

Clustering robuste des flux pickup–delivery et prévision de la demande pour la planification J–1 en logistique urbaine

Salma Janati¹, Frédéric Semet²

¹ Univ. Lille, CNRS, Inria, Centrale Lille, UMR 9189 CRISTAL, F-59000 Lille, France
`salma.janati@inria.fr`

² Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRISTAL, F-59000 Lille, France
`frederic.semet@inria.fr`

Mots-clés : *logistique urbaine, pickup–delivery, clustering robuste, optimisation, prévision de la demande*

1 Contexte et objectifs

La logistique urbaine du dernier kilomètre est confrontée à une forte variabilité spatiale et temporelle de la demande, ce qui complique significativement la planification quotidienne des transporteurs. Les approches existantes s'appuient généralement sur des zonages géographiques fondés sur les points de livraison ou sur des découpages spatiaux fixes, ainsi que sur des méthodes de prévision de la demande globales. Toutefois, ces travaux peinent à capturer conjointement la structure complète des flux pickup–delivery et leur évolution spatio-temporelle, en particulier dans un contexte opérationnel contraint où les décisions doivent être prises à l'échelle J–1 à partir d'informations partielles et incertaines.

Dans ce contexte, l'objectif de ce travail est de proposer un cadre méthodologique permettant, d'une part, de structurer géographiquement les flux pickup–delivery en zones opérationnelles cohérentes, robustes et exploitables, et, d'autre part, d'anticiper la charge future associée à ces zones sans chercher à prédire individuellement chaque requête. L'ambition est de transformer des données hétérogènes en informations directement exploitables afin de soutenir une planification J–1 fiable et adaptée aux contraintes du dernier kilomètre.

2 Construction de zones géographiques

La première contribution repose sur une méthode de construction de zones qui exploite directement la structure complète des requêtes de transport. Dans cette approche, chaque requête est considérée comme une unité indivisible formée d'un point de collecte et d'un point de livraison. L'idée centrale de RADII est de regrouper les requêtes présentant une cohérence géographique simultanée entre leurs points de ramassage et de livraison, de manière à capturer fidèlement la dynamique réelle des flux du dernier kilomètre.

La méthode commence par identifier les requêtes dites « atypiques », dont les points de collecte et de livraison sont anormalement éloignés par rapport au reste du système. Ces demandes sont isolées dans un groupe séparé afin de ne pas perturber la formation des zones principales. Cette étape, issue des analyses empiriques, permet d'éviter la création de clusters déformés. Une fois les requêtes cohérentes conservées, la méthode s'intéresse à la proximité spatiale entre les points de ramassage d'un côté et entre les points de livraison de l'autre. Pour appartenir au même cluster, les points de ramassage

doivent se situer dans un voisinage restreint, et les points de livraison doivent présenter une proximité comparable. Ces seuils sont déterminés à partir de la distribution réelle des distances observées dans les données, afin de regrouper des demandes dont les origines et les destinations évoluent ensemble. La méthode impose également que les clusters aient une taille raisonnable, garantissant qu'ils soient exploitables en pratique et représentatifs de sous-ensembles cohérents de demandes.

3 Approche alternative orientée stabilité

La seconde méthode adopte une perspective plus opérationnelle et plus proche des pratiques courantes. Au lieu de considérer les couples formés par un point de ramassage et un point de livraison comme une unité, elle commence par analyser exclusivement la géographie des points de livraison, qui structure fortement les tournées avec un algorithme balltree. Cette première étape permet de construire des zones stables et régulières, moins sensibles aux variations quotidiennes. Une fois ces zones établies, les points de ramassage sont affectés aux clusters en résolvant un programme linéaire en nombres entiers, permettant de prendre en compte la structure ramassage-livraison tout en préservant la stabilité de la partition. Cette approche produit ainsi des clusters plus simples et mieux adaptés à une utilisation quotidienne, même si leur cohérence origine–destination est légèrement moins fine que celle obtenue avec RADII.

4 Prédiction de la demande et planification J–1

À partir des segmentations spatiales obtenues, nous intégrons une composante de prédiction destinée à anticiper la charge future associée aux zones. Notre approche repose sur l'analyse de la série temporelle des flux journaliers : détection des valeurs aberrantes, stabilisation de la variance, extraction de tendance et de saisonnalité, puis utilisation de modèles séquentiels profonds (LSTM, GRU, BiLSTM). L'objectif n'est pas de prédire individuellement chaque requête, mais d'estimer un volume global fiable permettant d'anticiper les niveaux de charge et les pics éventuels. La demande future est ainsi estimée directement au niveau des clusters, afin de produire des informations exploitables pour la planification opérationnelle. Cette estimation distingue la part de la demande qui reste à l'intérieur des zones et celle qui transite entre zones, un enjeu essentiel puisque les flux inter-zones requièrent davantage de coordination et constituent une source majeure de coûts opérationnels.

5 Résultats et conclusion

Les méthodes proposées sont évaluées sur un jeu de données réel de logistique urbaine en Île-de-France, comprenant plus de 60 000 requêtes pickup–delivery collectées sur une période d'un an. Les résultats montrent que les deux approches de clustering sont complémentaires : la méthode RADII produit des zones très cohérentes du point de vue des flux origine–destination, tandis que l'approche fondée sur les livraisons fournit des clusters plus stables et plus facilement exploitables au quotidien. La composante de prévision améliore la capacité d'anticipation de la charge et fournit des informations directement utilisables pour la planification J–1. Dans son ensemble, ce travail propose un cadre intégré combinant segmentation géographique, affectation robuste et prévision de la demande, afin de soutenir une orchestration fiable et opérationnelle du dernier kilomètre.

Références

- [1] Ben-Hur, A., Elisseeff, A., & Guyon, I. (2002). A stability based method for discovering structure in clustered data. *Pacific Symposium on Biocomputing*, 6, 6–17.

- [2] Crainic, T. G., Ricciardi, N., & Storchi, G. (2010). Models for evaluating and planning city logistics systems. *Transportation Science*, 44(1), 1–18.
- [3] Melo, M. T., Nickel, S., & Saldanha-da-Gama, F. (2019). Facility location and supply chain management – A review. *European Journal of Operational Research*, 196(2), 401–412.
- [4] Bertsimas, D., Orfanoudaki, A., & Wiberg, H. (2017). Dynamic clustering and prediction in healthcare. *Operations Research*, 65(6), 1460–1480.
- [5] von Luxburg, U. (2010). Clustering stability : An overview. *Foundations and Trends in Machine Learning*, 2(3), 235–274.
- [6] Laage, G. (2021). Models for Integrated Demand Forecasting and Planning : Application to Large-scale Transportation Networks and Impact Assessment. Thèse de doctorat, Université de Montréal, GERAD.
- [7] Omohundro, S. M. (1989). Five Balltree Construction Algorithms. *International Computer Science Institute Technical Report*, 89-063, Berkeley, CA.
- [8] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn : Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.