A multi-objective approach for station clustering in bike sharing systems

Selin Ataç, Nikola Obrenović, Michel Bierlaire

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

September 14, 2021

Outline

1. Introduction
   - Previous work
   - Literature
2. Methodology
   - Clustering
3. Computational experiments
   - Case studies
   - Experiments
   - Results
4. Conclusion

Introduction

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

Introduction

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

Introduction

- 14 % of the global greenhouse gas emissions is due to transportation (Pachauri et al., 2014).
- More sustainable solutions
  - Carbon neutral fuel and electric cars
  - Ride-sharing and vehicle sharing (car, bike, e-scooter, etc.)
- Bike sharing systems (BSSs)
  - Short rentals
  - Higher bike and less parking utilization
  - Examples: PubliBike, nextbike, mobike

Motivation

It comes with operational challenges.

- In Ataç et al. (2020), we look into the relationship between rebalancing operations and demand forecasting.

Motivation

It comes with operational challenges.

- In Ataç et al. (2020), we look into the relationship between rebalancing operations and demand forecasting.

Real world

Discrete event simulations
to imitate the daily demand

Modeling flexible and
stochastic system behavior

Decision center

Mathematical models
to determine the routing of
rebalancing operations

More specific and
sometimes unrealistic decisions

Motivation

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, and thereby evaluate the demand forecasting by analyzing the trade-off between the two scenarios.

Motivation

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, and thereby **evaluate the demand forecasting** by analyzing the trade-off between the two scenarios.

Notation

Set	
V	the set of stations, $V = \{0, \ldots, N\}$, where $\{0\}$ is the depot
Parameters	
N	the number of stations
m	number of relocation vehicles available
Q	capacity of a relocation vehicle
$qCount$	number of stations to be visited
c_{ij}	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 variables	
x_{ij}	1 if arc (i,j) is used by a relocation vehicle, 0 otherwise, $i,j \in V$
θ_i	the load of a vehicle after it leaves node i, $i \in V$
u_i	auxiliary decision variable for the MTZ constraints, $i \in V$

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

$$
\begin{aligned}
(F1M) \qquad & \min \sum_{i\in V}\sum_{j\in V} c_{ij}x_{ij} && && (1)\\
& \text{s.to} \quad \sum_{i\in V} x_{ij} = 1 && \forall j \in V\setminus\{0\} && (2)\\
& \sum_{i\in V} x_{ji} = 1 && \forall j \in V\setminus\{0\} && (3)\\
& \sum_{j\in V} x_{0j} \le m && && (4)\\
& \sum_{j\in V\setminus\{0\}} x_{0j} = \sum_{j\in V\setminus\{0\}} x_{j0} && && (5)\\
& u_i - u_j + N * x_{ij} \le N-1 && \forall i,j \in V\setminus\{0\} && (6)\\
& 1 \le u_i \le N - qCount && \forall i \in V && (7)\\
& \min\{Q, Q+q_j\} \ge \theta_j \ge \max\{0, q_j\} && \forall j \in V && (8)\\
& \theta_j - \theta_i + M(1-x_{ij}) \ge q_j && \forall i \in V, j \in V\setminus\{0\} && (9)\\
& \theta_i - \theta_j + M(1-x_{ij}) \ge q_j && \forall i \in V\setminus\{0\}, j \in V && (10)\\
& x_{ij} + \sum_{h\in S(i,j)} x_{jh} \le 1 && \forall i,j \in V\setminus\{0\}, h \in S(i,j) && (11)\\
& \sum_{h\in S(i,j)} x_{hi} + x_{ij} \le 1 && \forall i,j \in V\setminus\{0\}, h \in S(i,j) && (12)\\
& \theta_0 = 0 && && (13)\\
& x_{ii} = 0 && \forall i \in V && (14)\\
& x_{ij} \in \{0,1\} && \forall i,j \in V && (15)
\end{aligned}
$$

The research question

The size of the rebalancing operations optimization model increases exponentially.

The research question

The size of the rebalancing operations optimization model increases exponentially.

- Tailor-made branch and cut algorithms
  - Dell'Amico et al. (2014), Erdogan et al. (2014), Chemla et al. (2013b)

The research question

The size of the rebalancing operations optimization model increases exponentially.

- Tailor-made branch and cut algorithms
  - Dell'Amico et al. (2014), Erdogan et al. (2014), Chemla et al. (2013b)
- Benders decomposition
  - Erdogan et al. (2014)

The research question

The size of the rebalancing operations optimization model increases exponentially.

- Tailor-made branch and cut algorithms
  - Dell'Amico et al. (2014), Erdogan et al. (2014), Chemla et al. (2013b)
- Benders decomposition
  - Erdogan et al. (2014)
- Neighborhood search
  - Ho and Szeto (2017), Cruz et al. (2017)

The research question

The size of the rebalancing operations optimization model increases exponentially.

- Tailor-made branch and cut algorithms
  - Dell'Amico et al. (2014), Erdogan et al. (2014), Chemla et al. (2013b)
- Benders decomposition
  - Erdogan et al. (2014)
- Neighborhood search
  - Ho and Szeto (2017), Cruz et al. (2017)
- Clustering based approaches
  - Schuijbroek et al. (2017), Liu et al. (2016), Boyaci et al. (2017), Feng et al. (2017), Ma et al. (2019), Lahoorpoor et al. (2019)

The research question

The size of the rebalancing operations optimization model increases exponentially.

- Tailor-made branch and cut algorithms
  - Dell'Amico et al. (2014), Erdogan et al. (2014), Chemla et al. (2013b)
- Benders decomposition
  - Erdogan et al. (2014)
- Neighborhood search
  - Ho and Szeto (2017), Cruz et al. (2017)
- **Clustering based approaches**
  - Schuijbroek et al. (2017), Liu et al. (2016), Boyaci et al. (2017), Feng et al. (2017), Ma et al. (2019), **Lahoorpoor et al. (2019)**

Performance measures

The performance measures are:

- (P1) the total in-cluster Manhattan distance,
- (P2) the deviation of the total in-cluster demand from zero, and
- (P3) the deviation of number of stations per cluster from the average number of stations per cluster.

Clustering methods

We look into the following approaches:

- **(C1)**: Agglomerative hierarchical clustering (AHC) with Ward linkage and **proximity of stations** as a similarity matrix
  - considers (P1)
- **(C2)**: AHC with Ward linkage and **number of trips between stations** as a similarity matrix adapted from Lahoorpoor et al. (2019)
  - considers (P2)
- **(C3)**: A mixed-integer non linear program
  - considers all performance measures
- **(C4)**: A mixed-integer linear program
  - considers all performance measures

Clustering methods - Notation

Parameters	
N	number of stations $(i,\ j \in \{1, \ldots, N\})$
C	number of clusters $(c \in \{1, \ldots, C\})$
d_{ij}	the distance from station i to station j, $i,j \in N$
q_i	the demand at each station, $i \in N$
$\alpha,\ \beta,\ \gamma$	weight of 1^{st}, 2^{nd} and 3^{rd} objective function, respectively
Decision variable	
s_{ic}	1 if station i is assigned to cluster c, 0 otherwise, $i \in N, c \in C$
Auxiliary decision variables	
$devSN_c^+,\ devSN_c^-$	the positive and negative deviation of number of stations in cluster c from the average number of stations per cluster, $c \in C$, respectively
$devD_c^+,\ devD_c^-$	the positive and negative deviation of total demand from 0 in cluster c, $c \in C$, respectively
$inClusterDist_c$	the total Manhattan distance between each pair of stations in cluster c, $c \in C$
m_{ijc}	1 if both i and j are in cluster c, 0 otherwise, $i,j \in N, c \in C$

Clustering methods - (C3N)

$$
\begin{aligned}
(C3N)\min \quad & \alpha \cdot \sum_{c \in C} inClusterDist_c && (16)\\
& +\beta \cdot \sum_{c \in C} (devD_c^+ + devD_c^-) && (17)\\
& +\gamma \cdot \sum_{c \in C} (devSN_c^+ + devSN_c^-) && (18)\\
\text{s.to} \quad & \sum_{c \in C: q_i \neq 0} s_{ic} = 1 \quad \forall i \in N && (19)\\
& \sum_{i,j \in N: j \geq i} s_{ic} \cdot s_{jc} \cdot d_{ij} = inClusterDist_c \quad \forall i,j \in N, \forall c \in C && (20)\\
& \sum_{i \in N} s_{ic} \cdot q_i = devD_c^+ - devD_c^- \quad \forall c \in C && (21)\\
& \sum_{i \in N} s_{ic} = \frac{N}{C} + devSN_c^+ - devSN_c^- \quad \forall c \in C && (22)\\
& s_{ic} \in \{0,1\} \quad \forall i \in N, c \in C && (23)\\
& devSN_c^+, devSN_c^- \geq 0 \quad \forall c \in C && (24)\\
& devD_c^+, devD_c^- \geq 0 \quad \forall c \in C && (25)\\
& inClusterDist_c \geq 0 \quad \forall c \in C && (26)
\end{aligned}
$$

Clustering methods - (C3N)

$$(C3N)\ \min \quad \alpha \cdot \sum_{c \in C} inClusterDist_c \tag{16}$$

$$+\beta \cdot \sum_{c \in C} (devD_c^+ + devD_c^-) \tag{17}$$

$$+\gamma \cdot \sum_{c \in C} (devSN_c^+ + devSN_c^-) \tag{18}$$

$$\text{s.to} \quad \sum_{c \in C : q_i \neq 0} s_{ic} = 1 \qquad \forall i \in N \tag{19}$$

$$\boxed{\sum_{i,j \in N : j \geq i} s_{ic} \cdot s_{jc} \cdot d_{ij} = inClusterDist_c \qquad \forall i,j \in N, \forall c \in C} \tag{20}$$

$$\sum_{i \in N} s_{ic} \cdot q_i = devD_c^+ - devD_c^- \qquad \forall c \in C \tag{21}$$

$$\sum_{i \in N} s_{ic} = \frac{N}{C} + devSN_c^+ - devSN_c^- \qquad \forall c \in C \tag{22}$$

$$s_{ic} \in \{0,1\} \qquad \forall i \in N, c \in C \tag{23}$$

$$devSN_c^+, devSN_c^- \geq 0 \qquad \forall c \in C \tag{24}$$

$$devD_c^+, devD_c^- \geq 0 \qquad \forall c \in C \tag{25}$$

$$inClusterDist_c \geq 0 \qquad \forall c \in C \tag{26}$$

Clustering methods - (C3)

$$
\begin{aligned}
(C3)\min \quad & \alpha \cdot \sum_{c \in C} inClusterDist_c && (27)\\
& +\beta \cdot \sum_{c \in C} (devD_c^+ + devD_c^-) && (28)\\
& +\gamma \cdot \sum_{c \in C} (devSN_c^+ + devSN_c^-) && (29)\\
\text{s.to} \quad & \sum_{c \in C : q_i \neq 0} s_{ic} = 1 \quad \forall i \in N && (30)\\
& m_{ijc} \leq s_{ic} \quad \forall i,j \in N, \forall c \in C && (31)\\
& m_{ijc} \leq s_{jc} \quad \forall i,j \in N, \forall c \in C && (32)\\
& m_{ijc} \geq s_{ic} + s_{jc} - 1 \quad \forall i,j \in N, \forall c \in C && (33)\\
& \sum_{i,j \in N : j \geq i} m_{ijc} \cdot d_{ij} = inClusterDist_c \quad \forall i,j \in N, \forall c \in C && (34)\\
& m_{ijc} \in \{0,1\} \quad \forall i,j \in N, \forall c \in C && (35)\\
& \sum_{i \in N} s_{ic} \cdot q_i = devD_c^+ - devD_c^- \quad \forall c \in C && (36)\\
& \sum_{i \in N} s_{ic} = \frac{N}{C} + devSN_c^+ - devSN_c^- \quad \forall c \in C && (37)\\
& s_{ic} \in \{0,1\} \quad \forall i \in N, c \in C && (38)\\
& devSN_c^+, devSN_c^- \geq 0 \quad \forall c \in C && (39)\\
& devD_c^+, devD_c^- \geq 0 \quad \forall c \in C && (40)\\
& inClusterDist_c \geq 0 \quad \forall c \in C && (41)
\end{aligned}
$$

Clustering methods - Additional notation

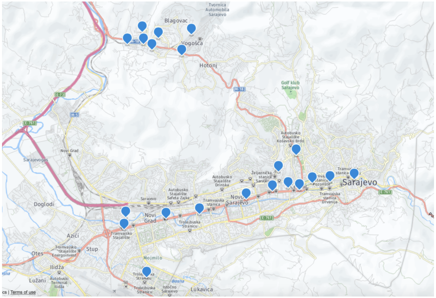
Parameters	
M	big-M value
lon_i, lat_i	the longitude and latitude of station i, $i \in N$, respectively
Auxiliary decision variables	
$lonC_c$, $latC_c$	the longitude and latitude of cluster c, $c \in C$, respectively
$diffLon_{ic}$	the distance in longitude between station i and cluster c, $i \in N, c \in C$
$diffLat_{ic}$	the distance in latitude between station i and cluster c, $i \in N, c \in C$
md_{ic}	the Manhattan distance between station i and cluster c, $i \in N, c \in C$

Clustering methods - (C4)

$$
\begin{aligned}
(C4)\,\min \quad & (27)+(28)+(29) & & \\
\text{s.to} \quad & (30) & & \\
& lon_i - lonC_c \le diffLon_{ic} & \forall i \in N, \forall c \in C \quad & (42)\\
& lonC_c - lon_i \le diffLon_{ic} & \forall i \in N, \forall c \in C \quad & (43)\\
& lat_i - latC_c \le diffLat_{ic} & \forall i \in N, \forall c \in C \quad & (44)\\
& latC_c - lat_i \le diffLat_{ic} & \forall i \in N, \forall c \in C \quad & (45)\\
& diffLon_{ic} + diffLat_{ic} \le md_{ic} + M \cdot (1 - s_{ic}) & \forall i \in N, \forall c \in C \quad & (46)\\
& \sum_{i \in N} md_{ic} \le inClusterDist_c & \forall c \in C \quad & (47)\\
& (36), (37), (38) & & \\
& diffLon_{ic}, diffLat_{ic}, md_{ic} \ge 0 & \forall i \in N, c \in C \quad & (48)\\
& lonC_c, latC_c \ge 0 & \forall c \in C \quad & (49)\\
& (39), (40), (41) & &
\end{aligned}
$$

Case studies

nextbike Sarajevo with 21 stations and approx. 120 bikes



nextbike Berlin with 298 stations and approx. 3000 bikes

The experiments

- Neither $(C3N)$ nor $(C3)$ are tractable, therefore they are not included in the experimentation.

The experiments

- Neither $(C3N)$ nor $(C3)$ are tractable, therefore they are not included in the experimentation.
- The computational experiments are done with $(C1)$, $(C2)$, and $(C4)$.

The experiments

- Neither $(C3N)$ nor $(C3)$ are tractable, therefore they are not included in the experimentation.
- The computational experiments are done with $(C1)$, $(C2)$, and $(C4)$.
- Lexicographic method is tried.

The experiments

- Neither $(C3N)$ nor $(C3)$ are tractable, therefore they are not included in the experimentation.
- The computational experiments are done with $(C1)$, $(C2)$, and $(C4)$.
- Lexicographic method is tried.
    - No solutions in real time

The experiments

- Neither $(C3N)$ nor $(C3)$ are tractable, therefore they are not included in the experimentation.
- The computational experiments are done with $(C1)$, $(C2)$, and $(C4)$.
- Lexicographic method is tried.
  - No solutions in real time
- $(C4)$ is experimented with the following two settings:

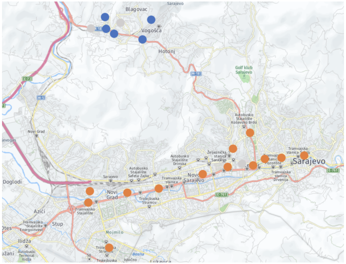
The experiments

- Neither $(C3N)$ nor $(C3)$ are tractable, therefore they are not included in the experimentation.
- The computational experiments are done with $(C1)$, $(C2)$, and $(C4)$.
- Lexicographic method is tried.
  - No solutions in real time
- $(C4)$ is experimented with the following two settings:
  - $(C4_{DD})$: $\beta \gg \alpha \gg \gamma$

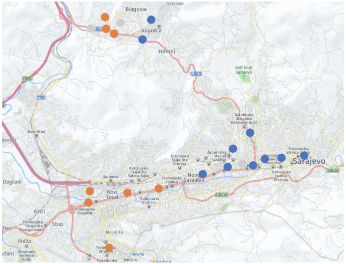
The experiments

- Neither $(C3N)$ nor $(C3)$ are tractable, therefore they are not included in the experimentation.
- The computational experiments are done with $(C1)$, $(C2)$, and $(C4)$.
- Lexicographic method is tried.
  - No solutions in real time
- $(C4)$ is experimented with the following two settings:
  - $(C4_{DD})$: $\beta \gg \alpha \gg \gamma$
  - $(C4_{ICD})$: $\alpha \gg \beta \gg \gamma$

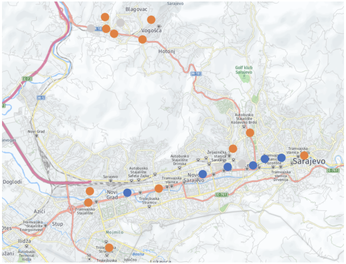
Sarajevo with 2 clusters



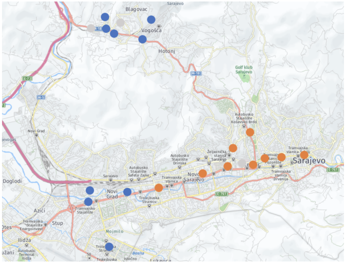
Clustering with $(C1)$



Clustering with $(C4_{DD})$

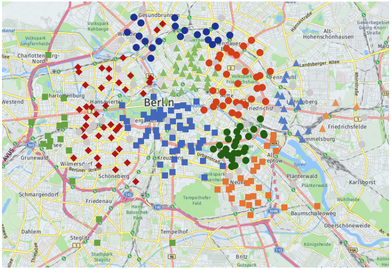


Clustering with $(C2)$

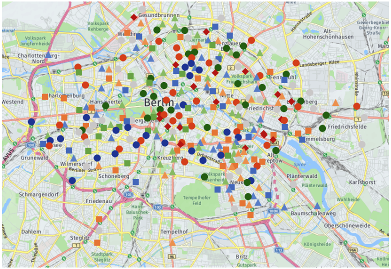


Clustering with $(C4_{ICD})$

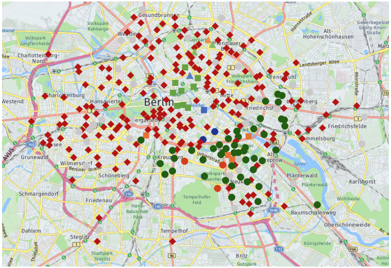
Berlin with 10 clusters



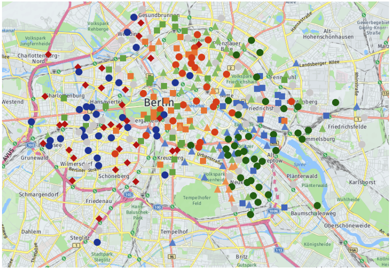
Clustering with $(C1)$



Clustering with $(C4_{DD})$



Clustering with $(C2)$



Clustering with $(C4_{ICD})$

Rebalancing cost (in kms) for all the clustering methods

Dataset	# of clusters	$(C1)$	$(C2)$	$(C4_{DD})$	$(C4_{ICD})$
Sarajevo	2	9.728	15.591	12.709	12.627
Berlin	10	75.351	372.332	139.880	126.983
	15	83.393	103.923	163.261	173.630
	20	90.471	120.289	159.271	197.483

Rebalancing cost (in kms) for all the clustering methods

Dataset	# of clusters	$(C1)$	$(C2)$	$(C4_{DD})$	$(C4_{ICD})$
Sarajevo	2	9.728	15.591	12.709	12.627
Berlin	10	75.351	372.332	139.880	126.983
	15	83.393	103.923	163.261	173.630
	20	90.471	120.289	159.271	197.483

- The costs resulted from application of (C2) increase compared to (C1).

Rebalancing cost (in kms) for all the clustering methods

Dataset	# of clusters	$(C1)$	$(C2)$	$(C4_{DD})$	$(C4_{ICD})$
Sarajevo	2	9.728	15.591	12.709	12.627
Berlin	10	75.351	372.332	139.880	126.983
	15	83.393	103.923	163.261	173.630
	20	90.471	120.289	159.271	197.483

- The costs resulted from application of (C2) increase compared to (C1).
    - Accumulation in a few stations

Rebalancing cost (in kms) for all the clustering methods

Dataset	# of clusters	$(C1)$	$(C2)$	$(C4_{DD})$	$(C4_{ICD})$
Sarajevo	2	9.728	15.591	12.709	12.627
Berlin	10	75.351	372.332	139.880	126.983
	15	83.393	103.923	163.261	173.630
	20	90.471	120.289	159.271	197.483

- The costs resulted from application of (C2) increase compared to (C1).
    - Accumulation in a few stations
- Compared to (C1), the kilometers traveled for both (C4DD) and (C4ICD) increases.

Rebalancing cost (in kms) for all the clustering methods

Dataset	# of clusters	$(C1)$	$(C2)$	$(C4_{DD})$	$(C4_{ICD})$
Sarajevo	2	9.728	15.591	12.709	12.627
Berlin	10	75.351	372.332	139.880	126.983
	15	83.393	103.923	163.261	173.630
	20	90.471	120.289	159.271	197.483

- The costs resulted from application of (C2) increase compared to (C1).
    - Accumulation in a few stations
- Compared to (C1), the kilometers traveled for both (C4DD) and (C4ICD) increases.
    - Overlapping clusters

Rebalancing cost (in kms) for all the clustering methods

Dataset	# of clusters	$(C1)$	$(C2)$	$(C4_{DD})$	$(C4_{ICD})$
Sarajevo	2	9.728	15.591	12.709	12.627
Berlin	10	75.351	372.332	139.880	126.983
	15	83.393	103.923	163.261	173.630
	20	90.471	120.289	159.271	197.483

- The costs resulted from application of (C2) increase compared to (C1).
  - Accumulation in a few stations
- Compared to (C1), the kilometers traveled for both (C4DD) and (C4ICD) increases.
  - Overlapping clusters
- The demand-based objective does not reduce the rebalancing cost.

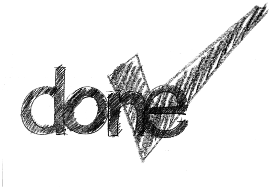
Rebalancing cost (in kms) for all the clustering methods

Dataset	# of clusters	$(C1)$	$(C2)$	$(C4_{DD})$	$(C4_{ICD})$
Sarajevo	2	9.728	15.591	12.709	12.627
Berlin	10	75.351	372.332	139.880	126.983
	15	83.393	103.923	163.261	173.630
	20	90.471	120.289	159.271	197.483

- The costs resulted from application of (C2) increase compared to (C1).
  - Accumulation in a few stations
- Compared to (C1), the kilometers traveled for both (C4DD) and (C4ICD) increases.
  - Overlapping clusters
- The demand-based objective does not reduce the rebalancing cost.
  - Large optimality gap

Conclusion

- Different clustering methods are assessed.



Conclusion

- Different clustering methods are assessed.
- AHC using proximity results in geographically collective clusters, but not zero demand deviation.

Conclusion

- Different clustering methods are assessed.
- AHC using proximity results in geographically collective clusters, but not zero demand deviation.
- AHC using number of trips brings about uneven distribution in number of stations.

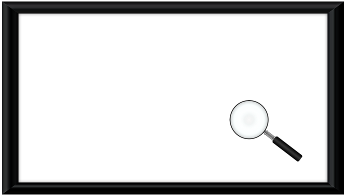
Conclusion

- Different clustering methods are assessed.
- AHC using proximity results in geographically collective clusters, but not zero demand deviation.
- AHC using number of trips brings about uneven distribution in number of stations.
- (C4) overcomes this yet the areas spanned by each cluster tend to overlap some other clusters.

Future work

The future work includes

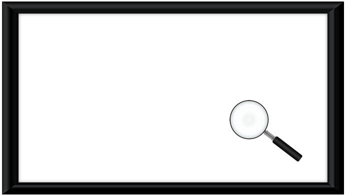
- A heuristic approach for multi-objective station clustering



Future work

The future work includes

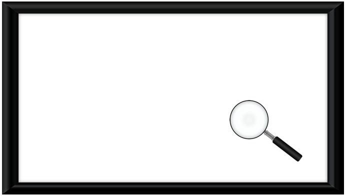
- A heuristic approach for multi-objective station clustering
- Testing the approaches for many days to achieve statistical significance



Future work

The future work includes

- A heuristic approach for multi-objective station clustering
- Testing the approaches for many days to achieve statistical significance
- The consideration of different data sets to derive conclusions about the relation between the city and demand structure



Questions and discussion

Selin Ataç

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

selin.atac@epfl.ch

References I

- Ataç, S., N. Obrenović and M. Bierlaire (2020) Vehicle sharing systems: Does demand forecasting yield a better service?, paper presented at the 20th Swiss Transport Research Conference.
- Boyaci, B., K. G. Zografos and N. Geroliminis (2017) An integrated optimization-simulation framework for vehicle and personnel relocations of electric carsharing systems with reservations, Transportation Research Part B: Methodological, 95, 214 - 237, ISSN 0191-2615.
- Chemla, D., F. Meunier and R. Wolfler Calvo (2013b) Bike sharing systems: Solving the static rebalancing problem, Discrete Optimization, 10 (2) 120-146, ISSN 1572-5286.
- Cruz, F., A. Subramanian, B. P. Bruck and M. Iori (2017) A heuristic algorithm for a single vehicle static bike sharing rebalancing problem, Computers & Operations Research, 79, 19 -33, ISSN 0305-0548.
- Dell'Amico, M., E. Hadjicostantinou, M. Iori and S. Novellani (2014) The bike sharing rebalancing problem: Mathematical formulations and benchmark instances, Omega, 45, 7-19, ISSN 0305-0483.
- Erdogan, G., G. Laporte and R. Wolfler Calvo (2014) The static bicycle relocation problem with demand intervals, European Journal of Operational Research, 238 (2) 451-457, ISSN 0377-2217.
- Feng, Y., R. C. Aonso and M. Zolghadri (2017) Analysis of bike sharing system by clustering: the velib? case, IFAC-PapersOnLine, 50 (1) 12422-12427, ISSN 2405-8963. 20th IFAC World Congress.
- Ho, S. C. and W. Szeto (2017) A hybrid large neighborhood search for the static multi-vehicle bike-repositioning problem, Transportation Research Part B: Methodological, 95, 340-363, ISSN 0191-2615.
- Lahoorpoor, B., H. Faroqi, A. Sadeghi-Niaraki and S.-M. Choi (2019) Spatial cluster-based model for static rebalancing bike sharing problem, Sustainability, 11 (11), ISSN 2071-1050.
- Liu, J., L. Sun, W. Chen and H. Xiong (2016) Rebalancing bike sharing systems: A multi-source data smart optimization, paper presented at the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1005-1014, 08 2016.
- Ma, X., R. Cao and Y. Jin (2019) Spatiotemporal clustering analysis of bicycle sharing system with data mining approach, Information, 10 (5), ISSN 2078-2489.

References II

- 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. (2014) 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.
- Schuijbroek, J., R. Hampshire and W.-J. van Hoeve (2017) Inventory rebalancing and vehicle routing in bike sharing systems, European Journal of Operational Research, 257 (3) 992-1004, ISSN 0377-2217.

Mathematical model - The base model (Dell'Amico et al., 2013)

$$
\begin{array}{lll}
(F3)\,\min & \displaystyle\sum_{i\in V}\sum_{j\in V} c_{ij}x_{ij} & \\
\text{s.to} & \displaystyle\sum_{i\in V} x_{ij} = 1 & \forall j \in V \setminus \{0\} \\
 & \displaystyle\sum_{i\in V} x_{ji} = 1 & \forall j \in V \setminus \{0\} \\
 & \displaystyle\sum_{j\in V} x_{0j} \le m & \\
 & \displaystyle\sum_{j\in V\setminus\{0\}} x_{0j} = \sum_{j\in V\setminus\{0\}} x_{j0} & \\
 & \displaystyle\sum_{i\in S}\sum_{j\in S} x_{ij} \le |S| - 1 & \forall S \subseteq V \setminus \{0\}, S \ne \emptyset \\
 & \min\{Q, Q + q_j\} \ge \theta_j \ge \max\{0, q_j\} & \forall j \in V \\
 & \theta_j - \theta_i + M(1 - x_{ij}) \ge q_j & \forall i \in V, j \in V \setminus \{0\} \\
 & \theta_i - \theta_j + M(1 - x_{ij}) \ge q_j & \forall i \in V \setminus \{0\}, j \in V \\
 & x_{ij} \in \{0,1\} & \forall i,j \in V
\end{array}
$$

Sarajevo - 2 clusters

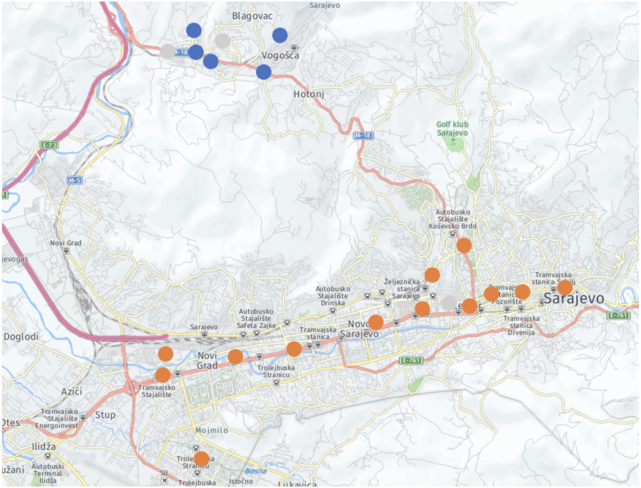


Figure: Clustering with $(C1)$

Sarajevo - 2 clusters

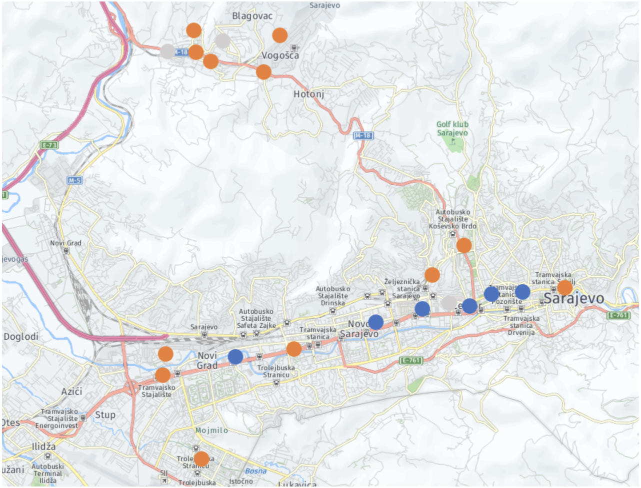


Figure: Clustering with $(C2)$

Sarajevo - 2 clusters

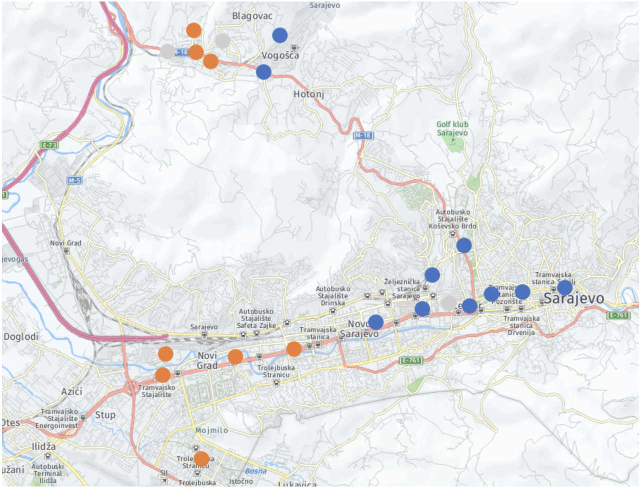


Figure: Clustering with (*C4DD*)

Sarajevo - 2 clusters

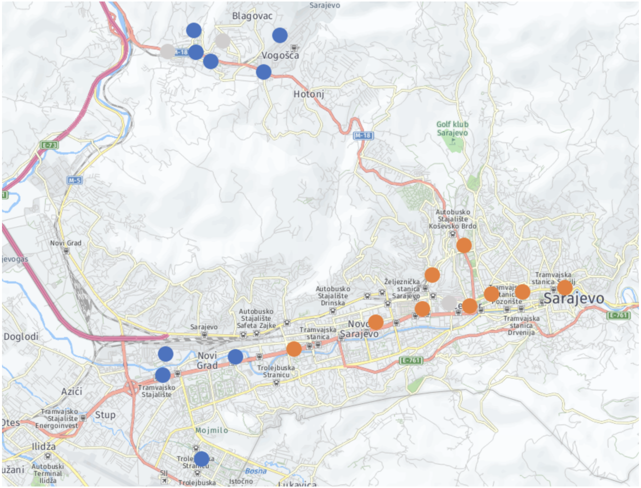


Figure: Clustering with ($C4ICD$)

Berlin - 10 clusters

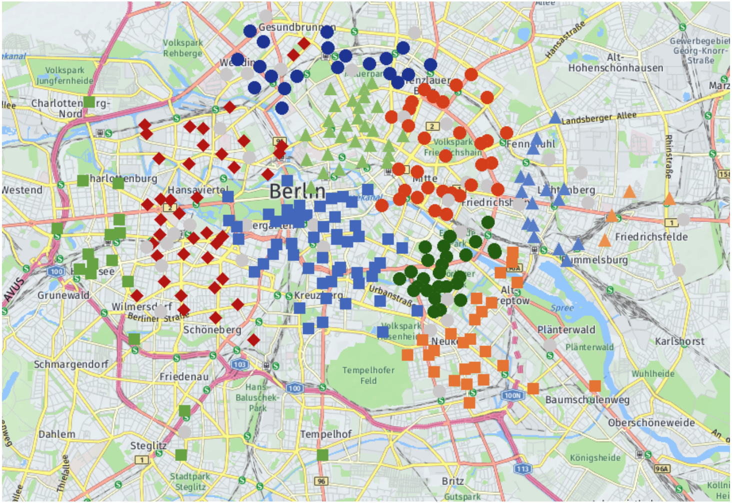


Figure: Clustering with $(C1)$

Berlin - 10 clusters

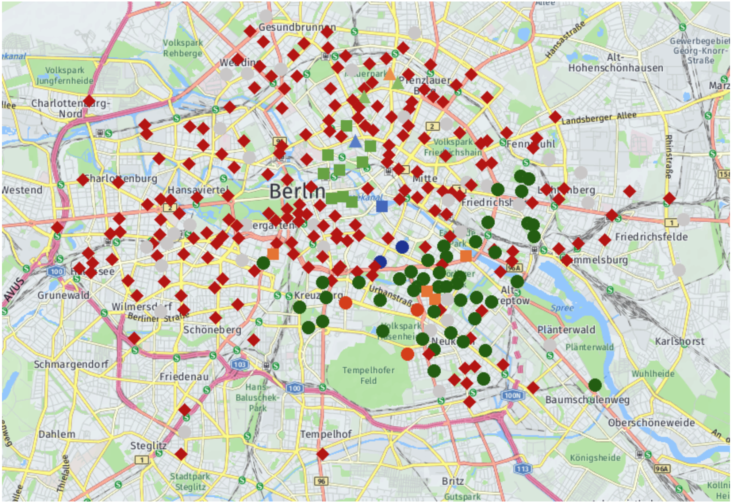


Figure: Clustering with $(C2)$

Berlin - 10 clusters

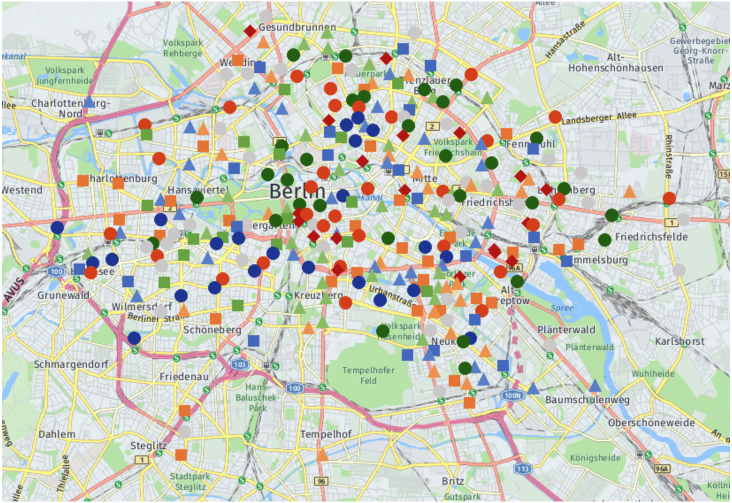


Figure: Clustering with $(C4DD)$

Berlin - 10 clusters

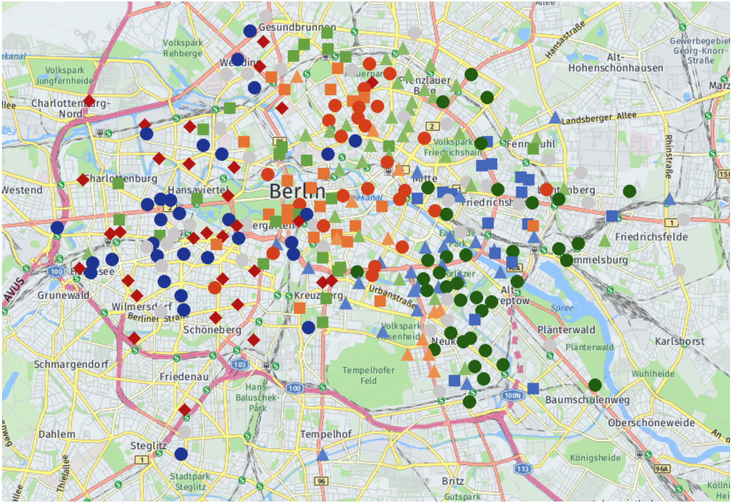


Figure: Clustering with (*C4ICD*)