A Continuum Approximation Approach to the Hub Location Problem in a Crowd-Shipping System

P. Stokkink[a,*], N. Geroliminis[a]

[a] École Polytechnique Fédérale de Lausanne (EPFL), Urban Transport Systems Laboratory (LUTS), Lausanne, Switzerland
patrick.stokkink@epfl.ch, nikolas.geroliminis@epfl.ch
[*] Corresponding author

Extended abstract submitted for presentation at the 11th Triennial Symposium on Transportation Analysis conference (TRISTAN XI)
June 19-25, 2022, Mauritius Island

April 4, 2022

Keywords: Crowd-Shipping, Hub Location Problem, Network Design, Last-Mile Delivery

1 INTRODUCTION

The growing demand for e-commerce has led to a substantial increase in the challenges faced in traditional delivery. Nowadays, the *sharing economy* allows to rapidly connect supply and demand which can be used to overcome these challenges. Such a system where last-mile delivery is outsourced to a large number of individuals is referred to as *crowd-shipping*. In a crowd-shipping system, individual couriers perform deliveries on their pre-existing route and thereby contribute to the last-mile delivery of small parcels. Crowd-shipping has numerous advantages for customers, retailers and the society. Both customers and retailers are offered a fast, flexible and cheap alternative of delivery. Society benefits mostly from the reduced environmental impacts as well as reduced traffic congestion. According to Generation IM *et al.* (2020), the vast majority of the emissions in the whole logistics chain are currently generated by last-mile delivery.

In the literature, substantial research has been done on the operational problems that arise in crowd-shipping systems. For a review on the recent academic research and an overview of the operational challenges, the reader is referred to Le *et al.* (2019) and Pourrahmani & Jaller (2021). These operational challenges mostly focus on the matching of parcels to crowd-shippers (Li *et al.*, 2014) and the joint problem of crowd-shipping and vehicle routing (Archetti *et al.*, 2016).

As the availability of supply is a key determinant of the performance of a crowd-shipping system, parcels may be stored at intermediate hub locations (or transshipment points) such that they are easily reachable by potential crowd-shippers. Wang *et al.* (2016) consider a fixed set of *pop-stations* distributed around the city where crowd-shippers can perform pickups. Raviv & Tenzer (2018) and Macrina *et al.* (2020) consider a crowd-shipping system where crowd-shippers can pickup parcels either from the depot or from transshipment points. Their results show the economic benefits of such transshipment nodes. Similarly, Yıldız (2021) also considers transshipment points but use a dynamic programming algorithm to solve their problem. Contrary to the fixed transshipment points in the previous works, Mousavi *et al.* (2020) consider mobile depots. They do not consider the routing of vehicles, but they determine the optimal location of these mobile depots under uncertainty in supply.

Contrary to the majority of the literature that have studied operational problems, in this paper we consider the strategic planning problem of network design. We develop a framework that allows to determine the best hub locations for a crowd-shipping system in a large urban area considering uncertainty in supply and demand. This is a bi-level problem which is specifically difficult as the optimal hub locations (upper level) depend on the potential to assign parcels to crowd-shippers (lower level). We solve the lower level assignment problem through a Continuum Approximation Approach (CAA), allowing us to solve the upper level with an efficient heuristic. We compare our approach to a simulation-optimization approach, to evaluate the objective and computation time needed to attain this objective. The performance is evaluated using a discrete event simulator based on a part of the city of Washington DC.

2 METHODOLOGY

2.1 Continuum Approximation of Lower Level

We consider a network split into R regions. Expected daily demand for small parcels in every region is equal to $\hat{d}_r$. Potential crowd-shippers travel between regions such that the daily average number of crowd-shippers travelling between regions i and j is equal to $\hat{\lambda}_{ij}$. Potential crowd-shippers are assumed to have a maximum detour τ they are willing to make to pickup and deliver a parcel. We define parameter e_{ijhr} which is equal to 1 if a crowd-shipper with origin i and destination j can pickup a parcel at hub h and deliver it to the final destination in region r, and 0 otherwise. When multiple hubs are opened, the main difficulty is that both demand and supply have to be split over the various opened hubs. We define $\tilde{e}_{ijr}$ identifying whether a crowd-shipper can pickup a parcel *from at least one open hub*. Specifically: $\tilde{e}_{ijr} = \min(1, \sum_{h \in H} e_{ijhr})$.

We define $s_{ij} = \sum_{r \in R} \tilde{e}_{ijr} \hat{d}_r$ as the total demand that can potentially be served by crowd-shippers going from i to j. For the sake of the approximation, we assume that a crowd-shipper is equally likely to choose any of the parcels he/she can feasibly deliver. Following from this, the probability that he/she picks a parcel with destination region r is equal to $\frac{\hat{d}_r}{s_{ij}}$ if $\tilde{e}_{ijr} = 1$ and 0 otherwise. We can then consider all potential crowd-shippers to obtain the following estimated served demand:

$$y_r = \sum_{i,j \in R} \tilde{e}_{ijr} \hat{\lambda}_{ij} \frac{\hat{d}_r}{s_{ij}} \qquad \forall r \in R. \tag{1}$$

It is possible that crowd-shippers with different origin-destination pairs are assigned to the same parcel-destination region r. As this could lead to an overestimation of served demand in that region ($y_r > d_r$), we take into account that at most $\hat{d}_r$ demand can be delivered to a region r. Therefore, we set the estimate to

$$z_r = \min(\hat{d}_r, y_r) \qquad \forall r \in R. \tag{2}$$

Especially if supply is high, by overestimating y_r in region r (i.e. $y_r > \hat{d}_r$), it is likely that $y_{r'}$ for another region $r' \neq r$ will be underestimated. Therefore, we use an iterative process to ensure that this overestimation is accounted for in the other regions. We consider the leftover demand $l_r = \max(0, y_r - \hat{d}_r)$ and split it evenly over the potential suppliers. Similar to the assignment of parcels to crowd-shippers, we assume that every crowd-shipper that can be feasibly assigned to a region r ($\tilde{e}_{ijr} = 1$), is equally likely to assigned to one of the leftover demand units in l_r. Therefore, the l_r leftover demand units are split over the origin-destination pairs proportional to the number of suppliers that could be feasibly assigned to region r. We define the leftover supply as follows:

$$\hat{\lambda}'_{ij} = \sum_{r \in R} l_r \frac{\tilde{e}_{ijr} \hat{\lambda}_{ij}}{\sum_{i,j \in R} \tilde{e}_{ijr} \hat{\lambda}_{ij}}. \tag{3}$$

Thereby, we define the unserved demand $\hat{d}'_r = \hat{d}_r - z_r$. All demand units that are already expected to be served by previously assigned crowd-shippers no longer need to be considered and are therefore disregarded. We then compute y_r according to Equation (1), but now using $\hat{\lambda}'_{ij}$ and $\hat{d}'_r$ as inputs in stead of $\hat{\lambda}_{ij}$ and $\hat{d}_r$. Using these we find an additional portion of demand which can be served and we update the estimated demand served and the leftover demand as follows:

$$z_r = \min(\hat{d}_r, z_r + y_r) \qquad \forall r \in R. \tag{4}$$

$$l_r = \max(0, y_r - \hat{d}'_r) \qquad \forall r \in R. \tag{5}$$

This iterative process can be repeated until the leftover demand l_r is zero for all regions $r \in R$.

2.2 Large Neighborhood Search for Upper Level

To find the best set of hub locations (i.e. to solve the upper level problem), a Large Neighborhood Search (LNS) heuristic is used to explore the search space. For every potential set of hub locations, the CA approach is used to evaluate the quality efficiently which is substantially faster than a simulation or optimization approach. Repair and destroy operators are based on the expected quality of a hub in a single-hub system and the similarity between hubs, that are all pre-computed and used as inputs to the heuristic. A multi-start heuristic is used to explore a large search space.

3 RESULTS

In this section we evaluate the performance of our CA approach to finding the optimal hub locations in an urban crowd-shipping system. Our results are obtained through a case study based on a part of the city-center of Washington DC consisting of 90 regions. An approximation of the population of every region has been made using Census Reporter (2021) data, which has been used as a proxy for demand. Historic system data from the Capital Bikeshare (2020) database has been used as a proxy for the supply of bicycle-based crowd-shippers.

3.1 Comparison of CA Approach to Simulation-Optimization Approach

We evaluate the performance of our algorithm by comparing the CA-based solution algorithm proposed in this work to a well-known simulation-optimization approach. For the simulation-optimization approach, we use the same algorithm but the continuum approximation is replaced by performing 2 simulations using a minimal detour assignment strategy. We compare the two methods in terms of objective value (i.e. number of parcels delivered by crowd-shippers) and computation time. The objective value is obtained through a run of 10 simulations using the hubs that are obtained by the algorithms.

The results are displayed in Figure 1. The left-hand panel displays that the objective of the CA method is extremely close to that of the simulation-optimization method. The computation times of each of the methods are displayed in the right-hand panel of Figure 1. We observe that the CA approach is significantly faster than the simulation-optimization approach. With the CA approach, we are able to determine the best set of hubs within approximately 1 minute. The computation time of the simulation-optimization method is between 10 and 25 times higher, depending on the supply level. For the simulation-optimization method, the computation time increases proportionally to the number of supply units, whereas the computation time of the CA approach is only marginally influenced by the number of supply units. The objective of the simulation-optimization method can be further improved by increasing the number of simulation runs, but this comes at a cost of even higher computation times.

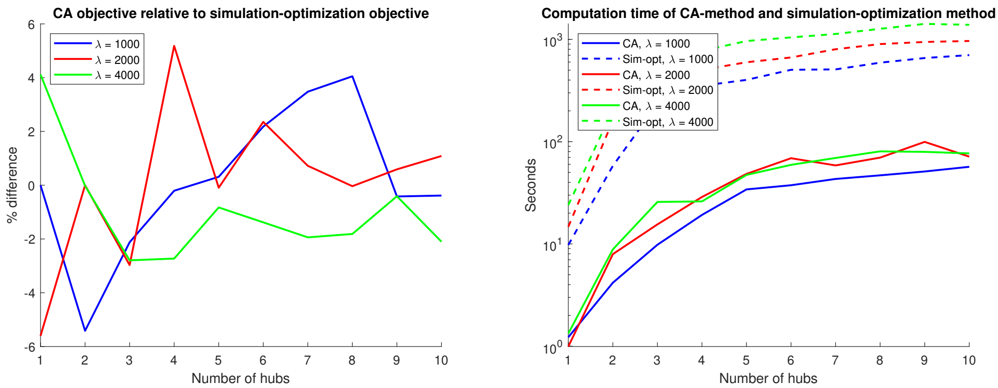

Figure 1 – *Comparison of CA and simulation-optimization method*

4 DISCUSSION

Our results show that the CA approach obtains hubs of similar quality compared to those obtained by a simulation-optimization approach, but the CA approach is much less costly in terms of computation time. Especially for large urban networks, where the number of regions and expected supply and demand are high, the CA approach shows to be extremely useful. Results on the network show the importance of incorporating supply flow in the decision of hub locations, as the optimal choice of hubs is not necessarily a central location, but more importantly a popular origin location of potential crowd-shippers, such as a train station.

In the full paper, we further evaluate the quality of our CA approach by comparing the objective value directly to that of a static and dynamic assignment problem. On top of this we perform extensive sensitivity analysis and use the CA approach to design a dynamic assignment strategy that outperforms existing dynamic assignment strategies.

References

Archetti, Claudia, Savelsbergh, Martin, & Speranza, M Grazia. 2016. The vehicle routing problem with occasional drivers. *European Journal of Operational Research*, **254**(2), 472–480.

Capital Bikeshare. 2020. *Capital Bikeshare, https://www.capitalbikeshare.com/system-data.*

Census Reporter. 2021. *Census Reporter, https://censusreporter.org/.* Accessed: 10-03-2021.

Generation IM, Preston, Felix, Kukrika, Nicholas, Matthews, H. Scott, & Martelo, Miguel A. Jaller. 2020. *The Carbon Footprint of Retail: ECommerce vs Bricks & Mortar.*

Le, Tho V, Stathopoulos, Amanda, Van Woensel, Tom, & Ukkusuri, Satish V. 2019. Supply, demand, operations, and management of crowd-shipping services: a review and empirical evidence. *Transportation Research Part C: Emerging Technologies*, **103**, 83–103.

Li, Baoxiang, Krushinsky, Dmitry, Reijers, Hajo A, & Van Woensel, Tom. 2014. The share-a-ride problem: People and parcels sharing taxis. *European Journal of Operational Research*, **238**(1), 31–40.

Macrina, Giusy, Pugliese, Luigi Di Puglia, Guerriero, Francesca, & Laporte, Gilbert. 2020. Crowd-shipping with time windows and transshipment nodes. *Computers & Operations Research*, **113**, 104806.

Mousavi, Kianoush, Bodur, Merve, & Roorda, Matthew J. 2020. *Stochastic Last-mile Delivery with Crowd-shipping and Mobile Depots.*

Pourrahmani, Elham, & Jaller, Miguel. 2021. Crowdshipping in Last Mile Deliveries: Operational Challenges and Research Opportunities. *Socio-Economic Planning Sciences*, 101063.

Raviv, Tal, & Tenzer, Eyal Z. 2018. *Crowd-shipping of small parcels in a physical internet.*

Wang, Yuan, Zhang, Dongxiang, Liu, Qing, Shen, Fumin, & Lee, Loo Hay. 2016. Towards enhancing the last-mile delivery: An effective crowd-tasking model with scalable solutions. *Transportation Research Part E: Logistics and Transportation Review*, **93**, 279–293.

Yıldız, Barış. 2021. Express package routing problem with occasional couriers. *Transportation Research Part C: Emerging Technologies*, **123**, 102994.