HAL open science

Classification des problèmes stochastiques et dynamiques de collectes et de livraisons par des véhicules intelligents

Guillaume Bono, Jilles Steeve Dibangoye, Laëtitia Matignon, Florian Pereyron, Olivier Simonin

To cite this version:

Guillaume Bono, Jilles Steeve Dibangoye, Laëtitia Matignon, Florian Pereyron, Olivier Simonin. Classification des problèmes stochastiques et dynamiques de collectes et de livraisons par des véhicules intelligents. Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes (JFPDA 2017), Jul 2017, Caen, France. hal-01576351

HAL Id: hal-01576351

https://hal.science/hal-01576351v1

Submitted on 23 Aug 2017

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.

Classification des problèmes stochastiques et dynamiques de collectes et de livraisons par des véhicules intelligents

Guillaume Bono[1,2], Jilles S. Dibangoye[1,2], Laëtitia Matignon[1,2,3], Florian Pereyron[4], Olivier Simonin[1,2]

[1] INSA de Lyon, laboratoire CITI
6 avenue des Arts, 69621 Villeurbanne
prenom.nom@insa-lyon.fr

[2] INRIA, équipe Chroma
prenom.nom@inria.fr

[3] Univ Lyon, Université Lyon 1, LIRIS, CNRS, UMR5205
Villeurbanne, F-69622, France
laetitia.matignon@univ-lyon1.fr

[4] Volvo Group, Advanced Technology and Research
99 route de Lyon, 69806 Saint-Priest
florian.pereyron@volvo.com

Résumé : Les problèmes de planification de tournées de véhicules présentent une très grande richesse et disposent de nombreux raffinements dans la littérature. Les progrès récents autour des véhicules autonomes ouvrent certaines perspectives quant à leur usage pour le transport de marchandises. Nous avons étudié la classification de cette famille de problèmes, et les modèles qui en découlent, pour tenter de nous positionner dans ce domaine avec cette nouvelle approche, intégrant l'utilisation d'une flotte de véhicules autonomes et intelligents.

Mots-clés : Recherche Opérationnelle, Systèmes Multi-Agents, Véhicules autonomes, Transport Intelligent Collaboratif, Optimisation dans l'incertain.

1 Introduction

Les métiers de la logistique et du transport routier sont aujourd'hui confrontés à la révolution numérique, et sont encore assez loin d'exploiter pleinement les récents progrès en informatique et télécommunication. Dans un avenir imminent, ils auront tout intérêt à intégrer une nouvelle « révolution » : celle du véhicule autonome. Les recherches gravitant autour de ces systèmes de transports intelligents aboutissent aujourd'hui de plus en plus à des expérimentations sur routes ouvertes et à des preuves de concepts d'exploitation réelle. Néanmoins, en France, elles visent surtout le transport de personnes, et ne se sont pas encore beaucoup tournées vers le transport de marchandises.

La littérature dans le domaine du transport logistique est extrêmement riche, grâce aux enjeux économiques importants auxquels elle tente d'apporter des solutions. En tant qu'extension du problème du voyageur de commerce, le problème de planification de tournées de véhicules, introduit dans (Dantzig & Ramser, 1959), est difficile. Il présente une complexité algorithmique non polynomiale (NP-hard), comme l'ont montré (Lenstra & Kan, 1981). De plus les instances réelles de ce problème rencontrées dans l'industrie du transport ont souvent des dimensions très importantes.

Il en existe de nombreux raffinements, dont le problème de collectes et de livraisons introduit dans (Wilson *et al.*, 1971) que nous décrirons plus en détails, et que nous formulerons en problème d'optimisation mathématique, dans la deuxième section. Puis nous explorerons quelques familles de méthodes de résolution appliquées dans la littérature, pour enfin conclure sur les perspectives futures de notre travail, potentiellement tourné vers de nouvelles approches d'optimisation et d'apprentissage, utilisant des systèmes multi-agents distribués.

2 Description du problème

2.1 Concepts

Le problème de collectes et de livraisons (ou PDP pour Pickup and Delivery Problem en anglais) est une extension du problème de tournées de véhicules (VRP), lui-même variante du problème du voyageur de commerce (TSP). Dans l'idéal, résoudre un PDP consiste à trouver le parcours optimal d'une flotte de m véhicules devant transporter un ensemble de n biens d'un point de collecte à un point de livraison à travers un réseau de routes.

Une partie des demandes de transport peut arriver en cours d'exécution, rendant le problème dynamique. La proportion de ces requêtes initialement inconnues constitue le degré de dynamisme (DoD) du problème, comme l'ont défini (Lund *et al.*, 1996). De plus, de nombreux paramètres incertains peuvent être pris en compte et anticipés lors de l'optimisation, comme la congestion du réseau routier, ou les lieux de collecte et de livraison des requêtes dynamiques. On obtient à ce stade un problème dynamique et stochastique de collectes et de livraisons (SDPDP).

À ceci s'ajoute la possibilité d'intégrer diverses contraintes et degrés de liberté, comme une capacité de charge limitée (cVRP) étudiée dans les travaux de (Ralphs *et al.*, 2003), et très souvent considérée de façon implicite dans d'autres instances du problème. Une autre contrainte classique est l'introduction de fenêtres temporelles, comme le fait (Solomon, 1987) restreignant le temps pendant lequel les véhicules peuvent procéder à une prise en charge ou une livraison (PDP-TW). On peut également s'intéresser au cas où l'autonomie énergétique des véhicules est limité (eVRP) comme (Artmeier *et al.*, 2010), ou encore à la possibilité de transférer des biens entre véhicules (PDP-T), traitée par (Cortés *et al.*, 2010). La figure 1 ci-dessous illustre les différents éléments intervenant dans la version étendue du problème.

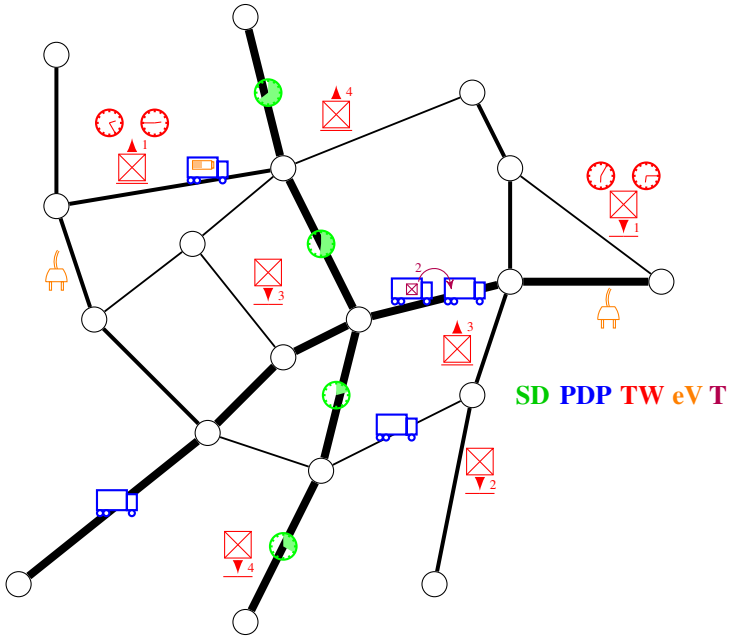

FIGURE 1 – Illustration du problème : 3 véhicules (en bleu) doivent effectuer 4 livraisons (en rouge) ayant des horaires de collecte et de livraison limités (cadrans rouges). Ils doivent tenir compte des temps de trajets variables sur chaque axe (cadrans verts), et de leur capacité énergétique limitée (pile orange), les forçant à se rendre régulièrement à une station de charge (prise orange).

Dans un contexte différent, le problème de transport à la demande (ou DARP, pour Dial-A-Ride Problem en anglais), qui s'applique au transport de passagers, est une forme de PDP où chaque personne n'occupe qu'un seul siège, et où certaines contraintes plus spécifiques sont prise en compte (temps de détour maximum par exemple). La Figure 2 ci-dessous détaille la construction de ces différents raffinements, et explicite les acronymes anglais.

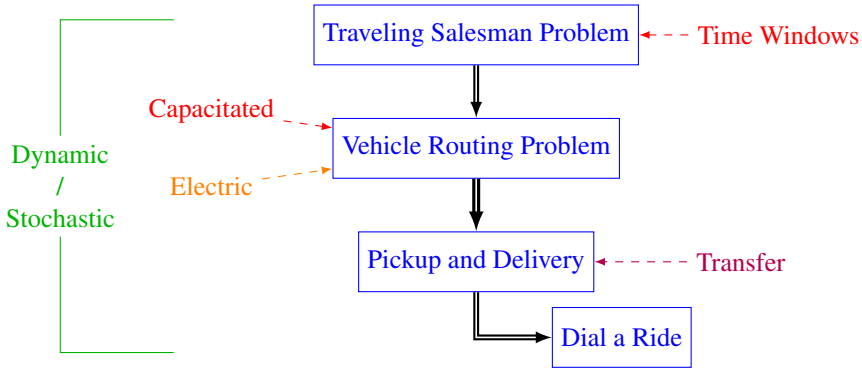

FIGURE 2 – Taxonomie des variantes du problème

2.2 Formulation mathématique

Dans cette section, nous reprendrons le formalisme utilisé dans leur chapitre de livre par (Desaulniers *et al.*, 2001). Nous essayerons de reconstruire de façon incrémentale le problème et ses raffinements pour mieux identifier les inéquations et les contraintes auxquelles elles correspondent. Nous tâcherons aussi de faire ressortir les variables de décision binaire x_{ijk}, définissant l'assignation des requêtes aux véhicules ainsi que l'enchaînement par lequel elles seront desservies, et les variables continues (par exemple L_{ik}, T_{ik}, E_{ik}) maintenant la cohérence de l'état du système d'un nœud à l'autre. Cette mise en valeur progressive des différentes contraintes vise à faciliter les parallèles avec d'autres formalisations. Nous montrerons également comment étendre le programme mathématique original des auteurs pour enrichir l'état du système en prenant pour exemple l'intégration de la contrainte d'autonomie énergétique limitée (eVRP).

2.2.1 VRP

Ramené au cas le plus simple du problème de planification de tournées où sont donnés l'ensemble des lieux à visiter $N = \{1, \ldots, n\}$ et les positions de départ et d'arrivée $o(k), d(k)$ de chaque véhicule $k \in K$, on construit un ensemble de $|K| = m$ graphes complets $\{G_k = (V_k, A_k)\}_{k \in K}$. Chaque nœud correspond à la visite d'un véhicule sur un lieu : $V_k = \{o(k), d(k)\} \cup N_k$, avec $N_k \subset N$ le sous-ensemble des lieux que le véhicule est autorisé à desservir. Chaque arc $(i, j) \in A_k$, pondéré par un coût c_{ijk}, représente le trajet par le véhicule $k \in K$ d'un lieu i à un lieu j. Outre la prise en compte des spécificités de chaque véhicule, cette modélisation en m graphes permet de visualiser plus simplement l'existence d'un arc par véhicule et par paire de nœud de service nécessaire à maintenir la cohérence de chaque route. L'optimisation repose alors sur le programme en nombres entiers (binaires) suivant :

$$\min_{x_{ijk}} \sum_{k \in K} \sum_{ij \in A_k} x_{ijk} c_{ijk} \tag{1}$$

$$\text{s.t.} \quad \sum_{k \in K} \sum_{j \in N_k \cup \{d(k)\}} x_{ijk} = 1 \quad \forall i \in N \tag{2}$$

$$\sum_{j \in N_k \cup \{d(k)\}} x_{o(k),j,k} = 1 \quad \forall k \in K \tag{3}$$

$$\sum_{j \in N_k \cup \{d(k)\}} x_{ijk} - \sum_{j \in N_k \cup \{o(k)\}} x_{jik} = 0 \quad \forall k \in K, \forall i \in N_k \tag{4}$$

$$\sum_{j \in N_k \cup \{o(k)\}} x_{j,d(k),k} = 1 \quad \forall k \in K \tag{5}$$

$$x_{ijk} \in \{0,1\} \quad \forall k \in K, \forall ij \in A_k$$

Les variables de décision binaires x_{ijk} indique si le véhicule $k \in K$ servira le nœud $j \in N_k \cup d(k)$ après avoir visiter le nœud $i \in N_k \cup o(k)$. La contrainte (2) garantie que chaque lieu est desservi une et une seule fois, en imposant qu'un seul arc sortant du nœud soit utilisé. Elle aurait pu s'écrire de façon équivalente en utilisant les arcs entrant dans le nœud. Les contraintes (3) à (5) assurent la cohérence des trajets par conservation du flux.

Le problème ainsi posé peut donner lieu à des résultats infaisables et absurdes, car rien n'empêche la solution de contenir des boucles. Il faudrait ajouter une variable entière pour chaque nœud, qui tienne compte de leur ordre de passage. Néanmoins dans la formulation complète de (Desaulniers *et al.*, 2001) que nous reconstruisons, ce rôle sera rempli par les variables permettant d'intégrer les fenêtres temporelles. Nous ne détaillerons donc pas d'avantage ici, pour ne pas surcharger les notations.

2.2.2 PDP

Pour étendre cette formulation au problème de collectes et de livraisons, l'ensemble N se décompose en un ensemble des lieux de collecte $P = \{1, \ldots, n\}$ et un ensemble de lieux de livraison correspondants $D = \{n+1, \ldots, 2n\}$. Les sous-ensembles $P_k \subset P$ et $D_k \subset D$ répondent aux mêmes limitations pour le véhicule $k \in K$ que $N_k \subset N$. Il faut ajouter la contrainte (6) qui assure que le même véhicule réalise la collecte et la livraison d'un bien donné :

$$\sum_{j \in N_k} x_{ijk} - \sum_{j \in N_k} x_{j,n+i,k} = 0 \quad \forall k \in K, \forall i \in P_k \tag{6}$$

Ici aussi, il faudrait maintenir un ordre de passage pour garantir que la livraison ait lieu après la collecte correspondante. Pour ne pas introduire de notations temporaires, nous résoudrons ce défaut dans la partie modélisant le respect de fenêtres temporelles.

2.2.3 cVRP

La prise en compte de la capacité maximale C_k de chaque véhicule $k \in K$ et de la charge l_i de chaque bien $i \in P$ (en posant pour des facilités de notations : $l_{n+i} = -l_i \quad \forall n+i \in D$) donne lieu aux contraintes suivantes :

$$x_{ijk} \cdot (L_{ik} + l_j - L_{jk}) = 0 \quad \forall k \in K, \forall ij \in A_k \tag{7}$$

$$l_i \leq L_{ik} \leq C_k \quad \forall k \in K, \forall i \in P_k \tag{8}$$

$$0 \leq L_{n+i,k} \leq C_k - l_i \quad \forall k \in K, \forall n+i \in D_k \tag{9}$$

$$L_{o(k),k} = 0 \quad \forall k \in K \tag{10}$$

Les variables d'état continues positives L_{ik} représentent la charge d'un véhicule $k \in K$ après avoir desservi le nœud $i \in N_k$. La contrainte (7) garantie une évolution cohérente de la charge, les contraintes (8) et (9) maintiennent les conditions limites, et la contrainte (10) donne leur valeur initiale aux variables.

2.2.4 PDP-TW

La restriction des collectes et des livraisons à des fenêtres temporelles strictes notées $[a_i, b_i] \quad \forall i \in N$, accompagnées d'un temps de service (charge ou décharge) s_i, requiert d'estimer le temps de trajet t_{ijk} pour chaque arc. Le départ et l'arrivée des véhicules peuvent aussi être munis d'une fenêtre temporelle pour limiter leur temps d'utilisation. Les nouvelles contraintes ci-dessous interviennent :

$$x_{ijk} \cdot (T_{ik} + s_i + t_{ijk} - T_{jk}) = 0 \quad \forall k \in K, \forall ij \in A_k \tag{11}$$

$$a_i \leq T_{ik} \leq b_i \quad \forall k \in K, \forall i \in V_k \tag{12}$$

$$T_{ik} + t_{i,n+i,k} \leq T_{n+i,k} \quad \forall k \in K, \forall i \in P_k \tag{13}$$

Les variables d'état continues positives T_{ik} représentent le temps de début de service du véhicule $k \in K$ au nœud $i \in V_k$. La contrainte (11) garantie une évolution cohérente de ces temps de passage, la contrainte (12) assure le respect des fenêtres temporelles, et la contrainte (13) empêche les livraisons d'être réalisées avant les collectes correspondantes.

2.2.5 eVRP

Des deux sous-sections précédentes se dégage une forme commune permettant d'intégrer diverses variables d'état et de maintenir leur cohérence par rapport aux décisions prises. On peut ainsi par exemple faire intervenir la notion d'autonomie énergétique limitée de véhicules électriques.

Pour cela, notons B_k la capacité des batteries de chaque véhicule $k \in K$, et attribuons à chaque arc $ij \in A_k$ un coût en énergie w_{ijk}. Un ensemble S de nouveaux nœuds représentant les stations de charge est intégré à tous les graphes G_k. Ces stations ne sont pas soumises à la contrainte de visite unique imposée pour les autres nœuds par l'équation (2). De plus, dans la contrainte (11), leurs visites ont un temps de service s_i variable, fonction du niveau d'énergie du véhicule : $\forall i \in S, s_i = (B_k - E_{ik})/\rho_i$, avec ρ_i la puissance de charge disponible. Chaque nœud est aussi affublé d'une variation d'énergie $q_i \leq 0$ négative ou nulle, sauf pour les stations, où cette variation est positive et vaut $q_i = (B_k - E_{ik})$, si on se limite au cas où le plan ne peut pas faire intervenir des rechargements incomplets.

$$x_{ijk} \cdot (E_{ik} + q_i - w_{ijk} - E_{jk}) = 0 \quad \forall k \in K, \forall ij \in A_k \tag{14}$$

$$|q_i| \leq E_{ik} \leq B_k \quad \forall k \in K, \forall i \in V_k \setminus S \tag{15}$$

$$E_{o(k),k} = B_k \quad \forall k \in K \tag{16}$$

Les variables d'état continues positives E_{ik} représentent le niveau d'énergie du véhicule $k \in K$ lorsqu'il atteint le nœud $i \in V_k$. La contrainte (14) garantie une évolution cohérente du niveau d'énergie, la contrainte (15) impose que le véhicule ait suffisamment d'énergie avant de servir un nœud autre qu'une station, et la contrainte (16) donne leur valeur initiale aux variables.

2.2.6 DPDP

Ajouter une dimension dynamique au problème signifie que l'ensemble des données d'entrée est amené à évoluer lors de l'exécution du plan. Dans le cas du problème de collectes et de livraisons, le principal degré de dynamisme pouvant être intégré au problème est le taux de requêtes inconnues en début de planification. L'ensemble des nœuds n'est donc pas connu dans son intégralité avant le début de la tournée de chaque véhicule.

Dans un cadre dynamique, il est difficile de définir l'optimalité d'une solution. En effet, le plan construit de façon incrémentale avec des requêtes découvertes pendant l'exécution a de fortes chances d'être sous optimal par rapport à celui qui aurait été élaboré en ayant connaissance de toutes les requêtes dès le départ.

Ce dynamisme met aussi d'autant plus en avant le fait qu'il n'existe peut-être pas de solution servant toutes les requêtes et respectant l'ensemble des contraintes. Il est alors utile d'introduire une mesure de qualité du service, pouvant par exemple prendre en compte : le taux de requêtes desservies (garantie de service), le retard sur les fenêtres temporelles, le temps de détour maximum par requête, l'écart à l'optimum (regret) par rapport au cas statique où tout est fixé à l'avance.

Pour permettre un service incomplet, la contrainte (2) doit être relâchée, en introduisant des variables d'écart ξ_i. De façon similaire, la contrainte (12) peut être relâchée par l'ajout de variables ζ_i pour permettre des retards. La fonction objectif est modifiée pour décourager l'utilisation de ces variables d'écart, avec deux nouveaux meta-paramètres $\alpha, \beta \in \mathbb{R}$ permettant d'équilibrer les coûts.

$$\min_{x_{ijk}} \sum_{k \in K} \sum_{ij \in A_k} x_{ijk} c_{ijk} + \alpha \sum_{i \in P} \xi_i + \beta \sum_{i \in N} |\zeta_i| \tag{17}$$

$$\text{s.t.} \quad \sum_{k \in K} \sum_{j \in N_k \cup \{d(k)\}} x_{ijk} + \xi_i = 1 \quad \forall i \in P \tag{18}$$

$$a_i \leq T_{ik} + \zeta_i \leq b_i \quad \forall k \in K, \forall i \in V_k \tag{19}$$

$$\xi_i \in \{0, 1\} \quad \forall i \in P$$

$$\zeta_i \in \mathbb{R} \quad \forall i \in N$$

2.2.7 SPDP

Le caractère stochastique du problème peut être introduit en considérant plusieurs sources d'incertitude :

— Les temps de trajets t_{ijk} dépendant de l'état de congestion du réseau routier et d'événements difficiles à modéliser de façon déterministe,

— La consommation d'énergie des véhicules w_{ijk} pouvant elle aussi s'éloigner fortement des modèles déterministes les plus simples,

— Les coûts c_{ijk}, liés aux deux points précédents,

— Les pannes sur des véhicules, pouvant invalider une partie du plan,

— La position des lieux de collectes et de livraisons, dont la forme de la distribution dépend de l'instance du problème considérée (potentiellement uniforme pour du transport de personnes, et peut-être plus ponctuelle pour du transport de marchandises...)

Bien que le problème puisse alors s'exprimer comme un programme stochastique (SP), comme dans les travaux de (Laporte & Louveaux, 1993), beaucoup d'auteurs préfèrent reformuler le problème sous la forme d'un Processus de Décision Markovien, comme (Simão *et al.*, 2009) ou (Cortés *et al.*, 2009). Dans ce modèle, un vecteur $s \in \mathcal{S}$ décrit l'état du système, à partir duquel une politique $\pi : \mathcal{S} \times \mathcal{A} \to [0, 1]$ indique les probabilités de choisir chaque action $a \in \mathcal{A}$. Une fonction de transition $f : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$ donne une distribution de probabilité sur l'état suivant du système s', tandis qu'une fonction de récompense $r : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to \mathbb{R}$ fournit la récompense associée à cette transition. L'optimisation de la politique est alors guidée par la maximisation de la récompense amortie, accumulée sur une suite de transitions constituant des épisodes. Les concepts d'apprentissage par renforcement utilisés pour optimiser la politique π dans notre contexte, comme l'approximation de la fonction de valeur, sont détaillés dans le livre de (Sutton & Barto, 1998).

2.2.8 PDP-T

Permettre les transferts de biens d'un véhicule à l'autre donne un degré de liberté supplémentaire au problème, et complexifie énormément l'optimisation. Deux cas peuvent être identifiés : celui où les lieux d'échange sont fixés à l'avance, et celui où les transferts peuvent avoir lieu n'importe où.

C'est dans cet espace d'actions plus large que (Coltin & Veloso, 2014) tente d'explorer la possibilité de transfert, en se limitant néanmoins dans leur expérimentation à identifier les intersections existantes entre les routes individuelles de chaque robot à l'itération courante, et à explorer de façon exhaustive tous les transferts pouvant avoir lieu à ce point, sans autoriser de détours.

Pour revenir à notre modèle, on peut imaginer dans le premier cas ajouter ces lieux de transfert connus aux nœuds des graphes, avec un nœud par bien à transporter et par véhicule, et ajouter des contraintes s'apparentant à celles déjà introduites pour ne pas transférer un bien avant de le collecter, pour respecter la capacité des véhicules et la cohérence temporelle.

Dans le deuxième cas, il est assez intuitif de visualiser un transfert comme la création d'un nouveau nœud de livraison intermédiaire pour le véhicule transportant le bien, et la création d'un nouveau nœud de collecte pour tous les autres véhicules de la flotte. Les possibilités de créer un nœud de transfert sont innombrables, cette décision est donc souvent guidée par des heuristiques.

3 Approches de résolution existantes

De nombreux passages en revue des différentes approches aux problèmes de collectes et de livraisons existent dans la littérature, comme celle mené récemment par (Ritzinger *et al.*, 2016), qui se focalise sur les méthodes intégrants les dimensions dynamiques et stochastiques du problème. Le Table 1 ci-dessous synthétise et étend ce passage en revue à d'autres exemples extrait de la littérature, et les algorithmes qui ont été appliqués, dont le nom sera donné en anglais pour ne pas perdre en spécificité.

	Traveling Salesman Problem	Vehicle Routing Problem	Pickup and Delivery Problem	⋯ with capacitated vehicles	⋯ with electric vehicles	⋯ with Time Window	⋯ with Transfers	Dynamic replanning	Stochastic demand	Stochastic customers	Stochastic travel times
	Problem flavors							*Var. properties*			
Branch-and-bound / cut		1,22		1,22	22	22					
Chance Constrained Programming		2,3		2,3					2		3
Stochastic Programming with recourse		3,4							4		3
Approximate Linear Programming	5								5		5
Approximate Dynamic Programming		8,9	6,7		(7)	6,7	7	9	8,9	6,7	6,7
Variable Neighboorhood Search		21	10	10,21	21	10,21					
Stochastic VNS		21	12	12		12		12		12	12
Multiple Plan Approach			11,12	11,12		11,12		11,12			
Multiple Scenario Approach			11,12	11,12		11,12		11,12		11,12	12
Genetic Algorithm		13	14	13,14		14		13,14		14	14
Particle Swarm Optimization			15,16	15,16		15		15	15		
Auction-based algorithm			17,18	17,18		17,18	18	17,18			
Monte-Carlo Tree Search	20	19		19		20					

TABLE 1 – Exemple de méthodes appliquées aux différentes variantes du problème

1	(Fisher, 1994)	9	(Novoa & Storer, 2009)	17	(Mes *et al.*, 2013)
2	(Dror *et al.*, 1993)	10	(Parragh *et al.*, 2010)	18	(Coltin & Veloso, 2014)
3	(Laporte *et al.*, 1992)	11	(Bent & Hentenryck, 2004)	19	(Mańdziuk & Nejman, 2015)
4	(Laporte & Louveaux, 1993)	12	(Schilde *et al.*, 2011)	20	(Rimmel *et al.*, 2011)
5	(Toriello *et al.*, 2014)	13	(Marinakis & Marinaki, 2010)	21	(Schneider *et al.*, 2014)
6	(Simão *et al.*, 2009)	14	(Sáez *et al.*, 2008)	22	(Desaulniers *et al.*, 2016)
7	(Bouzaiene-Ayari *et al.*, 2016)	15	(Cortés *et al.*, 2009)		
8	(Meisel *et al.*, 2011)	16	(Chen *et al.*, 2016)		

3.1 Recherche par voisinage

La méthode de recherches par voisinage (VNS) introduite par (Mladenović & Hansen, 1997) semble être une référence de pointe dans la littérature traitant de problèmes dynamiques. (Parragh *et al.*, 2010) applique cette méthode générale au problème du transport à la demande (DARP). À partir d'une solution initiale, la méthode repose sur diverses heuristiques d'échanges entre véhicules, d'insertions et de décalages de tâches dans leurs plans. Elles servent à construire des solutions « voisines » dont on conserve la plus prometteuse lors d'une recherche locale, avant d'explorer de nouveaux voisinages pour échapper aux minima locaux.

La variante stochastique (SVNS) de (Gutjahr *et al.*, 2007) utilise un modèle génératif pour estimer la valeur accumulée par une solution sur un court horizon. Elle a été adaptée au problème de transport à la demande par (Schilde *et al.*, 2011). Introduites par (Bent & Hentenryck, 2004), l'approche par plans multiples (MPA), et son équivalent stochastique l'approche par scénarios multiples (MSA), appliquent la recherche par voisinage à un ensemble de solutions qui évoluent au déclenchement d'événements comme l'arrivée d'une nouvelle requête, ou le départ d'un véhicule. Une fonction de consensus permet de générer un plan à exécuter en ligne à partir de l'ensemble de solutions.

(Schilde *et al.*, 2014) explorent ces différentes variantes (VNS, SVNS, MPA, MSA) combiné à un algorithme de planification par blocs et mettent en avant l'intérêt de prendre en compte les phénomènes dynamiques et d'anticiper les éléments stochastiques lors de l'optimisation.

3.2 Programmation dynamique approchée

La programmation dynamique approchée telle qu'elle est introduite par (Powell, 2007) présente le problème sous la forme d'un ensemble de ressources (dans notre problème, la flotte de véhicules) et de tâches (les requêtes de transport) dotées d'un vecteur d'attributs. L'état du système est un immense vecteur dont chaque élément correspond au nombre de ressources ayant une combinaison particulière d'attributs. La prise de décision consiste à affecter les tâches aux ressources. L'évolution du système intervient en deux temps, faisant d'abord évoluer les comptes des différents type de ressources suite à la décision, puis perturbant ce nouvel état par les phénomènes stochastiques issus de l'environnement. L'optimisation d'une politique d'assignation passe par l'estimation de la fonction de valeur par les méthodes classiques de l'apprentissage par renforcement comme l'approximation de la fonction de valeur ou les traces d'éligibilités (Sutton & Barto, 1998).

(Simão *et al.*, 2009) appliquent ce modèle au problème de la planification de tournée de flotte au sein d'une entreprise de transport américaine, donnant au final un véritable système d'aide à la décision utilisé dans l'industrie. Après avoir entrainé le modèle à reproduire les résultats historiques de l'entreprise grâce à de nombreuses traces enregistrées, ils démontrent l'intérêt de leur modèle en proposant des méthodes d'estimation de gains suite à un changement de composition de flotte, et une estimation de la valeur marginale de chaque type de ressources disponible. Le modèle passe à l'échelle de ce problème aux très larges dimensions grâce à un mécanisme d'agrégation réduisant efficacement la taille du vecteur d'état.

(Bouzaiene-Ayari *et al.*, 2016) proposent une autre mise en place de ce modèle pour aider à la planification de l'assignation de locomotives aux différents trajets à effectuer, en prenant en compte des contraintes très serrées concernant leur maintenance et la disponibilité réduite des ateliers capables de l'effectuer. Le système final est constitué de plusieurs algorithmes de résolutions travaillant à différentes échelles d'agrégation. Leurs résultats montrent l'influence des différents paramètres d'optimisation, et de fortes réductions des temps d'inactivité et des files d'attente aux ateliers.

L'utilisation de la programmation dynamique approchée et des techniques d'apprentissage par renforcement citées ci-dessus n'est bien sûr pas limité à cette formalisation en ressources et tâches issue des travaux de (Powell, 2007).

3.3 Recherche arborescente Monte-Carlo

Rencontrant beaucoup de succès dans le domaine des jeux (échecs, go...), cette méthode de recherche « meilleur d'abord » consiste à construire progressivement un arbre de décision en explorant en priorité les branches les plus prometteuses. Une politique d'arbre, comme celle utilisée dans l'algorithme UCT (pour Upper-Confidence bound 1 applied to Tree) présentée par (Kocsis & Szepesvári, 2006), dicte le choix d'actions jusqu'à un nœud feuille, à partir duquel une nouvelle action est explorée, créant un nouveau nœud en estimant l'état du système par un modèle génératif. Cet état est évalué en procédant à plusieurs simulations choisissant les actions par une politique plus simple (au hasard, par exemple) jusqu'à atteindre un état terminal. Autant de simulations que possible sont effectuées dans le temps de calcul imparti et la valeur obtenue est rétro-propagée jusqu'au nœud racine.

(Couetoux, 2013) présente dans sa thèse différentes extensions de la recherche arborescente classique permettant de travailler sur des espaces d'action continus et d'exploiter les valeurs obtenues dans des branches parallèles en comparant les états et actions en utilisant des fonctions noyaux. Il applique entre autre ces méthodes au problème de planification de tournées de véhicules avec fenêtres temporelles.

L'utilisation de cette méthode dans le cadre du transport logistique semble plutôt focalisé sur des problèmes statiques. Néanmoins l'engouement récent dont elle bénéficie et les extensions existantes pour passer à des états et des actions continus nous permettrons peut-être d'explorer cette approche dans des cas plus complexes. D'autant plus que dans une architecture de décision distribuée, des objectifs collaboratifs mais aussi compétitifs entre agents peuvent être identifiés, nous rapprochant d'un modèle de jeu stochastique pour lequel MCTS s'est illustré avec succès.

Les méthodes par recherche arborescente peuvent facilement bénéficier d'heuristiques et d'apport de connaissances expertes pour guider le choix d'actions à explorer dans les différentes phases de l'algorithme. Il n'est également pas nécessaire de poursuivre les simulations jusqu'à un état final, en utilisant des estimateurs de la fonction de valeur pour donner un score approximatif à l'état atteint.

4 Conclusion et discussions

Le problème de collectes et de livraisons dans sa forme dynamique et stochastique présente un grand nombre de défis à relever. Les grandes dimensions des instances du problème habituellement rencontrées et les contraintes temporelles de la prise de décision en ligne ont poussé les recherches en Recherche Opérationnelle dans la direction d'heuristiques efficaces et rapides, comme celles utilisées dans le VNS appliqué au DARP par (Parragh *et al.*, 2010). Nous souhaitons poursuivre les travaux visant une prise de décision en ligne exploitant l'information disponible dans un environnement dynamique et incertain. De plus, inspiré par les travaux récents autour des systèmes de transports intelligents collaboratifs (cITS), nous envisageons la possibilité de distribuer la prise de décision sur des agents autonomes.

Cette architecture distribuée prend d'autant plus de sens quand les agents peuvent prendre certaines décisions de façon indépendante. Cela pourrait être le cas dans le contexte de problèmes de planification de tournées si l'on suppose possible d'obtenir une « bonne » assignation initiale des requêtes aux différents véhicules. Chaque véhicule pourrait alors planifier sa tournée indépendamment avec son sous-ensemble de requêtes sauf lors de négociations ponctuelles avec les autres véhicules pour l'assignation de nouvelles requêtes, la détection de pannes ou de plans invalidés, ou encore le transfert de biens entre véhicules.

Une piste intéressante dans la direction d'un apprentissage par renforcement distribué au sein d'un système multi-agents relativement indépendants a été soulevées par (Melo & Veloso, 2009). En utilisant une action « de synchronisation » particulière permettant de se coordonner ponctuellement, ils parviennent à limiter l'empreinte mémoire de l'approximation de la fonction de Q-valeurs. D'autres pistes sont orientées vers des algorithmes d'optimisation par enchères, comme dans (Coltin & Veloso, 2014) ou (Mes *et al.*, 2013).

Cette article prospectif ouvre la voie dans la suite de nos travaux à l'application de ce genre de méthodes distribuées d'optimisation et d'apprentissage par renforcement. Elles ont fait leurs preuves dans des cadres dynamiques et stochastiques dans la littérature des transports intelligents et de la robotique. Nous souhaitons les appliquer à notre problème de planification de tournées, plutôt lié à la recherche opérationnelle, en intégrant un maximum de contraintes, dont les fenêtres temporelles et l'autonomie énergétique limitée, dans un environnement dynamique et incertain.

Références

ARTMEIER A., HASELMAYR J., LEUCKER M. & SACHENBACHER M. (2010). *The Shortest Path Problem Revisited : Optimal Routing for Electric Vehicles*, In R. DILLMANN, J. BEYERER, U. D. HANEBECK & T. SCHULTZ, Eds., *KI 2010 : Advances in Artificial Intelligence : 33rd Annual German Conference on AI, Karlsruhe, Germany, September 21-24, 2010. Proceedings*, p. 309–316. Springer Berlin Heidelberg : Berlin, Heidelberg.

BENT R. W. & HENTENRYCK P. V. (2004). Scenario-based planning for partially dynamic vehicle routing with stochastic customers. *Operations Research*, **52**(6), 977–987.

BOUZAIENE-AYARI B., CHENG C., DAS S., FIORILLO R. & POWELL W. B. (2016). From single commodity to multiattribute models for locomotive optimization : A comparison of optimal integer programming and approximate dynamic programming. *Transportation Science*, **50**(2), 366–389.

CHEN M.-C., HSIAO Y.-H., REDDY R. H. & TIWARI M. K. (2016). The self-learning particle swarm optimization approach for routing pickup and delivery of multiple products with material handling in multiple cross-docks. *Transportation Research Part E : Logistics and Transportation Review*, **91**, 208 – 226.

COLTIN B. & VELOSO M. (2014). Online pickup and delivery planning with transfers for mobile robots. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, p. 5786–5791.

CORTÉS C. E., MATAMALA M. & CONTARDO C. (2010). The pickup and delivery problem with transfers : Formulation and a branch-and-cut solution method. *European Journal of Operational Research*, **200**(3), 711 – 724.

CORTÉS C. E., SÁEZ D., NÚÑEZ A. & MUÑOZ-CARPINTERO D. (2009). Hybrid adaptive predictive control for a dynamic pickup and delivery problem. *Transportation Science*, **43**(1), 27–42.

COUETOUX A. (2013). *Monte Carlo Tree Search for Continuous and Stochastic Sequential Decision Making Problems*. Theses, Université Paris Sud - Paris XI.

DANTZIG G. B. & RAMSER J. H. (1959). The truck dispatching problem. *Management Science*, **6**(1), 80–91.

DESAULNIERS G., DESROSIERS J., ERDMANN A., SOLOMON M. M. & SOUMIS F. (2001). The vehicle routing problem. chapter VRP with Pickup and Delivery, p. 225–242. Philadelphia, PA, USA : Society for Industrial and Applied Mathematics.

DESAULNIERS G., ERRICO F., IRNICH S. & SCHNEIDER M. (2016). Exact algorithms for electric vehicle-routing problems with time windows. *Operations Research*, **64**(6), 1388–1405.

DROR M., LAPORTE G. & LOUVEAUX F. V. (1993). Vehicle routing with stochastic demands and restricted failures. *Zeitschrift für Operations Research*, **37**(3), 273–283.

FISHER M. L. (1994). Optimal solution of vehicle routing problems using minimum k-trees. *Operations Research*, **42**(4), 626–642.

GUTJAHR W. J., KATZENSTEINER S. & REITER P. (2007). *A VNS Algorithm for Noisy Problems and Its Application to Project Portfolio Analysis*, In J. HROMKOVIČ, R. KRÁLOVIČ, M. NUNKESSER & P. WIDMAYER, Eds., *Stochastic Algorithms : Foundations and Applications : 4th International Symposium, SAGA 2007, Zurich, Switzerland, September 13-14, 2007. Proceedings*, p. 93–104. Springer Berlin Heidelberg : Berlin, Heidelberg.

KOCSIS L. & SZEPESVÁRI C. (2006). Bandit based monte-carlo planning. In *European conference on machine learning*, p. 282–293 : Springer.

LAPORTE G., LOUVEAUX F. & MERCURE H. (1992). The vehicle routing problem with stochastic travel times. *Transportation Science*, **26**(3), 161–170.

LAPORTE G. & LOUVEAUX F. V. (1993). The integer l-shaped method for stochastic integer programs with complete recourse. *Operations Research Letters*, **13**(3), 133 – 142.

LENSTRA J. K. & KAN A. H. G. R. (1981). Complexity of vehicle routing and scheduling problems. *Networks*, **11**(2), 221–227.

LUND K., MADSEN O., RYGAARD J. & OF DENMARK. INSTITUTE OF MATHEMATICAL MODELLING T. U. (1996). *Vehicle Routing Problems with Varying Degrees of Dynamism*. IMM-REP. Technical Univ.

MAŃDZIUK J. & NEJMAN C. (2015). *UCT-Based Approach to Capacitated Vehicle Routing Problem*, In L. RUTKOWSKI, M. KORYTKOWSKI, R. SCHERER, R. TADEUSIEWICZ, L. A. ZADEH & J. M. ZURADA, Eds., *Artificial Intelligence and Soft Computing : 14th International Conference, ICAISC 2015, Zakopane, Poland, June 14-18, 2015, Proceedings, Part II*, p. 679–690. Springer International Publishing : Cham.

MARINAKIS Y. & MARINAKI M. (2010). A hybrid genetic – particle swarm optimization algorithm for the vehicle routing problem. *Expert Systems with Applications*, **37**(2), 1446 – 1455.

MEISEL S., SUPPA U. & MATTFELD D. (2011). *Serving Multiple Urban Areas with Stochastic Customer Requests*, In H.-J. KREOWSKI, B. SCHOLZ-REITER & K.-D. THOBEN, Eds., *Dynamics in Logistics : Second International Conference, LDIC 2009, Bremen, Germany, August 2009, Proceedings*, p. 59–68. Springer Berlin Heidelberg : Berlin, Heidelberg.

MELO F. S. & VELOSO M. (2009). Learning of coordination : Exploiting sparse interactions in multiagent systems. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, p. 773–780 : International Foundation for Autonomous Agents and Multiagent Systems.

MES M., VAN DER HEIJDEN M. & SCHUUR P. (2013). Interaction between intelligent agent strategies for real-time transportation planning. *Central European Journal of Operations Research*, **21**(2), 337–358.

MLADENOVIĆ N. & HANSEN P. (1997). Variable neighborhood search. *Computers and Operations Research*, **24**(11), 1097 – 1100.

NOVOA C. & STORER R. (2009). An approximate dynamic programming approach for the vehicle routing problem with stochastic demands. *European Journal of Operational Research*, **196**(2), 509 – 515.

PARRAGH S. N., DOERNER K. F. & HARTL R. F. (2010). Variable neighborhood search for the dial-a-ride problem. *Computers and Operations Research*, **37**(6), 1129 – 1138.

POWELL W. B. (2007). *Approximate Dynamic Programming : Solving the curses of dimensionality*, volume 703. John Wiley & Sons.

RALPHS T., KOPMAN L., PULLEYBLANK W. & TROTTER L. (2003). On the capacitated vehicle routing problem. *Mathematical Programming*, **94**(2), 343–359.

RIMMEL A., TEYTAUD F. & CAZENAVE T. (2011). Optimization of the Nested Monte-Carlo Algorithm on the Traveling Salesman Problem with Time Windows. In *Evostar*, Turin, Italy.

RITZINGER U., PUCHINGER J. & HARTL R. F. (2016). A survey on dynamic and stochastic vehicle routing problems. *International Journal of Production Research*, **54**(1), 215–231.

SCHILDE M., DOERNER K. & HARTL R. (2011). Metaheuristics for the dynamic stochastic dial-a-ride problem with expected return transports. *Computers and Operations Research*, **38**(12), 1719 – 1730.

SCHILDE M., DOERNER K. & HARTL R. (2014). Integrating stochastic time-dependent travel speed in solution methods for the dynamic dial-a-ride problem. *European Journal of Operational Research*, **238**(1), 18 – 30.

SCHNEIDER M., STENGER A. & GOEKE D. (2014). The electric vehicle-routing problem with time windows and recharging stations. *Transportation Science*, **48**(4), 500–520.

SIMÃO H. P., DAY J., GEORGE A. P., GIFFORD T., NIENOW J. & POWELL W. B. (2009). An approximate dynamic programming algorithm for large-scale fleet management : A case application. *Transportation Science*, **43**(2), 178–197.

SOLOMON M. M. (1987). Algorithms for the vehicle routing and scheduling problems with time window constraints. *Operations Research*, **35**(2), 254–265.

SUTTON R. & BARTO A. (1998). *Reinforcement Learning : An Introduction*. A Bradford book. Bradford Book.

SÁEZ D., CORTÉS C. E. & NÚÑEZ A. (2008). Hybrid adaptive predictive control for the multi-vehicle dynamic pick-up and delivery problem based on genetic algorithms and fuzzy clustering. *Computers and Operations Research*, **35**(11), 3412 – 3438. Part Special Issue : Topics in Real-time Supply Chain Management.

TORIELLO A., HASKELL W. B. & POREMBA M. (2014). A dynamic traveling salesman problem with stochastic arc costs. *Operations Research*, **62**(5), 1107–1125.

WILSON N. H. M., SUSSMAN J. M. & WONG H.-K. (1971). *Scheduling algorithms for a dial-a-ride system / by Nigel H.M. Wilson ... [et al.]*. MIT, Urban Systems Laboratory Cambridge, Mass.