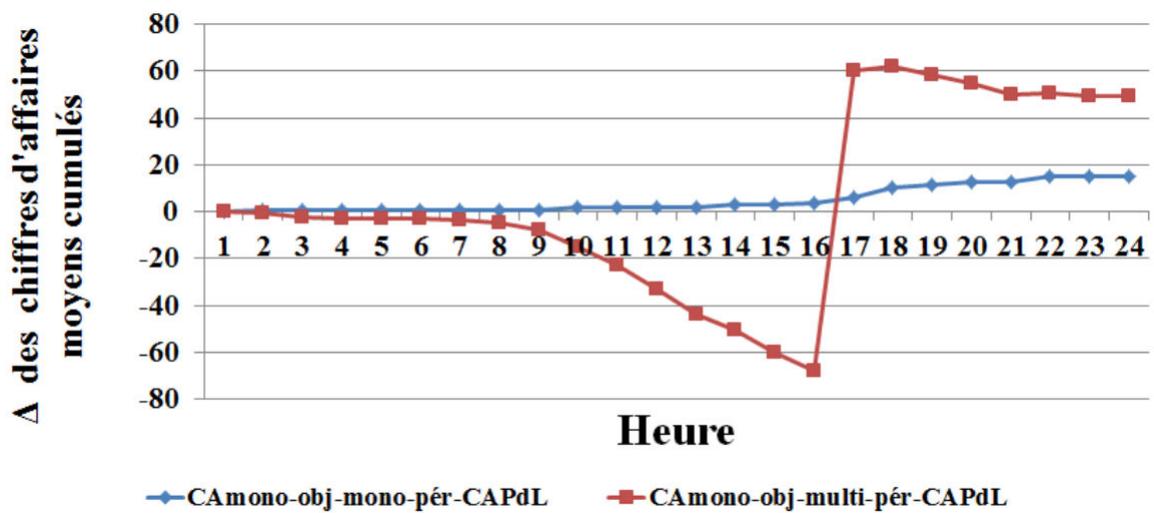


Figure 6.1 — Différences des chiffres d'affaires journaliers moyens cumulés.

De toute évidence, afin d'apprécier pleinement le gain apporté par le modèle d'optimisation mono-obj multi-pér, nous devons rejouer l'expérience sur une semaine (voir la section 6.2.2.3).

6.2.2.2 Intervalle de distribution

Expérience : Nos expériences ont pour objectif de déterminer l'intervalle de distribution de prospects Internet permettant d'obtenir une bonne optimisation. En d'autres termes, les modèles d'optimisation présentés précédemment effectuent à chaque instant t des affectations de prospects Internet aux commandes. Quel serait l'intervalle de temps approprié entre deux affectations afin d'obtenir de meilleurs résultats ?

Afin d'identifier l'intervalle de distribution le plus adéquat à nos modèles, nous simulons 30 jours avec les deux modèles d'optimisation mono-obj mono-pér et mono-obj multi-pér. Nous considérons deux intervalles de distribution à savoir 10 et 60 minutes. Les résultats de l'affectation sont ensuite comparés et la proportion du gain est calculée.

Résultats, description et interprétation : Les résultats expérimentaux sont présentés dans la table 6.5. Ils montrent que le modèle d'optimisation mono-obj mono-pér en considérant un

	10 mn	60 mn	Proportion moyenne
$CA_{mono-obj-mono-pér}$	2847,81€	2771,64€	2,70%
$CA_{mono-obj-multi-pér}$	2891,87€	2813,43€	2,80%

Tableau 6.5 — Chiffre d'affaires journalier moyen des modèles d'optimisation mono-obj mono-pér et mono-obj multi-pér, avec 10 et 60 minutes d'intervalle de distribution.

intervalle de distribution de 10 minutes, génère un chiffre d'affaires journalier moyen de 2,7% plus élevé que celui obtenu avec un intervalle de 60 minutes. Ils montrent également que le modèle d'optimisation mono-obj multi-pér en considérant un intervalle de distribution de 10 minutes, génère un chiffre d'affaires journalier moyen de 2,8% plus élevé que celui obtenu avec un intervalle de 60 minutes. Ceci peut être expliqué par l'expiration de la date de validité de certains prospects internet lorsque la distribution de ces derniers s'effectue toutes les heures. En effet, certaines commandes exigent des prospects Internet avec un certain âge : moins de 20 minutes, moins de 40 minutes ... ect. Par conséquent, si nous devons attendre une heure pour effectuer l'affectation, l'âge du prospect Internet dépasserait le critère des commandes et deviendrait donc non affectable.

6.2.2.3 Passage à l'échelle

Expérience : les expériences que nous proposons permettent d'étudier la relation entre la quantité de données (nombre de prospects Internet et de commandes) et le gain réalisé, en utilisant les modèles d'optimisation proposés. Plus précisément, nous cherchons à répondre aux questions expérimentales suivantes :

1. Existe t-il une relation entre la croissance de données et l'augmentation du gain ?
2. Les instances avec une quantité importante de données, sont-elles difficiles à résoudre ?

Les expériences sont menées comme suit. Dans un premier temps et dans l'optique de faire varier la taille des données, nous récupérons de la base de données PdL des données historiques sur 48 semaines. Par la suite, nous tirons aléatoirement parmi elles 5 semaines fictives. Les prospects Internet ainsi que les commandes de ces 5 semaines fictives construisent l'en-

semble *Données-réelles*. Afin de multiplier la taille de cet ensemble par 5 (par 10), nous tirons aléatoirement 5 jours, 35 fois (10 jours, 35 fois respectivement) à partir des 48 semaines de données historiques. Ainsi, nous construisons l'ensemble $Données-réelles_{\times 5}$ ($Données-réelles_{\times 10}$ respectivement). Dans un second temps, nous utilisons chacun des modèles d'optimisation précédemment présentés pour effectuer les affectations en considérant les différentes tailles de données. La proportion du gain en termes de chiffre d'affaires est étudiée. Le temps de calcul est également enregistré afin d'évaluer la traçabilité des modèles.

Résultats, description et interprétation : Les résultats expérimentaux sont présentés dans les tables 6.6 et 6.7. Ils indiquent la différence entre les chiffres d'affaires hebdomadaires moyens des modèles d'optimisation, en incrémentant progressivement la taille des données. Nous pouvons noter deux observations importantes à savoir :

1. Lorsque la taille des données est plus importante, la différence entre le chiffre d'affaires hebdomadaire moyen généré par le système PdL et ceux générés par les modèles d'optimisation mono-obj mono-pér et mono-obj multi-pér, est plus importante. Cette observation confirme que nos modèles d'optimisation sont évolutifs, dans le sens où plus la taille des données s'accroît plus le chiffre d'affaires généré est élevé.
2. Le modèle d'optimisation mono-obj multi-pér est passé de 2% de gain sur une journée (voir la table 6.4) à 5,48% sur une semaine (voir la table 6.7). Une telle évolution confirme le fait que le gain du modèle d'optimisation mono-obj multi-pér est plus appréciable sur une longue période.

	$CA_{mono-obj-mono-pér} - CA_{PdL}$	Proportion moyenne
<i>Données-réelles</i>	31,05€	0,14%
<i>Données-réelles</i> _{×5}	2376,34€	9,38%
<i>Données-réelles</i> _{×10}	2928,89€	9,81%

Tableau 6.6 — La différence entre le chiffre d'affaires hebdomadaire moyen généré par le système PdL et celui généré par le modèle d'optimisation mono-obj mono-pér, en fonction de la taille des données.

	$CA_{mono-obj-multi-pér} - CA_{PdL}$	Proportion moyenne
<i>Données-réelles</i>	1202,39€	5,48%
<i>Données-réelles</i> _{×5}	2444,18€	9,64%
<i>Données-réelles</i> _{×10}	3059,00€	10,25%

Tableau 6.7 — La différence entre le chiffre d'affaires hebdomadaire moyen généré par le système PdL et celui généré par le modèle d'optimisation mono-obj multi-pér, en fonction de la taille des données.

Nous avons également procédé à l'enregistrement du temps de calcul. Les résultats sont présentés dans la table 6.8 qui indique le temps de calcul moyen des modèles d'optimisation mono-obj mono-pér et mono-obj multi-pér, en faisant varier la taille des données à savoir *Données-réelles*, *Données-réelles*_{×5} et *Données-réelles*_{×10}. Les temps de calcul obtenus sont toujours inférieur à 2 minutes. Cela démontre que, d'une part, l'évolution des données ne génère pas des complications supplémentaires concernant le temps de résolution, d'autre part, tous les intervalles de distribution de prospects Internet proposés dans la section 6.2.2.2 sont réalisables.

	mono-obj mono-pér	mono-obj multi-pér
<i>Données-réelles</i>	3,54 sec	69,96 sec
<i>Données-réelles</i> _{×5}	5,34 sec	97,04 sec
<i>Données-réelles</i> _{×10}	5,73 sec	103,18 sec

Tableau 6.8 — Temps de calcul.

6.2.2.4 Périodicité du modèle d'optimisation multi-période

Expérience : une question expérimentale importante se pose concernant le nombre de périodes de temps ainsi que leur durée, considéré dans le modèle d'optimisation multi-période. En effet, le nombre des périodes de temps que nous considérons dans le modèle d'optimisation multi-période dépend de leur durée. À titre d'exemple, une durée d'une heure nécessite 24 périodes de temps afin d'inclure la réinitialisation journalière des quotas de commandes. Cependant, une durée de 2 heures (de 4) nécessite seulement 12 (6 respectivement) périodes de temps afin de prendre en compte la réinitialisation journalière des quotas de commandes. En

outre, si la durée de la période dépasse 1 heure, toutes les réinitialisations horaires des quotas de commandes ne sont pas prises en compte (voir la section 4.3.2). Par conséquent, nous étudions les deux questions suivantes :

1. La non considération de certaines réinitialisations horaires des quotas de commandes dans le modèle d'optimisation multi-période, exerce-t-elle une influence sur les résultats ?
2. Quel est le nombre de périodes de temps ainsi que leur durée adéquat au modèle d'optimisation multi-période ?

Pour étudier ces questions, nous simulons le modèle d'optimisation multi-période sur 30 jours avec différentes périodicités à savoir 24 périodes d'une heure ($24 \times 1h$), 12 périodes de 2 heures ($12 \times 2h$) et 6 périodes de 4 heures ($6 \times 4h$). Le chiffre d'affaires journalier moyen généré par chacune de ces périodicités est calculé afin d'étudier l'impact du choix de la durée et du nombre de périodes de temps sur le modèle d'optimisation multi-période.

Périodicité	Chiffre d'affaires journalier moyen
$6 \times 4h$	2894,70€
$12 \times 2h$	2892,00€
$24 \times 1h$	2891,87€

Tableau 6.9 — Chiffre d'affaires journalier moyen du modèle d'optimisation multi-période en fonction des différentes périodicités.

Résultats, description et interprétation : la table 6.9 présente le chiffre d'affaires journalier moyen du modèle d'optimisation multi-période selon trois périodicités différentes. Nous observons que les périodicités $6 \times 4h$, $12 \times 2h$ et $24 \times 1h$ génèrent un chiffre d'affaires journalier moyen de 2894 €, 2892 € et 2891 € respectivement. Une autre façon d'observer les résultats obtenus consiste à les analyser en termes de proportion moyenne de gain. La différence de gain entre $6 \times 4h$ et $12 \times 2h$ est de 0,093%. Par ailleurs, la différence de gain entre $6 \times 4h$ et $24 \times 1h$ est de 0,098%. Ainsi, la différence de gain entre les trois périodicités est négligeable. Cela revient au fait que les flux de prospects Internet arrivant sur la place de marché PdL n'atteignent jamais les quotas horaires des commandes. Par conséquent, la non

considération de certaines réinitialisations horaires des quotas de commandes dans le modèle d'optimisation multi-période (dans les deux cas $12 \times 2h$ et $6 \times 4h$) n'influence pas les résultats des affectations chez PdL. Cependant, les quotas journaliers sont souvent atteints chez PdL. Ainsi, une des solutions qui rend le modèle d'optimisation multi-période plus intéressant chez PdL, est l'implication des réinitialisations journalières des quotas de commandes. Pour ce faire, nous considérons des périodes de 24 heures afin d'inclure la réinitialisation journalière des quotas de commandes dans le modèle d'optimisation multi-période. Ainsi, toute durée de période atteignant ou dépassant les 24 heures est considérée comme convenable à l'optimisation multi-période chez PdL, par exemple : 5 périodes de 24 heures, 4 périodes de 48 heures, ... ect.

6.2.2.5 Frontière efficace du modèle d'optimisation bi-objectif mono-période

Expérience : Nous étudions la relation entre la taille des données (nombre de prospects Internet et commandes) et le nombre de points sur la frontière efficace. En outre, nous évaluons l'impact de la densité de la frontière efficace, sur le choix d'une bonne solution de compromis. Pour ce faire, nous faisons varier la taille des données. Par la suite, nous calculons la frontière efficace du modèle d'optimisation bi-obj mono-pér en utilisant la méthode " ε – *constraint*" décrite dans la section 2.2.4.2. Le temps de calcul est également évalué. Bien que les méthodes de génération sont les moins populaires en raison de leur effort de calcul (voir la section 2.2.4.2), nous adoptons la méthode ε – *constraint* pour résoudre notre modèle d'optimisation bi-obj mono-pér car les expériences réalisées ont montré que le nombre de points sur la frontière efficace est relativement réduit. Par ailleurs, nous abordons dans les perspectives de ce travail la possibilité de définir un modèle de préférences plus général basé sur les méthodes interactives (voir la section 8.2).

Résultats, description et interprétation : les figures 6.2 et 6.3 présentent les points de la frontière efficace des différentes instances calculées par les programmes mathématiques P_2 et P'_2 précédemment présentés dans la section 3.6.2. La table 6.10 présente :

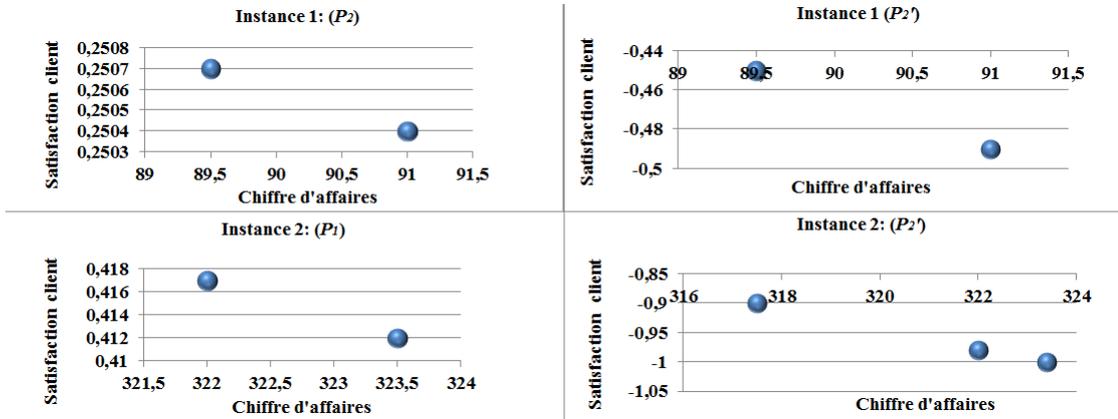


Figure 6.2 — Frontière efficace (instances 1 et 2).

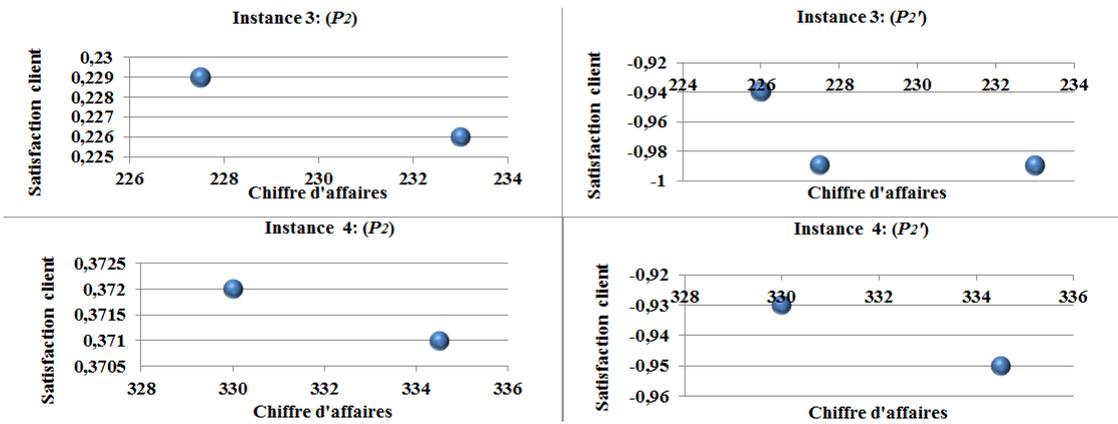


Figure 6.3 — Frontière efficace (instances 3 et 4).

1. Les détails des instances à savoir le nombre des commandes (Nb_C), le nombre des projets Internet (Nb_{PI}) et le nombre des variables de décision (Nb_{VD}).
2. Les temps de calcul générés par les programmes mathématiques P_2 et P_2' présentés par (TC_{P_2}) et $(TC_{P_2'})$ respectivement.

Nous observons que le nombre maximum de points sur les frontières efficaces est de 3. Concernant le temps de calcul, il est toujours inférieur à une seconde. Par conséquent, il est facile de calculer à chaque instant t : (1) la frontière efficace et (2) choisir un point qui peut représenter une bonne solution de compromis pour effectuer l'affectation.

	Nb_C	Nb_{PI}	Nb_{VD}	TC_{P_2}	$TC_{P'_2}$
Instance 1	30	95	150	0,01 sec	0,03 sec
Instance 2	40	123	288	0,02 sec	0,06 sec
Instance 3	25	50	67	0,00 sec	0,01 sec
Instance 4	30	295	382	0,03 sec	0,09 sec

Tableau 6.10 — Instances.

6.3 Évaluation du modèle de prévision pour l'optimisation multi-période

Nous rappelons que le modèle d'optimisation multi-période considère deux informations à savoir l'évolution continue des flux de prospects Internet et les mises à jour régulières des quotas de commandes, afin de fournir une meilleure optimisation. Dans la section 6.2, nous avons considéré dans l'étude expérimentale que le modèle d'optimisation multi-période ne possède que l'information concernant les mises à jour régulières des quotas de commandes. Cependant, dans cette section, nous proposons une nouvelle étude expérimentale qui considère que le modèle d'optimisation multi-période possède également une connaissance concernant l'évolution continue des flux de prospects Internet. En d'autres termes, cette section a pour but d'illustrer l'efficacité de la combinaison du modèle d'optimisation multi-période présenté dans la section 4.2.4 avec le modèle de prévision présenté dans la section 5.3.

Cette étude expérimentale est réalisée en utilisant des données réelles de la base de données PdL. Les calculs des séries chronologiques nécessaires à la prévision sont réalisés avec un programme *JAVA*. La prévision est établie sous le logiciel *R* (voir la section 2.3.6). La combinaison des deux modèles (optimisation multi-période et prévision) est réalisée avec le logiciel *CPLEX* et le langage *JAVA*. Les tests ont été réalisés sur le serveur de la place de marché PdL.

6.3.1 Schéma expérimental

Notre stratégie expérimentale a principalement pour objectif d'observer l'effet de l'intégration de la prévision dans le modèle d'optimisation multi-période. Un autre objectif est celui

d'étudier l'effet des erreurs de prévision sur l'optimisation, en faisant varier le taux d'erreur de prévision. En sus, cette étude expérimentale a aussi comme but de prospector le pas de temps de prévision adéquat au modèle d'optimisation multi-période. Nous considérons le schéma expérimental suivant afin de répondre aux points cités ci-dessus.

1. Nous récupérons, de la base de données PdL, les commandes et les prospects Internet arrivant sur la plateforme PdL sur 10 semaines différentes. Ainsi, nous aurons 10 instances de jeux de données sur lesquelles nous effectuerons notre étude expérimentale.
2. Nous utilisons l'algorithme défini dans la section 5.3.1 afin de : (i) construire les groupes de commande partageant les mêmes prospects Internet et (ii) calculer pour chaque groupe de commandes l'historique du nombre de prospects Internet satisfaisant simultanément toutes les commandes du groupe.
3. Nous simulons le modèle d'optimisation multi-période sur chacune des 10 semaines, sans considérer la prévision des flux de prospects.
4. Nous simulons encore une fois le modèle d'optimisation multi-période sur chacune des 10 semaines, mais en lui intégrant cette fois-ci la prévision exacte des flux de prospects Internet qui arrivent pour chaque groupe de commandes. La prévision exacte des flux de prospects de chaque groupe de commandes représente l'historique du nombre de prospects Internet satisfaisant simultanément toutes les commandes du groupe (voir l'étape 2 du schéma expérimental).
5. Nous appliquons le modèle de prévision présenté dans la section 5.3.2 afin de calculer la prévision réelle de chaque groupe de commandes déjà construit dans l'étape 2 du schéma expérimental. Par la suite, nous calculons le taux d'erreur de prévision de chaque groupe de commandes.
6. Nous simulons ensuite le modèle d'optimisation multi-période sur chacune des 10 semaines, en lui intégrant progressivement les prévisions des groupes de commandes en fonction de leurs taux d'erreur de prévision. En effet, nous commençons d'abord par intégrer les groupes de commandes ayant un taux d'erreur de prévision inférieur à 5%, ensuite

ceux inférieur à 10% et enfin ceux inférieur à 15%.

7. Nous appliquons le modèle de prévision présenté dans la section 5.3.2 afin de calculer les prévisions de chaque groupe de commandes, sur chacune des 10 semaines, en faisant varier le pas de temps de prévision. En effet, nous commençons d'abord par considérer un pas de temps de prévision de 1h, ensuite de 12h et enfin de 24h.
8. Nous simulons une dernière fois le modèle d'optimisation multi-période sur chacune des 10 semaines, en lui intégrant les prévisions calculées en considérant les trois pas de temps de prévision à savoir 1h, 12h et 24h.

6.3.2 Expériences et résultats

6.3.2.1 Prévision exacte

Expérience : cette première expérience est conçue afin d'étudier la capacité de la prévision à optimiser les résultats du modèle d'affectation multi-période en termes de chiffre d'affaires. Pour ce faire, nous effectuons les affectations sur 30 jours différents en utilisant : (i) le modèle d'optimisation multi-période sans prévision, appelé le système *Opt* et (ii) le modèle d'optimisation multi-période en lui intégrant une prévision exacte, appelé le système *Opt + Pré*. Nous calculons les chiffre d'affaires journaliers moyens générés par les deux systèmes *Opt + Pré* et *Opt*. Par la suite, nous évaluons les bénéfices induits par l'introduction d'une prévision sans erreur dans le modèle d'optimisation multi-période.

Résultats, description et interprétation : la figure 6.4 présente l'évolution moyenne, sur un jour, de la différence des chiffres d'affaires moyens des deux systèmes *Opt + Pré* et *Opt*. Aux premières heures, le système *Opt + Pré* génère un chiffre d'affaires moyen inférieur à celui généré par le système *Opt*. Cette situation est maintenue jusqu'à 4h30. À partir de ce moment, le système *Opt + Pré* entame une amélioration par rapport au système *Opt*. À 10h00, le système *Opt + Pré* récupère tout le chiffre d'affaires moyen perdu par rapport au système *Opt* et continue d'être meilleur jusqu'à 19h30. Néanmoins, à partir de 19h30, la performance

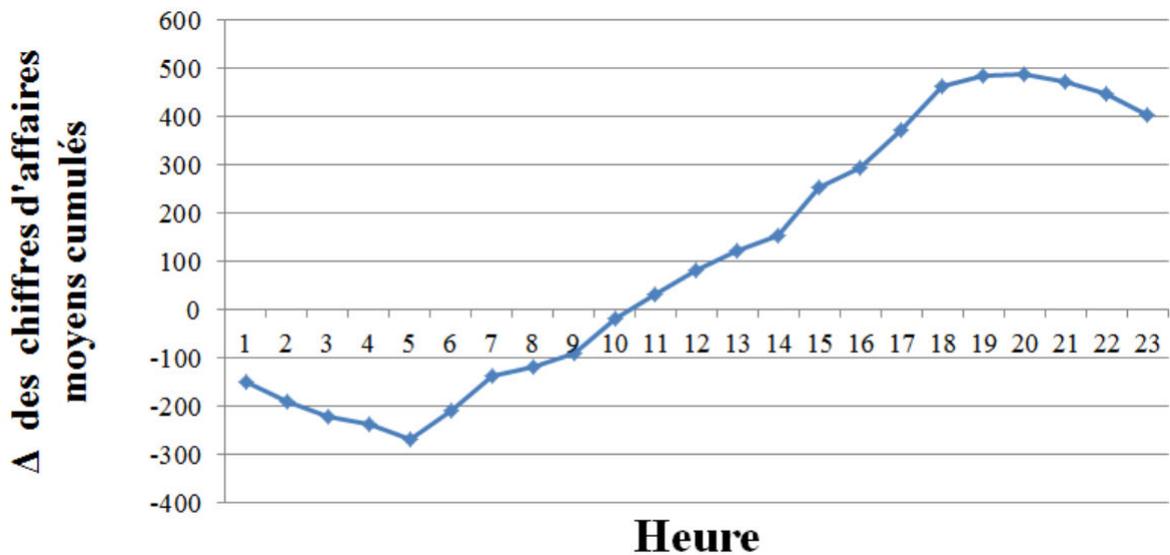


Figure 6.4 — Différence des chiffres d'affaires journaliers moyens cumulés, générés par les deux systèmes *Opt + Pré* et *Opt*.

du système *Opt + Pré* commence à diminuer par rapport au système *Opt*. Sur la totalité de la période concernée, le système *Opt + Pré* génère un chiffre d'affaires moyen plus important que celui généré par le système *Opt*.

Cette observation s'explique par le fait que le système *Opt* effectue, à chaque instant courant, toutes les affectations possibles. Cependant, le système *Opt + Pré* tire profit de sa connaissance sur l'évolution des flux de prospects Internet au cours des périodes de temps futures et préserve ainsi certains prospects Internet pour les vendre ultérieurement à des prix plus élevés. Nous observons ce phénomène en début et en fin de journée. De toute évidence et dans le but d'apprécier pleinement le gain du modèle d'optimisation mono-obj multi-pér, il est nécessaire d'étendre l'expérience sur une durée plus significative. Nous proposons ainsi de reproduire l'expérience sur une semaine. Un exemple est détaillé dans la figure 6.5.

Afin d'étudier l'impact de l'introduction de la prévision dans le modèle d'optimisation multi-période sur une longue durée, nous procédons au calcul des chiffres d'affaires sur 10 semaines différentes en simulant les deux systèmes *Opt + Pré* et *Opt*.

La figure 6.5 montre l'évolution moyenne, sur une semaine, de la différence des chiffres d'affaires

faibles moyens cumulés, générés par les deux systèmes *Opt + Pré* et *Opt*. Nous observons que le système *Opt + Pré* génère un chiffre d'affaire moyen de 300,1€ plus élevé que celui généré par le système *Opt* durant le 1^{er} jour de la semaine, de 300,3€ plus élevé durant le 2^{ème} jour, de 250,5€ plus élevé durant le 3^{ème} jour, de 450,1€ plus élevé durant le 4^{ème} jour, de 300,2€ plus élevé durant le 5^{ème} jour, de 200,3€ plus élevé durant le 6^{ème} jour et de 250,4€ plus élevé durant le 7^{ème} jour. En d'autres termes, la différence du gain en termes de chiffre d'affaires moyen cumulé passe de 300,1€ au 1^{er} jour de la semaine à 600,4€ au 2^{ème} jour, à 850,9€ au 3^{ème} jour, à 1301€ au 4^{ème}, à 1601,2€ au 5^{ème}, à 1801,5€ au 6^{ème} et à 2051,9€ au 7^{ème}.

Cette observation confirme les résultats déjà obtenus dans l'expérience précédente, dans le sens où la différence de gain en termes de chiffre d'affaires moyen des deux systèmes *Opt + Pré* et *Opt* évolue de 300,1€ généré au 1^{er} jour de la semaine à 2051,9€ généré en fin de semaine. Ceci confirme que le gain est plus conséquent sur une plus longue période.

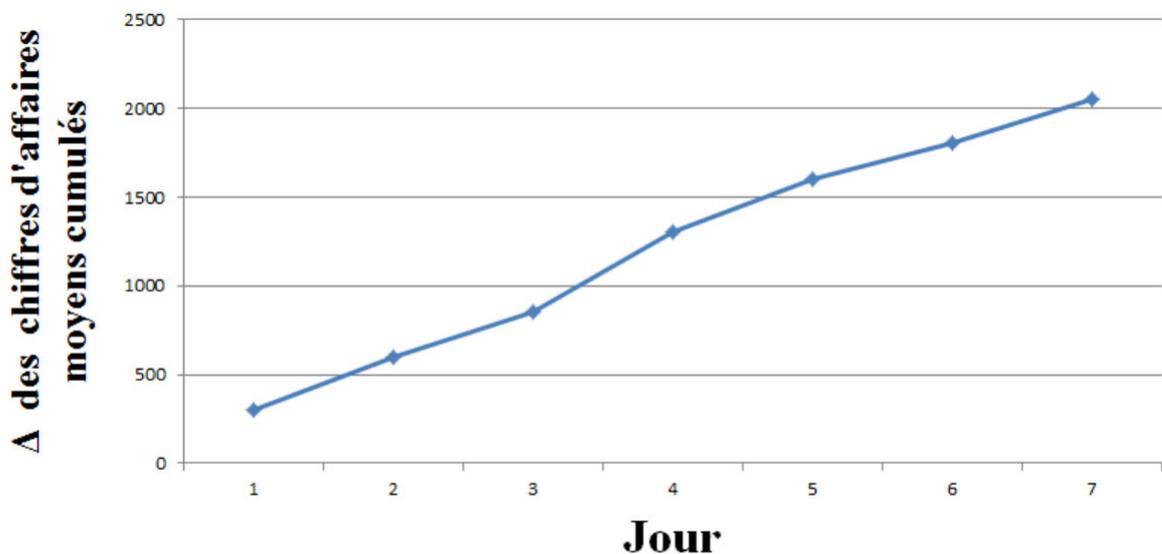


Figure 6.5 — Différence des chiffres d'affaires hebdomadaires moyens cumulés, générés par les deux systèmes *Opt + Pré* et *Opt*.

6.3.2.2 Prévision réelle

Expérience : dans les expériences précédentes, nous avons étudié la capacité de la prévision exacte d'optimiser les résultats du modèle d'affectation multi-période. À présent, nous examinons l'influence des erreurs de la prévision réelle sur le modèle d'affectation multi-période.

Pour ce faire, nous calculons d'abord la prévision réelle de chaque groupe de commandes partageant les mêmes prospects Internet. Par la suite, nous calculons le taux d'erreur de prévision de chaque groupe de commandes. Enfin, nous simulons le modèle d'optimisation multi-période sur 10 semaines différentes : (1) sans considérer la prévision puis (2) en lui intégrant :

1. La prévision des flux de prospects concernant les groupes de commandes ayant un taux d'erreur de prévision inférieur à 5%.
2. La prévision des flux de prospects concernant les groupes de commandes ayant un taux d'erreur de prévision inférieur à 10%.
3. La prévision des flux de prospects concernant les groupes de commandes ayant un taux d'erreur de prévision inférieur à 15%.

Nous calculons le chiffre d'affaires hebdomadaire moyen généré par chaque alternative. Ensuite, nous analysons l'influence des erreurs de la prévision sur le modèle d'optimisation multi-période.

Résultats, description et interprétation : La figure 6.6 montre que l'intégration de la prévision des flux de prospects concernant les groupes de commandes ayant un taux d'erreur de prévision inférieur à 10% améliore les résultats en termes de chiffre d'affaires. Le fait que le chiffre d'affaires moyen généré, en intégrant les groupes de commandes ayant un taux d'erreur inférieur à 10% est plus important que celui généré, en intégrant ceux ayant un taux d'erreur inférieur à 5%, est expliqué par le fait que la considération d'un taux d'erreur inférieur à 10% permet d'intégrer plus de groupes de commandes qu'en considérant un taux d'erreur inférieur à 5%. Cependant, la considération d'une prévision ayant un taux d'erreur inférieur à 15% permet d'augmenter encore plus le nombre de groupes de commandes inclus. Ceci dit, le

taux d'erreur élevé dégrade les résultats de l'optimisation. En d'autres termes, la considération d'un taux d'erreur inférieur à 15% permet d'intégrer d'autres groupes de commandes dont la prévision dégrade l'optimisation.

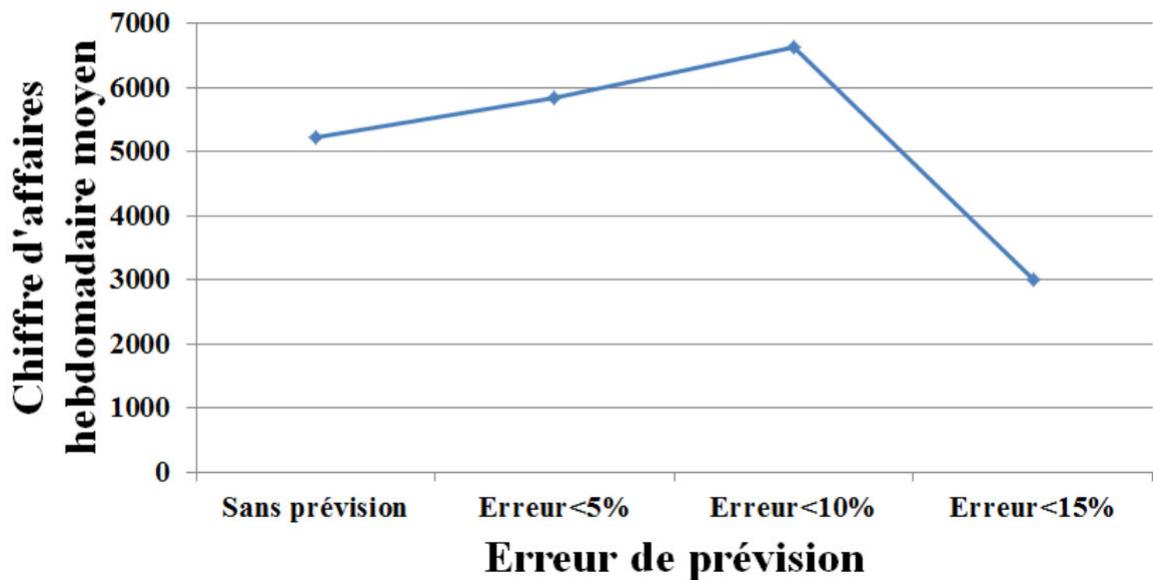


Figure 6.6 — L'influence des erreurs de prévision sur le chiffre d'affaires hebdomadaire moyen, généré par le modèle d'optimisation multi-période.

6.3.2.3 Pas de temps de prévision

Expérience : l'objectif est d'étudier l'impact du choix du pas de temps de prévision sur les résultats de l'optimisation.

Pour ce faire, nous simulons le modèle d'optimisation multi-période, sur 10 semaines différentes, en lui intégrant les prévisions calculées en considérant les trois pas de temps de prévision à savoir 1h, 12h et 24h.

Résultats, description et interprétation : La table 6.11 présente les chiffres d'affaires hebdomadaires moyens générés par le modèle d'optimisation multi-période en fonction des dif-

férents pas de temps de prévision à savoir $1h$, $12h$ et $24h$. Nous observons que plus le pas de temps de prévision est long plus le chiffre d'affaires moyen est important. Nous expliquons cela par le fait que la densité des flux de prospects Internet est plus importante sur une durée de $24h$ heures que sur une durée de $12h$ et $1h$. Et donc la prévision fournit de meilleurs résultats avec un pas de prévision de $24h$.

Nous rappelons que le pas de temps de prévision ainsi que la durée des périodes de temps du modèle d'optimisation multi-période doivent être les mêmes (voir section 5.2.3). Nous pouvons constater qu'une durée de périodes de $24h$ ne dégrade pas les résultats de l'optimisation malgré le fait qu'avec une telle durée toutes les réinitialisations horaires ne sont pas considérées dans le modèle d'optimisation multi-période. Nous expliquons ce résultat par le fait que les flux de prospects Internet arrivant sur la place de marché PdL n'atteignent jamais les quotas horaires des commandes. Par conséquent, la non considération de certaines réinitialisations horaires des quotas de commandes dans le modèle d'optimisation multi-période ne dégrade pas les résultats de l'optimisation chez PdL. Cependant, les quotas journaliers sont souvent atteints chez PdL. Ainsi, la considération des réinitialisations journalières dans le modèle d'optimisation multi-période améliore considérablement les résultats des affectations chez PdL. Par conséquent, la considération d'une durée de $24h$ comme un pas de temps de prévision ainsi que comme une durée de périodes pour le modèle d'optimisation multi-période, génère un meilleur résultat en termes de chiffre d'affaires chez PdL.

Pas de temps de prévision	Chiffre d'affaires hebdomadaire moyen
$1h$	631,94€
$12h$	2151,20€
$24h$	5964,73€

Tableau 6.11 — Chiffres d'affaires hebdomadaires moyens générés, en considérant les pas de temps de prévision $1h$, $12h$ et $24h$.

6.4 Conclusion

Ce chapitre a porté sur la validation de nos modèles proposés pour l'optimisation de l'adéquation offre/demande de prospects Internet ainsi que la prévision des flux de ces derniers. Nous avons ainsi cherché à démontrer que les solutions proposées dans cette thèse permettent de répondre à la problématique initialement formulée.

Dans un premier temps, nous avons commencé par tester l'efficacité des modèles d'optimisation que nous avons proposé et nous nous sommes intéressés à la qualité de la solution apportée. Globalement, les expériences réalisées ont confirmé l'applicabilité de nos modèles dans des systèmes réels. Plus précisément, il a été démontré que : (i) nos modèles d'optimisation sont toujours capables d'améliorer considérablement les résultats des affectations par rapport au système de la place de marché PdL et (ii) la croissance de données conduit à l'augmentation du gain sans apporter de difficultés supplémentaires quant à la résolution des instances ayant une quantité importante de données. L'intervalle de distribution ainsi que la périodicité adéquats au modèle d'optimisation multi-période ont également été étudiés dans le détail.

Dans un second temps, nous nous sommes penchés sur l'étude de l'efficacité de la combinaison d'un modèle d'optimisation multi-période avec un modèle de prévision. Ainsi, nous avons commencé par étudier la capacité de la prévision à optimiser les résultats du modèle d'affectation multi-période en lui intégrant une prévision exacte. En suite, nous avons examiné l'influence des erreurs de la prévision réelle sur le modèle d'optimisation multi-période. Enfin, nous avons étudié l'impact du choix du pas de temps de prévision sur les résultats de l'optimisation.

Les résultats obtenus sont très prometteurs et apportent une réelle valeur ajoutée à l'activité de la place de marché PdL.

Chapitre 7

Généralisation des modèles

7.1 Introduction

Dans les chapitres précédents, nous nous sommes intéressés à la problématique posée par la place de marché PdL et nous avons donc proposé différents modèles d'optimisation ainsi qu'un modèle de prévision afin de répondre aux besoins formulés. Dans ce chapitre, nous proposons une extension de ces modèles dans le but d'avoir un modèle générique qui pourra être appliqué à d'autres contextes similaires à celui de PdL. En effet, nous présentons :

1. Une définition plus générale de la satisfaction des clients.
2. Un modèle permettant de considérer l'anticipation des commandes et des clients dans un contexte où des nouveaux flux de la demande font évoluer en temps réel l'ensemble des commandes à satisfaire.
3. Une redéfinition du nombre de ventes non exclusives des prospects Internet afin de permettre à chaque client de définir son propre nombre de ventes souhaité, alors que chez PdL le nombre de ventes non exclusives est de 3 pour tous les clients.
4. Une extension multi-période de la satisfaction des clients.
5. Un modèle d'optimisation bi-objectif multi-période incluant toutes les généralisations ci-

tées ci-dessus.

7.2 Satisfaction des clients

Afin de répondre à l'un des besoins de la place de marché PdL, nous avons proposé dans la section 3.6.1 la fonction objectif 3.10 dont le but est de maximiser la satisfaction moyenne des commandes sur la journée.

La proposition d'une formalisation sous la forme d'une satisfaction moyenne des commandes répond bien aux besoins de PdL, cependant nous pensons qu'il est possible d'envisager d'autres formulations en vue d'une meilleure optimisation. En effet, dans le contexte de PdL, seules les commandes sont équitablement satisfaites alors que le fait d'être juste envers tous les clients pourrait être plus intéressant. Nous rappelons que chaque client peut effectuer plusieurs commandes de volumes différents. La satisfaction équitable des commandes ne garantit pas forcément une satisfaction équitable des clients. En effet, cette équité est proportionnelle au nombre de commandes effectué par chaque client. En d'autres termes, les clients ayant un nombre de commandes plus important seront mieux servis, ce qui génère un risque de perdre des petits clients (clients ayant un nombre de commandes moins important).

Dans cette section, nous proposons une fonction objectif qui maximise la satisfaction moyenne des clients. Le but est d'avoir une distribution journalière juste et équitable des prospects Internet vis-à-vis de tous les clients quelque soit leurs importances. Formellement,

$$\max \frac{1}{r} \left[\sum_{k:k \in C} \left[\sum_{i,j:x_{ij} \in X} (a_{ik} V_i + \sum_{i,j:x_{ij} \in X} a_{ik} x_{ij}) / Dc_k \right] \right] \quad (7.1)$$

Où, V_i représente le nombre de prospects Internet délivré à la commande i à partir du début de la journée.

Nous rappelons que :

- $P = \{j : j \in \{1, \dots, m\}\}$: ensemble des prospects Internet.
- $O = \{i : i \in \{1, \dots, n\}\}$: ensemble des commandes.

- $C = \{k : k \in \{1, \dots, r\}\}$: ensemble des clients.
- $X = \{x_{ij} : i \in O, j \in P\}$: ensemble des variables de décision, tel que :

$$x_{ij} = \begin{cases} 1 & \text{si le prospect Internet } j \text{ est affecté à la commande } i \\ 0 & \text{sinon} \end{cases}$$

- r présente le nombre de clients.
- a_{ik} présente une variable binaire décrite de la manière suivante :

$$a_{ik} = \begin{cases} 1 & \text{si la commande } i \text{ provient du client } k \\ 0 & \text{sinon} \end{cases}$$

- Dc_k présente le quota journalier du client k .

En outre, nous pouvons envisager une satisfaction moyenne horaire ou encore mensuelle des clients. Formellement,

$$\max \frac{1}{r} \left[\sum_{k:k \in C} \left[\sum_{i,j:x_{ij} \in X} (a_{ik} W_i + \sum_{i,j:x_{ij} \in X} a_{ik} x_{ij}) / Hc_k \right] \right] \quad (7.2)$$

Où, W_i représente le nombre de prospects Internet délivré à la commande i à partir du début de l'heure.

Nous rappelons que :

- Hc_k présente le quota horaire du client k .

$$\max \frac{1}{r} \left[\sum_{k:k \in C} \left[\sum_{i,j:x_{ij} \in X} (a_{ik} Y_i + \sum_{i,j:x_{ij} \in X} a_{ik} x_{ij}) / Mc_k \right] \right] \quad (7.3)$$

Où, Y_i représente le nombre de prospects Internet délivré à la commande i à partir du début du mois.

Nous rappelons que :

- Mc_k présente le quota mensuel du client k .

Les fonctions objectif 7.2 et 7.3 permettent d'avoir une distribution horaire et mensuelle respectivement, juste et équitable des prospects Internet vis-à-vis de tous les clients.

7.3 Nouveaux flux de commandes et de clients

Dans la section 4.2.4, nous avons défini un modèle d'optimisation multi-période P_3 qui permet d'établir une adéquation de l'offre et de la demande. Cette adéquation s'effectue en tirant profit de la connaissance de l'évolution du flux de prospects Internet sur des périodes de temps futures. Afin de répondre aux besoins de la place de marché PdL, nous avons considéré pour la définition de ce modèle que les commandes et les clients restent inchangés pour toutes les périodes de temps futures, alors qu'il est possible d'avoir des nouveaux flux de prospects Internet arrivant de façon non déterministe sur l'ensemble des périodes de temps futures.

À présent et afin d'avoir un modèle plus général de l'optimisation multi-période de l'adéquation offre/demande de prospects Internet, nous allons considérer qu'il est possible d'avoir des nouveaux flux de : (i) commandes, (ii) clients et (iii) prospects Internet, arrivant simultanément de façon non déterministe sur l'ensemble des périodes de temps futures. En effet, nous pouvons parfaitement envisager l'arrivée d'un nouveau client dans le système à tout moment et donc de nouvelles commandes.

Pour ce faire, nous considérons à chaque instant t , un ensemble T' de périodes de temps de sorte que $T' = \{t' : t' \in \{1, \dots, g\}\}$, où $t' = 1$ présente la période de temps courante et $t' \in \{2, \dots, g\}$ sont les périodes de temps futures. Sur chacune des périodes de temps futures t' , nous avons un nouveau flux de commandes, de clients ainsi que de prospects Internet arrivant de façon non déterministe. Ainsi, nous définissons :

- $O_{t'} = \{i : i \in \{1, \dots, n_{t'}\}\}$ représente l'ensemble des commandes à la période de temps t' et $n_{t'}$ est le nombre de commandes à la période de temps t' .
- $C_{t'} = \{k : k \in \{1, \dots, r_{t'}\}\}$, représente l'ensemble des clients à la période de temps t' et $r_{t'}$ est le nombre de clients à la période de temps t' .
- $P_{t'} = \{j : j \in \{1, \dots, m_{t'}\}\}$, représente l'ensemble des prospects Internet à la période de temps t' et $m_{t'}$ est le nombre de prospects Internet à la période de temps t' .

Nous définissons également l'ensemble des variables de décision $X_{t'} = \{x_{ijt'} : i \in O_{t'}, j \in P_{t'}, t' \in T'\}$, tel que :

$$x_{ijt'} = \begin{cases} 1 & \text{si le prospect Internet } j \text{ est affecté à la commande } i \text{ à la période de temps } t' \\ 0 & \text{sinon} \end{cases}$$

De plus, nous définissons les ensembles $E_{t'} = \{x_{ijt'} : i \in O_{t'}, j \in P_{t'}, t' \in T', i \text{ est une commande exclusive}\}$ et $\overline{E}_{t'} = \{x_{ij} : i \in O_{t'}, j \in P_{t'}, t' \in T', i \text{ est une commande non exclusive}\}$. Ainsi, $X_{t'} = E_{t'} \cup \overline{E}_{t'}$ tel que $t' \in T'$.

$$P_4 \left\{ \begin{array}{l} \max \quad f(X_{t'}) = \sum_{t'=1}^g \sum_{i,j:x_{ijt'} \in X_{t'}} Pr_i x_{ijt'} \\ s.t. \\ \sum_{j:x_{ijt'} \in X_{t'}} x_{ijt'} \leq Ro_i - \sum_{t''=1}^{t'-1} \sum_{j:x_{ijt''} \in X_{t''}} x_{ijt''}, \quad \forall i \in O_{t'}, \forall t' \in T' \\ \sum_{i,j:x_{ijt'} \in X_{t'}} a_{ik} x_{ijt'} \leq Rc_k - \sum_{t''=1}^{t'-1} \sum_{i,j:x_{ijt''} \in X_{t''}} a_{ik} x_{ijt''}, \quad \forall k \in C_{t'}, \forall t' \in T' \\ S_j + \sum_{t''=1}^{t'-1} \sum_{i:x_{ijt''} \in X_{t''}} x_{ijt''} + b \sum_{i:x_{ijt'} \in E_{t'}} x_{ijt'} + \sum_{i:x_{ijt'} \in \overline{E}_{t'}} x_{ijt'} \leq b, \quad \forall j \in P_{t'}, \forall t' \in T' \\ \sum_{t'=1}^g \sum_{i:x_{ijt'} \in X_{t'}} a_{ik} x_{ijt'} \leq 1, \quad \forall j \in P_{t'}, \forall k \in C_{t'} \\ x_{ijt'} \in \{0, 1\} \end{array} \right. \quad (7.4)$$

Nous rappelons que :

- Ro_i représente le nombre de prospects Internet restant à délivrer à la commande i , à la première période de temps.
- Rc_k représente le nombre de prospects Internet restant à délivrer au client k , à la première période de temps.
- S_j représente le nombre de ventes antérieures (avant la première période de temps) du prospect Internet j .

– b représente le nombre de reventes autorisées des prospects Internet non exclusifs.

Le modèle d'optimisation P_4 possède le même objectif que celui du modèle d'optimisation P_3 défini dans la section 4.2.4 à savoir la maximisation du chiffre d'affaires sur des périodes de temps futures. Il est également soumis aux mêmes contraintes déjà définies pour le modèle d'optimisation P_3 . Cependant, le programme mathématique P_4 optimise à un instant t donné l'affectation des prospects Internet aux commandes en considérant plusieurs périodes de temps futures, pour lesquelles nous avons simultanément des nouveaux flux de commandes, de clients et de prospects Internet.

À titre d'exemple, considérons que : (i) nous effectuons l'affectation sur 3 périodes de temps à savoir les périodes de temps 0, 1 et 2, (ii) nous disposons de 2 commandes et 1 prospect Internet (PI) à la période de temps 0 et (iii) la prévision pour les périodes de temps 1 et 2 présente l'arrivée de :

- 3 prospects Internet à savoir le prospect Internet 2 et 3 à la période de temps 1 et le prospect Internet 4 à la période de temps 2.
- 2 clients à savoir un premier client à la période de temps 1 qui effectuera une commande (commande 3) et un deuxième client à la période de temps 2 qui effectuera deux commandes (commandes 4 et 5).

Par conséquent, l'ensemble de toutes les affectations possibles de chacune des périodes de temps 0, 1 et 2 est présenté dans les tables 7.1, 7.2 et 7.3 respectivement. Ces ensembles sont construits afin d'être fournis au modèle d'optimisation multi-période à la période de temps 0. Par conséquence, la connaissance concernant les flux de commandes, clients et prospects Internet sur les 3 périodes de temps, est prise en compte dans le processus d'affectation afin de procéder à une meilleure optimisation.

t_0	PI 1
Commande 1	x_{110}
Commande 2	/

Tableau 7.1 — Ensemble de toutes les affectations possibles à la période de temps 0.

t_1	PI 1	PI 2	PI 3
Commande 1	x_{110}	prévision	prévision
Commande 2	/	prévision	prévision
Commande 3	prévision	prévision	prévision

Tableau 7.2 — Ensemble de toutes les affectations possibles à la période de temps 1.

t_2	PI 1	PI 2	PI 3	PI 4
Commande 1	x_{110}	prévision	prévision	prévision
Commande 2	/	prévision	prévision	prévision
Commande 3	prévision	prévision	prévision	prévision
Commande 4	prévision	prévision	prévision	prévision
Commande 5	prévision	prévision	prévision	prévision

Tableau 7.3 — Ensemble de toutes les affectations possibles à la période de temps 2.

7.4 Nombre de ventes non exclusives des prospects Internet

Dans la section 3.3.3, nous avons défini la contrainte 3.5 liée au nombre de ventes et à l'exclusivité. Il s'agit de considérer deux types de ventes à savoir : (i) une vente exclusive où un prospect Internet est vendu au plus une fois et (ii) une vente non exclusive où un prospect Internet est vendu au plus trois fois. Le fait que le nombre de ventes non exclusives des prospects Internet soit le même pour toutes les commandes est spécifique à la place de marché PdL. Dans le sens où, si un prospect Internet est vendu en non exclusivité à une commande i , il pourra être encore revendu jusqu'à 3 fois à d'autres commandes n'exigeant pas l'exclusivité. Toutefois, nous pouvons personnaliser le nombre de ventes non exclusives des prospects Internet. En effet, il est possible d'envisager un nombre de ventes non exclusives des prospects Internet différent pour chaque commande.

À titre d'exemple, supposons que nous possédons trois commandes : (i) une première commande (C_1) exigeant que si un prospect Internet lui a été vendu en non exclusivité, il pourra être revendu encore à d'autres commandes non exclusives jusqu'à 4 fois, (ii) une deuxième commande (C_2) jusqu'à 5 fois et (iii) une dernière commande (C_3) jusqu'à 6 fois. Ainsi, nous devons repérer pour chaque prospect Internet non exclusif, toutes les commandes pour

lesquelles il a été vendu. Par la suite, nous limitons le nombre de ventes autorisées pour ce prospect Internet au minimum des nombres de ventes autorisées par chacune des commandes pour lesquelles il a déjà été vendu. Concernant notre exemple, si nous supposons qu'un prospect Internet j a déjà été vendu en non exclusivité aux commandes C_1 et C_3 , alors le nombre de ventes autorisées pour le prospect Internet j est limité au minimum des nombres de ventes autorisées par chacune des commandes C_1 et C_3 qui est égal au minimum de l'ensemble $\{4, 6\}$. Formellement, nous définissons pour chaque commande $i \in O$, un nombre de ventes non exclusives autorisées Q_i .

$$S_j + F_j \sum_{i:x_{ij} \in E} x_{ij} + \sum_{i:x_{ij} \in \bar{E}} x_{ij} \leq F_j, \quad \forall j \in P \quad (7.5)$$

Où,

– F_j représente le nombre de ventes non exclusives autorisées pour le prospect Internet j .

Ainsi, F_j est décrit comme suit :

$$F_j = \min \{ Q_i : i \in O, i \text{ est une commande pour laquelle le prospect Internet } j \text{ a déjà été vendu} \}.$$

Nous rappelons que :

– $P = \{ j : j \in \{1, \dots, m\} \}$: ensemble des prospects Internet.

– $O = \{ i : i \in \{1, \dots, n\} \}$: ensemble des commandes.

– $X = \{ x_{ij} : i \in O, j \in P \}$: ensemble des variables de décision, tel que :

$$x_{ij} = \begin{cases} 1 & \text{si le prospect Internet } j \text{ est affecté à la commande } i \\ 0 & \text{sinon} \end{cases}$$

– $E = \{ x_{ij} : i \in O, j \in P, i \text{ est une commande exclusive} \}$ et $\bar{E} = \{ x_{ij} : i \in O, j \in P, i \text{ est une commande non exclusive} \}$.

– S_j présente le nombre de ventes antérieures du prospect Internet j .

La contrainte 7.5 permet de vérifier qu'un prospect Internet exclusif est vendu au plus une fois, et qu'un prospect Internet non exclusif n'est pas vendu plus que F_j fois. L'extension

multi-période de la contrainte 7.5 est donnée par l'équation 7.6.

$$S_j + \sum_{t''=1}^{t'-1} \sum_{i: x_{ijt''} \in X_{t''}} x_{ijt''} + F_j \sum_{i: x_{ijt'} \in E_{t'}} x_{ijt'} + \sum_{i: x_{ijt'} \in \overline{E_{t'}}} x_{ijt'} \leq F_j, \quad \forall j \in P_{t'}, \forall t' \in T' \quad (7.6)$$

Où,

- $O_{t'} = \{i : i \in \{1, \dots, n_{t'}\}\}$ représente l'ensemble de commandes à la période de temps t' et $n_{t'}$ est le nombre de commandes à la période de temps t' .
- $C_{t'} = \{k : k \in \{1, \dots, r_{t'}\}\}$, représente l'ensemble de clients à la période de temps t' et $r_{t'}$ est le nombre de clients à la période de temps t' .
- $P_{t'} = \{j : j \in \{1, \dots, m_{t'}\}\}$, représente l'ensemble de prospects Internet à la période de temps t' et $m_{t'}$ est le nombre de prospects Internet à la période de temps t' .
- $X_{t'} = \{x_{ijt'} : i \in O_{t'}, j \in P_{t'}, t' \in T'\}$ représente l'ensemble des variables de décision, tel que :

$$x_{ijt'} = \begin{cases} 1 & \text{si le prospect Internet } j \text{ est affecté à la commande } i \text{ à la période de temps } t' \\ 0 & \text{sinon} \end{cases}$$
- $E_{t'} = \{x_{ijt'} : i \in O_{t'}, j \in P_{t'}, t' \in T', i \text{ est une commande exclusive}\}$ et $\overline{E_{t'}} = \{x_{ij} : i \in O_{t'}, j \in P_{t'}, t' \in T', i \text{ est une commande non exclusive}\}$.
- Q_i est le nombre de reventes non exclusives autorisées pour toute commande $i \in O_{t'}$.
- F_j représente le nombre de ventes non exclusives autorisées pour le prospect Internet j . Il est décrit comme suit : $F_j = \min\{Q_i : i \in O_{t'}, t' \in T', i \text{ est une commande pour laquelle le prospect Internet } j \text{ a déjà été vendu}\}$.
- S_j représente le nombre de ventes antérieures du prospect Internet j .

7.5 Extension multi-période de la satisfaction des commandes et des clients

La fonction objectif de la satisfaction des commandes (3.10) (des clients (7.1)) définie dans la section 3.6.1 (la section 7.2, respectivement) permet de maximiser la satisfaction journalière moyenne des commandes (des clients, respectivement), en considérant une seule période de

temps. Dans cette section, nous proposons une extension multi-période de ces deux fonctions objectif.

$$\max \frac{1}{n} \left[\sum_{t'=1}^g \sum_{i: x_{ijt'} \in X_{t'}} (V_i + \sum_{t'=1}^g \sum_{j: x_{ijt'} \in P_{t'}} x_{ijt'}) / Do_i \right] \quad (7.7)$$

$$\max \frac{1}{r} \left[\sum_{t'=1}^g \sum_{k: k \in C_{t'}} \left[\sum_{t'=1}^g \sum_{i, j, t': x_{ijt'} \in X_{t'}} (a_{ik} V_i + \sum_{t'=1}^g \sum_{i, j, t': x_{ijt'} \in X_{t'}} a_{ik} x_{ijt'}) / Dc_k \right] \right] \quad (7.8)$$

Où :

- $O_{t'} = \{i : i \in \{1, \dots, n_{t'}\}\}$ représente l'ensemble de commandes à la période de temps t' et $n_{t'}$ est le nombre de commandes à la période de temps t' .
- $C_{t'} = \{k : k \in \{1, \dots, r_{t'}\}\}$, représente l'ensemble de clients à la période de temps t' et $r_{t'}$ est le nombre de clients à la période de temps t' .
- $P_{t'} = \{j : j \in \{1, \dots, m_{t'}\}\}$, représente l'ensemble de prospects Internet à la période de temps t' et $m_{t'}$ est le nombre de prospects Internet à la période de temps t' .
- $X_{t'} = \{x_{ijt'} : i \in O_{t'}, j \in P_{t'}, t' \in T'\}$ représente l'ensemble des variables de décision, tel que :

$$x_{ijt'} = \begin{cases} 1 & \text{si le prospect Internet } j \text{ est affecté à la commande } i \text{ à la période de temps } t' \\ 0 & \text{sinon} \end{cases}$$

- V_i représente le nombre de prospects Internet délivré à la commande i à partir du début de la journée.
- Do_i représente le quota journalier de la commande i .
- a_{ik} représente une variable binaire décrite de la manière suivante :

$$a_{ik} = \begin{cases} 1 & \text{si la commande } i \text{ provient du client } k \\ 0 & \text{sinon} \end{cases}$$

- Dc_k représente le quota journalier du client k .

La fonction objectif (7.7)((7.8)) permet d'avoir une distribution journalière juste et équitable des prospects Internet vis-à-vis de toutes les commandes (tous les clients, respectivement), en considérant plusieurs périodes de temps futures t' , sur lesquelles nous avons de nouveaux

flux de commandes, de clients ainsi que de prospects Internet. Par conséquent, nous pouvons optimiser les résultats de l'affectation en tirant profit de la connaissance concernant l'évolution du système sur les périodes de temps futures.

7.6 Modèle d'optimisation bi-objectif multi-période

Dans les sections précédentes, nous avons proposé différentes extensions et généralisations des fonctions objectif ainsi que des contraintes déjà étudiées dans les chapitres précédents. À présent et en s'appuyant sur ce qui a déjà été défini dans les sections précédentes, nous proposons un modèle d'optimisation générique qui pourra être adapté à d'autres contextes analogues à celui de la place de marché PdL. À titre d'exemple, les places de marché de réservation de taxis par des clients en ligne ou encore les affectations des missions de prestation de services dans les SSII (sociétés de services en ingénierie informatique). Ainsi, nous définissons :

- $O_{t'} = \{i : i \in \{1, \dots, n_{t'}\}\}$ représente l'ensemble des commandes à la période de temps t' et $n_{t'}$ est le nombre de commandes à la période de temps t' .
- $C_{t'} = \{k : k \in \{1, \dots, r_{t'}\}\}$, représente l'ensemble des clients à la période de temps t' et $r_{t'}$ est le nombre de clients à la période de temps t' .
- $P_{t'} = \{j : j \in \{1, \dots, m_{t'}\}\}$, représente l'ensemble des prospects Internet à la période de temps t' et $m_{t'}$ est le nombre de prospects Internet à la période de temps t' .

7.6.1 Variables de décision

Pour chacune des périodes de temps t' de l'ensemble T' , l'affectation d'un prospect Internet à une commande est représentée par la variable binaire suivante :

$$x_{ijt'} = \begin{cases} 1 & \text{si le prospect Internet } j \text{ est affecté à la commande } i \text{ à la période de temps } t' \\ 0 & \text{sinon} \end{cases}$$

Nous définissons $X_{t'} = \{x_{ijt'} : i \in O_{t'}, j \in P_{t'}, t' \in T'\}$ l'ensemble des variables de décision. Nous définissons également les ensembles $E_{t'} = \{x_{ijt'} : i \in O_{t'}, j \in P_{t'}, t' \in T', i \text{ est une commande}$

exclusive } et $\overline{E}_{t'} = \{x_{ij} : i \in O_{t'}, j \in P_{t'}, t' \in T', i \text{ est une commande non exclusive}\}$. Ainsi, $X_{t'} = E_{t'} \cup \overline{E}_{t'}$ tel que $t' \in T'$.

7.6.2 Contraintes

Contrainte liée aux volumes des commandes :

La contrainte (7.9) vérifie que le nombre de prospects Internet affecté à chaque commande i respecte bien son volume. Cela est vérifié, à chaque période de temps de l'ensemble T' , pour les commandes déjà présentes dans le système ainsi que celles prévues pour les périodes de temps futures.

$$\sum_{j:x_{ijt'} \in X_{t'}} x_{ijt'} \leq Ro_i - \sum_{t''=1}^{t'-1} \sum_{j:x_{ijt''} \in X_{t''}} x_{ijt''}, \quad \forall i \in O_{t'}, \forall t' \in T' \quad (7.9)$$

Tel que, Ro_i représente le nombre de prospects Internet restant à délivrer à la commande i , à la période de temps 1.

Contrainte liée aux volumes des clients :

La contrainte (7.10) vérifie que le nombre de prospects Internet affecté à chaque client k respecte bien son volume. Cela est vérifié, à chaque période de temps de l'ensemble T' , pour les clients déjà présents dans le système ainsi que ceux prévus pour les périodes de temps futures.

$$\sum_{i,j:x_{ijt'} \in X_{t'}} a_{ik} x_{ijt'} \leq Rc_k - \sum_{t''=1}^{t'-1} \sum_{i,j:x_{ijt''} \in X_{t''}} a_{ik} x_{ijt''}, \quad \forall k \in C_{t'}, \forall t' \in T' \quad (7.10)$$

Tel que :

– a_{ik} représente une variable binaire décrite de la manière suivante :

$$a_{ik} = \begin{cases} 1 & \text{si la commande } i \text{ provient du client } k \\ 0 & \text{sinon} \end{cases}$$

– Rc_k représente le nombre de prospects Internet restant à délivrer au client K , à la période de temps 1.

Contrainte liée au nombre de ventes et à l'exclusivité :

La contrainte (7.11) exprime le fait que, sur toutes les g périodes de temps prises en compte dans l'optimisation, tout prospect Internet exclusif peut être vendu au plus une fois et que tout prospect Internet non exclusif peut être vendu au plus F_j fois. Cela est vérifié, à chaque période de temps de l'ensemble T' , pour les prospects Internet déjà présents dans le système ainsi que ceux prévus pour les périodes de temps futures.

$$S_j + \sum_{t''=1}^{t'-1} \sum_{i: x_{ijt''} \in X_{t''}} x_{ijt''} + F_j \sum_{i: x_{ijt'} \in E_{t'}} x_{ijt'} + \sum_{i: x_{ijt'} \in \bar{E}_{t'}} x_{ijt'} \leq F_j, \quad \forall j \in P_{t'}, \forall t' \in T' \quad (7.11)$$

Tel que :

- S_j représente le nombre des ventes antérieures (avant la première période de temps) du prospect Internet j .
 - Q_i représente le nombre des ventes non exclusives autorisées pour toute commande $i \in O_{t'}$.
 - F_j représente le nombre des ventes non exclusives autorisées pour le prospect Internet j .
- Ainsi, F_j est décrit comme suit : $F_j = \min\{Q_i : i \in O_{t'}, t' \in T, i \text{ est une commande pour laquelle le prospect Internet } j \text{ a déjà été vendu}\}$.

Contrainte liée à la duplication des prospects Internet :

La contrainte (7.12) garantit le fait qu'un client k ne reçoit pas un prospect Internet j plus d'une fois, sur toutes les g périodes de temps prises en compte dans l'optimisation. Cela est vérifié, à chaque période de temps de l'ensemble T , pour les clients et les prospects Internet déjà présents dans le système ainsi que ceux prévus pour les périodes de temps futures.

$$\sum_{t'=1}^g \sum_{i: x_{ijt'} \in X_{t'}} a_{ik} x_{ijt'} \leq 1, \quad \forall k \in C_{t'}, \forall j \in P_{t'} \quad (7.12)$$

7.6.3 Fonctions objectif

Maximisation du chiffre d'affaires

L'équation (7.13) permet de maximiser le chiffre d'affaires sur toutes les g périodes de temps prises en compte dans l'optimisation, en considérant les nouveaux flux de commandes, de clients ainsi que de prospects Internet.

$$\max \sum_{t'=1}^g \sum_{i,j:x_{ijt'} \in X_{t'}} Pr_i x_{ijt'} \quad (7.13)$$

Tel que : Pr_i représente le prix de vente de la commande i .

Maximisation de la satisfaction des clients

L'équation (7.14) permet de maximiser la satisfaction journalière moyenne des clients sur toutes les g périodes de temps prises en compte dans l'optimisation, en considérant les nouveaux flux de commandes, de clients ainsi que de prospects Internet.

$$\max \frac{1}{r} \left[\sum_{t'=1}^g \sum_{k:k \in C_{t'}} \left[\sum_{t'=1}^g \sum_{i,j,t':x_{ijt'} \in X_{t'}} (a_{ik} V_i + \sum_{t'=1}^g \sum_{i,j,t':x_{ijt'} \in X_{t'}} a_{ik} x_{ijt'}) / Dc_k \right] \right] \quad (7.14)$$

Tel que :

- V_i représente le nombre de prospects Internet délivré à la commande i à partir du début de la journée.
- Rc_k représente le nombre de prospects Internet restant à délivrer au client k , à la première période de temps.

7.6.4 Programme mathématique bi-objectif multi-période

Dans cette section, nous présentons un modèle d'optimisation bi-objectif multi-période P_5 en utilisant la programmation mathématique linéaire avec des variables binaires.

$$\begin{array}{l}
 \left. \begin{array}{l}
 \max \quad f_1(X_{t'}) = \sum_{t'=1}^g \sum_{i,j:x_{ijt'} \in X_{t'}} Pr_i x_{ijt'} \\
 \max \quad f_2(X_{t'}) = \frac{1}{r} \left[\sum_{t'=1}^g \sum_{k:k \in C_{t'}} \left[\sum_{t'=1}^g \sum_{i,j,t':x_{ijt'} \in X_{t'}} (a_{ik} V_i + \sum_{t'=1}^g \sum_{i,j,t':x_{ijt'} \in X_{t'}} a_{ik} x_{ijt'}) / Dc_k \right] \right] \\
 s.t. \\
 \sum_{j:x_{ijt'} \in X_{t'}} x_{ijt'} \leq Ro_i - \sum_{t''=1}^{t'-1} \sum_{j:x_{ijt''} \in X_{t''}} x_{ijt''}, \quad \forall i \in O_{t'}, \forall t' \in T' \\
 \sum_{i,j:x_{ijt'} \in X_{t'}} a_{ik} x_{ijt'} \leq Rc_k - \sum_{t''=1}^{t'-1} \sum_{i,j:x_{ijt''} \in X_{t''}} a_{ik} x_{ijt''}, \quad \forall k \in C_{t'}, \forall t' \in T' \\
 S_j + \sum_{t''=1}^{t'-1} \sum_{i:x_{ijt''} \in X_{t''}} x_{ijt''} + F_j \sum_{i:x_{ijt'} \in E_{t'}} x_{ijt'} + \sum_{i:x_{ijt'} \in \bar{E}_{t'}} x_{ijt'} \leq F_j, \quad \forall j \in P_{t'}, \forall t' \in T' \\
 \sum_{t'=1}^g \sum_{i:x_{ijt'} \in X_{t'}} a_{ik} x_{ijt'} \leq 1, \quad \forall j \in P_{t'}, \forall k \in C_{t'} \\
 x_{ijt'} \in \{0, 1\}
 \end{array} \right\} P_5
 \end{array} \tag{7.15}$$

Le modèle mathématique P_5 permet d'optimiser l'affectation des prospects Internet aux commandes en termes de chiffre d'affaires et de satisfaction moyenne des clients. En effet, à un instant t donné et en considérant plusieurs périodes de temps futures, le modèle P_5 maximise le chiffre d'affaires tout en cherchant une distribution journalière juste et équitable des prospects Internet vis-à-vis de tous les clients quelque soit leurs importances. Ainsi, la connaissance concernant les nouveaux flux de : (i) commandes, (ii) clients et (iii) prospects Internet sur les périodes de temps futures est prise en compte dans le processus d'affectation afin d'aboutir à

une meilleure optimisation.

Les fonctions objectif de ce modèle sont données par les équations (7.13) et (7.14). Concernant les contraintes, elles sont décrites par les équations (7.9), (7.10), (7.11) et (7.12).

7.7 Conclusion

Le but de ce chapitre était de présenter un aperçu de certaines extensions possibles des modèles proposés précédemment. En effet, dans un premier temps, nous nous sommes intéressés à la satisfaction des clients de manière globale sur l'heure, le jour et enfin le mois au lieu de se concentrer uniquement sur la satisfaction des commandes sur la journée.

Dans un deuxième temps, nous avons proposé une extension du modèle d'optimisation mono-objectif multi-période afin d'inclure les futurs flux de commandes et de clients simultanément.

Dans un troisième temps, nous avons redéfini le nombre de ventes non exclusives des prospects Internet afin de pouvoir personnaliser le nombre de ventes non exclusives pour chacune des commandes. Par la suite, nous nous sommes penchés sur l'extension multi-période de la satisfaction des commandes et des clients.

Enfin, nous avons proposé une formalisation générique d'un modèle d'optimisation bi-objectif multi-période pour qu'il puisse être adapté à d'autres contextes proches du fonctionnement de la place de marché PdL.

Dans le chapitre suivant, nous allons développer différentes expériences permettant de tester les modèles et les algorithmes proposés dans cette thèse.

Conclusion et perspectives

L'objectif de cette thèse était de répondre à la problématique complexe de PdL à savoir "l'affectation des prospects Internet aux clients". Plus précisément, établir une correspondance en temps réel entre l'offre et la demande de prospects Internet tout en considérant les différentes caractéristiques de ces derniers ainsi que les personnalisations des commandes. Cette correspondance s'effectue dans un contexte caractérisé par des flux continus faisant évoluer en temps réel l'ensemble des prospects Internet disponibles. L'objectif était, d'une part, de construire un cadre formel pour répondre à la problématique de la place de marché PdL, d'autre part, de proposer une formalisation la plus générique possible afin de répondre à toute autre problématique ayant des caractéristiques semblables à celles de PdL.

Ainsi, ce chapitre présente un bilan de cette thèse en revenant sur les motivations des choix effectués ainsi que les modèles et les algorithmes proposés. En outre, nous décrivons certaines perspectives pour des travaux futurs.

8.1 Contributions

Dans ce travail, nous avons proposé plusieurs modèles visant à répondre au problème de l'adéquation de l'offre et de la demande en temps réel. Plus précisément, nous avons proposé des

modèles d'affectation qui s'appuient à la fois sur l'optimisation et la prévision.

Le chapitre 2, nous a permis de présenter une étude bibliographique afin de rappeler les différents concepts et notions utilisés dans nos travaux. En effet, nous avons présenté deux formalismes, le premier est dédié à la modélisation des problèmes d'optimisation et le second est consacré à la modélisation des problèmes de prévision.

Cette thèse a été l'occasion de voir dans quelle mesure il était possible d'articuler deux modèles différents mais complémentaires pour offrir une meilleure politique d'affectation à la place de marché PdL.

Pour ce faire, nous avons commencé par présenter dans la section 3.3 une structure formelle du problème de l'affectation de l'offre à la demande en temps réel, tout en prenant en compte leurs différentes spécificités. Par la suite, nous avons proposé dans la section 3.6 un modèle d'optimisation bi-objectif mono-période de l'affectation offre/demande de prospects Internet. L'un des enjeux forts de cette affectation est de prendre en compte instantanément les nouvelles offres et demandes de prospects Internet et leur affectation en temps réel. En sus, l'objectif de la place de marché est de maximiser ses profits tout en conservant ses clients. En effet, l'objectif recherché à travers cela est une attitude de justice envers tous les clients, dans la mesure où la favorisation de certains se traduit par la perte de ceux qui ont été marginalisés, d'où l'intérêt de satisfaire équitablement tous les clients. Par conséquent, l'enjeu est double : (i) optimiser le chiffre d'affaires et (ii) optimiser la satisfaction des clients.

Nous avons introduit dans la section 4.2 une extension multi-période du modèle d'affectation. Ainsi, nous avons proposé un modèle d'optimisation multi-période qui étend le problème sur plusieurs périodes de temps futures. Cette extension conduit à optimiser l'affectation en considérant les éventuelles futures mises à jour du système à savoir nouveaux flux d'offre et réactualisation des quotas de la demande. En effet, l'une des caractéristiques principales de PdL réside dans le fonctionnement continu de son système. Ainsi, la formalisation mono-période du problème qui consiste à optimiser l'affectation des prospects Internet en considérant uniquement la période de temps courante, ne permet pas de réaliser des affectations de façon pertinente. En effet, l'optimisation mono-période effectue des affectations à chaque instant

sans prendre en compte l'évolution du système sur les périodes de temps futures. Cependant, la considération d'une période de temps continue conduit à optimiser les résultats de l'affectation en tirant profit de la connaissance de l'évolution du système sur les périodes de temps futures.

Afin de fournir au modèle d'optimisation multi-période la connaissance concernant l'évolution des flux de prospects Internet sur les périodes de temps futures, nous avons défini dans la section 5.3 un système de prévision des flux de prospects Internet tout en tenant compte des caractéristiques du modèle d'optimisation multi-période. En effet, le défi consistait à définir les paramètres du modèle de prévision qui conviendront au fonctionnement du modèle d'optimisation multi-période afin de pouvoir combiner les deux modèles. Par conséquent, nous avons étudié dans un premier temps l'articulation prévision/optimisation afin de pouvoir définir dans un second temps un algorithme qui génère l'ensemble des séries chronologiques. Ces derniers servent comme données d'entrée pour le modèle de prévision défini dans la section 5.3.2.

Après avoir intégré la prévision dans le modèle d'optimisation multi-période et défini ainsi une modélisation de l'optimisation avec anticipation répondant aux besoins de PdL, nous nous sommes focalisés sur la généralisation des modèles proposés. Le but d'une telle généralisation est de présenter un formalisme aussi générique que possible pouvant être appliqué à toute autre place de marché ayant les mêmes caractéristiques que celles de PdL. Par conséquent, nous avons introduit dans un premier temps une formalisation permettant de satisfaire tous les clients de manière équitable. Dans un deuxième temps, nous nous sommes intéressés aux flux de la demande afin de permettre au modèle d'optimisation multi-période d'intégrer la connaissance concernant les flux des clients et des commandes sur les périodes de temps futures. Dans un troisième temps, nous avons redéfini la contrainte liée au nombre de ventes et à l'exclusivité. Cette redéfinition a pour objectif de pouvoir personnaliser le nombre de ventes non exclusives des prospects Internet de chaque commande. Enfin, nous avons introduit dans la section 7.6 un modèle général de l'optimisation bi-objectif multi-période permettant de maximiser le chiffre d'affaires tout en cherchant une adéquation équitable de l'offre et de la demande vis-à-vis de tous les clients quelque soit leur importance. Ce modèle d'optimisation est réalisé en considé-

rant plusieurs périodes de temps futures afin de prendre en compte la connaissance concernant les nouveaux flux de commandes, de clients et de prospects Internet.

Au-delà des contributions méthodologiques, nous avons souhaité construire un environnement facilitant l'implémentation et l'expérimentation des modèles proposés. Cela nous semble important, d'une part, pour encourager l'utilisation de ces modèles qui restent parfois développés de manière théorique uniquement, et d'autre part, pour valider les résultats obtenus et ainsi permettre à PdL de réutiliser directement les modèles proposés.

Concernant ce dernier aspect, nous avons consacré beaucoup d'énergie à analyser les difficultés liées à l'implémentation des modèles mathématiques proposés dans ce document. C'est un travail qui est malheureusement parfois laissé de côté par la communauté scientifique, mais sans lequel les modèles sont difficilement utilisables.

8.2 Perspectives de recherche

Nous considérons des perspectives à court et moyen terme ainsi qu'à long terme, à savoir :

1. Perspectives à court et moyen terme : dans le cadre de cette thèse, nous nous sommes intéressés à l'optimisation de deux objectifs par le biais du processus de l'affectation des prospects Internet à savoir le chiffre d'affaires et la satisfaction des clients. Cependant, dans le cadre de Place des leads, une affectation vise à permettre la prise en compte simultanée de nombreux objectifs. En effet, la considération de différents objectifs à optimiser au niveau du traitement des prospects Internet constitue un besoin réel pour la place de marché PdL. Ainsi, dans un premier temps, une des perspectives de ce travail est de considérer d'autres objectifs à optimiser au niveau du traitement des prospects Internet. Un des objectifs à réaliser par le biais de l'affectation de l'offre à la demande est la satisfaction des fournisseurs de prospects Internet. En effet, de nombreux fournisseurs envoient une quantité importante de prospects Internet à PdL. Cette dernière se charge ensuite de les revendre en temps réel à des clients finaux. Ainsi, satisfaire équitablement tous les fournisseurs constitue un enjeu majeur pour PdL afin qu'elle puisse être alimentée

continuellement de prospects Internet. Un deuxième objectif à satisfaire par le processus de l'affectation est la fraîcheur des prospects Internet. En effet, un des critères définissant la topologie du prospect Internet est son âge qui est compté à partir de sa date de création. Les clients de PdL se basent sur ce critère pour définir la fraîcheur des prospects Internet qu'ils souhaitent recevoir. Par conséquent, plus le prospect Internet est jeune plus sa vente est facile. L'innovation envisagée par PdL consiste à revendre des prospects Internet jeunes.

Nous avons brièvement abordé dans les sections 2.2.4.1 et 2.2.4.2 la question de l'expression des préférences par un décideur. À présent, nous développons quelques réflexions concernant cette question. Notre approche pour traiter le problème de l'adéquation de l'offre et de la demande en temps réel de sorte à maximiser simultanément deux objectifs, consistait à le formuler par un programme mathématique bi-objectif dont la résolution conduit à identifier la surface de pareto des solutions pareto-optimales. Une des perspectives de ce travail concerne la prise en compte d'un modèle de préférences dans le modèle d'optimisation bi-objectif afin de sélectionner sur la surface de pareto une bonne solution de compromis représentant les arbitrages souhaités par le décideur. Nous pensons qu'une piste de recherche de solutions de bon compromis revient à agréger les critères à l'aide d'une fonction scalarisante f_γ , dite aussi fonction d'agrégation. L'indice γ de la fonction scalarisante fait référence aux paramètres qui modélisent les préférences du décideur. Une solution de compromis est donc obtenue en résolvant le problème (8.1).

$$\min_{x \in X} f_\gamma(z_1(x), \dots, z_p(x)) \quad (8.1)$$

Par la suite, nous adoptons une méthode interactive. Cette dernière consiste en une alternance d'étapes de calculs et de dialogue entre un analyste et un décideur. Ce dialogue doit prendre en compte le fait que : (1) l'affectation des flux de l'offre et de la demande s'effectue en temps et (2) l'intervalle de temps entre deux affectations est relativement réduit. La première étape du calcul fournit une première solution optimale pour le problème 8.1. Cette solution est présentée au décideur, qui réagit en apportant des informations supplé-

mentaires concernant ses préférences que l'analyste transcrit généralement en modifiant les paramètres γ . Une nouvelle solution de compromis est calculée et présentée à nouveau au décideur. Le processus interactif s'arrête lorsque le décideur est satisfait.

Un enjeu important du problème posé consiste à intégrer les modèles d'optimisation et de prévision dans un système de pilotage de PdL. Pour y parvenir, il faut concevoir un système d'aide à la décision permettant à un utilisateur humain de piloter le processus temps réel de l'affectation des prospects Internet aux commandes. Ce pilotage s'effectue en fonction de l'évolution de la performance du système selon les différents objectifs à optimiser. Il s'agit de permettre la modulation d'un compromis entre les différents objectifs et notamment de clarifier l'articulation entre les modules d'optimisation développés et l'intervention de l'opérateur du système. Nous envisageons différents niveaux de fonctionnement : automatique, semi-automatique, ... etc.

2. Perspectives à long terme : au-delà des perspectives de recherche présentées ci-dessus, nous souhaitons esquisser des perspectives à plus long terme. Ces perspectives concernent la méthode de prévision. Afin de mettre en oeuvre le système de prévision des prospects Internet proposé dans cette thèse, nous avons utilisé l'ensemble des modèles de prévision *ETS* implémenté sur le logiciel *R*. Ainsi, une piste intéressante serait de développer une méthode de prévision des arrivées des flux de prospects Internet propre au contexte de la place de marché PdL. Cette dernière présente un contexte assez particulier dans la mesure où elle reçoit l'offre des prospects Internet de manière continue. En d'autres termes, PdL est une place de marché en ligne ce qui signifie que les prospects Internet peuvent arriver sur sa plateforme en temps réel et 24h/24. Ainsi, les flux de prospects Internet sont caractérisés par une densité différente en fonction du moment de la journée. Par exemple, le nombre de prospects Internet que reçoit PdL de 10h00 à 11h00 est beaucoup plus important de ce qu'elle reçoit entre 2h00 et 3h00. En outre, la prévision s'effectue pour des périodes de temps très courtes alors que PdL ne possède pas forcément un flux de prospects Internet important et significatif sur toutes les périodes de la journée pour pouvoir effectuer une prévision pertinente. Ceci rend le problème de prévision des flux de pros-

pects Internet chez PdL assez complexe et nécessite une méthode de prévision particulière prenant en compte toutes ces caractéristiques.

En effet, d'une part, la méthodologie générale sera fondée sur une analyse très fine de séries chronologiques au niveau hebdomadaire, journalier et horaire. Elle combinera donc une approche endogène (décomposition de chronique) et exogène (recherche de variables explicatives externes). Il ne s'agira pas de décomposer des prévisions journalières en prévisions horaires, mais bien de déterminer des saisonnalités propres à chaque composante. D'autre part, cette méthodologie devrait considérer les groupes de commandes pour lesquels nous possédons un historique de flux de prospects Internet composé de valeurs soit très faibles, soit nulles, nous parlons alors d'une offre de prospects Internet erratique.

Enfin, l'approche de prévision devrait s'intéresser à l'impact de la prévision sur le modèle d'optimisation. En effet, l'enjeu des erreurs de prévision est d'autant plus important lorsque la prévision dépasse la réalisation que lorsque la prévision est dépassée par la réalisation. Illustrons ces propos par des exemples :

La prévision dépasse la réalisation

Supposons qu'à un instant t , nous avons les commandes et les prospects Internet définis par les tables 8.1 et 8.2 respectivement.

	Exclusivité	Âge	Prix	Profession	Code postal
Commande 1	oui	2h	7€	cadre	{75,78}

Tableau 8.1 — Table des commandes à l'instant t

	Nb Reventes	Âge	Profession	Code postal
Prospect Internet dispo	0	59min	cadre	{75}
Prospect Internet prévu	0	0min	cadre	{75}

Tableau 8.2 — Table des prospects Internet à l'instant t

Supposons que l'affectation optimale des prospects Internet aux commandes, consiste à affecter le prospect Internet prévu à la commande 1. Cependant, ce prospect Internet

prévu n'arrivera jamais. Ainsi, le prospect Internet disponible ne pourra plus être affecté. Cela génère une perte de 7 € par rapport au modèle d'optimisation sans anticipation.

En revanche, si à un instant t , la réalisation dépasse la prévision, ce n'est pas une perte mais un manque de gain. Nous expliquons cela par l'exemple suivant.

La prévision est dépassée par la réalisation

Supposons qu'à un instant t , nous avons les commandes et les prospects Internet définis par les tables 8.3 et 8.4 respectivement.

	Exclusivité	Âge	Prix	Profession	Code postal
Commande 1	oui	2h	7€	cadre	{75,78}
Commande 2	non	1h	2€	–	{75,76}
Commande 3	non	1h	2€	–	{75,92}

Tableau 8.3 — Table des commandes à l'instant t

	Nb Reventes	Âge	Profession	Code postal
Prospect Internet 1	0	59min	cadre	{75}

Tableau 8.4 — Table des prospects Internet à l'instant t

Dans cette configuration, il n'est possible que d'affecter exclusivement le prospect Internet 1 à la commande 1 pour un bénéfice de 7 €. Supposons que nous enregistrons 5 minutes plus tard l'arrivée d'un prospect Internet 2 non prévu appartenant au département 78. Donc ce prospect Internet vérifie les caractéristiques de la commande 1 mais pas celles des commandes 2 et 3. Il ne peut donc être affecté. Ainsi, l'anticipation de l'arrivée de ce prospect Internet aurait pu générer un gain de 11 €, en affectant le prospect Internet 1 aux commandes 2 et 3 et le prospect Internet 2 à la commande 1. Nous aurons donc un manque de gain de 4 €.

Par conséquent, le fait de prévoir moins que la réalisation ne génère aucune perte. Cependant, il génère un manque de gain. Idéalement, l'approche de prévision devrait essayer de trouver un bon compromis entre les deux.

Parmi les perspectives ouvertes par ce travail, l'élaboration d'un modèle de prévision pour les flux des demandes. En effet, dans la section 7.3 et dans le but de définir un modèle plus général de l'optimisation multi-période de l'adéquation offre/demande de prospects Internet, nous avons considéré qu'il est possible d'avoir des nouveaux flux de commandes arrivant de façon continue et non déterministe. Ainsi, il serait nécessaire de définir un modèle de prévision afin de fournir au modèle d'optimisation multi-période la connaissance concernant l'évolution des flux de commandes sur l'ensemble des périodes de temps prises en compte dans le modèle d'optimisation multi-période. Un tel modèle de prévision devrait prendre en compte les caractéristiques du modèle d'optimisation multi-période. En sus, ses paramètres devraient convenir au nombre et à la durée des périodes de temps du modèle d'optimisation multi-période.

Bibliographie

- A. Adeyefa and M. Luhandjula. Multiobjective stochastic linear programming : An overview. *American Journal of Operations Research*, 1 :203–213, 2011.
- R. Ahuja, T. Magnanti, and J. Orlin. Network flows : Theory, algorithms and applications. In *Prentice Hall, New Jersey*, 1993.
- J. Albert. Bayesian computation with r. In *Springer Science et Business Media*, 2009.
- D. Avis and L. Devroye. An analysis of a decomposition heuristic for the assignment problem. *Operations Research Letters*, pages 279–283, 1985.
- M.L. Balinski. A competitive (dual) simplex method for the assignment problem. *Mathematical Programming*, pages 125–141, 1986.
- R.S. Barr, F. Glover, and D. Klingman. The alternating basis algorithm for assignment problems. *Mathematical Programming*, pages 1–13, 1977.
- R. Bourbonnais and J.C. Usunier. Pr evision des ventes : Th eorie et pratique. In *Economica*, 2007.
- G.E.P. Box, G.M. Jenkins, and G.C. Reinsel. Time series analysis : Forecasting and control. In *Wiley-Blackwell*, 2008.
- R. Burkard, M. Dell’ Amico, and S. Martello. Assignment problems (revised version). In *SIAM Society for Industrial and Applied Mathematics, Philadelphia*, 2012.

- G.M. Campbell and M. Diaby. Development and evaluation of an assignment heuristic for allocating cross-trained workers. *European Journal of Operational Research*, 138 :9–12, 2002.
- J.F. Campell and A. Langevin. The snow disposal assignment problem. *The Journal of the Operational Research Society*, 48 :919–929, 1995.
- V. Chankong and Y.Y. Haimes. Multiobjective decision making : Theory and methodology. In *North Holland, New York*, 1983.
- T.P. Chartier, V. Ellison, and A.N. Langville. A davidson college multi-objective assignment problem : a case study. *A Quarterly Journal of Operations Research*, 12 :379–401, 2014.
- C. Chegireddy and H. Hamacher. Algorithms for finding k-best perfect matchings. *Discrete Applied Mathematics*, 18 :155–165, 1987.
- D.S. Chen, R.G. Batson, and Y. Dang. Applied integer programming : Modeling and solution. In *Wiley, New York*, 2010.
- J.L. Cohon. Multiobjective programming and planning. In *Academic Press, New York*, 1978.
- P. Dalgaard. Introductory statistics with r. In *Springer Science et Business Media*, 2008.
- M.M. Dessouky and B.A. Kijowski. Production scheduling of single-stage multi-product batch chemical processes with fixed batch sizes. *IIE Transactions*, 29 :399–408, 1997.
- C.W. Duin and A. Volgenant. On weighting two criteria with a parameter in combinatorial optimization problems. *European Journal of Operational Research*, 221 :1–6, 2012.
- S.P. Eberhardt, T. Duad, A. Kerns, T.X. Brown, and A.P. Thakoor. Competitive neural architecture for hardware solution to the assignment problem. *Neural Networks*, pages 431–442, 1991.
- M. Ehrgott and D.M. Ryan. Constructing robust crew schedules with bicriteria optimization. *Journal of Multi-Criteria Decision Analysis*, 11 :139–150, 2002.

- B.E. Emrah and R.S. Edis. An integer programming model for the conference timetabling problem. *C.B.U. Journal of Science*, 2 :55–62, 2013.
- J.J. Faraway. Extending the linear model with r. In *CRC Press*, 2005.
- A.S. Foulkes. Applied statistical genetics with r. In *Springer Science et Business Media*, 2009.
- M. Gondran and M. Minoux. Graphes et algorithmes. In *Tec et Doc Lavoisier*, 2009.
- H.W. Hamacher, C.R. Pedersen, and S. Ruzika. Finding representative systems for discrete bicriterion optimization problems. *Operations Research Letters*, 35 :336–344, 2007.
- M.S. Hung and W.O. Rom. Solving the assignment problem by relaxation. *Operations Research*, pages 969–982, 1980.
- C.L. Hwang and A. Masud. Multiple objective decision making. methods and applications : A state of the art survey. In *Springer-Verlag Berlin Heidelberg*, 1979.
- R.J. Hyndman, A.B. Koehler, R.D. Snyder, and S. Grose. A state space framework for automatic forecasting using exponential smoothing methods. *International J. Forecasting*, pages 439–454, 2002.
- R.J. Hyndman, Md. Akram, and B. Archibald. The admissible parameter space for exponential smoothing models. *Annals of Statistical Mathematics*, pages 407–426, 2008.
- O. Karsu and M. Azizoglu. Bicriteria multiresource generalized assignment problem. *Naval Research Logistics*, 61 :621–636, 2014.
- R.L. Keeney and H. Raiffa. Decisions with multiple objectives : Preference and value tradeoffs. In *New York : Cambridge University Press*, 1993.
- G. Koole. Call center optimisation. In *MG books*, 2013.
- H.W. Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, pages 83–97, 1955.

- J. Larusic and A.P. Punnen. The balanced traveling salesman problem. *Computers and Operations Research*, 38(5) :868–875, 2011.
- M. Laumanns, L. Thiele, and E. Zitzler. An efficient adaptive parameter variation scheme for metaheuristics based on the epsilon-constraint method. *European Journal of Operational Research*, 169 :932–942, 2006.
- L.J. Leblanc and K. Farhangian. Efficient algorithm for solving elastic demand traffic assignment problem and mode split-assignment problem. *Transport. Sci*, pages 306–317, 1981.
- S.G. Makridakis, S.C. Wheelwright, and R.J. Hyndman. Forecasting : Methods and applications. In *Wiley*, 1997.
- G. Mavrotas. Effective implementation of the e-constraint method in multi-objective mathematical programming problems. *Applied Mathematics and Computation*, 213 :455–455, 2009.
- G. Mavrotas and D. Diakoulaki. Multi-criteria branch and bound : A vector maximization algorithm for mixed 0-1 multiple objective linear programming. *Applied Mathematics and Computation*, 171 :53–71, 2005.
- S.L. Maxon and J. Bhadury. An ms-excel implementation of a multi-period assignment problem with repetitive tasks. In *Proceedings of the 13th Annual CSU-POM Conference, California State University San Bernardino*, pages 39–48, 2001.
- L.F. McGinnis. Implementation and testing of a primal-dual algorithm for the assignment problem. *Operations Research*, pages 277–291, 1983.
- P. Mckeown and B. Workman. A study in using linear programming to assign students to schools. *Interfaces*, 6 :96–101, 1976.
- G. Mélard and M. Carbon. Méthodes de prévision à court terme. In *Université de Bruxelles*, 2008.

- P. Lafaye De Micheaux, R. Drouilhet, and B. Liquet. Le logiciel r. In *Springer Verlag France*, 2014.
- K.M. Miettinen. Nonlinear multiobjective optimization. In *Kluwer Academic Publishers, Boston*, 1999.
- M. Miller, H. Stone, and I. Cox. Optimizing murty’s ranked assignment method. *IEEE Transactions on Aerospace and Electronic Systems*, 33 :851–862, 1997.
- K. Murty. An algorithm for ranking all the assignments in order of increasing cost. *Operations Research*, 16 :682–687, 1968.
- D. Nace, L.N. Doan, O. Klopfenstein, and A. Bashllari. Max-min fairness in multi-commodity flows. *Computers and Operations Research*, 35 :557–573, 2008.
- M. Pascoal, M. Captivo, and J. Climaco. A note on a new variant of murty’s ranking assignments algorithm. *Quarterly Journal of the Belgian, French and Italian Operations Research Societies*, 1 :243–255, 2003.
- C. Pedersen, L. Nielsen, and K. Andersen. An algorithm for ranking assignments using reoptimization. *Computers et Operations Research*, 35 :3714–3726, 2008.
- D.W. Pentico. Assignment problems : A golden anniversary survey. *European Journal of Operational Research*, 176(2) :774–793, 2007.
- M.D. Pettypool and P. Karathanos. An equity check. *European Journal of Operational Research*, 157 :465–470, 2004.
- R.F. Potthoff and S.J. Brams. Scheduling of panels by integer programming : results for the 2005 and 2006 new orleans meetings. *Public Choice*, 131 :465–468, 2007.
- B. Roy. Multicriteria methodology for decision analysis. In *Kluwer Academic Publishers*, 1996.

- M. Sakarovitch. Optimisation combinatoire : Programmation discrète. In *Collection Enseignement des sciences, Hermann*, 1984.
- S.E. Sampson. Practical implications of preference-based conference scheduling. *Production and Operations Management*, 13 :205–215, 2004.
- R.H. Shumway and D.S. Stoffer. Time series analysis and its applications : With r examples. In *Springer Science et Business Media*, 2010.
- F. Soumis, J. Ferland, and J. Rousseau. A model for large-scale aircraft routing and scheduling problems. *Transport. Res. Part B : Meth*, 14 :191–201, 1980.
- R.E. Steuer. Multiple criteria optimization : theory, computation, and application. In *Malabar, FL : Krieger*, 1986.
- G.A. Suer and I.S. Bera. Optimal operator assignment and cell loading when lot-splitting is allowed. *Computers and Industrial Engineering*, 35 :431–434, 1998.
- J. Teghem. Recherche opérationnelle - tome 1 : Méthodes d’optimisation. In *Ellipses Marketing*, 2012.
- L. Turner, A.P. Punnen, Y.P. Aneja, and H.W. Hamacher. On generalized balanced optimization problems. *Mathematical Methods of Operations Research*, 73 :19–27, 2011.
- S. Wood. Generalized additive models : An introduction with r. In *CRC Press*, 2006.
- G. Yu, M. Arguello, G. Song, S.M. McCowan, and A. White. A new era for crew recovery at continental airlines. *Interfaces*, 33 :5–22, 2003.
- X. Zhang and J.F. Barad. A multi-period machine assignment problem. *European Journal of Operational Research*, 170 :398–415, 2006.

Annexe **A**

Implémentation des modèles d'optimisation

Dans cette première partie, nous détaillons l'implémentation des trois modèles d'optimisation présentés dans cette thèse, à savoir :

1. Le modèle d'optimisation mono-objectif mono-période proposé dans la section 3.4.
2. Le modèle d'optimisation bi-objectif mono-période présenté dans la section 3.6.2.
3. Le modèle d'optimisation mono-objectif multi-période défini dans la section 4.2.4.

A.1 Lecture et structuration des données

Comme nous avons mentionné dans le chapitre 6, nous avons effectué une étude numérique par l'utilisation des données réelles récupérées de la base de données PdL. En effet, les données sont récupérées depuis la base de données PdL dans un fichier CSV. Afin de pouvoir traiter ces données, nous développons un programme *JAVA* (voir la figure A.1) permettant la lecture de ces données.

```

public class MonoCrit {

    public static void main(String[] args) throws Exception {

        Matriccee obj = new Matriccee();// Vector des commandes
        String nom = null ;
        JFileChooser ff= new JFileChooser() ;// 1 instantiation
        int v = ff.showOpenDialog(null);// afficher
        if(v == JFileChooser.APPROVE_OPTION){
            nom=ff.getSelectedFile().getAbsolutePath();
        }
        else{ System.exit(0);// pour sortir du pgm}

        String fichier = nom;
        InputStream ips = new FileInputStream(fichier);
        InputStreamReader ipsr = new InputStreamReader(ips);
        BufferedReader br = new BufferedReader(ipsr);
        String ligne = null;
    
```

Figure A.1 — Lecture des données

La figure A.2 représente la structure utilisée afin de stocker les commandes lues depuis le fichier CSV.

```

while ((ligne = br.readLine()) != null)
{ if (bbl1 == true)    // Pour ne pas considérer la première ligne du CSV
  String[] data = ligne.split(",");
  String texte = data[0]; // une ligne du CSV
  Pattern separateur = Pattern.compile(";");

  String[] items = separateur.split(texte);
  String markagep1 = items[0];
  Integer mark1 = new Integer(markagep1);
  int markage1 = mark1.intValue(); //conversion d'Integer en int

  String markagep2 = items[1];
  Integer mark2 = new Integer(markagep2);
  int markage2 = mark2.intValue(); //conversion d'Integer en int

  String markagep3 = items[2];
  float markage3 = Float.valueOf(markagep3.trim()).floatValue();

  String markagep4 = items[3];
  Integer mark4 = new Integer(markagep4);
  int markage4 = mark4.intValue(); //conversion d'Integer en int

  String markagep40 = items[4];
  Integer mark40 = new Integer(markagep40);
  int markage40 = mark40.intValue(); //conversion d'Integer en int

  String markagep41 = items[5];
  Integer mark41 = new Integer(markagep41);
  int markage41 = mark41.intValue(); //conversion d'Integer en int

  Commandeee c1 = new Commandeee(markage1,markage2, markage3 ,markage4,markage40,markage41);
    
```

Figure A.2 — Structure des Commandes

A.2 Implémentation du modèle d'optimisation mono-objectif mono-période

Les modèles d'optimisation que nous avons définis dans les différents chapitres de cette thèse ont été implémentés en utilisant le logiciel *CPLEX* et le langage *JAVA*. En effet, dans la section A.1, nous avons montré la manière avec laquelle les données sont lues et stockées. Dans cette section, nous présentons un code implémentant les différentes contraintes ainsi que la fonction objectif du modèle d'optimisation mono-objectif mono-période proposé dans la section 3.4.

A.2.1 Création du projet *CPLEX* avec variables de décision binaires

Afin de pouvoir implémenter le modèle d'optimisation mono-objectif proposé dans la section 3.4, nous commençons par créer un projet *CPLEX* (voir la figure A.3).

```
//Projet Cplex
try {
    IloCplex cplex = new IloCplex();

    int co=0;
    ListAffichage LisAff = new ListAffichage();// pour l'affichage
    for (int v1 = 0; v1 < obj.getListCom().size(); v1++) {
        Commandeee c2 = (Commandeee) obj.getListCom().get(v1);
        for (int w = 0; w < c2.getListLead().size(); w++) {
            Lead l2 = (Lead) c2.getListLead().get(w);
            if (l2.getValue() == 1)
            {
                Affichage aff1 = new Affichage(c2.getId_comm(),l2.getId_lead());// pour l'affichage
                LisAff.ajouterAff(aff1);// pour l'affichage
                co=co+1;
            }
        }
    }

    IloIntVar[] x = cplex.intVarArray(co, 0, 1);// co variables entières toutes avec bornes 0 et 1
    double [] objvals = new double[co]; //construit un tableau de co doubles, numérotés de 0 à co-1.
}
```

Figure A.3 — Création du projet *CPLEX*

Après avoir créer le projet *CPLEX*, nous définissons un vecteur x de variables de décision binaires et de taille co . Pour ce faire, nous utilisons le codage suivant : `IloIntVar[] x = cplex.intVarArray(co,0,1);` (voir la figure A.3).

A.2.2 Implémentation des contraintes et de la fonction objectif

À partir du projet *CPLEX* créé dans la section A.2.1, nous définissons les différentes fonctions permettant d'implémenter les contraintes et la fonction objectif de ce modèle d'optimisation mono-objectif. L'implémentation des contraintes et de la fonction objectif présentées dans les sections 3.3.3 et 3.3.4 est donnée par les figures A.4, A.5, A.6, A.7 et A.8.

```
// constraint 3: quota des commandes
for (int i = 0; i < c3_lines.length; i++) {

    String linne;
    String[] itemss;
    String[] params;
    String[] params2;

    IloLinearNumExpr lin1 = cplex.linearNumExpr(); // Construire une expression lin1 pour un lead i
    linne = c3_lines[i];
    itemss = linne.split(";");
    params = itemss[0].split("\\|");
    for (int j = 1; j < itemss.length; j++) {
        params2 = itemss[j].split("\\|");
        for (int u = 0; u < index.length; u++) {
            if (index[u] == Integer.parseInt(params2[0])) { // params2[0] = index recherché
                lin1.addTerm(x[u], Integer.parseInt(params2[1])); // params2[1] = coef //test
            }
        }
    }
    cplex.addLe(lin1, Integer.parseInt(params[1])); // Ajouter la contrainte du lead i // params[1] = la limite
}
}
```

Figure A.4 — Implémentation de la contrainte liée aux volumes des commandes

```
// constraint 4: quota des clients
for (int i = 0; i < c4_lines.length; i++) {

    String linne;
    String[] itemss;
    String[] params;
    String[] params2;

    IloLinearNumExpr lin1 = cplex.linearNumExpr(); // Construire une expression lin1 pour un lead i
    linne = c4_lines[i];
    itemss = linne.split(";");
    params = itemss[0].split("\\|");
    for (int j = 1; j < itemss.length; j++) {
        params2 = itemss[j].split("\\|");
        for (int u = 0; u < index.length; u++) {
            if (index[u] == Integer.parseInt(params2[0])) { // params2[0] = index recherché
                lin1.addTerm(x[u], Integer.parseInt(params2[1])); // params2[1] = coef //test
            }
        }
    }
    cplex.addLe(lin1, Integer.parseInt(params[1])); // Ajouter la contrainte du lead i // params[1] = la limite
}
}
```

Figure A.5 — Implémentation de la contrainte liée aux volumes des clients

```

// constraint 1: exclusivité
for (int i = 0; i < c1_lines.length; i++) {

    String linne;
    String[] itemss;
    String[] params;
    String[] params2;

    IloLinearNumExpr lin1 = cplex.linearNumExpr(); // Construire une expression lin1 pour un lead i
    linne = c1_lines[i];
    //System.out.println("linne= " +linne);// test
    itemss = linne.split(";");
    params = itemss[0].split("\\|");
    for (int j = 1; j < itemss.length; j++) {
        params2 = itemss[j].split("\\|");
        for (int u = 0; u < index.length; u++) {
            if ( index[u] == Integer.parseInt(params2[0])) { // params2[0] = index recherché
                lin1.addTerm(x[u], Integer.parseInt(params2[1])); // params2[1] = coef //test
            }
        }
    }
    cplex.addLe(lin1, Integer.parseInt(params[1]) ); // Ajouter la contrainte du lead i // params[1] = la limite
}

```

Figure A.6 — Implémentation de la contrainte liée au nombre de ventes et l'exclusivité

La figure A.4 implémente la contrainte 3.1, la figure A.5 implémente la contrainte 3.3, la figure A.6 implémente la contrainte 3.5, la figure A.7 implémente la contrainte 3.7 et la figure A.8 implémente la fonction objectif 3.8.

```

// constraint 2: duplication des prospects
for (int i = 0; i < c2_lines.length; i++) {

    String linne;
    String[] itemss;
    String[] params;
    String[] params2;

    IloLinearNumExpr lin1 = cplex.linearNumExpr(); // Construire une expression lin1 pour un lead i
    linne = c2_lines[i];
    itemss = linne.split(";");
    params = itemss[0].split("\\|");
    for (int j = 1; j < itemss.length; j++) {
        params2 = itemss[j].split("\\|");
        for (int u = 0; u < index.length; u++) {
            if ( index[u] == Integer.parseInt(params2[0])) { // params2[0] = index recherché
                lin1.addTerm(x[u], Integer.parseInt(params2[1])); // params2[1] = coef //test
            }
        }
    }
    cplex.addLe(lin1, Integer.parseInt(params[1]) ); // Ajouter la contrainte du lead i // params[1] = la limite
}

```

Figure A.7 — Implémentation de la contrainte liée à la duplication des prospects Internet

```
// Fonction objectif
double [] objvals = new double[lp_col]; // tableau des prix de lp_col doubles, numérotés de 0 à (lp_col - 1).
int compt2 = 0;
for (int i = 0; i < nb_order; i++) {
    for (int j = 6; j < nb_lead + 6; j++) {
        //lp_col_index = Integer.parseInt((matrix[i][j][0]).substring(0, 1));
        lp_col_index = Integer.parseInt((matrix[i][j][0]));
        if (lp_col_index > 0) {
            objvals[compt2] = Double.parseDouble(matrix[i][2][0]); // lead price
            compt2++;
        }
    }
}
cplex.addMaximize(cplex.scalProd(x, objvals)); //Fonction objectif
```

Figure A.8 — Implémentation de la fonction objectif : maximisation du chiffre d'affaires

A.2.3 Résolution du programme mathématique et affichage des résultats

La résolution du programme mathématique défini à travers le projet *CPLEX* est donnée par la figure A.9. L'affichage des résultats est réalisé avec la méthode donnée par la figure A.10.

```
// Résolution
if ( cplex.solve() ) {
    double CA=0;
    double[] val = cplex.getValues(x);
    System.out.println("The decision variables are: ");
    ecrire("nomFic.txt", "The decision variables are: " + "\r\n");
    for (int j = 0; j < lp_col; ++j)
    {
        String result = "order_lead = " + idOrderLead[j]+ "    decision variable = " + String.valueOf(val[j]) + "
        ecrire("nomFic.txt", result);
        System.out.println("order_lead = " +idOrderLead[j]+ "    decision variable = " +val[j]+ "    turnover =
        CA = CA + (objvals[j] * val[j]);
    }

    System.out.println("Number of orders = " + nb_order);
    ecrire("nomFic.txt", "Number of orders = " + String.valueOf(nb_order)+ "\r\n");

    System.out.println("Number of leads = " + nb_lead);
    ecrire("nomFic.txt", "Number of leads = " + String.valueOf(nb_lead)+ "\r\n");

    System.out.println("Number of decision variables = " + lp_col);
    ecrire("nomFic.txt", "Number of decision variables = " + String.valueOf(lp_col)+ "\r\n");

    System.out.println("The total turnover = " + CA);
    ecrire("nomFic.txt", "The total turnover = " + String.valueOf(CA)+ "\r\n");
}

cplex.end();
}
catch (IloException e) {
    System.err.println("Concert exception '" + e + "' caught");
}
```

Figure A.9 — Résolution du programme mathématique mono-objectif

```

// Ecriture dans un fichier
static void ecrire(String nomFic, String texte)
{
    String adressedufichier = System.getProperty("user.dir") + "/" + nomFic;

    try
    {
        FileWriter fw = new FileWriter(adressedufichier, true);
        BufferedWriter output = new BufferedWriter(fw);

        output.write(texte);
        output.flush();

        output.close();
        //System.out.println("fichier créé");
    }
    catch(IOException ioe){
        System.out.print("Erreur : ");
        ioe.printStackTrace();
    }
}

```

Figure A.10 — Affichage des résultats

A.3 Implémentation du modèle d'optimisation bi-objectif mono-période

Dans cette section, nous illustrons l'implémentation du modèle d'optimisation bi-objectif mono-période présenté dans la section 3.6.2. Afin de pouvoir récupérer les données, nous utilisons les méthodes définies dans la section A.1. Quant à la création du projet *CPLEX* et la définition des variables de décision, nous utilisons le même code défini dans la section A.2.1.

La différence entre les implémentations des deux modèles d'optimisation, à savoir le modèle mono-objectif mono-période et le modèle bi-objectif mono-période réside dans la fonction objectif ainsi que la méthode de résolution. En effet, pour résoudre le modèle d'optimisation bi-objectif, nous utilisons la méthode ϵ -constraint dont l'implémentation nécessite la considération des points suivants.

D'une part, les deux fonctions objectif données par les équations (3.8) et (3.10) doivent être concaténées. D'autre part, nous rajoutons deux principales instructions, à savoir : (i) une condition d'arrêt pour la méthode ϵ -constraint et (ii) une contrainte spécifique à la méthode

ε -constraint. Ainsi, nous proposons une nouvelle présentation du modèle d'optimisation bi-objectif mono-période (voir la figure A.11).

$$P \left\{ \begin{array}{l} \max \quad f(X) = \sum_{i,j:x_{ij} \in X} Pr_i x_{ij} + \varepsilon \left[\frac{1}{n} \left[\sum_{i:x_{ij} \in X} (V_i + \sum_{j:x_{ij} \in X} x_{ij}) / Do_i \right] \right] \\ \text{s.t.} \\ \sum_{j:x_{ij} \in X} x_{ij} \leq Ro_i, \quad \forall i \in O \\ \sum_{i,j:x_{ij} \in X} a_{ik} x_{ij} \leq Rc_k, \quad \forall k \in C \\ S_j + b \sum_{i:x_{ij} \in E} x_{ij} + \sum_{i:x_{ij} \in \bar{E}} x_{ij} \leq b, \quad \forall j \in P \\ \sum_{i:x_{ij} \in X} a_{ik} x_{ij} \leq 1, \quad \forall j \in P, \forall k \in C \\ \left[\sum_{i,j:x_{ij} \in X} Pr_i x_{ij} \right]_t \leq \left[\sum_{i,j:x_{ij} \in X} Pr_i x_{ij} \right]_{t-1} - 0.5 \quad \forall i \in O \\ x_{ij} \in \{0,1\} \end{array} \right.$$

Figure A.11 — Redéfinition du programme mathématique bi-objectif

Algorithme A.1 — ε -constraint()

Data :

t : itération courante.

$t - 1$: itération antérieure.

$P = \{j : j \in \{1, \dots, m\}\}$ est l'ensemble de prospects Internet à l'instant t .

$O = \{i : i \in \{1, \dots, n\}\}$ est l'ensemble de commandes à l'instant t .

$C = \{k : k \in \{1, \dots, r\}\}$ est l'ensemble de clients à l'instant t .

$X = \{x_{ij} : i \in O, j \in P, x_{ij} \text{ est une variable binaire} \}$ est l'ensemble des variables de décision.

- 1 **while** $\left[\frac{1}{n} \left[\sum_{i:x_{ij} \in X} (V_i + \sum_{j:x_{ij} \in X} x_{ij}) / Do_i \right] \right]_t >$
 $\left[\frac{1}{n} \left[\sum_{i:x_{ij} \in X} (V_i + \sum_{j:x_{ij} \in X} x_{ij}) / Do_i \right] \right]_{t-1}$ **do**
 - 2 Appeler le programme mathématique $P(X)$.
 - 3 **end**
-

À chaque itération de l'algorithme A.1, un point de la frontière efficace est calculé. L'implémentation de l'algorithme A.1 est détaillée ci-dessous.

Les deux fonctions objectif données par les équations (3.8) et (3.10) doivent être concaténées (voir la figure A.12).

```

// la fonction objectif
IloLinearNumExpr lin555 = cplex.linearNumExpr();
for (int k = 0; k < co; k++){
    lin555.addTerm(objvals[k], x[k]); // fonction objectif du CA
}
// Construire le fonction objectif de la SMC
IloNumExpr neww2 = cplex.numExpr(); // somme alpha i
for (int v5 = 0; v5 < lc1.getListCompte().size(); v5++) {
    Comprttee cm2 = (Comprttee) lc1.getListCompte().get(v5);

    IloLinearNumExpr lin44 = cplex.linearNumExpr();
    for (int iu = cm2.getMin(); iu <= cm2.getMax(); ++iu){
        lin44.addTerm(1, x[iu]);
    }
    Commandeee com2 = (Commandeee) obj.getListCom().get(v5);
    IloNumExpr neww = cplex.numExpr();
    neww = cplex.sum(com2.getQuotaFourni(), lin44); // quotaF + somme Xij
    IloNumExpr neww1 = cplex.numExpr();
    neww1 = cplex.prod(neww, 1.0/com2.getQuotaJour()); // un alpha i
    neww2 = cplex.sum(neww2, neww1);
}
IloNumExpr neww3 = cplex.prod(0.001/y, neww2); // fonction objectif de la SMC
IloNumExpr neww4 = cplex.sum(lin555, neww3); // fonction objectif finale
cplex.addMaximize(neww4); // insertion de la fonction objectif finale

```

Figure A.12 — Fonctions objectif

Nous rajoutons : (i) une condition d'arrêt pour la méthode ε -constraint (voir la figure A.13) et (ii) une contrainte spécifique à la méthode ε -constraint (voir la figure A.14). Afin de pouvoir mettre en oeuvre ces deux instructions, nous devons calculer le chiffre d'affaires ainsi que la satisfaction moyenne des clients pour chaque point de la frontière efficace (voir la figure A.15). Par conséquent, chaque point de la frontière efficace est donné par la résolution du projet *CPLEX* (voir la figure A.16).

```
//condition d'arrêt: pas d'amélioration pour la satisfaction des clients
if(ListElem.getListAf().size() != 0){
    for (int jj = 0; jj < ListElem.getListAf().size(); jj++){
        Elem eel = (Elem) ListElem.getListAf().get(jj);
        if (eel.getValeurZ() >= sommf){
            bbol1=true;
        }
    }
}
```

Figure A.13 — Condition d'arrêt pour la méthode ϵ constraint

```
//contrainte de epsilon constraint
if(bbol==true){
    IloLinearNumExpr lin550 = cplex.linearNumExpr();
    for (int k = 0; k < co; k++){
        lin550.addTerm(objvals[k], x[k]);
    }
    cplex.addLe(lin550, caCal - 0.5);
}
```

Figure A.14 — Contrainte pour la méthode ϵ constraint

```
//calcul de la satisfaction moyenne des clients
double sommf = 0; // la somme des sommf
for (int v5 = 0; v5 < lc1.getListCompte().size(); v5++) {
    double somm = 0; // somme Xij
    double sommf1 = 0; // (quotaF + somme Xij) / quotaJ
    Compttee cm2 = (Compttee) lc1.getListCompte().get(v5);
    for (int iu = cm2.getMin(); iu <= cm2.getMax(); ++iu){
        somm=somm + val[iu];
    }
    Commandeee com2 = (Commandeee) obj.getListCom().get(v5);
    sommf1 = (com2.getQuotaFourni() + somm)/com2.getQuotaJour();
    sommf = sommf + sommf1;
}
double sommf = sommf / y; // la satisfaction moyenne des clients
System.out.println("*****SMC***** " + sommf);
```

Figure A.15 — Calcul du chiffre d'affaires et de la satisfaction moyenne des clients

```
// Résolution
if ( cplex.solve() ) {
    cplex.output().println("Solution status = " + cplex.getStatus());
    cplex.output().println("Solution value = " + cplex.getObjValue());

    double[] val = cplex.getValues(x);
}
```

Figure A.16 — Résolution du programme mathématique bi-objectif

A.4 Implémentation du modèle d'optimisation mono-objectif multi-période

Nous rappelons que le modèle d'optimisation mono-objectif mono-période optimise l'affectation des prospects Internet à un instant donné et en considérant une seule période de temps. En effet, cette première formalisation est mono-période, ce qui signifie que l'optimisation considère uniquement la période de temps courante sans prendre en compte l'évolution du système sur les périodes de temps futures.

Nous rappelons également qu'afin de surmonter ces limitations, nous avons proposé un modèle d'optimisation multi-période qui étend le problème sur plusieurs périodes de temps futures. En effet, le modèle d'optimisation multi-période permet d'établir des affectations à un instant donné et en considérant plusieurs périodes de temps futures. Ainsi, le modèle d'optimisation mono-objectif multi-période défini dans la section 4.2.4 consiste à étendre le modèle d'optimisation mono-objectif mono-période proposé dans la section 3.4 sur plusieurs périodes de temps.

Afin de pouvoir implémenter le modèle d'optimisation mono-objectif mono-période, nous avons récupéré les données concernant la période de temps en question sous forme matricielle. Ensuite, l'implémentation du modèle d'optimisation mono-objectif mono-période a été réalisée en se basant sur cette matrice de données.

Notre idée pour implémenter le modèle d'optimisation mono-objectif multi-période consiste à récupérer les données de chaque période de temps prise en compte dans l'optimisation sous forme matricielle (voir la figure A.17). Dans un second temps, ces matrices vont être concaténées tout en gardant un indice sur la période. L'union des matrices de toutes les périodes de temps considérées dans l'optimisation constituera la matrice de données pour le modèle d'optimisation mono-objectif multi-période (voir la figure A.18).

Ainsi, l'implémentation du modèle d'optimisation mono-objectif multi-période restera la même que celle du modèle d'optimisation mono-objectif mono-période. Grace à l'indice de

la période, nous pouvons distinguer celles concernées par l'affectation ou pas d'un prospect Internet donné à une commande donnée (voir la figure A.19).

```
public class MonoCrit
{
    public static void main(String[] arg)
    {
        String root_path= new String();

        root_path = "C:\\Users\\maamar\\Desktop\\MonoPeriod\\";

        // extract data from files
        String[] p_lines = extract_lines(root_path + "matrix.txt");
        String[] c1_lines = extract_lines(root_path + "c1_matrix.txt");
        String[] c2_lines = extract_lines(root_path + "c2_matrix.txt");
        String[] c3_lines = extract_lines(root_path + "c3_matrix.txt");
        String[] c4_lines = extract_lines(root_path + "c4_matrix.txt");
        String[] c5_lines = extract_lines(root_path + "c5_matrix.txt");
        String[] c6_lines = extract_lines(root_path + "c6_matrix.txt");
        String[] c7_lines = extract_lines(root_path + "c7_matrix.txt");
        String[] c8_lines = extract_lines(root_path + "c8_matrix.txt");
        String[] c9_lines = extract_lines(root_path + "c9_matrix.txt");
    }
}
```

Figure A.17 — Matrices de données pour le modèle mono-objectif multi-période

```
String[][][] matrix = new String[nb_order][nb_lead + nb_col_before_lead_info][5];

for (int i = 0; i < nb_order; i++) {
    line = p_lines[i];
    items = line.split(";");
    matrix[i][0][0] = items[0];
    matrix[i][1][0] = items[1];
    matrix[i][2][0] = items[2];
    matrix[i][3][0] = items[3];
    for (int j = nb_col_before_lead_info; j < nb_lead + nb_col_before_lead_info; j++) {
        String lead = items[j];
        String[] lead_items = lead.split("\\|");
        if (lead_items[2].equals("1")) {
            lp_col++;
            matrix[i][j][0] = lead_items[3];
        } else {
            matrix[i][j][0] = "0";
        }
        matrix[i][j][1] = lead_items[0];
        matrix[i][j][2] = lead_items[1];
        matrix[i][j][3] = lead_items[2];
        matrix[i][j][4] = lead_items[4];
    }
}
```

Figure A.18 — Matrice unique pour le modèle mono-objectif multi-période

```

System.out.println("The total turnover = " + CA);
ecrire("nomFic.txt", "The total turnover = " + String.valueOf(CA) + "\r\n");
System.out.println("The turnover of the current period = " + CAC);
ecrire("nomFic.txt", "The turnover of the current period = " + String.valueOf(CAC) + "\r\n");
System.out.println("The turnover of the period 1 = " + CAC1);
ecrire("nomFic.txt", "The turnover of the period 1 = " + String.valueOf(CAC1) + "\r\n");
System.out.println("The turnover of the period 2 = " + CAC2);
ecrire("nomFic.txt", "The turnover of the period 2 = " + String.valueOf(CAC2) + "\r\n");
System.out.println("The turnover of the period 3 = " + CAC3);
ecrire("nomFic.txt", "The turnover of the period 3 = " + String.valueOf(CAC3) + "\r\n");
System.out.println("The turnover of the period 4 = " + CAC4);
ecrire("nomFic.txt", "The turnover of the period 4 = " + String.valueOf(CAC4) + "\r\n");
System.out.println("The turnover of the period 5 = " + CAC5);
ecrire("nomFic.txt", "The turnover of the period 5 = " + String.valueOf(CAC5) + "\r\n");
System.out.println("The turnover of the period 6 = " + CAC6);
ecrire("nomFic.txt", "The turnover of the period 6 = " + String.valueOf(CAC6) + "\r\n");
System.out.println("The turnover of the period 7 = " + CAC7);
ecrire("nomFic.txt", "The turnover of the period 7 = " + String.valueOf(CAC7) + "\r\n");
System.out.println("The turnover of the period 8 = " + CAC8);
ecrire("nomFic.txt", "The turnover of the period 8 = " + String.valueOf(CAC8) + "\r\n");
System.out.println("The turnover of the period 9 = " + CAC9);
ecrire("nomFic.txt", "The turnover of the period 9 = " + String.valueOf(CAC9) + "\r\n");
System.out.println("The turnover of the period 10 = " + CAC10);
ecrire("nomFic.txt", "The turnover of the period 10 = " + String.valueOf(CAC10) + "\r\n");
}

cplex.end();
}

```

Figure A.19 — Résultat du modèle mono-objectif multi-période

Annexe **B**

Implémentation du modèle de prévision

Dans cette deuxième partie, nous détaillons la mise en oeuvre du modèle de prévision proposé dans le chapitre 5 de cette thèse. En effet, afin de réaliser une étude prévisionnelle, nous procédons comme suit :

1. Nous commençons par récupérer les historiques des flux de prospects Internet ainsi des commandes depuis la base de données de PdL.
2. En suite, nous construisons les groupes de commandes partageant les mêmes prospects Internet.
3. Dans un troisième temps, nous calculons la prévision sous le logiciel *R*.
4. Enfin, nous calculons les erreurs de prévision.

L'extraction des données ainsi que les différentes méthodes implémentant le calcul des groupes de commandes, de la prévision et des erreurs de prévision sont détaillées ci-dessous.

B.1 Structure du prospect Internet

La classe "Lead" (voir la figure B.1) définit la structure dans laquelle nous sauvegardons le prospect Internet récupéré depuis la base de données PdL.

```

public class Lead {

    private ArrayList <Integer> AL = new ArrayList<Integer>();
        //Va contenir l'Id des Commandes satisfaites par le Lead

    public Lead (ArrayList <Integer> A) {
        for (int i=0; i<A.size();i++)
            this.AL.add(A.get(i));
    }

    // Pour afficher dans la console la liste des commandes satisfaites par le lead
    public void Affichage () {
        for (int i=0; i<this.AL.size();i++){
            if (i<this.AL.size()-1) System.out.print(this.AL.get(i)+", ");
            else System.out.print(this.AL.get(i));
        }
    }

    //Pour afficher la liste des commandes satisfaites par le lead dans un fichier externe
    public String Afficher(){
        String s="";
        for (int i=0; i<this.AL.size();i++){
            if (i<this.AL.size()-1) s+=this.AL.get(i)+",";
            else s+=this.AL.get(i);
        }
        return s;
    }

    public int nbCommandes () {return this.AL.size();} // rend le nombre de Commandes satisfaites par le Lead
    public int getCom (int i) {return this.AL.get(i);} // rend l'Id dela i eme Cde satisfaites par le Lead

    boolean Equal (Lead Le){ // Compare les commandes satisfaites du Lead avec un autre Lead
        boolean bool = false;
        int j=-1;
        if (this.AL.size()==Le.nbCommandes()){
            for (int i=0; i<this.AL.size();i++){
                if (this.AL.get(i)!=Le.getCom(i)) {break;}
                j=i+1;
            }
            if (j==Le.nbCommandes()) {bool=true;}
        }
        return bool;
    }
}

```

Figure B.1 — Structure du prospect Internet

B.2 Extraction des données

Les données historiques sont récupérées depuis la base de données PdL dans un fichier texte nommé "Données.txt". La classe "Extraction-des-données" (voir la figure B.2) définit une structure permettant de lire les données depuis le fichier "Données.txt" et les stocker dans une matrice "prospects-commandes" que nous appelons "matrice".

```

public class Extraction_des_Données {
    public static void main(String[] args) {
        // TODO Auto-generated method stub
        //***** Sauvegarde des Données dans une structure *****
        String [] items;
        String line;
        String [] p_lines = extract_lines("Données.txt"); // Lecture de la matrice Lead-C
        ArrayList<Integer> IdCommandes = new ArrayList<Integer>();

        line=p_lines[0];
        items=line.split("\t");
        for (int i=1; i<items.length;i++){
            IdCommandes.add(Integer.parseInt(items[i]));
        }

        int i1=Integer.parseInt(items[0]);
        line=p_lines[p_lines.length-1];
        items=line.split("\t");
        int i2=Integer.parseInt(items[0]);
        int Nbjours=i2-i1+1;

        int [] [] matrice = new int [p_lines.length-1][IdCommandes.size()];

        for (int i=0; i<(p_lines.length-1);i++){
            line=p_lines[i+1];
            items=line.split("\t");
            for (int j=0; j<items.length-1;j++){
                matrice[i][j]=Integer.parseInt(items[j+1]);
            } //System.out.println();
        }
        ArrayList<Integer> A = new ArrayList<Integer>();
        ArrayList <Lead> Ld = new ArrayList<Lead>(); // va contenir les commandes satisfa

        for (int i=0; i<(p_lines.length-1);i++){
            for (int j=0; j<IdCommandes.size();j++){
                if (matrice[i][j]==1) { //System.out.print(IdCommandes.get(j)+"\t");
                    A.add(IdCommandes.get(j));
                }
            }
            A=Ordonner(A); // Ordonne l'id des commandes par ordre croissant
            Ld.add(new Lead(A));
            A.removeAll(A);
        }
    }
}

```

Figure B.2 — Extraction des données

B.3 Ordonnement des commandes

La méthode "Ordonner" (voir la figure B.3) permet de trier les commandes par rapport à leurs identifiants. En effet, afin de faciliter le traitement des commandes, nous les trions dans un ordre croissant en se basant sur leurs identifiants.

```
// Fonction Ordonner : Ordonne un ArrayList
static ArrayList<Integer> Ordonner (ArrayList<Integer> AL){
    int Tampon;
    for (int i=0; i<AL.size()-1;i++){
        for (int j=i+1; j<AL.size();j++){
            if (AL.get(i)>AL.get(j)){
                Tampon=AL.get(i);
                AL.set(i, AL.get(j));
                AL.set(j, Tampon);                }}}
    return AL; }
```

Figure B.3 — Ordonnement des commandes

B.4 Construction des groupes de commandes

La méthode définie ci-dessous (voir la figure B.4) permet de construire les ensembles de commandes partageant les mêmes prospects Internet. Ensuite, ces groupes de commandes sont stockés dans une structure nommée "Sous-Ens-Commandes".

Ainsi la figure B.4 implémente l'algorithme 5.1 qui calcule l'ensemble des séries chronologiques. En effet, l'algorithme 5.1 présenté dans la section 5.3.1 est mis en oeuvre par la figure B.4. Ainsi, nous calculons à la fois : (i) l'historique des groupes de commandes partageant les mêmes prospects Internet et (ii) pour chaque groupe de commandes, l'historique du nombre de prospects Internet satisfaisant simultanément toutes les commandes de ce groupe.

```
// ***** Cr ation des Sous ensembles de Commandes *****
ArrayList <Lead> Sous_Ens_Commandes = new ArrayList <Lead> ();
int [][] nbLead_Par_Commande = new int [Nbjours][p_lines.length-1];

// initialisation de la matrice
for (int i=0; i<Nbjours; i++){
    for (int j=0; j<p_lines.length-1;j++){
        nbLead_Par_Commande[i][j]=0;
    }
}

Lead L = Ld.get(0);
Sous_Ens_Commandes.add(L);
nbLead_Par_Commande[0][0]=1;

int k=1; int l=0; int item_1=i1;
for (int i=1; i<Ld.size(); i++){
    line=p_lines[i+1];
    items=line.split("\t");
    if ((Integer.parseInt(items[0]))!=item_1) { l=l+(Integer.parseInt(items[0])-item_1);
    L=Ld.get(i);
    for (int j=0; j<Sous_Ens_Commandes.size();j++){
        if (L.Equal(Sous_Ens_Commandes.get(j))==true){
            nbLead_Par_Commande[l][j]=nbLead_Par_Commande[l][j]+1;
            break;
        }
        k=j+1;
    }
    if (k==Sous_Ens_Commandes.size()) {
        nbLead_Par_Commande[l][Sous_Ens_Commandes.size()]=1;
        Sous_Ens_Commandes.add(L);
    }
    item_1=Integer.parseInt(items[0]);
}
}
```

Figure B.4 — Construction des groupes de commandes

B.5 Ordonnancement des groupes de commandes

En examinant l'historique des flux de prospects Internet pour les diff rents groupes de commandes, nous constatons que certains groupes de commandes poss dent un nombre tr s r duit de flux de prospects Internet. Ainsi et dans le but d'avoir de bons r sultats de pr vision, l' tude pr visionnelle concerne uniquement les groupes de commandes ayant re u un nombre

important de prospects Internet. Pour ce faire, nous ordonnons les groupes de commandes en fonction du nombre total de prospects Internet reçus sur tout l'historique pris en compte dans cette étude prévisionnelle (voir la figure B.5).

```

//***** Sous ensembles de Commandes par ordre décroissant en terme de leads recu *****
int [] Nbre = new int [Sous_Ens_Commandes.size()];
int S=0;
for (int i=0; i<Sous_Ens_Commandes.size(); i++){
    S=0;
    for (int j=0; j<Nbjours; j++){
        S+=nbLead_Par_Commande[j][i];
    }
    Nbre[i]=S;
}

int Tampon;
int [] Tchou = new int [Nbjours];

for (int i=0; i<Sous_Ens_Commandes.size()-1;i++){

    for (int j=i+1; j<Sous_Ens_Commandes.size();j++){

        if (Nbre[i]<Nbre[j]){
            Tampon=Nbre[i];
            Nbre[i]=Nbre[j];
            Nbre[j]=Tampon;

            L=Sous_Ens_Commandes.get(i);
            Sous_Ens_Commandes.set(i,Sous_Ens_Commandes.get(j));
            Sous_Ens_Commandes.set(j, L);

            for (int h=0; h<Nbjours; h++){
                Tchou[h]=nbLead_Par_Commande[h][i];
            }

            for (int h=0; h<Nbjours;h++){
                nbLead_Par_Commande[h][i]=nbLead_Par_Commande[h][j];
            }

            for (int h=0; h<Nbjours;h++){
                nbLead_Par_Commande[h][j]=Tchou[h];
            }

        }
    }
}

```

Figure B.5 — Ordonnancement des groupes de commandes

B.6 Affichage des groupes de commandes

Dans cette sections, nous présentons une méthode (voir la figure B.6) permettant d’afficher les groupes de commandes générés dans un fichier texte.

```
// ***** Affichage des sous ensembles de données *****

String Ecriture="";
for (int i=0; i<Sous_Ens_Commandes.size();i++){
    L=Sous_Ens_Commandes.get(i);
    Ecriture+="{"+L.Afficher()+"}"+"\t";
}
Ecriture=Ecriture.substring(0,Ecriture.length()-1);
ecrire("Résultats.txt", Ecriture+"\n");

for (int i=0; i<Nbjours;i++){
    Ecriture="";
    for (int j=0; j<Sous_Ens_Commandes.size();j++){
        Ecriture+=nbLead_Par_Commande[i][j]+" \t";
    }
    Ecriture=Ecriture.substring(0,Ecriture.length()-1);
    if (i<Nbjours-1) {ecrire("Résultats.txt", Ecriture+"\n");}
    else {ecrire("Résultats.txt", Ecriture);}
}
}
```

Figure B.6 — Affichage des groupes de commandes

B.7 Prévision sous le logiciel R

Après avoir déterminer les groupes de commandes concernés par la prévison (voir la section B.5), nous présentons dans cette section le programme qui permet de calculer la prévision des flux de prospects Internet pour chaque groupe de commandes partageant les mêmes prospects Internet. Ci-dessous (voir la figure B.7) le programme permettant d’effectuer cette prévision sous le logiciel R.

```

2 # Série chronologique des arrivées des leads par sous ensembles de commandes
3 data=scan(file="donnee.txt",skip=1)
4 y <- ts (data,frequency=7)
5
6 library(fpp) #Package contenant la fonction ETS
7
8 ETS <- function() {
9   ets0 <- ets(y, model="ZZZ", damped=NULL, alpha=NULL, beta=NULL,
10             gamma=NULL, phi=NULL, additive.only=FALSE, lambda=NULL,
11             lower=c(rep(0.0001,3), 0.8), upper=c(rep(0.9999,3),0.98),
12             opt.crit=c("lik","amse","mse","sigma","mae"), nmse=3,
13             bounds=c("both","usual","admissible"),
14             ic=c("aicc","aic","bic"), restrict=TRUE)
15   forecast(ets0, h=7, level=c(80,95), fan=FALSE, simulate=FALSE, bootstrap=FALSE)
16 }
17
18 ARIMA <- function() {
19   aaa <- auto.arima(y, stepwise=FALSE, approximation=FALSE)
20   forecast(aaa, h=7, level=c(80,95), fan=FALSE, simulate=FALSE, bootstrap=FALSE)
21 }
    
```

Figure B.7 — Prévision sous le logiciel R

B.8 Calcul des erreurs de prévision

Les erreurs de prévision sont calculées avec un programme *JAVA*. Ainsi, nous définissons une classe "WAPE" (voir la figure B.8) qui représente une structure permettant de définir les différentes erreurs de prévision, à savoir, $WAPE$, $WAPE^+$ et $WAPE^-$.

À partir de la classe présentée dans la figure B.8, nous définissons des instanciations des différentes fonctions *JAVA* permettant de calculer les valeurs des $WAPE$, $WAPE^+$ et $WAPE^-$ (voir la figure B.9).

```

public class WAPE {

    public static void main(String[] args) {
        // TODO Auto-generated method stub
        String [] items;
        String line;
        String [] p_lines = extract_lines("WAPE WAPE.txt");

        // Sauvegarde des réalisation et des Prévisions
        int [] Réalisation = new int [7];
        int [] Prévision = new int [7];

        for (int i=1; i<p_lines.length; i++){
            line = p_lines[i];
            items = line.split("\t");
            Réalisation[i-1] = Integer.parseInt(items[0]);
            Prévision[i-1] = Integer.parseInt(items[1]);
        }

        // Sauvegarde des variation des taux d'erreurs de prévisions +/- (10%, 20%, 30%)
        int[] PrévM30 = new int [7];
        int[] PrévM20 = new int [7];
        int[] PrévM10 = new int [7];
        int[] PrévP10 = new int [7];
        int[] PrévP20 = new int [7];
        int[] PrévP30 = new int [7];

        for (int i=0; i<7; i++){
            PrévM30[i] = (int)((0.7*Prévision[i])+0.5);
            PrévM20[i] = (int)((0.8*Prévision[i])+0.5);
            PrévM10[i] = (int)((0.9*Prévision[i])+0.5);
            PrévP10[i] = (int)((1.1*Prévision[i])+0.5);
            PrévP20[i] = (int)((1.2*Prévision[i])+0.5);
            PrévP30[i] = (int)((1.3*Prévision[i])+0.5);
        }

        // Sortie : les WAPE (+/-)
        System.out.println(WAPE(Réalisation, PrévM30)+"\t"+WAPE(Réalisation, PrévM20)+"\t".
        System.out.println(WAPE_Plus(Réalisation, PrévM30)+"\t"+WAPE_Plus(Réalisation, PrévM20)+"\t".
        System.out.println(WAPE_Minus(Réalisation, PrévM30)+"\t"+WAPE_Minus(Réalisation, PrévM20)+"\t".
    }
}

```

Figure B.8 — Calcul des erreurs de prévision

```

// Calcul du WAPE
static double WAPE(int[] Réa, int[] Prév){
    int [] abs = new int[7];
    int S=0;
    double Somme=0;
    for (int i=0; i<7; i++){
        abs[i]=Math.abs(Prév[i]-Réa[i]);
        S=S+Réa[i];
    }
    for (int i=0; i<7; i++) { Somme = Somme + abs[i]; }
    return Somme/S;
}

// Calcul du WAPE +
static double WAPE_Plus(int [] Réa, int[] Prév){
    int [] abs = new int[7];
    int S=0;
    double Somme=0;
    for (int i=0; i<7; i++){
        abs[i]=Prév[i]-Réa[i];
        if (abs[i]>=0) { S=S+Réa[i]; }
    }
    for (int i=0; i<7; i++) { if (abs[i]>=0) { Somme = Somme + abs[i]; } }
    if (S!=0) return Somme/S;
    else return 0;
}

// Calcul du WAPE-
static double WAPE_Minus(int [] Réa, int[] Prév){
    int [] abs = new int[7];
    int S=0;
    double Somme=0;
    for (int i=0; i<7; i++){
        abs[i]=Prév[i]-Réa[i];
        if (abs[i]<0) { S=S+Réa[i]; }
    }
    for (int i=0; i<7; i++) { if (abs[i]<0) { Somme = Somme + abs[i]; } }
    if (S!=0) return Somme/S;
    else return 0.0;
}

```

Figure B.9 — Méthodes de calcul des WAPE, WAPE⁺ et WAPE⁻