Optimisation basée sur les données d'un problème de localisation d'entrepôts et tournées de véhicules

Guillaume Massonnet[1], Juliette Medina[2]

[1] LS2N, IMT Atlantique, Nantes, France
guillaume.massonnet@imt-atlantique.fr
[2] CRC Services, Nantes, France
j.medina@crc-services.com

Mots-clés : *Optimisation stochastique et robuste, location and routing problem*

1 Motivation et description du problème

Les chaînes d'approvisionnement de produits frais fonctionnent généralement en flux tendu afin d'être en mesure de livrer leurs clients dans des délais très courts, même dans le cas où ces derniers passe leur commande seulement quelques heures à l'avance. Pour cela, les fournisseurs doivent anticiper une organisation de transport aux niveaux stratégique et tactique efficace et réactive, capable d'absorber les variations induites par cette incertitude. Ce travail est motivé par le cas d'un industriel français de l'agro-alimentaire qui sous-traite à des transporteurs externes l'expédition quotidienne de différents types de produits depuis ses usines vers les entrepôts de la grande distribution. L'objectif est de proposer une affectation pré-établie des commandes de produits aux usines et un plan de transport avec des tournées fixes, répétées chaque jour (en général vers deux à trois clients), afin de garantir un niveau de service minimum. En particulier, la solution définie par le fournisseur doit permettre de satisfaire la demande des clients avec une forte probabilité, tout en minimisant les coûts d'ouverture des entrepôts et des tournées de véhicules.

Nous désignons ce problème sous le nom de *Multi Products Capacitated Location Routing Problem (MCLRP)* avec des demandes stochastiques , qui généralise le problème de localisation et de tournées de véhicules (voir [1] pour une liste plus détaillée de références). Nous modélisons la cible du fournisseur en terme de niveau de service comme une probabilité prédéfinie de satisfaire l'ensemble de contraintes capacitaires. En pratique, les données historiques ne donnent accès qu'à une estimation empirique de la véritable distribution sous-jacente des demandes aléatoires. Cette limitation impacte les performances d'approches standards telles que le Sample Average Approximation (SAA), car elle dégrade notamment l'estimation de probabilité de violation des contraintes nécessaire pour identifier des solutions robustes. Pour palier ce problème, nous utilisons une formulation robuste en distribution qui exploite des ensembles d'ambiguité basée sur la distance de Wasserstein (voir [2, 3]), qui améliore la cohérence des résultats même lorsque la taille de l'échantillon de données disponible est limitée. Nous évaluons son efficacité sur des cas pratiques dérivés de données réelles.

2 Formulation mathématique et méthode de résolution

Soit $\mathcal{I} = \{1, \ldots, I\}$ l'ensemble des usines disponibles, $\mathcal{K} = \{1, \ldots, K\}$ l'ensemble des produits et $\mathcal{J} = \{1, \ldots, J\}$ l'ensemble des clients. Pour tout client $j \in \mathcal{J}$, on note $\mathcal{K}_j$ l'ensemble des produits qui peuvent être commandés par j. L'ensemble des demandes des clients pour chaque produit est modélisé par un vecteur stochastique $\tilde{\boldsymbol{\xi}} = (\tilde{\xi}_{11}, \ldots, \tilde{\xi}_{1K}, \ldots, \tilde{\xi}_{J1}, \ldots, \tilde{\xi}_{JK})$ dont la distribution jointe $\mathbb{P}$ est inconnue. On suppose néanmoins disposer d'un historique des flux commandés sur N jours passés $\hat{\boldsymbol{\xi}}^{(1)}, \ldots, \hat{\boldsymbol{\xi}}^{(N)}$ tirés indépendamment selon $\mathbb{P}$. Ouvrir une usine $i \in \mathcal{I}$ induit un coût f_i et permet de produire tous les types de produits jusqu'à une capacité journalière S_i. Les clients sont livrés via un ensemble de tournées de véhicules connu $\mathcal{R} = \{1, \ldots, R\}$, où chaque route part d'une usine $i \in \mathcal{I}$ et livre les clients $\mathcal{J}_r \subseteq \mathcal{J}$. La demande $\tilde{\xi}_{jk}$ peut être satisfaite grâce à des quantités issues de une ou plusieurs usines ouvertes, en utilisant un ou plusieurs véhicules. Enfin, le nombre véhicules affectés à une route $r \in \mathcal{R}$ ne peut dépasser une limite V_r, chaque véhicule ayant une capacité Q_r et un coût d'utilisation c_r.

On définit trois types de variable de décision : le nombre $x_r \in \{0 \ldots, V_r\}$ de véhicules qui utilisent la route $r \in \mathcal{R}$, une variable booléenne $y_i \in \{0; 1\}$ qui indique l'ouverture de l'usine i et une proportion $z^r_{jk} \in [0, 1]$ de la demande client j pour le produit k livrée par un véhicule via la route r. Le domaine admissible correspond à l'ensemble suivant :

$$\mathcal{X} = \left\{ (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \in \mathbb{N}^R_+ \times \{0,1\}^I \times \mathbb{R}^{RJK}_+ : \begin{array}{rl} x_r \leq V_r y_i & \forall i \in \mathcal{I}, \forall r \in \mathcal{R}_i \\ z^r_{jk} \leq x_r & \forall r \in \mathcal{R}, \forall j \in \mathcal{J}_r, \forall k \in \mathcal{K}_j \\ \sum_{r \in \mathcal{R}} z^r_{jk} = 1 & \forall j \in \mathcal{J}, \forall k \in \mathcal{K}_j \end{array} \right\}$$

Pour un ensemble de variables admissible $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \in \mathcal{X}$, on note

$$\mathcal{S}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = \left\{ \boldsymbol{\xi} \in \mathbb{R}^{JK}_+ : \begin{array}{rll} \sum_{j \in \mathcal{J}_r} \sum_{k \in \mathcal{K}_j} z^r_{jk} \xi_{jk} & \leq Q_r x_r & \forall r \in \mathcal{R} \\ \sum_{r \in \mathcal{R}_i} \sum_{j \in \mathcal{J}_r} \sum_{k \in \mathcal{K}_j} z^r_{jk} \xi_{jk} & \leq S_i & \forall i \in \mathcal{I} \end{array} \right\} \tag{1}$$

l'ensemble de réalisations de demandes telles que les contraintes capacitaires sur les usines et les véhicules sont respectées. Le programme ci-dessous permet d'obtenir une solution de coût minimum telle que la probabilité de satisfaire la totalité des demandes est supérieure à $1 - \epsilon$:

$$\min_{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}} \sum_{i \in \mathcal{I}} f_i y_i + \sum_{r \in \mathcal{R}} c_r x_r \tag{2}$$

$$\text{s.t } \mathbb{P}\left(\tilde{\boldsymbol{\xi}} \in \mathcal{S}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})\right) \geq 1 - \epsilon \tag{3}$$

$$(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \in \mathcal{X} \tag{4}$$

A partir de cette formulation, nous proposons une contrepartie robuste en distribution en définissant un ensemble d'ambiguïté sous la forme d'une boule de Wasserstein à partir de l'échantillon historique $\hat{\boldsymbol{\xi}}^{(1)}, \ldots, \hat{\boldsymbol{\xi}}^{(N)}$. En se basant sur [4], nous renforçons cette formulation initiale, puis nous proposons une approche combinant une décomposition du problème et une génération de contraintes pour accélérer la résolution et obtenir des solutions sur des instances basées sur des données réelles.

Remerciements

Ce travail est financé par l'Agence Nationale de la Recherche (ANR) dans le cadre du projet LabCom CRC Lab.

Références

[1] Michael Drexl et Michael Schneider. "A survey of variants and extensions of the location-routing problem". In : *Eur. J. of Oper. Res.* 241.2 (2015), p. 283-308.

[2] Zhi Chen, Daniel Kuhn et Wolfram Wiesemann. "Data-Driven Chance Constrained Programs over Wasserstein Balls". In : *Oper. Res.* 0.0 (0), null.

[3] Weijun Xie. "On distributionally robust chance constrained programs with Wasserstein distance". In : *Math. Prog.* 186.1-2 (2021), p. 115-155.

[4] Nam Ho-Nguyen et al. "Strong formulations for distributionally robust chance-constrained programs with left-hand side uncertainty under Wasserstein ambiguity". In : *INFORMS J. on Optim.* 5.2 (2023), p. 211-232.