An integrated optimization-simulation framework for vehicle and personnel relocations of electric carsharing systems with reservations

Burak Boyacı [a,*], Konstantinos G. Zografos [a], Nikolas Geroliminis [b]

[a] Centre for Transport and Logistics (CENTRAL), Lancaster University Management School, Lancaster, LA1 4YX, UK
[b] Urban Transport Systems Laboratory (LUTS), Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

A B S T R A C T

One-way electric vehicle carsharing systems are receiving increasing attention due to their mobility, environmental, and societal benefits. One of the major issues faced by the operators of these systems is the optimization of the relocation operations of personnel and vehicles. These relocation operations are essential in order to ensure that vehicles are available for use at the right place at the right time. Vehicle availability is a key indicator expressing the level of service offered to customers. However, the relocation operations, that ensure this availability, constitute a major cost component for the provision of these services. Therefore, clearly there is a trade-off between the cost of vehicle and personnel relocation and the level of service offered. In this paper we are developing, solving, and applying, in a real world context, an integrated multi-objective mixed integer linear programming (MMILP) optimization and discrete event simulation framework to optimize operational decisions for vehicle and personnel relocation in a carsharing system with reservations. We are using a clustering procedure to cope with the dimensionality of the operational problem without compromising on the quality of the obtained results. The optimization framework involves three mathematical models: (i) station clustering, (ii) operations optimization and (iii) personnel flow. The output of the optimization is used by the simulation in order to test the feasibility of the optimization outcome in terms of vehicle recharging requirements. The optimization model is solved iteratively considering the new constraints restricting the vehicles that require further charging to stay in the station until the results of the simulation are feasible in terms of electric vehicles' battery charging levels. The application of the proposed framework using data from a real world system operating in Nice, France sheds light to trade-offs existing between the level of service offered, resource utilization, and certainty of fulfilling a trip reservation.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The emergence of carsharing as a new transportation mean between public and private transport is quite recent. People benefiting from this service have to register and then can use vehicles located within a designated service area. Carsharing

* Corresponding author.
E-mail addresses: b.boyaci@lancaster.ac.uk (B. Boyacı), k.zografos@lancaster.ac.uk (K.G. Zografos), nikolas.geroliminis@epfl.ch (N. Geroliminis).

is an approach to increase vehicle utilization while decreasing the cost encountered to each user. Expenses associated to car ownership are nowadays quite high linked to purchase, gasoline, insurance, parking and maintenance. Avoiding these expenses may be a motivation for a driver to adopt carsharing as a regular transportation mean if the system is able to provide high quality of service (accessibility and availability of vehicles both at origin and destination of his/her trip at reasonable cost). In addition to users' benefits, carsharing systems are producing broader societal and environmental benefits such as reduction of congestion, air pollution, and the urban space needed for parking.

Carsharing systems can be classified in different categories according to the rental conditions. Free-floating systems allow people to park the vehicles anywhere in the covered area, whereas non-floating impose restrictions to users to park the vehicles at stations with limited number of parking spots. Another differentiating feature is the "one-way/round-trip" characteristic: Round-trip systems force the user to return the car to the location where it was picked-up whereas one-way systems allow drop-off at any station. The type of vehicle (combustion, electric, etc.) affects also the system's use.

Rental operations naturally induce imbalances in the spatial and temporal distribution of vehicles of one-way carsharing systems. To maximize the demand served, vehicle distribution needs to be rebalanced by performing relocations to maximize vehicle availability. There may be a wide variety of demand types depending on their reservation pattern: some reservations are made at the last minute while others are made long time in advance when a vehicle is needed for a scheduled activity. Both types of demand introduce constraints for the operator. Thus, the operational management of carsharing systems is complex due to the stochastic and dynamic nature of demand in time and space, and the limited availability of information. The operator of such a system has to manage it in a way that maximizes the use of the system while at the same time all operational and business constraints are satisfied.

We focus in this paper on non-floating one-way electric carsharing systems because of their increased flexibility towards the user compared to round-trip systems and their eco-friendly characteristics. In particular, the objective of this paper is to develop and apply an integrated framework for optimizing operational decisions related to vehicle and relocation personnel relocation decisions. The developed framework could be extended to free-floating one-way systems.

2. Previous related research

Carsharing systems have been studied intensively in recent years. The existing body of literature addresses issues related to carsharing systems characteristics and types, assessment of carsharing system impacts, and modelling of strategic, tactical and operational decisions. For a more comprehensive literature review the reader refers to Jorge and Correia (2013), Shaheen and Cohen (2013) and Boyacı et al. (2015). Given the focus of this paper on operational decisions we limit our literature review to papers that are mostly relevant to the modelling of operational decisions for one-way carsharing systems. These decisions relate to the allocation and re-allocation of carsharing vehicles and relocation personnel to stations. Optimization and/or simulation are the methodological approaches that have been used to support one-way carsharing operational decisions.

Barth et al. (1999) developed a queuing based discrete event simulation model to evaluate operational decisions for a shared vehicle system of a resort community in Southern California. The types of decisions addressed include vehicle availability, vehicle distribution, and energy management. This model uses a variety of performance measure, such as total average wait time, number of customers waiting, number of relocations, and average battery state of charge, to assess the performance of the system.

Kek et al. (2006) introduced a time-stepping simulation model for assessing relocation operations using shortest time and inventory balancing criteria. Zero-vehicle time, full-port time, and number of relocations were the three indicators used to assess the performance of the relocation operations. Kek et al. (2009) used a three-phase optimization-trend-simulation (OTS) approach to develop a decision support system (DSS), for relocation operations. The optimization phase is based on a mixed integer programming (MIP) formulation which is used to determine the resources needed to operate the carsharing system at minimum cost. In the second phase the output of the optimization phase is "filtered" to produce the parameters needed to simulate the operation of the carsharing system. The simulator developed in Kek et al. (2006) is used in the third step of the proposed DSS.

Nair and Miller-Hooks (2011) proposed a stochastic mixed integer programming (MIP) model to generate minimum cost vehicle redistribution plans to satisfy a proportion of the near-term stochastic demand of the carsharing system. The proposed model was demonstrated on a real-world carsharing system in Singapore.

Correia and Antunes (2012) developed an optimization approach for station location and trip selection schemes to maximize profit for one-way systems with stations. This work showed that offering a high level of service to one-way systems without involving relocation requires a very large fleet to cope with the demand and high involved cost. Thus, integrating a relocation module is necessary to deal with spatial and temporal asymmetry of demand. Correia et al. (2014) considered how information about vehicle availability in stations and flexibility of users in renting from a station nearby, but not the closest station, can improve the demand served. No relocation strategies have been integrated in this formulation.

Relocation strategies and inventory management for bike-sharing systems already exist in the literature (Benchimol et al., 2011; Nair et al., 2013; Raviv and Kolka, 2013; Raviv et al., 2013; Sayarshad et al., 2012). While bike- and carsharing systems share plenty of similarities both at the tactical and operational level, a significant difference is that relocation is usually performed with trucks that are able to reposition simultaneously a large amount of bicycles, while for carsharing systems, these actions are more time- and personnel-consuming.

Weikl and Bogenberger (2015a, 2015b) provided relocation strategies for free-floating systems for pick up and drop-off. They combine a macroscopic relocation optimization policy of moving vehicles between zones, with a rule based heuristic for station to station relocations. Specific movements of relocation personnel, frequency of relocation policies and detailed cost estimates are not provided in the paper. This approach was tested in a free-floating system in Munich and is one of the few reported tests in the literature.

Kaspi et al. (2014) considered a system with reservations and model it as a dynamic system with Markovian characteristics. The authors showed that under specific demand profiles, reservations can improve the performance of the system compared to free-access sharing systems. Nevertheless, the destination spot is reserved at the beginning of the trip, which might create under utilization of the system. Instead, if the return time of the vehicle is known to some extend, more advanced relocation policies can guarantee a free spot at the destination and at a desired time. We utilize this approach in our current work.

Pfrommer et al. (2014) developed a dynamic redistribution model for a public bicycle sharing (PBS) system. The proposed system is based on predictive control principles which are used to dynamically calculate rewards for system users. The goal of this model is to provide incentives to users to return the rented bicycles to under-utilized stations in order to reduce bicycle repositioning costs.

Bruglieri et al. (2014) addressed the vehicle relocation problem for an one way electric carsharing system in Milan. A mixed integer programming (MIP) model was developed to determine the optimum sequence of pick-up and delivery operations that should be performed by vehicle-relocation personnel.

Jorge et al. (2015) presented a mathematical programming model for maximizing profitability of a carsharing system optimizing relocation operations for one-way carsharing system. In addition a discrete-event time-driven simulation model was proposed to evaluate alternative relocation policies.

Nourinejad and Roorda (2014) introduced an optimization-simulation framework for supporting decisions for one way carsharing systems. The proposed framework includes a static benchmark module and a dynamic module. The static component determines the optimum number of vehicles needed to serve the entire demand for the system. The dynamic relocation component involves a vehicle relocation (VRO) and a parking inventory model (PIO). The vehicle relocation model seeks to optimize system revenues by relocating vehicles from the destination of user i to the origin of user j. The parking inventory model aims to minimize the total cost of parking by determining the optimal relocation times for the vehicles determined by the vehicle relocation optimization model. The VRO and PIO models are run in conjunction to a discrete-event simulation model where an event is defined by a new user arrival. The solution of the VRO and PIO determine the customer requests that will be satisfied, the vehicles that will be relocated, and the time that these relocations will take place.

Jorge et al. (2015) influenced vehicle imbalances by integrating a pricing scheme with a demand choice model that charges high prices for the trips that contribute to imbalances and lower prices for the ones that bring the system to more equilibrium states. This analysis showed that pricing without any relocation strategies can improve the balance of the system, but with less demand served, which might create less attractiveness in the service. Investigating trip pricing combined with relocation policies should be investigated.

The model proposed in this paper differs in many respects from the papers mentioned above and complements the existing literature. In particular, in this paper we are introducing an integrated optimization-simulation framework for vehicle and personnel relocations for one-way electric carsharing systems with reservations.

The proposed optimization framework is a multi-objective mixed integer linear programming problem (MMILP) with hierarchical objective functions. While in many multi-objective problems, the objective functions are combined in a single objective function after applying variable weights, we prefer to choose a hierarchical approach to ensure the satisfaction of the objectives according to the importance assigned to them by the decision makers. Our results highlight that there is a large multi-optimal solution set for each objective function that allows a meaningful application of the hierarchical approach. More discussion is provided later. In order of importance, we intend to maximize the demand of trips served and to minimize the relocation cost after rental.

Demand served is prioritized over relocation cost and profit because we would like to provide high quality of service and maximize the vehicle availability for the users especially when we are dealing with a system with reservations. The treatment of a multi-objective problem allows for interesting policy implications and different types of relocation strategies to be studied.

The problem formulation allows to determine the optimal temporal and spatial distribution of vehicles in stations and also the personnel responsible for the relocation. To deal with the computational complexity of the problem, physical characteristics are considered when designing the mathematical formulation. A simulation module is integrated in the optimization in order to ensure the feasibility of the optimization outcome in terms of vehicle recharging requirements. A direct integration of charging levels into the optimization model will not allow for solving large problem instances. As we formulate the problem as a network with nodes and arcs, we are not using individual stations to assign vehicles and personnel. Instead we consider clusters of stations based on travel times in arcs, that will aggregate stations with similar characteristics during the optimization. More details about clustering are provided in Section 3.5. Station clustering is used to reduce the computational complexity of the problem. The remainder of the paper is organized as follows: Section 3 presents the methodological framework involving an optimization and a simulation module. Section 4 provides the results of implementing and solving the problem with data from a real case-study for different system realizations and relocation policies. The last section provides a discussion and further future directions.
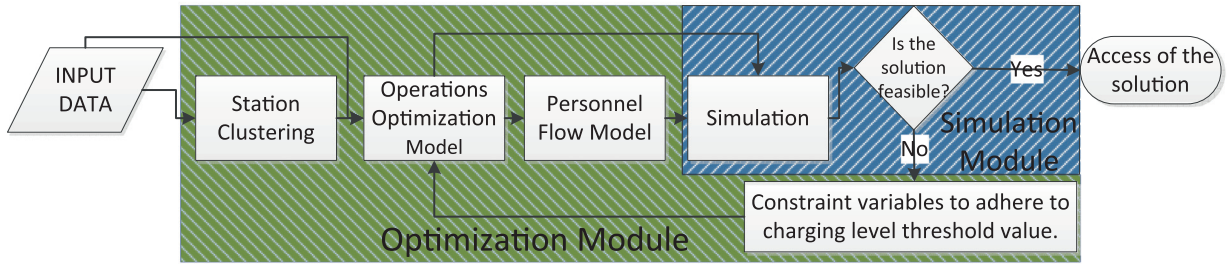
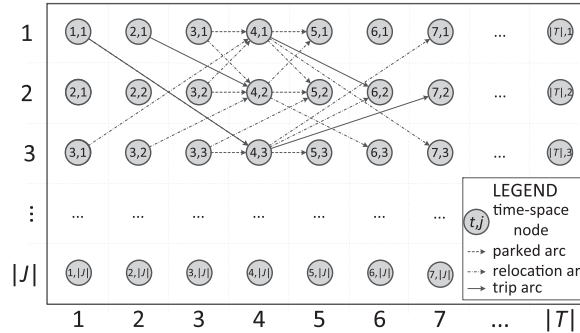**Fig. 1.** Flow chart for the solution framework.

Fig. 2. Time augmented vehicle network for a set of stations (J) and time intervals (T). Different types of arcs (i.e. park, trip and relocation) are shown starting and ending from the nodes at time interval 4.

3. Methodological framework

The proposed framework consists of the following two modules: (i) optimization, and (ii) simulation. Within the optimization module, a clustering algorithm is developed which creates clusters of stations in order to reduce the dimensionality of the optimization module.

The optimization module involves three mathematical models: (i) station clustering, (ii) operations optimization and (iii) personnel flow. The operations optimization formulation does not explicitly account for battery levels and this is addressed indirectly through the simulation module. The output of the optimization module is used by the simulation module in order to test the feasibility of the optimization outcome in terms of vehicle recharging requirements. After the first iteration of the optimization process, if the results of the simulation are feasible then the optimization process is stopped. If the simulation results are not feasible, the vehicles that require further charging are restricted to stay in the station and the optimization model is solved again considering the new constraints. The structure of the proposed model and the interrelationship between its modules are illustrated in Fig. 1.

3.1. Modelling assumptions and concepts

The assumptions underlying the operations of the system are:

- The carsharing system is modelled as a network of stations. The state of each station at each time interval (described in the next bullet) is the number of vehicles parked.
- The operating time of the carsharing system is divided into equal time intervals (15 min). Using the network of stations, a time-augmented network is generated by including all station-time interval pairs and repeating the network of stations for each time interval.
- Every trip and relocation with/without vehicle takes at least one time interval.
- Each vehicle can have only one state at each time interval: parked, serving trip (called trip from now on) or under relocation (called relocation from now on). These 3 states appear as arcs in Fig. 2. These 3 states are defined as follows:

  Parked: Vehicle is available waiting at one of the stations.
  Trip: Vehicle is picked-up by one of the users and will be unavailable until it is dropped-off.
  Relocation: Vehicle is relocated by one of the personnel and will be unavailable until its relocation is ended.

- Each (relocation) personnel can have only one state at each time interval: idle, driving or moving.

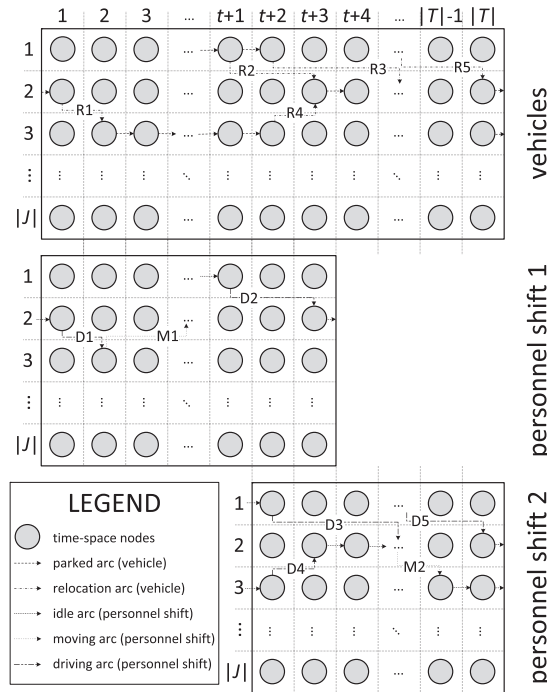 **Idle:** Personnel is available waiting at one of the stations.

Fig. 3. Time augmented vehicle and personnel shift networks for a set of stations (*J*) and time intervals (*T*). Different types of arcs are used for different operations in different networks. Parked (P) and relocation (R) arcs are used in vehicle's network. Idle (I), moving (M) and driving (D) arcs are used personnel shifts network. Trip arcs for the vehicles are excluded from the figure to keep it simpler.

Driving: Personnel is relocating one of the vehicles and will be unavailable until the relocation is completed.

Moving: Personnel is moving from his/her previous station to another station and will be unavailable until the moving is finalized

These assumptions allow the realistic representation of the problem and facilitate its mathematical formulation. In addition to the stated assumptions the following concepts underpin the mathematical formulation of the problem.

- The operations of vehicles and personnel are described in time-augmented networks, a single network for cars and separate networks, one for each personnel shift, are created. Interactions between the networks are described in the next bullet points.
- In the time-augmented networks, each node is defined with two attributes, i.e. station and time interval, and represents the number of vehicles or personnel depending on the network, for given time interval (Figs. 2 and 3). The arcs of the networks represent the states of vehicles and personnel. In Fig. 2, only the arcs starting and ending from nodes at time interval 4 are shown. For example, vehicle is rented from station 1 at time interval 1 and is returned to station 3 at time interval 4, where it is parked afterwards. Note that relocation arcs in the vehicle network of Fig. 3 are always associated with a driving arc in personnel networks of Fig. 3. Time augmented networks of personnel contain only nodes related to the time intervals that the personnel works (Fig. 3).
- Since every operation (e.g. trip, driving, moving) starts at the beginning and ends at the end of a time interval, a vehicle (or a personnel) starting the operation at time interval *t* from node *j* has to be at node *j* at the end of the previous time interval $t - 1$. Under the same reasoning, a vehicle (or a personnel) ending the operation at time interval *u* at node *l* is available for a new operation the earliest at time interval $u + 1$.
- When a trip is accepted to be served from station *j* to *l*, starting at time interval *t* and ending at time interval *u*, a unit flow is send from nodes (*j, t*) to (*l, u*) in the vehicles' network. In Fig. 3, 5 trips are shown with arcs T1-T5.
- When a driving is executed for a vehicle at station *j* to *l*, starting at time interval *t* and ending at time interval *u*, a unit flow is send from nodes (*j, t*) to (*l, u*) in both the vehicle and personnel networks. In Fig. 3, 5 drivings are shown with arcs R1-R5 in the vehicle network. Those drivings are executed by personnel from shift 1 (R1-R2) and shift 2 (R3-R5). Note that, there are driving arcs (D1-D2 and D3-D5 respectively) starting exactly from the same nodes in the personnel network by whom the driving is executed.
- In addition to driving, relocation personnel can change station by moving (i.e. they can ride to a vehicle driven by someone else). This type of operation is labelled with M1 and M2 in Fig. 3. When a member of staff from personnel shift *s* is moving from station *j* to *l* starting at time interval *t* and ending at time interval *u*, a unit flow is sent from (*j, t*) to (*l*,

u) in the personnel network of the shift s. If a personnel from shift s does not change its location at time t and stays in station j, a unit flow is sent from stay arc from (j, t) to (j, $t+1$).

- In the exact operations optimization model, all possible relocation, i.e. driving and moving, arcs are created. Note that, the duration of all these operations are parameters of the model. In other words, for any given relocation, driving and moving arc from station j to station l starting at time interval t, there is only one end time interval u.
- In order to decrease the number of variables in the operations optimization model (see Fig. 1), stations are clustered for driving and moving operations. A second layer is added for both driving and moving operations and each of these operations is handled at the cluster level. In other words, in the operations optimization model, driving (moving) from station j to station l from time interval t to u is done in three "artificial" states: (i) from origin station j to origin station's cluster b at time interval t, (ii) from origin station's cluster b to destination station's cluster d starting at time interval t and ending at time interval u and (iii) from destination stations cluster d to destination station l at time interval u. Any relocation between a cluster and a station is named *intra-relocation* and any relocation between two clusters is named *inter-relocation*.
- Clusters of stations are created using an approach similar to k-Medoid algorithm (Kaufman and Rousseeuw, 1987). Further information about the algorithm is given in Section 3.5. Since clustering depends only on the network topology, the clustering algorithm is the very first algorithm to be executed and is not updated with time. The clustering algorithm is applied for driving and moving arcs separately since different speeds and routes are used in these two operations (see Fig. 1).
- After solving the operations optimization model, the personnel flow model is used to create personnel assignments from flows between clusters and between clusters and stations for any given time interval.
- Instead of forcing every vehicle to be fully charged before rental, charging levels of vehicles are controlled with a simulator. The operations optimization model, the personnel flow model, and the simulation model are run iteratively until a charging level feasible solution is obtained. If there is an infeasibility in the charging level of any vehicle, this vehicle is forced to be charged before it is assigned to its next task (i.e. relocation or trip). The flow chart of the entire solution framework can be seen in Fig. 1.

3.2. Operations optimization model

3.2.1. Sets and indices
$i \in I$: trips
j and $l \in J$: stations
t, u and $w \in T$: time intervals
$s \in S$: working shifts
b and $d \in B$: station clusters

3.2.2. Parameters
start/end(s): start/end time intervals of working shift s
start/end(i): start/end time intervals of trip i
origin/dest(i): origin/destination stations of trip i
c.end(i): the last interval that vehicle of trip i can be charged
d.end/m.end(t, b, d): the time interval that driving/moving from cluster b to d ends if it starts at time interval t
PC_s: cost of a relocation personnel for working shift s
RC_j: intra-relocation cost from/to station j
$\widetilde{RC}_{bd}$: inter-relocation cost from a station in cluster b to a station in cluster d
SG: (safety gap) number of time intervals the drop-off spot is kept reserved at the destination station of a trip starting from its stated end time
I_c: set of trips after which the vehicle serving these trips should be charged

3.2.3. Variables
z_i: binary variable indicating if trip i is served or not
v_s: number of personnel used from shift s
n_j^t: number of vehicles in station j at the beginning of time interval t
$\tilde{n}_j^t$: number of vehicles in station j forced to be charged during time interval t
m_{sj}^t: number of personnel from shift s in station j at the beginning of time interval t
r_{sj}^t: number of vehicles relocated by personnel of shift s from station j starting at time interval t
$\bar{r}_{sj}^t$: number of vehicles relocated by personnel of shift s to station j finishing at time interval t
p_{sj}^t: number of personnel of shift s moved from station j starting at time interval t
$\bar{p}_{sj}^t$: number of personnel of shift s moved to station j finishing at time interval t
$\tilde{r}_{sbd}^{t(u)}$: number of vehicles relocated by personnel of shift s from a station in cluster b to a station from cluster d started at time interval t and ending at time interval u

$\tilde{p}_{sbd}^{t(u)}$: number of personnel of shift s moved from a station in cluster b to a station from cluster d started at time interval t and ending at time interval u

3.2.4. Formulation

$$\max \sum_i z_i \tag{1}$$

$$\min \quad \overbrace{\sum_{b,d} \widetilde{RC}_{bd} \tilde{r}_{sbd}^{t(u)}}^{\text{inter-relocation cost}} + \overbrace{\sum_{j,t} RC_j \left(r_{sj}^t + \bar{r}_{sj}^t \right)}^{\text{intra-relocation cost}} + \overbrace{\sum_s PC_s v_s}^{\text{personnel cost}} \tag{2}$$

$$\text{s.t.} \quad n_j^{t+1} = n_j^t - \sum_{\substack{i:\text{start}(i)=t \\ \text{origin}(i)=j}} z_i - \sum_s \bar{r}_{sj}^t \sum_{\substack{i:\text{end}(i)=t \\ \text{dest}(i)=j}} z_i + \sum_s r_{sj}^t \quad \forall j, t \backslash t_{\text{last}} \tag{3}$$

$$n_j^t \geq \sum_{\substack{i:\text{start}(i)=t \\ \text{origin}(i)=j}} z_i + \sum_s \bar{r}_{sj}^t \quad \forall j, t \tag{4}$$

$$\sum_j m_{sj}^{\text{start}(s)} = v_s \quad \sum_j m_{sj}^{\text{end}(s)} = v_s \quad \forall s \tag{5}$$

$$m_{sj}^{t+1} = \overline{m}_{sj}^t - \overline{p}_{sj}^t - \bar{r}_{sj}^t + p_{sj}^t + r_{sj}^t \quad \forall j, s, t \backslash t_{\text{last}} \tag{6}$$

$$m_{sj}^t \geq \overline{p}_{sj}^t + \bar{r}_{sj}^t \quad \forall j, s, t \tag{7}$$

$$\sum_d \tilde{r}_{sbd}^{t(u)} = \sum_{j \in J_b} \bar{r}_{sj}^t \sum_d \tilde{r}_{sdb}^{u(t)} = \sum_{j \in J_b} r_{sj}^t \quad \forall s, b, t \tag{8}$$

$$\sum_d \tilde{p}_{sbd}^{t(u)} = \sum_{j \in J_b^s} \overline{p}_{sj}^t \sum_d \tilde{p}_{sdb}^{u(t)} = \sum_{j \in J_b^s} p_{sj}^t \quad \forall s, b, t \tag{9}$$

$$n_j^t \leq CAP_j^t \quad \forall j, t \tag{10}$$

$$n_j^t - \sum_{\substack{i:\text{start}(i)=t \\ \text{origin}(i)=j}} z_i - \sum_s \bar{r}_{sj}^t \geq \sum_{\substack{i:\,\text{dest}(i)=j,\, i \in I_c \\ \text{end}(i) < t \leq c.\text{end}(i)}} z_i \quad \forall j, t \tag{11}$$

$$n_j^t - \sum_{\substack{i:\text{start}(i)=t \\ \text{origin}(i)=j}} z_i - \sum_s \bar{r}_{sj}^t \geq \sum_{\substack{i:\,\text{dest}(i)=j \\ \text{end}(i) < t \leq \text{end}(i)+SG}} z_i \forall j, t \tag{12}$$

$$z_i \in \{0, 1\}; n_j^t, v_s \in \mathcal{N} \quad \forall i, j, t, s \tag{13}$$

$$m_{sj}^t, r_{sj}^t, \bar{r}_{sj}^t, p_{sj}^t, \overline{p}_{sj}^t, \tilde{r}_{sbd}^{t(u)}, \tilde{p}_{sbd}^{t(u)} \in \mathcal{N} \forall s, j, t, b \tag{14}$$

The proposed optimization model is a MMILP with lexicographically ordered objective functions. The model, involves in order of importance, the following two objectives: (i) maximize number of trips served and (ii) minimize relocation cost.

The first-priority objective of the model (Eq. (1)) is to maximize the number of trip requests served.

The second-priority objective of the model (Eq. (2)) is to minimize total relocation cost. This objective function has three components:

- Inter-relocation cost includes the fuel cost associated with relocation operations between clusters. Note that, when there is a relocation between two stations, it is assumed that each relocation is performed in three artificial states: (i) from origin station to origin station's cluster, (ii) from origin station's cluster to destination station's cluster and (iii) from destination station's cluster to destination station. This function is related to state (ii), i.e. relocation between clusters.
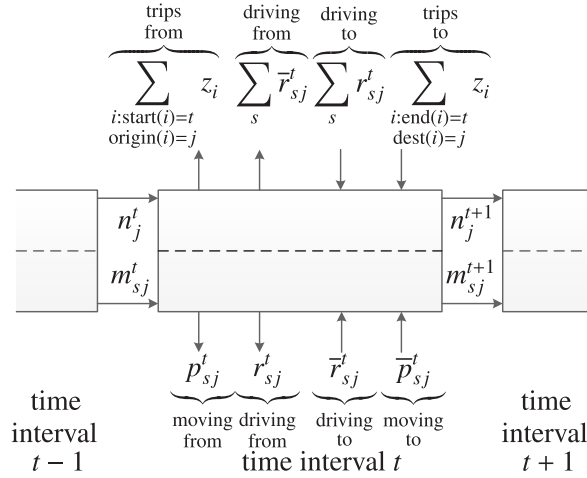
Fig. 4. Flow diagram of vehicles (top) and personnel shift s (bottom) in station j at time interval t.
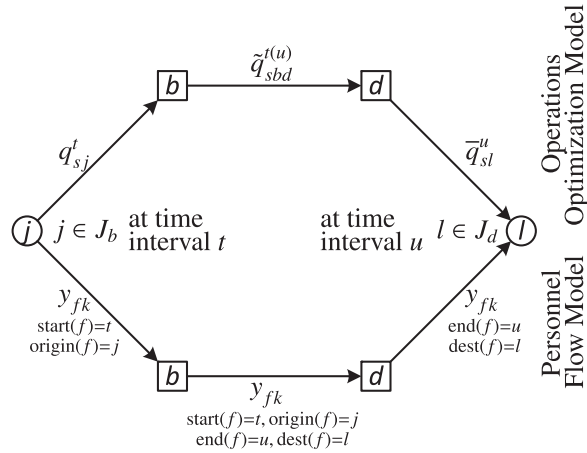

Fig. 5. The representation of flows (driving/moving) in two mathematical models. In the operations optimization model, there is at least one flows from station j to l from time interval t to u executed by personnel shift s. The count of these flows are kept with three variables: q_{sj}^t, $\tilde{q}_{sbd}^{t(u)}$ and $\bar{q}_{sl}^u$. In personnel flow model, this flow is performed by personnel k from personnel shift s. Only one binary variable is needed to describe this flow, y_{fk}.

- Intra-relocation cost includes the fuel cost associated with relocation operations between stations and clusters. In other words, these are the costs related to artificial states (i) and (iii) of the relocation operations described before.
- Personnel cost is the total cost of personnel used. It is assumed that there is a number of working shifts and the model determines the number of employees needed per shift.

Demand served is prioritized over relocation cost because we would like to maximize the availability of the system to the users especially when we assume that the trip has been already accepted and service is offered to the customer.

Constraints 3 are the flow conservation equations for vehicles for each station and time interval. For all stations and time intervals, the total number of vehicles at station j at the beginning of time interval $t + 1$ (n_j^{t+1}) is equal to the number of vehicles at the beginning of time interval t (n_j^t) minus the number of vehicles that have left station j at time interval t by trip ($\sum_{\substack{i:\text{start}(i)=t \\ \text{origin}(i)=j}} z_i$) and relocation ($\sum_s \bar{r}_{sj}^t$) plus the number of vehicles that have arrived to station j at time interval t by trip ($\sum_{\substack{i:\text{end}(i)=t \\ \text{dest}(i)=j}} z_i$) and relocation ($\sum_s r_{sj}^t$). Flow on the top of the rectangle at time interval t of Fig. 4 represents this relationship.

Constraints 4 ensure that there are enough number of vehicles to start trips and relocations at each station at each time interval. These constraints prevent vehicles to start trips and relocations at the time interval at which they end their trips and relocations. With these constraints, it is guaranteed that every vehicle can have one and only one state (i.e. parked, under rental, under relocation) at each time interval. For all stations and time intervals, the number of vehicles in station j at the beginning of time interval t (n_j^t) should be more than or equal to the number of vehicles that have left station j at time interval t by trip ($\sum_{\substack{i:\text{start}(i)=t \\ \text{origin}(i)=j}} z_i$) and relocation ($\sum_s \bar{r}_{sj}^t$).

Constraints 5 keep track of the number personnel allocated in each shift. In the first set of constraints, for all shifts, the total number of personnel started working at different stations at the beginning of the start time interval of shift s should be equal to the number of personnel allocated in the same shift. Similarly for the second set of constraints, for all shifts, the total number of personnel ended working at different stations at the end of the end time interval of shift s should be equal to the number of personnel allocated in the same shift.

Constraints 6 are the flow conservation equations for each shift, each station and time interval. For all shifts, stations and time intervals, the total number of personnel from shift s at station j at the beginning of time interval $t+1$ (m_{sj}^{t+1}) is equal to the number of vehicles at the beginning of time interval t (m_{sj}^{t}) minus the number of personnel that have left station j at time interval t by moving ($\overline{p}_{sj}^{t}$) or driving ($\overline{r}_{sj}^{t}$) plus the number of personnel that arrived the station j at time interval t by moving (p_{sj}^{t}) or driving (r_{sj}^{t}). Flow on the bottom rectangle at time interval t of Fig. 4 illustrates this relationship.

Constraints 7 ensure that there are enough number of personnel from each shift to execute moving and driving at each station at each time interval. These constraints prevent personnel to start drivings and movings at the time interval at which they end their drivings and movings. With these constraints, it is guaranteed that every personnel can have one and only one state (i.e. idle, moving, driving) at each time interval. For all shifts, stations and time intervals, the number of personnel from shift s in station j at the beginning of time interval t (m_{sj}^{t}) should be more than or equal to the number of personnel that have left station j at time interval t by driving ($\overline{p}_{sj}^{t}$) and moving ($\overline{r}_{sj}^{t}$).

Constraints 8 are used to preserve flow conservation between clusters and stations of clusters for driving operations. In the first set of constraints, for all shifts, clusters and time intervals, the total number of personnel of shift s started driving from the stations of cluster b at time interval t should be equal to the number of personnel driving from all stations of cluster b (J_b^s) at the same time interval. Similarly, in the second constraint, for all shifts, clusters and time intervals, the total number of personnel of shift s ended driving at the stations of cluster b ending at time interval t should be equal to the number of personnel driving to all stations of cluster b (J_b^s) at the same time interval.

Constraints 9 are used to preserve flow conservation between clusters and the stations of clusters for moving operations. In the first set of constraints, for all shifts, clusters and time intervals, the total number of personnel of shift s started moving from the stations of cluster b starting at time interval t should be equal to the number of personnel moving from all stations of cluster b (J_b^s) at the same time interval. Similarly, in the second set of constraints, for all shifts, clusters and time intervals, total number of personnel of shift s ended moving at the stations of cluster b ending at time interval t should be equal to the number of personnel moving to all stations of cluster b (J_b^s) at the same time interval.

Constraints 10 set the capacity restrictions for each station j at each time interval t.

Constraints 11 are used to keep track of charging vehicles after rentals. As it is stated in Section 3, we use simulation to check the charging level feasibility of the solution. The output of the simulation is used to update the set I_c. Feedback from the simulation adds the trips with infeasible starting charging levels to this set. The right hand sides of these constraints are only applicable to the trips in this set. In the mathematical model, we force vehicles to be charged after these trips ($i \in I_c$) if they are served (i.e. $z_i = 1$). The constraints are not binding for unserved trips (i.e. $z_i = 0$). We satisfy these constraints by forcing the number of parked vehicles after these trips at the destination stations during the charging period (end(i) $< t \le$ c.end(i)). Note that, for a given station and time interval, the trips and relocations departing from the station are subtracted from the number of vehicles at the beginning of the time interval at the left hand side of the constraint. Without this update, the model can assign any charging vehicles for a trip or a relocation and cover its place with another vehicle arriving to the same station without violating these constraints.

Constraints 12 are used to apply the safety gap of vehicle returning time provided for the users. These constraints are used to reserve destination stations for the duration of the safety gap (SG) which is defined in the units of time intervals. For instance, if SG is 2 and the time interval length is 15 min then the destination spots of every accepted trips (i.e. $z_i = 1$) is reserved starting from the end time of the trip for 30 min. The constraints are not binding for unserved trips (i.e. $z_i = 0$). With these constraints we guarantee two things: (i) There will always be a spot available for the vehicle starting from its stated trip end time for a period of the safety gap. (ii) The returned vehicle will not be relocated or rented until the end of the safety period, i.e. trip end time plus safety gap. We satisfy these two conditions by forcing the vehicle to stay at the station from the end of the trip for a duration of the safety gap. These constraints work very similar to constraints 11. Constraints 12 are applied to all trips for a period of safety gap whereas 11 are applied to only trips in set I_c until the charging end time.

Constraints 13–14 are used to define the domains of each variable: z_i's are restricted to be binary variables, the rest of them are positive integers.

3.3. Personnel flow model

For each working shift s, the following model is solved with the values acquired from the operations optimization model (Eqs. (15)–(20)).

3.3.1. Sets and indices
$k \in K$: personnel
$j \in J$: stations

b and $d \in B$: clusters

$t \in T$: time intervals

$f \in F$: flows

3.3.2. Parameters

m_{sj}^t: number of personnel of shift s in station j at the beginning of time interval t

r_{sj}^t: number of personnel of shift s starting driving (state D) from station j at time interval t

$\bar{r}_{sj}^t$: number of personnel of shift s finishing driving (state D) to station j at time interval t

p_{sj}^t: number of personnel of shift s starting moving (state M) from station j at time interval t

$\bar{p}_{sj}^t$: number of personnel of shift s finishing moving (state M) to station j at time interval t

$\tilde{r}_{sbd}^{t(u)}$: number of personnel of shift s driving from a station in cluster b to a station from cluster d starting at time interval t and ending at time interval u

$\tilde{p}_{sbd}^{t(u)}$: number of personnel of shift s moving from a station in cluster b to a station from cluster d starting at time interval t and ending at time interval u

v_s: number of personnel used from shift s

start/end(f): start/end time interval of flow f

origin/dest.(f): origin/destination stations of flow f

3.3.3. Variables

x_{jk}^t: binary variable indicating if personnel k is in node j at the beginning of time interval t or not

y_{fk}: binary variable indicating if personnel k follows flow f or not

3.4. Formulation

A prerequisite for setting-up the personnel flow model is to run the operations optimization model in order to obtain the optimal values of the flows r_{sj}^t, $\bar{r}_{sj}^t$, $\tilde{r}_{sbd}^{t(u)}$, p_{sj}^t, $\bar{p}_{sj}^t$ and $\tilde{p}_{sbd}^{t(u)}$. A flow contains two intra-flows and one inter-flow of the same type, i.e. driving or moving. In other words, in order to generate a flow from two intra- and one inter-flows, (i) the first inter-flow's time interval should be the starting time interval of the intra-flow, (ii) the station of the first intra-flow should be element of the origin cluster of the intra-flow, (iii) the second inter-flow's time interval should be ending time interval of the intra-flow, and (iv) the station of the second intra-flow should be element of the destination cluster of the intra-flow. All these generated flows constitute set *F*. In order to simplify representation, all flows, drivings (r) and movings (p), are represented with q.

$$\min \sum_{f,k} \text{dist}(f) y_{fk} \tag{15}$$

$$\text{s.t.} \sum_{k} x_{jk}^t = m_{sj}^t \quad \sum_{j} x_{jk}^t \le 1 \quad \forall j, k, t \tag{16}$$

$$x_{jk}^{t+1} = x_{jk}^t - \sum_{\substack{f: \text{start}(f)=t \\ \text{origin}(f)=j}} y_{fk} + \sum_{\substack{f: \text{end}(f)=t \\ \text{dest}(f)=j}} y_{fk} \quad \forall j, k, t \tag{17}$$

$$\sum_{\substack{f: \text{start}(f)=t \\ \text{origin}(f)=j}} y_{fk} = \bar{q}_{sj}^t \quad \sum_{\substack{f: \text{end}(f)=t \\ \text{dest}(f)=j}} y_{fk} = q_{sj}^t \quad \forall j, t \tag{18}$$

$$\sum_{\substack{f: \text{start}(f)=t, \text{origin}(f) \in J_b \\ \text{end}(f)=u, \text{dest}(f) \in J_d}} y_{fk} = \tilde{q}_{sbd}^{t(u)} \quad \forall b, d, t \tag{19}$$

$$x_{jk}^t, y_{fk} \in \{0, 1\} \quad \forall f, j, k, t \tag{20}$$

The objective of the personnel flow model (Eq. (15)) aims to minimize total distance travelled. We can minimize the total difference between real relocation time and assumed relocation time with clustering, which is discussed later.

First set of constraints 16 count the number of personnel from shift s at each station j at each time interval t and sets this number equal to the number calculated by the operations optimization model, i.e. x_{jk}^t. Second set of constraints 16 require that, every personnel can be at most in one station at each time interval t. Note that, the LHS of these constraints not necessarily equal to 1 since there may be driving and moving operations which take more than one time interval.

Constraints 17 are the flow conservation equations for individual personnel. If personnel k is at node j at the beginning of time interval t and he/she does not start driving or moving to another station during that period, then he/she should be in station j at the beginning of time interval $t + 1$. On the other hand, if personnel k is not at node j at the beginning of time interval t and he/she performs driving or moving to station j from any station finalizing at time interval t, he/she should be in station j at the beginning of time interval $t + 1$.

Constraints 18 and 19 are used to cover the same flow operations found by operations optimization model in the previous step. Constraints 18 ensure that, the total number of flows assigned to all personnel from shift s should exactly cover all first and second intra-flows with/without vehicle. Similarly, constraints 19 ensure that the total number of flows assigned to all personnel from shift s should exactly cover all inter-flows with/without vehicle.

3.5. Clustering algorithm

Although it is possible to develop a mathematical model without using clusters, this would be computationally very demanding. Since we allow dynamic relocations throughout the day, we need to create in the order of $O(|T||J|^2)$ number of variables for driving and moving operations only in which T is the set of time intervals and J is the set of stations. In order to decrease the number of variables and solution time as a result, we develop clusters and assume relocations are happening between clusters. This approach decreases the number of driving and moving variables to the order of $O(|T||C|^2 + 2|T||C|)$ in which C is the set of clusters. Note that, if every cluster is set with one and only one stations, the proposed model is equivalent to the mathematical model without clusters.

A clustering model is developed to find clusters that minimize the error created by aggregating stations to a given number of clusters. Since our model works with time intervals, the objective of the clustering model is to minimize the sum of squares of the deviation of relocation time of every pair of different stations. If $T(j, l)$ is defined as the travel time from station j to station l and $\left(T_{min}^B(j, l), T_{max}^B(j, l)\right]$ is the interval of the duration of relocation from cluster of station j to cluster of station l; positive (dev_B^+) and negative (dev_B^-) deviations' can be calculated using Eqs. (21)–(22). Note that if travel time from origin to destination station is in the interval of the duration of relocation from the cluster of origin to the cluster of destination station, both positive and negative deviations are zero.

$$dev_B^+(j, l) = max\left\{T(j, l) - T_{max}^B(j, l), 0\right\} \tag{21}$$

$$dev_B^-(j, l) = max\left\{T_{min}^B(j, l) - T(j, l), 0\right\} \tag{22}$$

The algorithm developed for clustering is similar to k-Medoid algorithm (Kaufman and Rousseeuw, 1987). The algorithm starts with an initial solution and iterates until the stopping criterion, i.e. iteration count, is met. This problem is quite complex to solve with an exact algorithm. We prefer k-Medoid instead of k-means since former aims to minimize dissimilarities within clusters whereas latter minimizes sum of squared Euclidean distances. Since the time is discretized in the mathematical model, our "similarity" function is discrete and k-Medoid algorithm allows us to reflect this on our objective function.

Initial solution is created step by step by assigning each station consecutively to the limited number of clusters. In every iteration, the station is added to the cluster that minimizes the objective function, the sum of positive and negative deviations.

After the initial solution is created, the configuration that improves the objective function is found with consecutive exchange of stations to different clusters. Each exchange is executed in two steps. First, we remove each station from their clusters iteratively to see which removal of the station improves the objective function most. Then this station is added to each of the clusters iteratively except its previous cluster. We calculate the objective function improvement at each addition to see to which cluster we should add the station to improve objective function most. After each exchange operations, best solution is updated with the current solution if current solution has smaller objective function than the best solution so far. The best solution is reported when one of the stopping criterion is met. The pseudocode of the clustering algorithm can be seen in Algorithm 1 for the set of stations (J), the maximum number of clusters (n_B), the number of iterations (n_c) and the objective function value for clusters B ($\mathbb{F}(B)$). The set of clusters ($B^\star$) with the best objective function is returned at the end of the algorithm. Results of the clustering algorithm are presented in Section 4 of the paper.

3.6. Simulation module

For carsharing systems with combustion engine vehicles, the solutions generated by the optimization module are always feasible. However, this is not the case for systems with electric vehicles. The charging level of each vehicle at each time interval should be controlled in order to satisfy charging level feasibility of the solution. The charging level feasibility means that a given vehicle has enough battery to allow it to fulfill the next trip. For this purpose, a simulation module is used by the solution framework.

The purpose of the simulation module is to check the feasibility of the solution generated by the optimization module. If the solution is feasible, it is accepted. If it is not feasible, the simulation module returns the infeasible vehicle with its assigned rentals. The optimization module puts a charging restriction for this vehicle and forces it to be charged. In other

---

Algorithm 1: Pseudocode for the clustering algoritm described in Section 3.5.

Input: J, n_B, n_c and $\mathbb{F}(B)$
Output: $B^\star$
// Initialization
for $b \leftarrow 1$ **to** n_B **do**
  | $B_b \leftarrow \varnothing$ // initialize clusters
end
for $j \leftarrow 1$ **to** $|J|$ // for all stations
do
  | $Z \leftarrow \infty$ // initialize objective
  | **for** $b \leftarrow 1$ **to** n_B // for all clusters
  | **do**
  | | $B_b \leftarrow B_b \cup \{j\}$ // add
  | | **if** $\mathbb{F}(B) < Z$ **then**
  | | | $b^\star \leftarrow b$ and $Z \leftarrow \mathbb{F}(B)$ // update selected
  | | **end**
  | | $B_b \leftarrow B_b \setminus \{j\}$ // remove
  | **end**
  | $B_{b^\star} \leftarrow B_{b^\star} \cup \{j\}$ // assign to selected cluster
end
$B^\star \leftarrow B$ and $Z^\star \leftarrow \mathbb{F}(B)$ // initialize best
// Iteration
for $c \leftarrow 1$ *to* n_c **do**
  | // Find station to be removed
  | $Z \leftarrow \infty$ // initialize c.best objective
  | **for** $b \leftarrow 1$ **to** n_B // for all clusters
  | **do**
  | | **foreach** $j \in B_b$ // for all stations
  | | **do**
  | | | $B_b \leftarrow B_b \setminus \{j\}$ // remove
  | | | **if** $\mathbb{F}(B) < Z$ **then**
  | | | | $j^\star \leftarrow j, b^\star \leftarrow b$ and $Z \leftarrow \mathbb{F}(B)$ // update
  | | | **end**
  | | | $B_b \leftarrow B_b \cup \{j\}$ // add
  | | **end**
  | **end**
  | $B_{b^*} \leftarrow B_{b^*} \setminus \{j^*\}$ // remove from selected
  | // Find cluster to be added
  | $Z \leftarrow \infty$ // initialize objective
  | **for** $b \leftarrow 1$ **to** n_B **do**
  | | **if** $b = b^\star$ **then**
  | | | **continue** // skip if the same
  | | **end**
  | | $B_b \leftarrow B_b \cup \{j^\star\}$ // add
  | | **if** $\mathbb{F}(B) < Z$ **then**
  | | | $b^\star \leftarrow b$ and $Z \leftarrow \mathbb{F}(B)$ // update cluster
  | | **end**
  | | $B_b \leftarrow B_b \setminus \{j^\star\}$ // remove
  | **end**
  | $B_{b^\star} \leftarrow B_{b^\star} \cup \{j^\star\}$ // assign to selected
  | // Update best solution
  | **if** $Z^\star < \mathbb{F}(B)$ **then**
  | | $B^\star \leftarrow B$ and $Z^\star \leftarrow \mathbb{F}(B)$
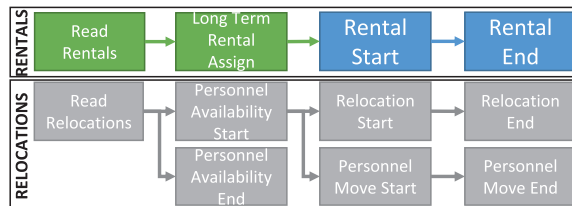 | **end**
end

Fig. 6. Simplified structure of events and their links to each other in the simulator.
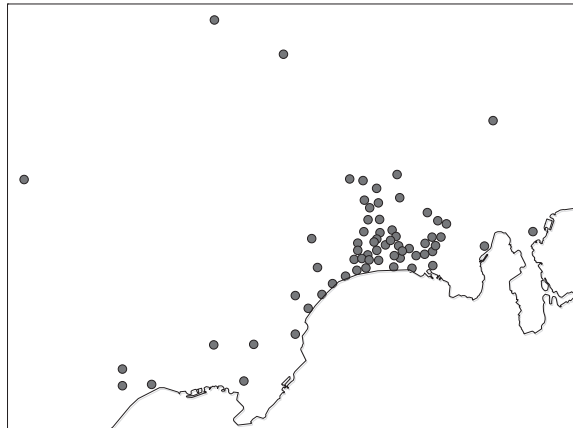

Fig. 7. Stations of the carsharing system of the city of Nice, France.

words, if there is a vehicle to be rented with a charging level below threshold, we force the vehicle which serves the same demand to stay longer at the station that was previously rented in order to be charged properly. It is worth noting that this strategy terminates within a limited number of iterations. The maximum number of iterations is not more than the total number of requested trip, but during the implementation we notice that occurs in reality with a much smaller number of iterations in most cases.

In order to find the optimal charging strategy, the model needs to test all possible charging restrictions which requires exponential number of runs. That is the reason, we select the vehicle with infeasible charging and apply forced charging to it. Furthermore, the strategy always returns a feasible solution. The duration of the charging is just enough to replenish the electric energy consumed in the last trip. Note that, forced charging constraint is active only if a demand for a trip is served. In the worst case, there might be a request for charging after every trip. In such a case, any solution would be charging level feasible since this will be a model with all forced charging restrictions. A system with a restriction of fully charging after every trip would be a carsharing system which only allows fully charged vehicles to be rented. This strategy is not preferred at the very first step because it sharply decreases the capacity of the system (the availability of vehicles).

We have developed a discrete event simulator, which is described in more details in Repoux et al. (2015). The general structure is included in Fig. 6. Rentals and relocations are fed to the simulator (see Read Rentals and Read Relocations events) as parameters. Every rental and relocation of vehicles, and movements of personnel are simulated with the help of this simulator.

4. Model application

The solution framework represented with the flow chart in Section 3 (see Fig. 1) was applied to support operational decisions for the one-way electric carsharing system in Nice, France operated by VENAP Auto Bleue (see Fig. 7). VENAP Auto Bleue has been operating a round-trip system with 66 stations since 2011. In 2014, the system provided service to one-way trips. In our model application, anonymised rental and station location data[1] provided by VENAP Auto Bleue are used. The demand data from individual days are aggregated. For different demand levels, subsets with different number of trips are generated randomly (for also confidentiality purposes).

The whole model including the simulator, clustering algorithm and mathematical models were implemented in C# .NET environment. IBM ILOG Cplex Version 12.6 with Concert Technology was used for solving MILPs.

---

[1] The full dataset cannot be shared publicly as it contains sensitive commercial and private data. Readers could contact the authors to obtain a subset of data for research purpose.

The duration of each time interval is set to 15 min. It is assumed that relocation personnel move between stations either with a car (if a vehicle is relocated) or with a bike (if a vehicle is not relocated). The driving and cycling distances were estimated using Google Maps API (Google, 2015). Driving and cycling speeds are set to 30 and 15 km per hour respectively. Different speeds could be integrated in the optimization module if historical (or real time) traffic data can be provided to test the effect of traffic congestion in the system operations.

4.1. Clustering stations

First, separate clusters for driving (D) and moving (M) are created. The aim of the clustering algorithm is to increase the efficiency of the optimization algorithms by reducing the dimensionality of the problem. Note that, the difference between the durations of relocations with (driving) and without (moving) vehicles results in two separate clusters. The clustering algorithm described in Section 3.5 (see Algorithm 1) is applied for different number of clusters.

The results for the different number of clusters are shown in Fig. 8 for both driving (Fig. 8(a)–(e)) and moving (Fig. 8(f)) operations. In these figures, the numbers on the first row and column show the station numbers. Values in the upper triangle show the travel duration between the two stations (see the first row and column) in seconds. This value is the largest of the two travel durations between every pair of stations. Values in the lower triangle are the number of time intervals needed to travel between the two stations, simply the smallest integer larger than the value symmetric to the diagonal divided by 900, i.e. the time interval length in seconds. To facilitate the visualization of the figure, colour of each cell shows the number of time intervals needed to travel between each pair of nodes.

Each partition is separated with horizontal and vertical lines. For instance, in Fig. 8(a) there are four clusters. The first cluster has 50 stations starting with station 0 and ending with station 59; the second cluster has seven stations: 14, 15, 54, 55, 62, 63 and 65; and so on and so forth.

In the runs, we assume that the travel time between two stations is equal to the biggest difference of the travel time between the pairs of stations of the two clusters. With this assumption, the clustering algorithm always produces feasible relocation operations. We have opted to overestimate relocation durations instead of ending up with infeasible solutions due to the underestimation of relocation durations. As a result we never have a positive difference (21) for any station pairs. Our objective only minimizes the sum of negative differences (22) between all pairs. For instance, any travel between the stations of cluster 1 and the stations of cluster 2 takes three time intervals since for some pairs of stations from cluster 1 and 2, travelling takes three time intervals (see orange cells). Travels from the stations of cluster 2 to the stations of cluster 2 takes four time intervals because of some red cells, for station pairs {15, 62} and {55, 62}.

The score of the cluster in Fig. 8(a), which is calculated with the help of algorithm described in Section 3.5, is 3302.7. This algorithm calculates the sum of square errors expressed in terms of time intervals squared between all pairs of different stations because of the clustering which is represented by the vertical axes in Fig. 9.

There is a trade-off between the efficiency and the accuracy in selecting the number of clusters to be used for relocations with/without vehicles. Higher number of clusters results in less overestimation in relocation durations but increases the number of variables and as a result the model complexity. Lower number of clusters results in higher overestimation in relocation times but decreases the number of variables and results in shorter run times. In order to decide the most suitable cluster sizes, the clustering algorithm is executed for various number of clusters. Scores of each run for relocations with (Fig. 9(a)) and without (Fig. 9(b)) are shown in Fig. 9. We have selected the elbows in both cases as the best trade-off points that are able to decrease 90% of the error compared to 2 clusters: 6 clusters for relocations with vehicles (Fig. 8(b)) and 10 clusters for relocations without vehicles (Fig. 8(f)). In Fig. 8, the results for the different number of clusters can also be seen. The maximum value in the *x*-axes (39 for driving and 66 for moving case) shows that number of clusters needed to have 0 error in the relocation times.

After all experiments, the error of using 6 and 10 clusters for relocations with/without vehicle for each run is also calculated. On average, the duration of relocations with vehicles is overestimated by 5.97% more. The 95th percentile of the error at each instance was 13.82%. These values were 10.93% and 19.10% respectively for relocations without vehicles. Regarding the problem size and the efficiency of the method, these errors are acceptable. Note also that, this error could be decreased if more clusters were used but this was not preferred to have solutions within reasonable run times. Definitely there is a trade-off between the accuracy of the solution obtained and the time needed to solve the problem. For this particular case we have opted to use a number of clusters that yields a fairly accurate solution in order to have a reasonable solution time for the operational decisions. In general it is up to the decision maker to select the number of clusters that will be used and the associated accuracy.

4.2. Case study settings

After deciding on the clusters, demand samples with different number of rentals were created. We tested samples of size 200, 250 and 300 rentals per day. For each demand level, 30 different samples were created. Each sample is tested with two different running configurations (Type 1 and 2). In Type 1 configuration, the entire demand is considered known a priori and it was fed to the model all at once and solved all in a single step. In Type 2 configuration, the demand is received by the model one by one and accepted/rejected after each iteration without reconsidering the past decisions. Every rental request is accepted or rejected in the iteration it is received. These decisions are ultimate and cannot be changed in further
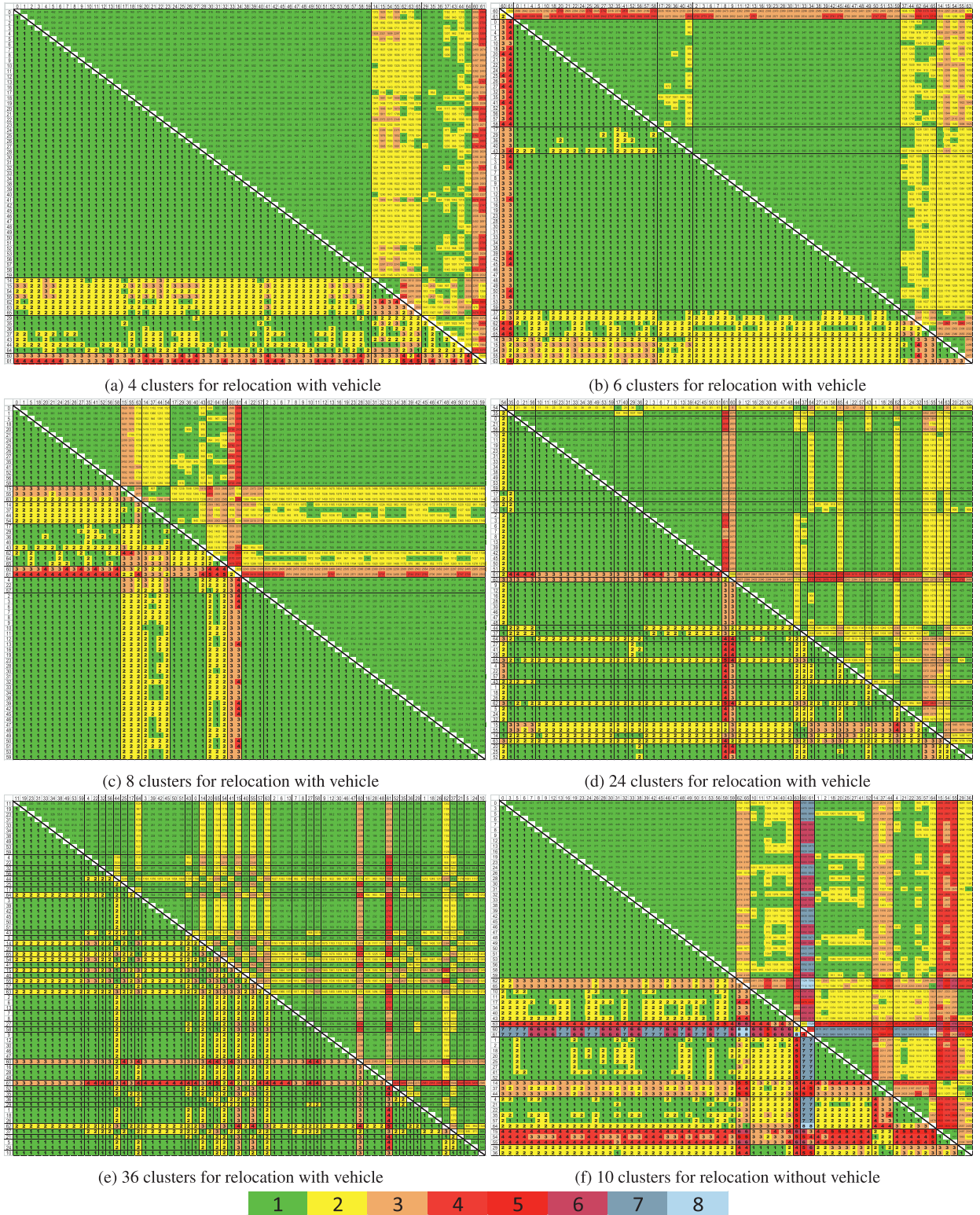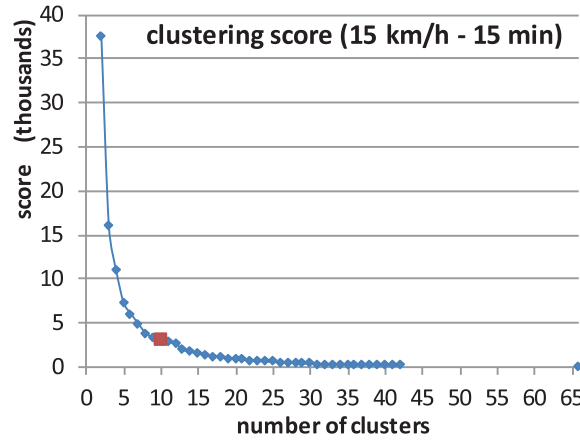
(a) 4 clusters for relocation with vehicle

(b) 6 clusters for relocation with vehicle

(c) 8 clusters for relocation with vehicle

(d) 24 clusters for relocation with vehicle

(e) 36 clusters for relocation with vehicle

(f) 10 clusters for relocation without vehicle

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

Fig. 8. Output of clustering algorithm for different number of clusters. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(a) Relocations with vehicle
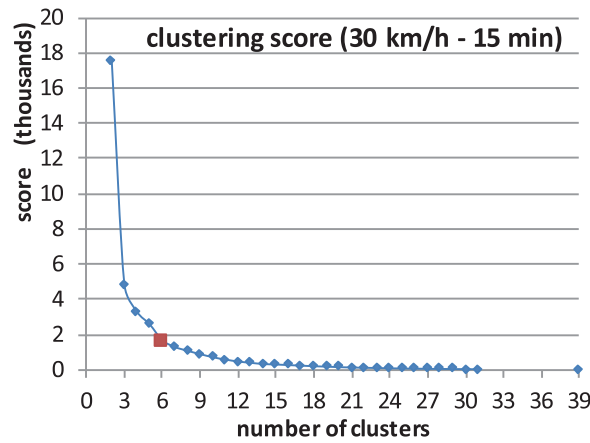
(b) Relocations without vehicle

Fig. 9. Clustering score for different number of clusters.

iterations. Type 1 configuration implies that all requests for trips are known through an advance reservation system that has a cut-off time for accepting reservations (e.g. 24 hours in advance). Type 2 configuration sits on the assumption that the demand is handled as it arrives in the reservation system.

In all scenarios of type 1 and type 2 configurations, it was assumed that there are three parking spots at each station one of which has a fully charged parked vehicle at the beginning of the runs. Since there were not many demand realized from 22:00 in the evening to 6:00 in the morning, all scenarios were started at 6:00 and ended at 22:00. We set three different working shifts throughout the day: (1) morning shift from 6:00 to 12:00, (2) afternoon shift from 11:00 to 17:00 and (3) evening shift from 16:00 to 22:00. Personnel cost was set to € 18 per hour which makes the cost of each personnel in any shift € 108 per person.

The system operating in Nice is a system with electric vehicles. Range of a vehicle with full charge is around 120 km. It takes 8 hours to fully charge a vehicle with an empty battery, i.e. for every kilometres of driving, 4 min of charging is required. It is assumed that rented vehicles were driven 30 km/h. The maximum trip length is set to 40 km. Some important values of the parameters used in the model representing the operating environment of our runs are summarized in Table 1.

4.3. Results

The comparison of different performance measures for runs with different demand levels (200, 250 and 300 demands per day), safety gaps (0, 15, 30 and 45min) and configurations (type 1 and 2) is shown in Fig. 10. Each set of parameters was repeated with 30 different samples. In all figures, the red dots show individual runs. In all figures, *x*-axes show the configuration and the level of demand. *y*-axes take different values depending on the performance measures they represent. These values are counts (in Figs. 10(b) and 10(c)), percentages (in Figs. 10(a), 10(e) and 10(f)) or ratios (in Fig. 10(d)).
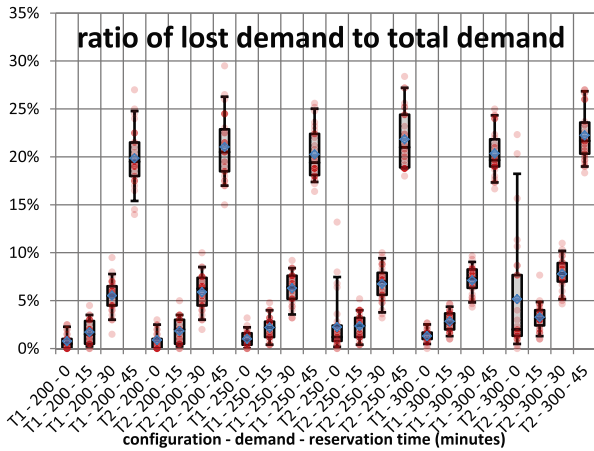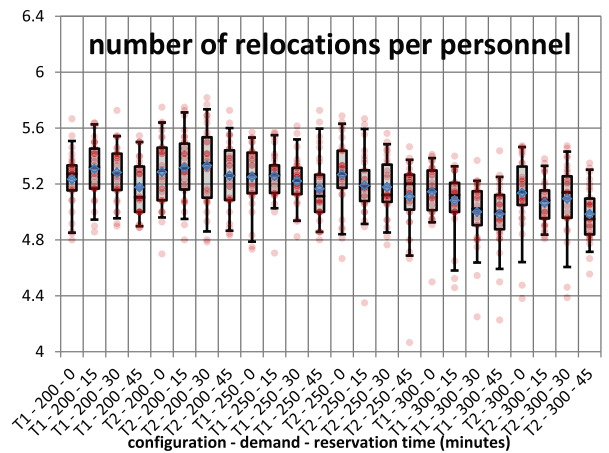
(a) Ratio of lost demand to total demand.
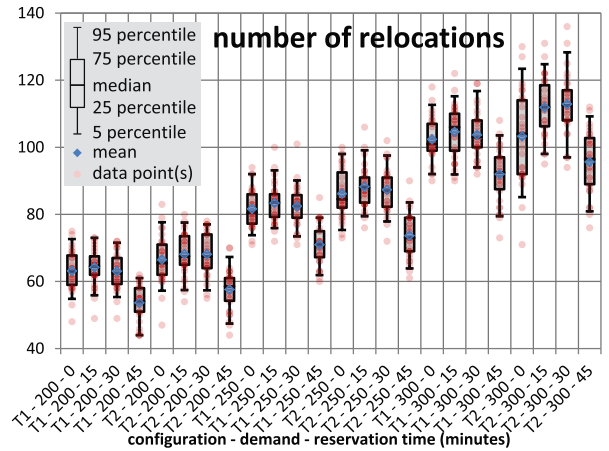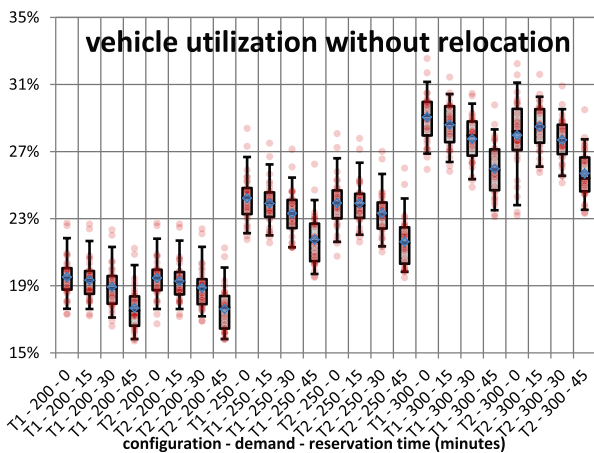

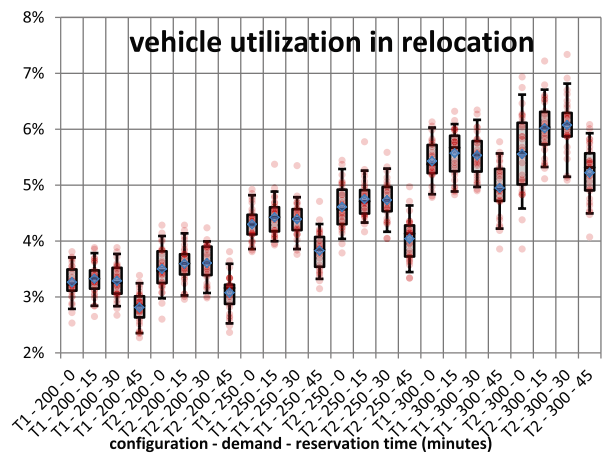(b) Total number of relocations.
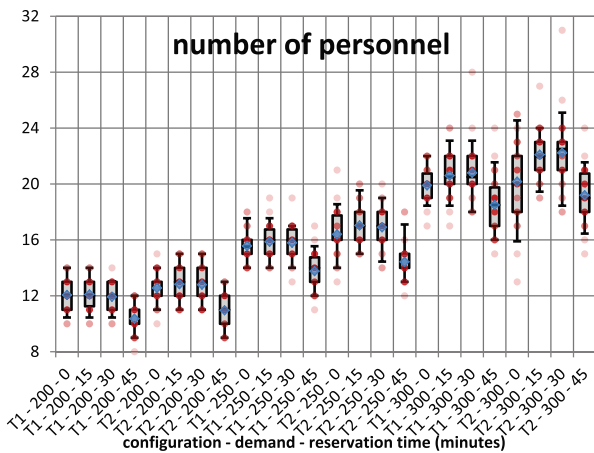
(c) Total number of personnel.

(d) Number of relocations per personnel.

(e) Average vehicle utilization without relocation.

(f) Average vehicle utilization in relocation.

Fig. 10. Some important performance measures of runs with different request counts (200, 250, 300) and configurations (type 1 and type 2). Upper and lower whiskers show the boundaries for the 95th and the 5th percentiles whereas the box shows the values between the 25th and the 75th percentiles. The median is shown with a line in the box and the mean with a blue diamond.

Table 1

Some important values of the parameters used in the model.

Relocation personnel cost (€ /shift):	108
Relocation fuel cost (€ /km):	.02
Maximum charging length (sec):	28,800
Charging length (sec/km):	240
Maximum range without charging (km):	120
Min-max (mean) trip distance (km):	3–40 (18)
Min-max (mean) trip duration (hour):	.2–10 (.8)
Relocation speed with vehicle (km/h):	30
Relocation speed without vehicle (km/h):	15
Number of rentals per sample:	200–300
Minimum trip start charging level:	40%

Fig. 10(a) shows the percentage of unserved demand to total demand for different configurations, safety gaps and demand levels. In runs with demand 200 or less, there is no difference between the runs of type 1 and 2. If the same set of demand requests is accepted to be served by the three configurations, the solutions of three runs should be the same. Because of that, for the lower level of demand, the algorithm returns the same solutions for both types of running configurations. Having the information of all demand requests before deciding in type 1 configuration gives the algorithm freedom to choose rental requests which are matching better to each other. In other words, since type 1 runs have additional information about the requests, type 2 runs cannot outperform type 1 runs if all the other settings are the same. Note that the uncertainty of the future demand for type 2 runs create a system which is less robust. Even if the average fraction of loss demand is similar (3% instead of 2% for demand 300), there are a few days with significant losses (in the range of 10–15%). This analysis suggests that different pricing strategies could be implemented in an operating system for short- and long-term reservations. This should be a research priority. For type 2 runs, the accepted requests to be served in previous iterations affect the decision to serve or not to serve an additional demand. On the other hand, type 1 runs try to find the largest subset of demands that can be served with the current system. Because of that, as the system is more stressed with increased demand levels, type 1 runs perform slightly better than type 2 runs in terms of number of requests served. However, it should be stressed that this is on the expense of the certainty provided to the user that his/her request for service will be granted.

The effect of the parameters on the number of relocations is shown in Fig. 10(b). In these figures, we see a trend of increase of the number of relocations with the increase in demand. The relationship between the number of relocations and the number of requests show almost a linear relationship: For every 8–10 rental requests, approximately 3 relocation operations are needed. Slightly more relocations are needed with higher levels of demand. As the system becomes congested and parking spaces and/or vehicles are not available at the right stations, more resources are used for relocations. The safety gap does not influence significantly relocations up to a value of 30 min. Nevertheless, we observe a significant decrease in relocations if a safety gap of 45 min is applied. This comes with a significant cost for the loss demand and highlights that high safety gaps should not be allowed (or late returns should be penalized).

In the same figure, we also observe some differences between two configurations. As expected, type 1 runs required less number of relocations compared to type 2 runs for the same level of demand requests. Type 2 runs usually serve less demand when demand is high. The difference in the number of relocations between configurations shows the effect of selecting better subsets of requests: If the requests are selected with more information, more demand can be served with less number of relocations. As expected, more relocation means more number of personnel. In Fig. 10(c), the total number of personnel from all shifts is shown. Very similar trends with Fig. 10(b) are observed.

In Fig. 10(d), we see the average number of relocations per personnel in different scenarios. While this value does not vary significantly (ranges between 5 and 5.3), it is clear that for higher safety gaps that increase the ratio of unserved demand (in Fig. 10(a)), a slightly smaller number of relocations per personnel is performed. This is an interesting observation as the operator can plan in advance the personnel needs if he has a good prediction of the demand level.

In the last two figures, Fig. 10(e) and (f), utilization of vehicles under rental and relocation are shown respectively: An increase in the demand of about 50% (from 200 to 300 requests) increases the vehicle utilization for about the same amount (from 19% to 28%). This is reasonable given that for the safety gap not more than 30 min, unserved demand is little. Vehicle utilization during relocation increases more sharply, from 3% to 6% as demand increases from 200 to 300. This highlights the need for a larger number of relocations (and personnel) to keep the quality of service high. As expected, the increase in the number of demand requests increases the vehicle utilization under rental and relocation. The difference between utilization for rental operations is more evident with the demand level 250 and 300. But this difference is not as severe as Fig. 10(a). This may be because of the way we selected the first objective function (see Objective 1) in the operations minimization model (see Section 3.2.4). In our first objective function, the duration of the requests was not considered in the objective. Demand requests with longer duration were not given higher weight as compared to shorter ones. As a result, type 1 runs are inclined to select shorter requests in case of high demand levels to serve more demand with the same system resources. Final figure, Fig. 10(f), shows the power of freedom to choose demand subset against accepting/rejecting requests with limited information. Especially for higher demand levels, although more demand requests were served, type 1 runs keep vehicles less busy with relocations.

The remaining analysis focuses on a specific instance with 300 demand and type 1 run. In Fig. 11, the movements of vehicles between stations are shown. In Fig. 12, movements of relocation personnel on different working shifts can be seen. In Fig. 13, two separate scenarios are applied on the same problem instance with additional constraints. Constraints 23 were applied to limit the maximum number of personnel at each shift (see Fig. 13(a)). Constraint 24 was applied to type 1 runs to limit the number of demand served.

$$\nu_s \leq \text{maximum number of personnel per shift} \quad \forall s \qquad (23)$$

$$\sum_i z_i = \text{demand served} \qquad (24)$$

In Fig. 11, the two leftmost vertical columns show the station number and the number of spots of this station (3 per station). The uppermost horizontal numbers are the starting hour and minute of the intervals. Each colour shows a different state for the parking spot. The legend for colour coding is shown in the legend on the same figure. Note that, for all states except "spot available", the number inside the cell identifies the vehicle affected with this state. For this instance, since only the 0th spot of every station has an available vehicle at the beginning of the day, each vehicle is numbered with the station parked at the beginning of the day.

Fig. 11 also enables to see the detailed relocation operations. Each orange line shows the relocations starting/ending time and origin/destination stations. One of the expected results observed from this figure is to see the relocations of vehicles from stations with no/low demand to stations with high demand at the beginning of the day and the opposite movement at the end of the day. Vehicles from stations without demand (i.e. 61, 62, 63, 64 and 65) were relocated to stations with high demand (i.e. 14, 35 and 55 then 42 and 55) at the beginning.

The identification of numbers enables to track the daily assignments of each vehicle. For instance, vehicle 0 started from the 0th spot of Station 0. Between 8:00–8:15, it was relocated to the 1st spot of Station 25. The vehicle was charged for one time interval and rented at 8:45. The vehicle was returned to the 0th spot of Station 54 at 9:45. At 12:00, it was rented again with a partial charge. At 12:30, it was returned to the 0th spot of Station 30. The vehicle stayed and fully charged there until it was rented again at 19:15. At 20:45, it was returned to the 1st spot of Station 29 and stayed there until the end of the planning period.

To see the detailed movements (i.e. relocations with/without vehicle) of each personnel, separate charts for each working shift are created (see Fig. 12). Each set of the connected quadrants represents a separate station. Area of each quadrant is directly proportional to the number of demand originating from (gray, left top), demand destined to (blue, right top), relocation destined to (green, bottom left) and relocation destined from (red, right bottom) the station they are representing. Each line set with different colors represents different personnel. Solid lines show relocations with vehicles and dotted lines show relocations without vehicles. For instance, the person represented with red firstly relocated a vehicle from an outer station (from Fig. 11 we can see it is Station 61) to Station 14. Then he moved to Station 10 without vehicle. After that, he relocated a vehicle to Station 16. Then he moved to Station 23 and relocated a vehicle to Station 33. He then relocated a vehicle from Station 33 to 36. And finally he moved back to Station 29 and relocated another vehicle to Station 36. Regarding the size of each quadrant per station, note that there are some repetitive trends by looking at the aggregated values. For example, stations with high demand of origins, usually have a high number of relocations destined to this station. Stations that have high demand both as origins and destinations do not have a clear trend with respect to relocations as the temporal distribution of demand matters.

In Fig. 13(a) type 1 and type 2 runs are compared for the same single instance with 300 demand. Different than the previous runs, the maximum number of available relocation personnel at each shift is limited to different values from 0 to 8. The dark (i.e. black) lines and (darker) columns (on the left) show the run results for type 1 runs. The light (i.e. gray) lines and (lighter) columns (on the right) show results for type 2 runs. The left vertical axis shows the number of personnel (columns) and the number of relocations per personnel (lines with triangle markers). The right vertical axis is for the total demand served (lines with square markers) and the demand per personnel (lines with diamond markers).

Note the difference between the number of demand served with type 1 and type 2 runs. This difference shows that a carsharing system can serve more demand if it is able to choose a specific subset. This analysis suggest that different business models can be used for offering carsharing services with different implications for the users in terms of service provision and associated costs. Specifically a business model that requires advance reservation and notification regarding acceptance or not of the requested services at the end of the reservation period (type 1 configuration) may provide benefits to the operator on the expense of uncertainty to the user's request. In contrast, a system that will provide immediate (within minutes) response if a trip request is granted or not (type 2 configuration) requires more resources for the same level of demand. These results can also create intuition with respect to the development of pricing strategies to deal with the emerging business models. It is also clear that short term reservations can contribute in system imbalances and loss demand. In other words, if the demand to be served is selected with full knowledge of total demand, the productivity of the system can be improved for the same level of resources. In our case, we see an improvement ranging from 0 to 60% depending on the maximum number of personnel available per shift.

In Fig. 13(b), the type 1 run has given freedom to choose a number of demand to be served (170 to 300 in every 5) from a set of 300 demand. 170 is selected as the starting point because from Fig. 13(a) we can see that 174 rental requests of the
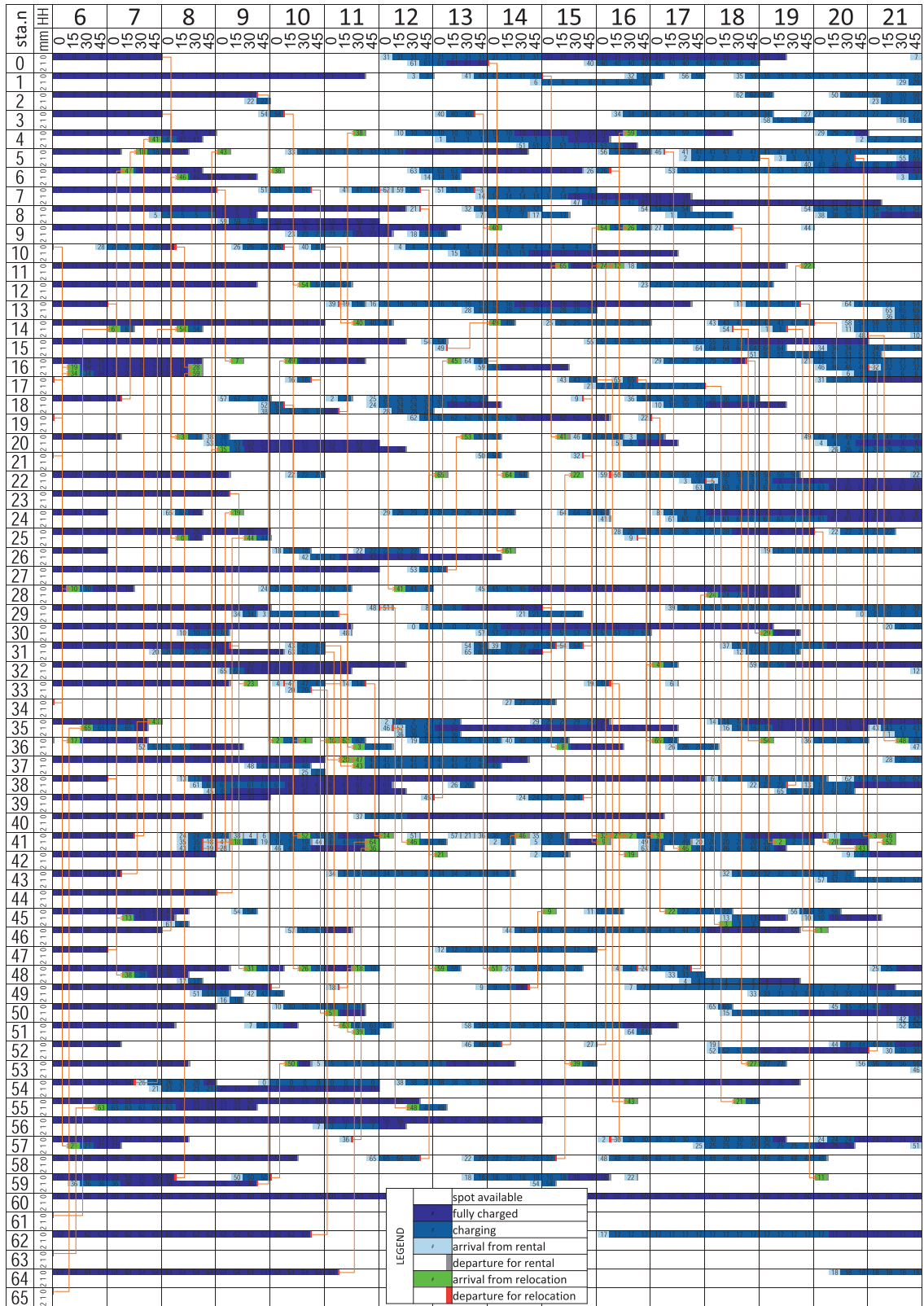
Fig. 11. Chart showing movements of vehicles and relocation routes for the case with 300 rental requests.
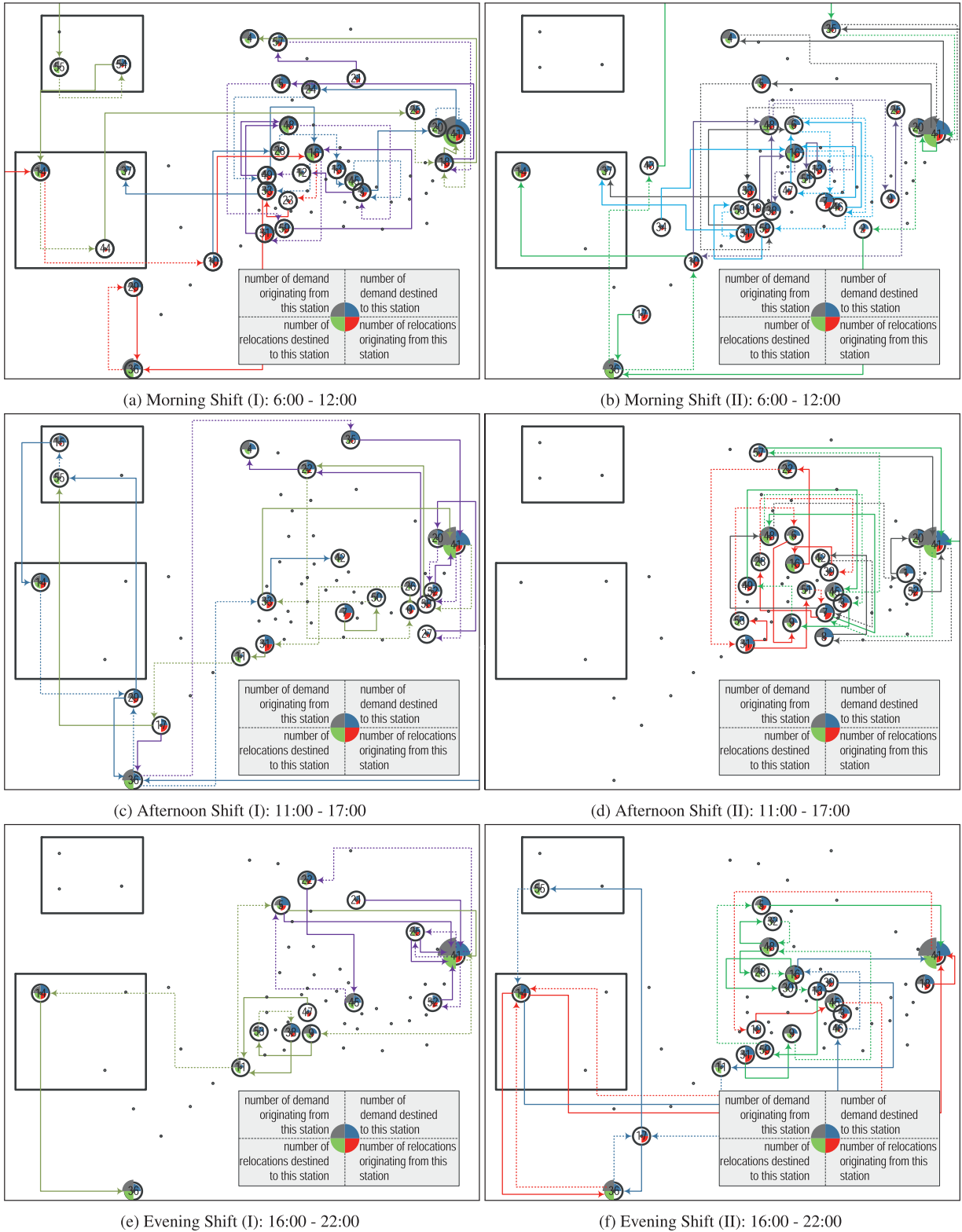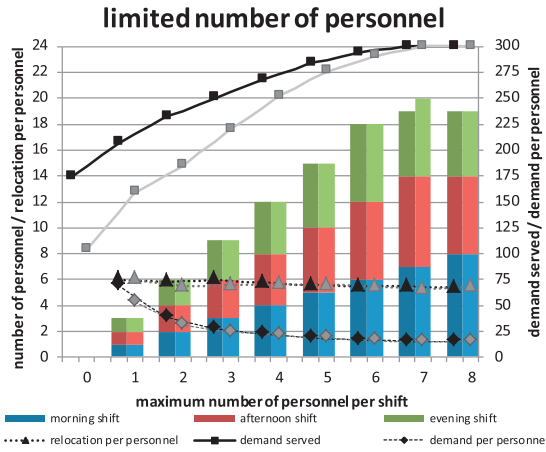
(a) Morning Shift (I): 6:00 - 12:00

(b) Morning Shift (II): 6:00 - 12:00

(c) Afternoon Shift (I): 11:00 - 17:00

(d) Afternoon Shift (II): 11:00 - 17:00

(e) Evening Shift (I): 16:00 - 22:00

(f) Evening Shift (II): 16:00 - 22:00

Fig. 12. Movements of different personnel with (straight line) and without (dashed line) vehicles at each shift for the selected case with 300 rental requests.

(a) Comparison of the type 1 (shown with black lines and left columns) and type 2 (shown with gray lines and right columns) runs when the number of personnel is limited at each shift.
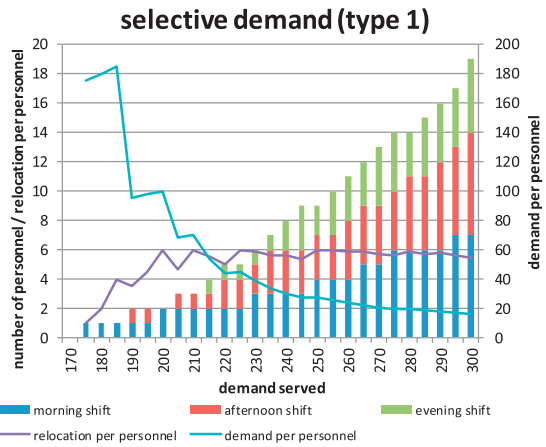
(b) Different performance measures when the number of demand served is fixed for the type 1 run.

Fig. 13. The effect of limited number of personnel and number of demand served on the selected case with 300 rental requests.

selected instance can be served without any relocation personnel. We could not have the same test for the other running configuration since in type 2 runs, the requests become available and are accepted/rejected iteratively as they appear in the system. In this figure, the horizontal axis shows the number of rental requests that needs to be served among 300 demand. Columns show the number of personnel from different shifts. The purple line shows the number of relocations per personnel. Both the number of personnel and relocations per personnel are shown on the left vertical axis. The blue line shows the demand per personnel with the values on the right vertical axis.

From this figure, we can see how the demand per personnel converges to a specific value. Although, this value is highly related to the system parameters, it is worth noting such a value may exist for carsharing systems. This ratio could help in carsharing systems to determine the number of personnel needed for any demand level. Similarly, with the increase in the number of demands served, relocation per personnel reaches to 6. In other words, every member of staff needs one hour per relocation on average since in these runs we assumed personnel shifts of 6 hours. Note that, these 6 hours include not only relocations with vehicles but also relocations without vehicles. Between every two relocations with vehicles, the personnel in charge of the relocations needs to travel from the destination station of the previous relocation to the origin station of the next relocation. This value probably depends on the topology of the system and the relocation strategy, but if known (or estimated), it can be utilized to develop simpler formulations for the tactical and strategic problems of the system (see for example Boyacı et al., 2015).

In Fig. 14, various system states during the day are compared for the two different types of runs. Figures on the left show the number of stations with a specific number of vehicles (0 to 3) for each time period. Figures on the right, show the number of vehicles providing service, under relocation and the charging level of vehicles at stations over time. The same
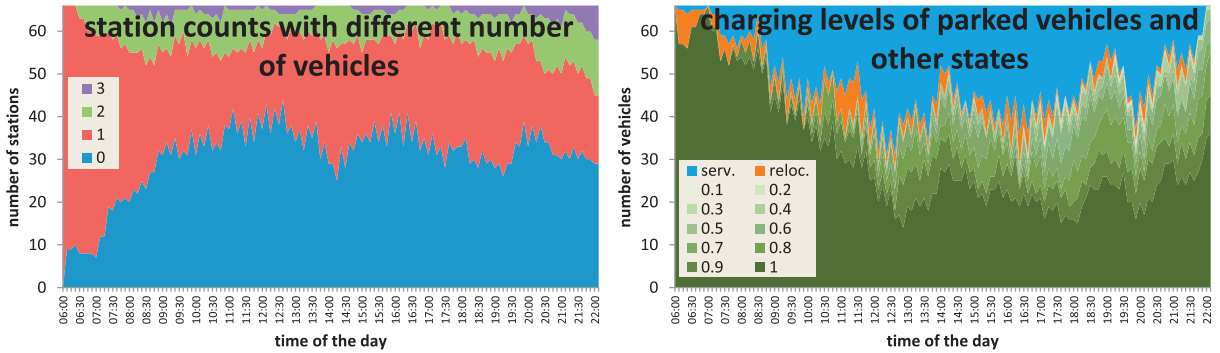
(a) Stations with different number of vehicles for the type 1 run of the selected case with 300 demand.

(b) Number of vehicles with different charging levels for the type 1 run of the selected case with 300 demand.
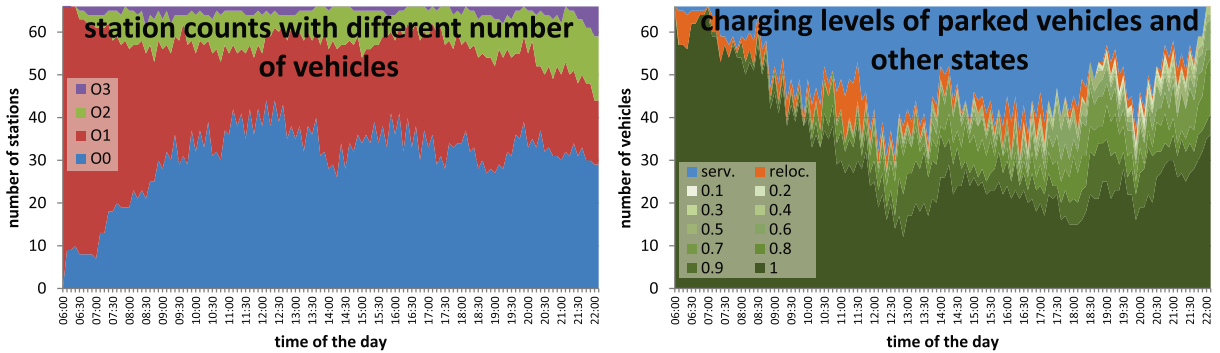
(c) Stations with different number of vehicles for the type 2 run of the selected case with 300 demand.

(d) Number of vehicles with different charging levels for the type 2 run of the selected case with 300 demand.

Fig. 14. Comparison of the states of vehicles and stations for three different approaches for the selected case with 300 rental requests. Left figures represent the stations with different number of vehicles whereas right figures represent the number of vehicles with different charging levels.

300-demand set is used for all cases. In all figures, *x*-axes show the time of the day during simulation. *y*-axes represents either the number of vehicles (Fig. 14(b) and (d)) or the number of stations (Fig. 14(a) and (c)). Note that the distribution of the number of vehicles per station converges to very similar values independently of the run type.

5. Conclusions and future work

In this paper, we developed an integrated optimization-simulation framework to deal with the operational decisions of one-way (electric) carsharing systems with reservations. We developed a clustering algorithm to decrease the computational complexity of the model, an event-based simulator to test the feasibility of charging levels during the optimization procedure and MILP models to handle the operational decisions. The three modules are interacting iteratively to guarantee close-to-optimal solutions. Experimental results showed that, the developed framework is efficient. We have tested two types of runs related to reservation confirmation policy: Type 1 is with full information about the system, as it might happen in the case of reservations in advance. In type 2, requests are appearing one per time and the system does not know the future demand during optimization. We also tested various demand levels to see the performance differences and applicability range for different settings.

Experimental results showed the importance of efficient algorithms for relocation operations. Comparison between different configurations also showed that especially in congested systems, forecasting the future demand or optimizing the initial locations of vehicles in the beginning of each day is quite important to utilize system resources efficiently. To serve higher number of requests without increasing system resources, demand should be managed through the use of incentives/disincentives (e.g. pricing). Specifically, business models that require reservations with an advanced cut-off point for confirming them may lead to a more efficient use of the system resources on the expense of not receiving immediate confirmation for serving the expressed demand (less flexible). The business models that require reservations based on type 2 runs provide higher flexibility to users in terms of making advanced reservations, but they are less efficient in terms of the utilization of resources.

A field test to investigate the optimization results in real conditions is also under preparation. Future work is looking to expand these models to systems without reservations. While research work (e.g. Correia et al., 2014) dealt with system imbalances through trip pricing, a combination of pricing and relocation strategies should be another research priority.

Acknowledgements

References

Barth, M., Todd, M., Xue, L., 1999. Simulation model performance analysis of a multiple station shared vehicle system. Transport Res. Part C 7, 237–259. doi:10.1016/S0968-090X(99)00021-2.

Benchimol, M., Benchimol, P., Chappert, B., de la Taille, A., Laroche, F., Meunier, F., Robinet, L., 2011. Balancing the stations of a self service "bike hire" system. RAIRO – Operat. Res. 45, 37–61. doi:10.1051/ro/2011102.

Boyacı, B., Zografos, K.G., Geroliminis, N., 2015. An optimization framework for the development of efficient one-way car-sharing systems. Eur. J. Oper. Res 240 (3), 718–733. doi:10.1016/j.ejor.2014.07.020.

Bruglieri, M., Colorni, A., Luè, A., 2014. The vehicle relocation problem for the one-way electric vehicle sharing: an application to the milan case. Procedia - Social Behav. Sci. 111, 18–27. Transportation: Can we do more with less resources? - 16th Meeting of the Euro Working Group on Transportation - Porto 2013. 10.1016/j.sbspro.2014.01.034.

Correia, G.H.A., Antunes, A.P., 2012. Optimization approach to depot location and trip selection in one-way carsharing systems. Trans. Res. Part E 48 (1), 233–247. doi:10.1016/j.tre.2011.06.003.

Correia, G.H.A., Jorge, D.R., Antunes, A.P., 2014. The added value of accounting for users' flexibility and information on the potential of a station-based one-way car-sharing system: an application in lisbon, portugal. J. Intell. Transp. Syst. Technol. Plann. Oper. 18 (3), 299–308. doi:10.1080/15472450.2013.836928.

Google, 2015. Google Maps API. https://developers.google.com/maps/.

Jorge, D., Correia, G.H., 2013. Carsharing systems demand estimation and defined operations: a literature review. Eur. J. Transp. Infrastructure Res. 13 (3), 201–220.

Jorge, D., Molnar, G., de Almeida Correia, G.H., 2015. Trip pricing of one-way station-based carsharing networks with zone and time of day price variations. Transport. Res. Part B: Methodological 81 (2), 461–482. doi:10.1016/j.trb.2015.06.003.

Kaspi, M., Raviv, T., Tzur, M., 2014. Parking reservation policies in one-way vehicle sharing systems. Transport. Res. Part B 62, 35–50. doi:10.1016/j.trb.2014.01.006.

Kaufman, L., Rousseeuw, P., 1987. Clustering by Means of Medoids. North-Holland, pp. 405–416.

Kek, A.G.H., Cheu, R.L., Chor, M.L., 2006. Relocation simulation model for multiple-station shared-use vehicle systems. Transport. Res. Record 1986, 81–88. doi:10.3141/1986-13.

Kek, A.G.H., Cheu, R.L., Meng, Q., Fung, C.H., 2009. A decision support system for vehicle relocation operations in carsharing systems. Transport. Res. Part E 45 (1), 149–158. doi:10.1016/j.tre.2008.02.008.

Nair, R., Miller-Hooks, E., 2011. Fleet management for vehicle sharing operations. Transport. Sci. 45 (4), 524–540. doi:10.1287/trsc.1100.0347.

Nair, R., Miller-Hooks, E., Hampshire, R.C., Bušić, A., 2013. Large-scale vehicle sharing systems: analysis of vélib'. Int. J. Sustainable Transport. 7 (1), 85–106. doi:10.1080/15568318.2012.660115.

Nourinejad, M., Roorda, M.J., 2014. A dynamic carsharing decision support system. Transport. Res. Part E 66, 36–50. doi:10.1016/j.tre.2014.03.003.

Pfrommer, J., Warrington, J., Schildbach, G., Morari, M., 2014. Dynamic vehicle redistribution and online price incentives in shared mobility systems. Intell. Transport. Syst. 15 (4), 1567–1578. doi:10.1109/TITS.2014.2303986.

Raviv, T., Kolka, O., 2013. Optimal inventory management of a bike-sharing station. IIE Trans. 45 (10), 1077–1093. doi:10.1080/0740817X.2013.770186.

Raviv, T., Tzur, M., Forma, I.A., 2013. Static repositioning in a bike-sharing system: models and solution approaches. EURO J. Transport. Logist. 2 (3), 187–229. doi:10.1007/s13676-012-0017-6.

Repoux, M., Boyacı, B., Geroliminis, N., 2015. Simulation and optimization of one-way car-sharing systems with variant relocation policies. In: 94th Annual Meeting of the Transportation Research Board, Washington D.C. http://trid.trb.org/view.aspx?id=1337340.

Sayarshad, H., Tavassoli, S., Zhao, F., 2012. A multi-periodic optimization formulation for bike planning and bike utilization. Appl. Math. Model 36 (10), 4944–4951. doi:10.1016/j.apm.2011.12.032.

Shaheen, S.A., Cohen, A.P., 2013. Carsharing and personal vehicle services: worldwide market developments and emerging trends. Int. J. Sustainable Trans. 7 (1), 5–34. doi:10.1080/15568318.2012.660103.

Weikl, S., Bogenberger, K., 2015. An integrated relocation model for free-floating carsharing systems: field trial results. In: Transportation Research Board 94th Annual Meeting. http://trid.trb.org/view.aspx?id=1338255.

Weikl, S., Bogenberger, K., 2015. A practice-ready relocation model for free-floating carsharing systems with electric vehicles - mesoscopic approach and field trial results. Trans. Res. Part C 57, 206–223. doi:10.1016/j.trc.2015.06.024.