Evaluation of demand forecasting in bike sharing systems: A general framework and selected case studies

Selin Ataç
Nikola Obrenović, Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
École Polytechnique Fédérale de Lausanne

11th Triennial Symposium on Transportation Analysis (TRISTAN XI)
Mauritius, June 23, 2022

Outline

1. Introduction
2. Motivation
3. Methodology
   - Considered system and the framework
   - Simulation
   - Mathematical model
4. Computational experiments
   - Scenarios
   - Case studies
   - Results
5. Conclusion

Introduction

- 14 % of the global greenhouse gas emissions is due to transportation (Pachauri et al., 2014)[1].

[1] Pachauri, R. K., M. R. Allen, V. R. Barros, J. Broome, W. Cramer, R. Christ, J. A. Church, L. Clarke, Q. Dahe, P. Dasgupta et al. Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change, Ipcc., 2014.

Introduction

- 14 % of the global greenhouse gas emissions is due to transportation (Pachauri et al., 2014)[1].
- More sustainable solutions
  - Carbon neutral fuel and electric cars
  - Ride-sharing and vehicle sharing (car, bike, e-scooter, etc.)

[1] Pachauri, R. K., M. R. Allen, V. R. Barros, J. Broome, W. Cramer, R. Christ, J. A. Church, L. Clarke, Q. Dahe, P. Dasgupta et al. Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change, Ipcc., 2014.

Introduction

What is a Vehicle Sharing System (VSS)?

- A VSS enables users to use the available vehicles generally for short period of time.
- It allows higher vehicle and less parking utilization.
- It introduces challenges such as vehicle imbalance, pricing, and demand modeling.

The framework

To understand how these challenges are related, we propose a management framework for VSSs (Ataç et al., 2021)[2].

- From decision maker point of view
- Applies to any kind of VSS
- Three dimensional classification
  - **Decision levels:** Strategic, Tactical, and Operational
  - **Actors:** Supply and Demand
  - **Layers:** Data, Models, and Actions
- Relations between the components

[2] Ataç, S., N. Obrenović, and M. Bierlaire. Vehicle sharing systems: A review and a holistic management framework. EURO Journal on Transportation and Logistics, 10, 100033, 2021.

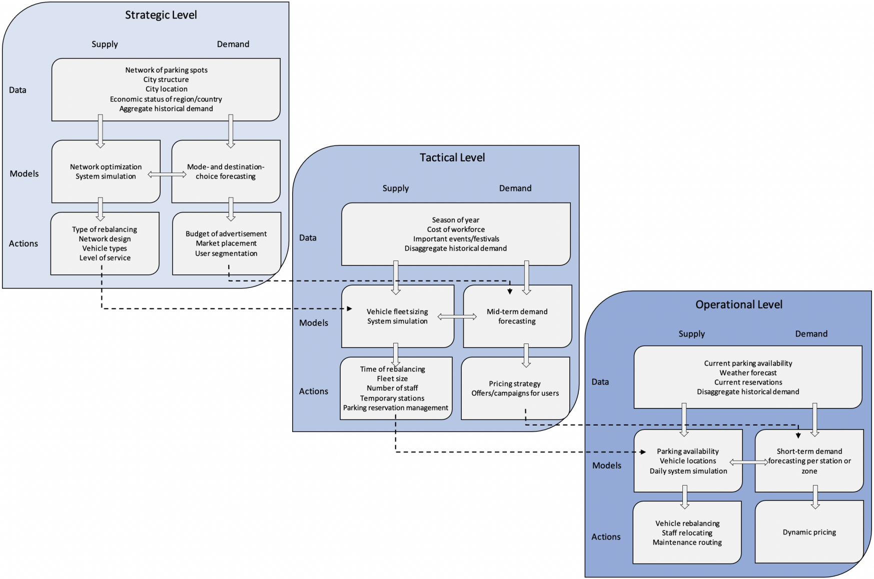


Figure: Holistic framework and inter-relations

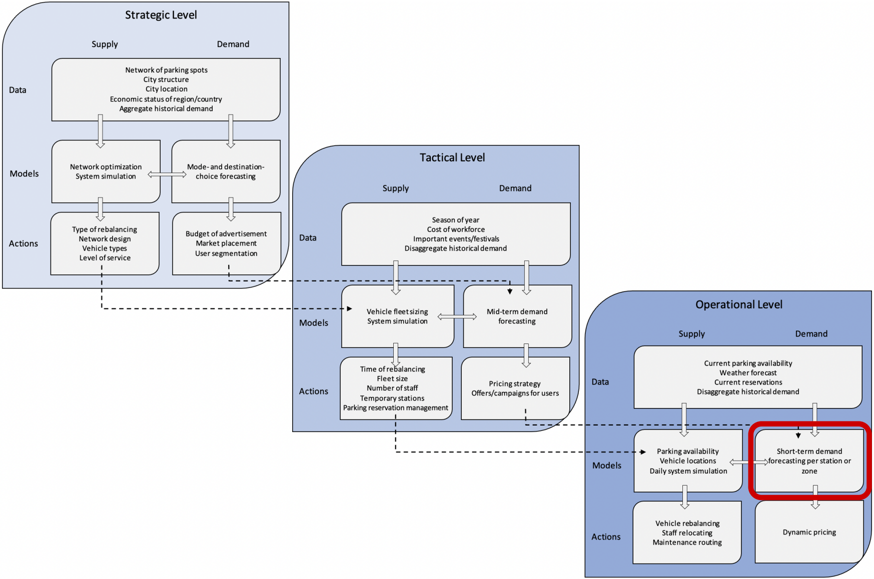


Figure: Holistic framework and inter-relations

Motivation

Is shared mobility as sustainable as we think?

- Reck et al. (2022)[3] claim that personalized micro-mobility is more sustainable than the shared one.
- One reason is costly rebalancing operations.

[3] Reck D. J., H. Martin, and K. W. Axhausen. Mode choice, substitution patterns and environmental impacts of shared and personal micro-mobility. Transportation Research Part D: Transport and Environment, 102:103134, 2022.

[4] Shu J., M. Chou, Q. Liu, C. Teo, and I-L. Wang. Models for effective deployment and redistribution of bicycles within public bicycle-sharing systems. Operations Research, 61:1346-1359, 11 2013.

Motivation

Is shared mobility as sustainable as we think?

- Reck et al. (2022)[3] claim that personalized micro-mobility is more sustainable than the shared one.
- One reason is costly rebalancing operations.

What is the added-value of bike rebalancing in bike sharing systems?

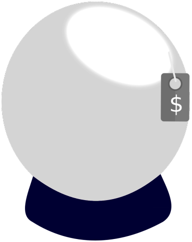
- Shu et al. (2013)[4] find that the number of substituted trips change as a function of number of bicycles and number of redistributions per day.
- Periodic and frequent rebalancing operations are not necessary for some configurations of the system.

[3] Reck D. J., H. Martin, and K. W. Axhausen. Mode choice, substitution patterns and environmental impacts of shared and personal micro-mobility. Transportation Research Part D: Transport and Environment, 102:103134, 2022.

[4] Shu J., M. Chou, Q. Liu, C. Teo, and I-L. Wang. Models for effective deployment and redistribution of bicycles within public bicycle-sharing systems. Operations Research, 61:1346-1359, 11 2013.

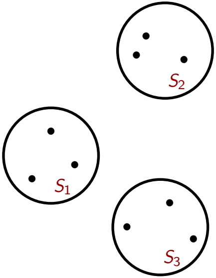
Big picture - revisited

- VSS related literature mainly focuses on rebalancing problems and their solutions by formulating them as VRP or TSP.
- Modeling the demand is also studied, but the added value of constructing such a model in bike sharing systems is not investigated.



The idea

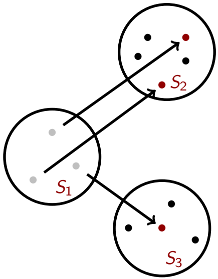
Real world



Decision center

The idea

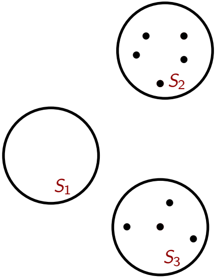
Real world



Decision center

The idea

Real world

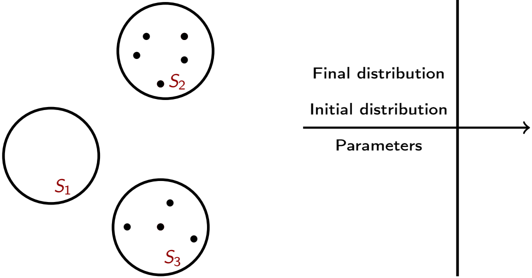


Decision center

The idea

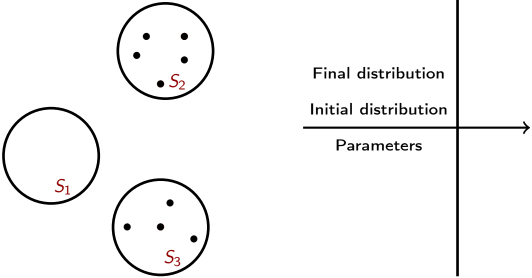
Real world

Decision center



The idea

Real world

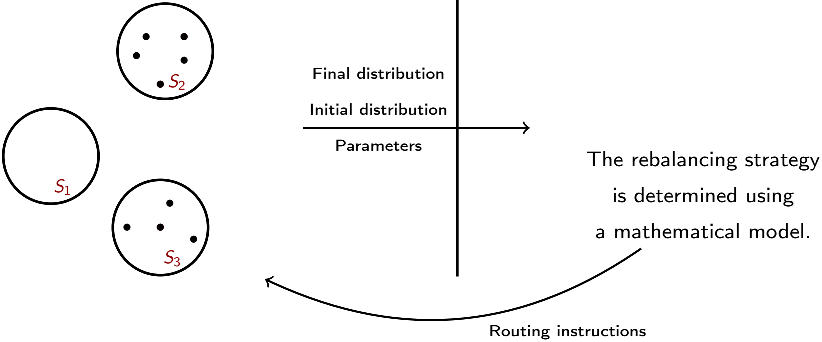


Decision center

The rebalancing strategy
is determined using
a mathematical model.

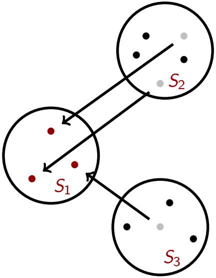
The idea

Real world | Decision center



The idea

Real world



Final distribution

Initial distribution

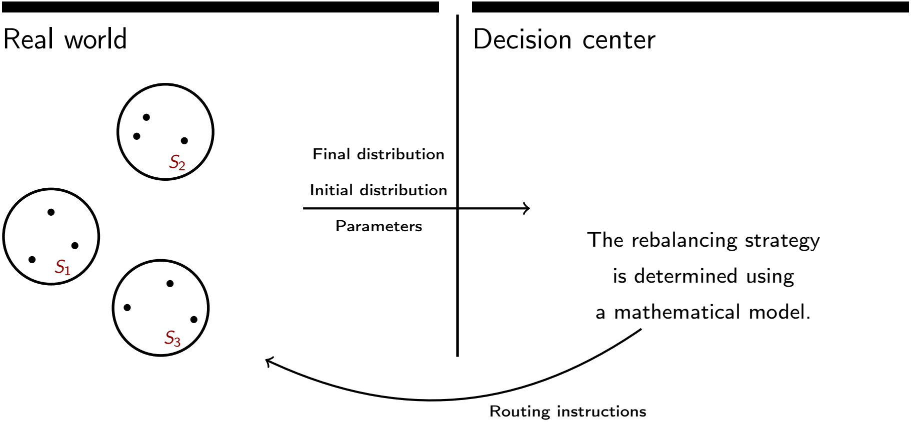
Parameters

Decision center

The rebalancing strategy
is determined using
a mathematical model.

Routing instructions

The idea


Real world
Decision center
S_1
S_2
S_3
Final distribution
Initial distribution
Parameters
The rebalancing strategy
is determined using
a mathematical model.
Routing instructions

The idea

Real world

Discrete event simulations:

1- the daily demand

2- the rebalancing operations

Modeling flexible and
stochastic system behavior

Decision center

Mathematical models
to determine the routing of
rebalancing operations

More specific and
sometimes unrealistic decisions

The framework and research question

What is the added value of the module "Trip demand forecasting"?

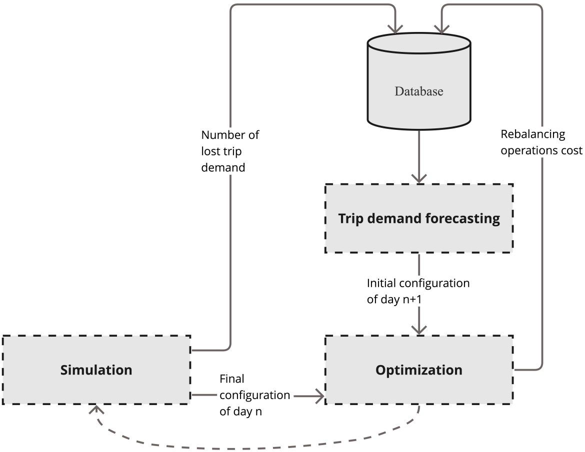
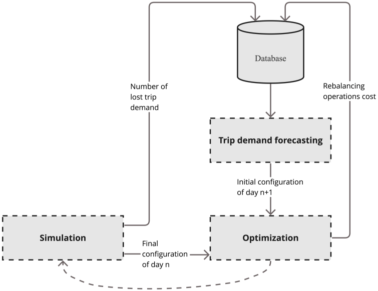
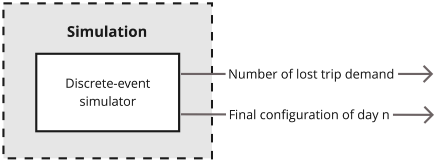
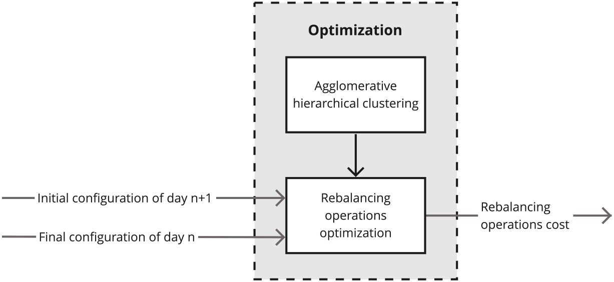


Figure: General framework


Number of lost trip demand
Database
Rebalancing operations cost
Trip demand forecasting
Initial configuration of day n+1
Simulation
Final configuration of day n
Optimization


Simulation
Discrete-event simulator
Number of lost trip demand
Final configuration of day n


Optimization
Agglomerative hierarchical clustering
Rebalancing operations optimization
Initial configuration of day n+1
Final configuration of day n
Rebalancing operations cost

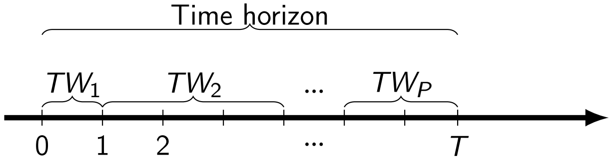
Real world - Discrete event simulator

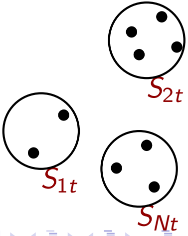
State variables:

- t: time,
- S_{it}: vehicle availability at station i at time t,
- Location of the orders in the system.

Parameters:

- C_i: the capacity of a station i, $i = 1, ..., N$,
- λ_p: the number of O-D pair requests per hour for time window p, $p = 1, ..., P$.





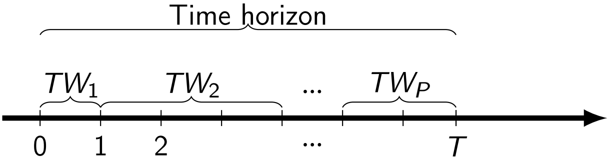
Real world - Discrete event simulator

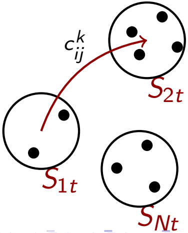
State variables:

- t: time,
- S_{it}: vehicle availability at station i at time t,
- Location of the orders in the system.

Parameters:

- C_i: the capacity of a station i, $i = 1, ..., N$,
- λ_p: the number of O-D pair requests per hour for time window p, $p = 1, ..., P$.





Real world - Discrete event simulator

Indicators:

- The realized travel time from origin to destination and from pick-up station to drop-off station,
- Number of users using the system,
- The number of lost demand.

Assumptions:

- After T, only the events in the system are served and no new requests are accepted.
- Reserving a vehicle is not possible.
- The O-D pair requests are spatially and temporally uniformly distributed.

Real world - Discrete event simulator

Indicators:

- The realized travel time from origin to destination and from pick-up station to drop-off station,
- Number of users using the system,
- The number of lost demand.

Assumptions:

- After T, only the events in the system are served and no new requests are accepted.
- Reserving a vehicle is not possible.
- **The O-D pair requests are spatially and temporally uniformly distributed.**

Decision center - Rebalancing operations optimization

Set:

- V: the set of stations, $V = \{0, ..., N\}$, where $\{0\}$ is the depot.

Parameters:

- m: the number of relocation vehicles available,
- Q: the capacity of a relocation vehicle,
- c_{ij}: the length of the shortest path between i and j, $\forall i, j \in V$,
- q_i: the difference between the number of bikes at station i at the end of the previous day and the number of bikes desired at the beginning of the next day, $\forall i \in V$.

Decision variable:

$$x_{ij} = \begin{cases} 1, & \text{if arc } (i,j) \text{ is used by a relocation vehicle} \\ 0, & \text{otherwise} \end{cases} \quad \forall i, j \in V, \tag{1}$$

Decision center - Modified model (Dell'Amico et al., 2013)

minimize cost

subject to

- all non zero demand stations are visited
- number of trucks is not exceeded
- the load of the trucks is not exceeded
- MTZ constraints
- valid inequalities
- domain constraints

Decision center - Modified model (Dell'Amico et al., 2013)

minimize cost subject to

- all non zero demand stations are visited
- number of trucks is not exceeded
- the load of the trucks is not exceeded
- **MTZ constraints**
- **valid inequalities**
- domain constraints

Scenarios - Trip demand forecasting

Two cases are investigated:

- **Unknown demand:** we rebalance the system to the same initial state every day.
- **Known demand:** we assume that we perfectly know the trip demand of the following day. The initial state of the next day is determined by considering the pick-up and drop-offs at a station throughout the time horizon of the following day.

The main idea is to see how the cost of rebalancing operations and the number of lost demand differ between the two cases.

Case studies

Synthetic case study

- PubliBike Lausanne, Switzerland
  - 35 stations, 175 bikes

O-D trip requests are generated.

Real-life case studies

- nextbike Sarajevo
  - 21 stations, 120 bikes
- nextbike Berlin
  - 298 stations, 3000 bikes
- Divvy Chicago
  - 681 stations, 6000 bikes
- Citi Bike New York
  - 1361 stations, 22000 bikes

O-D trip requests are obtained from their system.

Case studies

Synthetic case study

- PubliBike Lausanne, Switzerland
  - 35 stations, 175 bikes

O-D trip requests are generated.

Real-life case studies

- **nextbike Sarajevo**
  - **21 stations, 120 bikes**
- **nextbike Berlin**
  - **298 stations, 3000 bikes**
- **Divvy Chicago**
  - **681 stations, 6000 bikes**
- **Citi Bike New York**
  - **1361 stations, 22000 bikes**

O-D trip requests are obtained from their system.

Results - Clustering

- Four different clustering methods are considered (Atac et al., 2021)[6].
  - Agglomerative hierarchical clustering (AHC) with Ward linkage
    - Proximity as a similarity matrix
    - Number of trips as a similarity matrix
  - Multi-objective mathematical model approach
    - Mixed integer non linear model
    - Mixed integer linear model

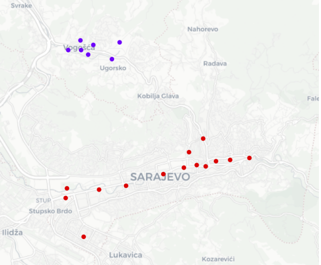
[6] Ataç, S., N. Obrenovic, and M. Bierlaire. A multi-objective approach for station clustering in bike sharing systems. 21st Swiss Transport Research Conference, 2021.

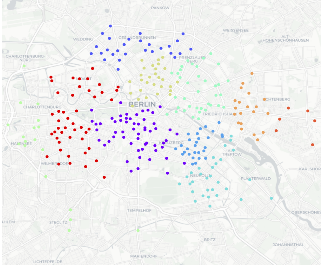
Results - Clustering

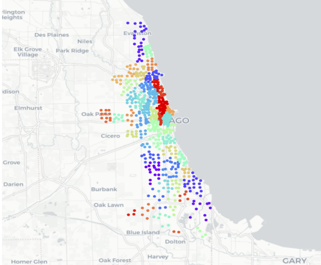
- Four different clustering methods are considered (Atac et al., 2021)[6].
  - **Agglomerative hierarchical clustering (AHC) with Ward linkage**
    - **Proximity as a similarity matrix**
    - Number of trips as a similarity matrix
  - Multi-objective mathematical model approach
    - Mixed integer non linear model
    - Mixed integer linear model

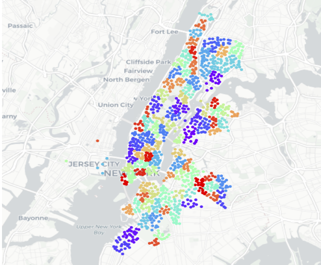
[6] Ataç, S., N. Obrenovic, and M. Bierlaire. A multi-objective approach for station clustering in bike sharing systems. 21st Swiss Transport Research Conference, 2021.

Results - Clustering









Comparison of two cases

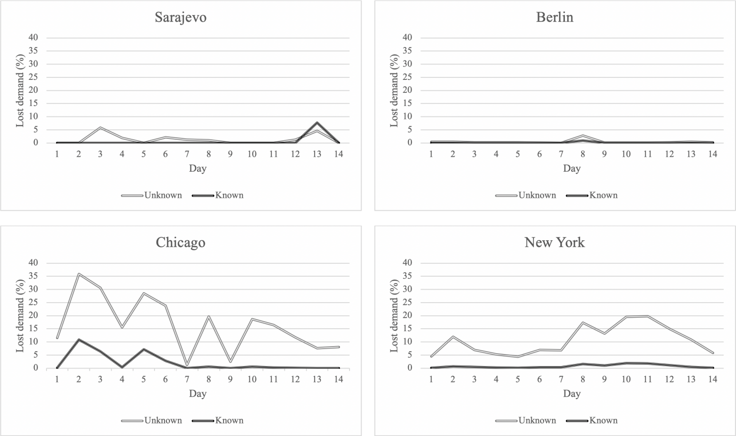


Figure: Lost demand percentage for both cases

Comparison of two cases

- Thanks to the extended framework, we are able to analyze larger instances.
- The rebalancing operations cost does not significantly change from one case to the other for any of the cases and case studies.
  - Routes tend to be the similar.
  - Demand forecasting does not affect rebalancing operations.

Intermediate case - 7-day scenarios

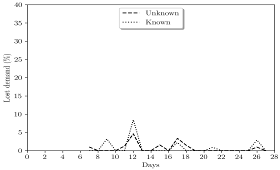


Figure: Sarajevo

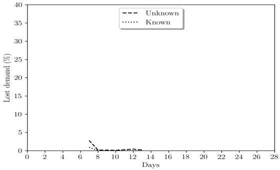


Figure: Berlin

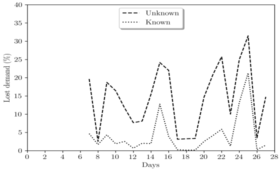


Figure: Chicago

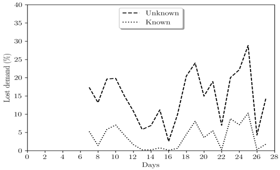


Figure: New York City

Intermediate case - 7-day scenarios

- The behavior is again different for smaller and larger case sudies.
  - For Sarajevo and Berlin case studies, the added-value of demand forecasting is not consistent through days.
  - For Chicago and New York City, it is possible to see persistent added value.

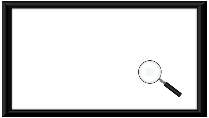
Conclusions and future work

- A generic framework to evaluate the added value of demand forecasting in BSSs is presented.
- Experiments on four case studies show interesting results.
  - The positive effect of demand forecasting is more visible in larger scale case studies.
  - Rebalancing operations routes tend to be similar to each other, hence no improvement.

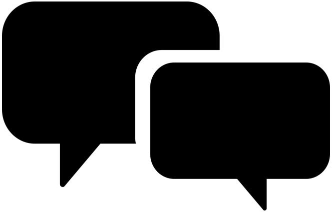
Conclusions and future work

- A generic framework to evaluate the added value of demand forecasting in BSSs is presented.
- Experiments on four case studies show interesting results.
  - The positive effect of demand forecasting is more visible in larger scale case studies.
  - Rebalancing operations routes tend to be similar to each other, hence no improvement.

- The next steps include
  - investigating intermediate cases by looking different levels of knowledge and
  - testing the effect of different values of input parameters.



Questions and discussion



Selin Ataç

Transport and Mobility Laboratory (TRANSP-OR)
École Polytechnique Fédérale de Lausanne (EPFL)

selin.atac@epfl.ch

Decision center - Modified model (Dell'Amico et al., 2013)

$$(F1_M)\min \quad \sum_{i\in V}\sum_{j\in V} c_{ij}x_{ij} \tag{2}$$

s.to

$$\sum_{i\in V} x_{ik} = 1 \qquad \forall k \in N \tag{3}$$

$$\sum_{i\in V} x_{ki} = 1 \qquad \forall k \in N \tag{4}$$

$$\sum_{j\in V} x_{0j} \le m \tag{5}$$

$$\sum_{k\in N} x_{0j} - \sum_{k\in N} x_{k0} = 0 \tag{6}$$

$$u_k - u_l + |N| * x_{kl} \le |N| - 1 \qquad \forall k,l \in N \tag{7}$$

$$1 \le u_i \le |N| - qCount \qquad \forall i \in V \tag{8}$$

$$x_{ii} = 0 \qquad \forall i \in V \tag{9}$$

$$\theta_j \ge \max\{0, q_j\} \qquad \forall j \in V \tag{10}$$

$$\theta_j \le \min\{Q, Q + q_j\} \qquad \forall j \in V \tag{11}$$

$$\theta_k - \theta_i + M(1 - x_{ik}) \ge q_k \qquad \forall i \in V, k \in N \tag{12}$$

$$\theta_k - \theta_j + M(1 - x_{kj}) \ge -q_j \qquad \forall k \in N, j \in V \tag{13}$$

$$x_{kl} + \sum_{h\in S(k,l)} x_{lh} \le 1 \qquad \forall k,l \in N, h \in S(k,l) \tag{14}$$

$$\sum_{h\in S(k,l)} x_{hk} + x_{kl} \le 1 \qquad \forall k,l \in N, h \in S(k,l) \tag{15}$$

$$\theta_0 = 0 \tag{16}$$

$$x_{ij} \in \{0,1\} \qquad \forall i,j \in V \tag{17}$$