Analysis of the value of demand forecasting within vehicle sharing systems

Jasso Espadaler Clapés

Ecole Polytechnique Fédérale de Lausanne - EPFL

Lausanne, January 28 2020

Content

Content

Introduction - Motivation

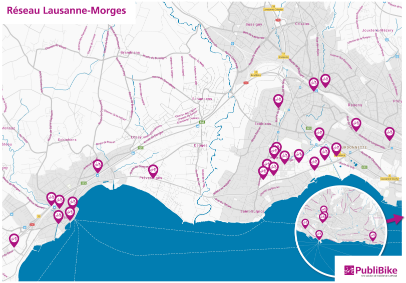
- The introduction of new VSS should be properly addressed, in particular the demand forecasting process.
- Collecting data to establish a demand model can become an exhausting process in terms of time and money.
- Does the demand knowledge in a VSS translate into a better service?

$\downarrow$

Goal: estimate the value of demand forecasting in terms of the trade-off between rebalancing cost and lost demand

Introduction - Case study

Network: PubliBike bike sharing system
Location: Lausanne-Morges agglomeration (VD), Switzerland



The network is composed of 35 stations, or 20 in its reduced version excluding Morges, Pully and some stations in the case of several in the same neighborhood.

Content

Literature review - Mobility related questions

(Perboli et al., 2017)

- Users would find dynamic pricing strategies very convenient.
- A big percentage of respondents are not able to correctly estimate the economic burden of car ownership

(Jorge and Correia, 2013)

- Do CSS have a greater effect in reducing the use of private cars or in reducing the number public transport users?
- Demand in such systems is difficult to estimate because of the existing relation between availability and number of trips.

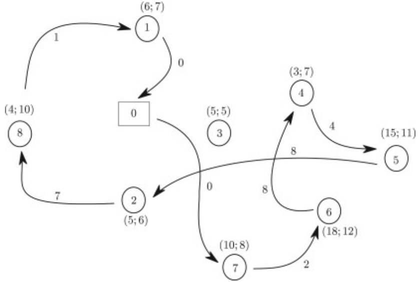
(Laporte et al., 2018)

- Challenges: station location, fleet dimensioning, station inventory, rebalancing initiatives and vehicle repositioning.

Literature review - Rebalancing operations

Rebalancing operations refer to the optimization problem dealing with the routing of trucks to supply bikes in stations that require additional bikes or pickup in those with a surplus. Several models are possible to shape the Bike Rebalancing Problem:

1. Capacitated Vehicle Routing Problem (CVRP)
2. Pickup and Delivery Problem (PDP)
3. Multiple Travelling Salesman Problem (m-TSP)



In most of the papers describing such models, demand at every node is assumed to be only positive.

Content

Methodology

What is needed:

1. Simple bike rebalancing model that can be easily solved (with special attention to subtour elimination constraints).
2. Simulation tool of the system.

In order to estimate the value of demand, combine these ingredients distinguishing two different contexts:

a) Unknown demand

Given the output of the simulation (= bikes in every station at the end of the period), the goal is to always return to the same initial configuration.

b) Known demand

We use the demand pattern of the following day to calculate the next day's initial configuration. Some nodes/station require more attention than others.

The value of demand is observed as the trade-off between rebalancing cost and lost demand

Content

Mathematical description

Let us consider a graph $G = (V, A)$ where $V = \{0, 1, 2, ..., n\}$ is the set of nodes, $\{0\}$ being the depot, and the set $A = \{(i, j) : i, j \in V, i \neq j\}$ is the arc set connecting the vertices. Each arc connecting two nodes is associated with a cost c_{ij}.

- Parameters:
  - There are m available trucks, each with a capacity Q.
  - Each node is associated with a certain demand or supply q_j (if $q_j \geq 0$ bikes should be removed from station j and if $q_j < 0$ bikes should be delivered to station j).
- The objective is to minimize the routing cost, i.e.

$$\text{Min} \sum_{(i,j) \in A} c_{ij} x_{ij}$$

where $x_{ij} = \begin{cases} 1 & \text{if arc (i,j) is used by a vehicle} \\ 0 & \text{otherwise} \end{cases}$

Dell'Amico's F1 formulation

The authors introduce a new variable θ_i to account for the load of the vehicle after leaving node i.

$$\text{Min} \sum_{i=0}^{n} \sum_{j=0}^{n} c_{ij} x_{ij} \qquad (1)$$

$\sum_{i \in V} x_{ij} = 1, \quad j \in V \setminus \{0\}$	(2)	Every node is visited once except depot
$\sum_{i \in V} x_{ji} = 1, \quad j \in V \setminus \{0\}$	(3)	Every node is visited once except depot
$\sum_{j \in V} x_{0j} \leq m$	(4)	At most m vehicles leave the depot
$\sum_{j \in V \setminus \{0\}} x_{0j} = \sum_{j \in V \setminus \{0\}} x_{j0}$	(5)	All vehicles return to depot
$\sum_{i \in S} \sum_{j \in S} \leq \lvert S \rvert - 1 \quad S \subseteq V \setminus \{0\}, S \neq \emptyset$	(6)	Classical subtour elimination constraints
$x_{ij} \in \{0,1\}, \quad i,j \in V$	(7)	Binary variable
$\max(0, q_j) \leq \theta_j \leq \min(Q, Q + q_j), \quad j \in V$	(8)	Bounds on the auxiliary variable
$\theta_j \geq \theta_i + q_j - M \cdot (1 - x_{ij}), \quad i \in V, j \in V \setminus \{0\}$	(9)	Flow conservation
$\theta_i \geq \theta_j - q_j - M \cdot (1 - x_{ij}), \quad i \in V \setminus \{0\}, j \in V$	(10)	Flow conservation

Other subtour elimination approaches? How to boost the model?

Subtour elimination constraints

There are 3 main subtour elimination formulations in a typical routing problem. Again, let us consider a graph $G = (V, A)$ where $V = \{0, 1, 2, ..., n\}$ is the set of nodes, $\{0\}$ being the depot, and the set $A = \{(i,j) : i,j \in V, i \neq j\}$ is the arc set connecting the vertices.

1) Dantzig-Fulkerson-Johnson (DFJ) $\implies O(2^{n+1})$

$$\sum_{i \in S} \sum_{j \in S} x_{ij} \leq |S| - 1, \quad S \subseteq V \setminus \{0\}, S \neq \varnothing \tag{11}$$

2) Miller-Tucker-Zemlin (MTZ) $\implies O((n+1)^2)$

$$u_i - u_j + n \cdot x_{ij} \leq n - 1, \quad i,j \in V \setminus \{0\} \tag{12}$$
$$1 \leq u_i \leq n, \quad i \in V \setminus \{0\}$$

3) Desrochers-Laporte (DL) $\implies O((n+1)^2)$

$$u_i - u_j + n \cdot x_{ij} + (n-2) \cdot x_{ji} \leq n - 1, \quad i,j \in V \setminus \{0\} \tag{13}$$
$$1 \leq u_i \leq n, \quad i \in V \setminus \{0\}$$

Valid inequalities

Valid inequalities are extra constraints that boost the computation of the models. The authors in Dell'Amico et al. (2014) propose to consider the following subsets in the graph:

$$S(i,j) = \{h \in V \setminus \{0\}, h \neq i, h \neq j : |q_i + q_j + q_h| > Q\}$$

Basically, remove routes including 3 consecutive nodes whose demand exceeding the truck capacity.

Then, the following inequalities are valid for the model:

$$x_{ij} + \sum_{h \in S(i,j)} x_{jh} \leq 1, \quad i,j \in V \setminus \{0\}, \quad h \in S(i,j) \quad (14)$$

$$\sum_{h \in S(i,j)} x_{hi} + x_{ij} \leq 1, \quad i,j \in V \setminus \{0\}, \quad h \in S(i,j) \quad (15)$$

Adaptation of Dell'Amico F1 formulation

Replacement of subtour elimination constraints with MTZ formulation and incorporation of valid inequalities and additional constraints.

$\text{Min} \sum_{i=0}^{n} \sum_{j=0}^{n} c_{ij} x_{ij}$	(1)	Objective function
$\sum_{i \in V} x_{ij} = 1, \quad j \in V \setminus \{0\}$	(2)	Every node is visited once except depot
$\sum_{i \in V} x_{ji} = 1, \quad j \in V \setminus \{0\}$	(3)	Every node is visited once except depot
$\sum_{j \in V} x_{0j} \leq m$	(4)	At most m vehicles leave the depot
$\sum_{j \in V \setminus \{0\}} x_{0j} = \sum_{j \in V \setminus \{0\}} x_{j0}$	(5)	All vehicles return to depot
$u_i - u_j + n \cdot x_{ij} \leq n - 1, \quad 1 \leq u_i \leq n, \quad i, j \in V \setminus \{0\}$	(12)	MTZ subtour elimination constraints
$x_{ij} \in \{0, 1\}, \quad i, j \in V$	(7)	Binary variable
$\max(0, q_j) \leq \theta_j \leq \min(Q, Q + q_j), \quad j \in V$	(8)	Bounds on the auxiliary variable
$\theta_j \geq \theta_i + q_j - M \cdot (1 - x_{ij}), \quad i \in V, j \in V \setminus \{0\}$	(9)	Flow conservation
$\theta_i \geq \theta_j - q_j - M \cdot (1 - x_{ij}), \quad i \in V \setminus \{0\}, j \in V$	(10)	Flow conservation
$x_{ij} + \sum_{h \in S(i,j)} x_{jh} \leq 1, \quad i, j \in V \setminus \{0\}, \quad h \in S(i,j)$	(14)	Valid inequality
$\sum_{h \in S(i,j)} x_{hi} + x_{ij} \leq 1, \quad i, j \in V \setminus \{0\}, \quad h \in S(i,j)$	(15)	Valid inequality
$x_{ii} = 0 \quad \forall i \in V$	(16)	Same node is not visited twice in a row

Content

Simulation for the unknown demand case

Algorithm 1: Framework of the PubliBike network for the unknown demand case

Result: List of lost demands and routing costs

Initial Configuration = 5 bikes per station ;

for *days = 1,2,..,N* **do**

    Lost Demand, Final Configuration = SimulationOfTheSystem() ;

    Next Day Initial Configuration ← Initial Configuration ;

    Routing Cost = RebalancingOperations() ;

end

Simulation for the known demand case

Algorithm 2: Framework of the PubliBike network for the known demand case

Result: List of lost demands and routing costs

Determine the initial configuration for day 1;

for *days = 1,2,..,N* **do**

    Lost Demand, Final Configuration = SimulationOfTheSystem() ;

    Next Day Initial Configuration ← (Days+1) Initial Configuration ;

    Routing Cost = RebalancingOperations() ;

end

Content

General remarks

- Simulation provides lost demand and final configuration in few seconds.
- The sensitivity of the rebalancing operations stems from the size of the graph.
- Solutions are very dependent on the number of trucks used and their capacity.
    - Capacity fixed to 25.

Lost demand

The system can generally secure more customers with a forecast of demand (known case). At most, there is a difference of 5 lost customers between the two cases (values very dependent on the simulation parameters).

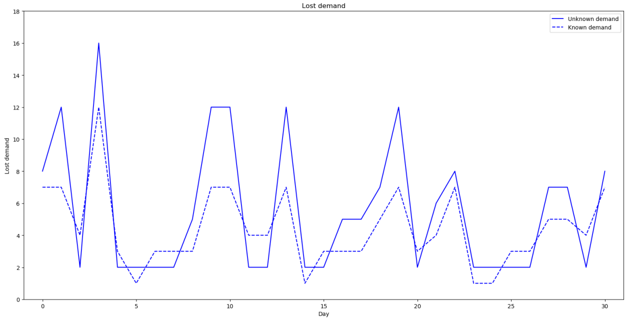


Figure 1: Lost demand over a random month

Rebalancing costs

The imposition of only 2 trucks comes at a very high cost. Moreover, the costs for both cases is exactly the same in a big proportion of the random days. It is impossible to find patterns to distinguish higher or lower costs for known and unknown demand.

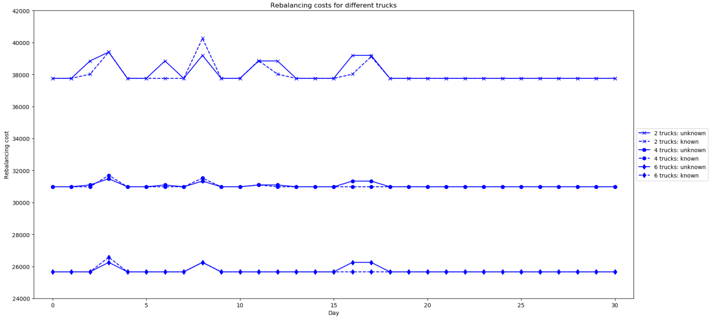


Figure 2: Rebalancing costs for different number of trucks over a random month

Relation between lost demand and rebalancing costs

There is no relation between the lost demand and rebalancing costs for this case study: small lost demand is not related to high rebalancing costs nor the fact of having a demand model influences this relation.

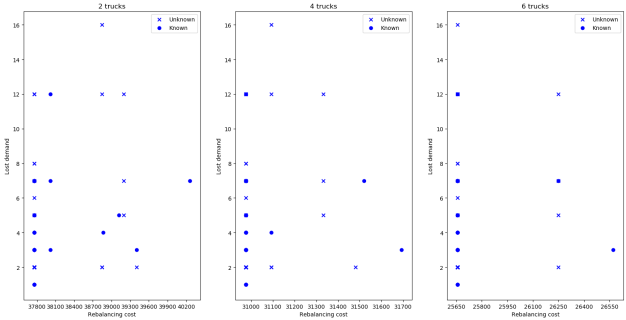


Figure 3: Relation between lost demand and rebalancing costs

Computational time

While the computational time of the simulation of the day is the same rebalancing operations are different. The solver takes more time to reach the optimal solution if the number of trucks used is low. Remarkably, the known demand case presents lower computational times.

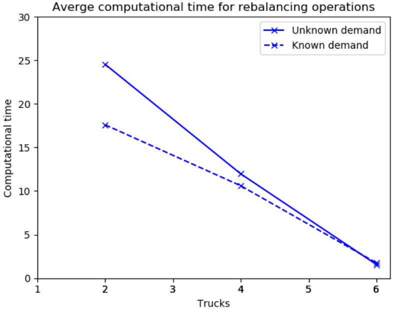


Figure 4: Computational times

Content

Conclusion

Results show that under the case study conditions:

1. The resolution of the rebalancing operations is a very sensitive problem and their cost depends essentially on the number of trucks used and their capacity.
2. The knowledge of the demand of the following day helps the manager to distribute the bikes in such a way that he can mitigate lost demand.
3. The rebalancing operations to distribute the bikes do not come at an extra cost.
4. There is no relation between the lost demand and rebalancing costs.

However, these results are very context specific.

References I

G. Perboli, B. Caroleo, and S. Musso. Car-sharing: Current and potential members behavior analysis after the introduction of the service. In *2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)*, volume 2, pages 771–776, July 2017. doi: 10.1109/COMPSAC.2017.82.

D. Jorge and G. Correia. Carsharing systems demand estimation and defined operations: A literature review. *European Journal of Transport and Infrastructure Research*, 13:201–220, 03 2013.

G. Laporte, F. Meunier, and R. Calvo. Shared mobility systems: an updated survey. *Annals of Operations Research*, 271, 12 2018. doi: 10.1007/s10479-018-3076-8.

References II

M. Dell'Amico, E. Hadjicostantinou, M. Iori, and S. Novellani. The bike sharing rebalancing problem: Mathematical formulations and benchmark instances. *Omega*, 45:7 – 19, 2014. ISSN 0305-0483. doi: https://doi.org/10.1016/j.omega.2013.12.001. URL http://www.sciencedirect.com/science/article/pii/S0305048313001187.

G. Dantzig, R. Fulkerson, and S. Johnson. Solution of a large-scale traveling-salesman problem. *Journal of the Operations Research Society of America*, 2(4):393–410, 1954. ISSN 00963984. URL http://www.jstor.org/stable/166695.

C. E. Miller, A. W. Tucker, and R. A. Zemlin. Integer programming formulation of traveling salesman problems. *J. ACM*, 7(4):326–329, October 1960. ISSN 0004-5411. doi: 10.1145/321043.321046. URL http://doi.acm.org/10.1145/321043.321046.

M. Desrochers and G. Laporte. Improvements and extensions to the miller-tucker-zemlin subtour elimination constraints. *Operations Research Letters*, 10(1):27 – 36, 1991. ISSN 0167-6377. doi: https://doi.org/10.1016/0167-6377(91)90083-2. URL http://www.sciencedirect.com/science/article/pii/0167637791900832.

Analysis of the value of demand forecasting within vehicle sharing systems

Jasso Espadaler Clapés

EPFL

Ecole Polytechnique Fédérale de Lausanne - EPFL

Lausanne, January 28 2020