Decision support for tactical resource allocation in bike sharing systems

Patrick Vogel[a], Jan F. Ehmke[b], Dirk C. Mattfeld[a]

[a]Decision Support Group, Technische Unveristaet Braunschweig, Muehlenpfordtstrasse 23, 38106 Braunschweig, Germany

[b]Department Information Systems, Freie Univeristaet Berlin, Garystrasse 21, 14195 Berlin, Germany

Abstract

This paper presents an integrated approach of intelligent data analysis and mathematical optimization to support tactical resource allocation in bike sharing systems. Bike sharing has recently enabled sustainable means of shared mobility by automated rental stations in metropolitan areas. Spatio-temporal variation of bike rentals leads to imbalances in the distribution of bikes causing full or empty stations in the course of a day. Ensuring the reliable provision of bikes and bike racks is crucial for the viability of these systems. This can be done within the scope of tactical resource allocation. Tactical planning requires the anticipation of operational decisions and a suitable aggregation of input data.

We present a linear programming formulation aiming at cost-efficient allocation of bikes to stations given a predefined service level for different scenarios of bike demand. In particular, tactical decisions regard the fulfilment of time-dependent target fill levels of bikes at stations. Target fill levels serve as input for the detailed planning of relocation tours on the operational level. Different scenarios of bike demand are considered as realizations of typical bike flows between stations in terms of time-dependent origin / destination matrices. We employ an intelligent data analysis approach to generate aggregated input data from individual trips recorded in Bike Sharing Systems. Intelligent data analysis produces spatio-temporal distributions of bike flows and helps determining typical trip purposes in combination with methods from the field of traffic modeling.

The proposed methodology is exemplified based on two years of trip data from Vienna's "Citybike Wien". Computational experiments show changing appropriate target fill levels for a potential increase of demand. Furthermore, determined characteristics of relocation regimes can support BSS operators in the planning and implementation of relocation.

Keywords: Decision support, Resource allocation, Linear programming, Intelligent data analysis, Shared mobility, Bike sharing systems

1 Bike sharing systems

Emerging metropolitan areas need efficient and sustainable mobility services in order to ensure their attractiveness, quality of life, and economic power. The more crowded a metropolitan area becomes, the more inefficient and expensive is the realization of trips with private vehicles. Municipalities have thus begun to implement innovative shared mobility systems (SMS), such as car and bike sharing systems, accommodating the mobility needs of their citizens, while ensuring sustainability and flexibility of transportation. Bike sharing systems (BSS) have become exceptionally popular. The number of implemented BSS is impressive; in Europe, about 400 BSS have been introduced in the last ten years (Büttner and Petersen 2011). Markets in America and Asia are catching up (Shaheen et al. 2010).

BSS provide an individual, but likewise public means of transportation for inner city trips (Midgley 2011). They are characterized by a high density of service facilities in heavily populated areas, e.g. an average distance of 300 meters between bike stations (Büttner and Petersen 2011). Municipalities typically engage advertising companies for the operation of BSS (DeMaio 2009). Short bike rentals are often free of charge, and revenue is indirectly generated from a license to advertise on street furniture. Other sources of revenue are subscription fees. Information systems play a crucial role in the management and operation of BSS. Rental, return and maintenance processes are automated, enabling fast and easy access as well as one-way use and short rental times through unattended stations. Every trip is recorded for tracking and billing purposes.

While the usage of BSS is often simple, inexpensive and convenient from a user point of view, the design, management and operation of BSS remain even more challenging. Imbalances in the distribution of bikes affect the service level, i.e., the provision of bikes and free bike racks when demanded. These imbalances are caused by spatio-temporal variation of bike rentals. Demand for bike rentals varies strongly, following typical traffic patterns in the course of day and week caused by e.g. commuter, leisure or tourist trips. Furthermore, one-way rentals intensify imbalances in the distribution of bikes. Due to limited capacity at stations, rentals are impossible at empty stations, and returns are impossible at full stations. BSS operators aim to ensure a service level which is self-stipulated or stipulated by municipalities. For instance, a tendering for the Arlington BSS requests that "stations shall not be full of bicycles for more than 60 minutes during the hours of 8am - 6pm and 180 minutes during the hours of 6pm - 8am" (Zahory 2009). In order to ensure such a service level, bike imbalances can be handled by means of strategic, tactical or operational planning tasks. On the strategic level, decisions on the size of stations have to be made. Acquiring a high number of bike racks at stations increases the probability of successful returns. On the tactical level, the allocation of bikes needs to be determined in order to compensate varying bike demand in the course of day. Allocating a large number of bikes increases the probability of successful rentals, for example, while decreasing the probability of successful returns at particular stations. On the operational level, relocation of bikes from rather full to rather empty stations helps maintaining the service level. Manual relocation with the help of a service fleet, however, results in significant costs affecting the viability of BSS (DeMaio 2009). Planning decisions are interdependent: reasonable sizing of stations and allocation of bikes may reduce relocation efforts, whereas high relocation efforts may compensate insufficient sizing and bike allocation. Hence, distinct optimization of the planning levels may lead to suboptimal decisions.

In this paper, we propose an integrated approach of intelligent data analysis and mathematical optimization supporting tactical resource allocation in BSS. The presented mathematical linear program (LP) determines optimal target fill levels at stations by minimizing the expected costs of relocation. It ensures a given service level for different scenarios based for a mid-term planning horizon. Scenarios are considered in terms of bike flows that are described by time-dependent origin-destination (OD) matrices. The required information is derived from the aggregation of observed customer trips in combination with well-known traffic modeling approaches. We present an information model that abstracts from observed trip data by means of intelligent data analysis. The information model provides trip purposes, which allow for the generation of different bike demand scenarios. The bike demand scenarios serve as input for resource allocation.

The remainder is organized as follows. A literature overview on the analysis of SMS and related optimization approaches is presented in Section 2. We discuss the integrated approach comprising the information model representing trip purposes, its integration into a traffic modeling approach and prediction of demand scenarios in Section 3. This section also presents an optimization model for tactical resource allocation. The proposed methodology is exemplified with the help of a case study including two years of trip data from Vienna's BSS "Citybike Wien" (Section 4). Future work is subject of Section 5.

2 Design, management and operation of shared mobility systems

Design, management and operation of SMS can be supported by data analysis and optimization approaches for strategic, tactical and operational planning tasks. We propose the classification of planning tasks and corresponding data flows as shown in Fig. 1. The classification provides background information on the planning tasks and helps to clarify our perspective on the tactical planning level. Decisions on a specific planning level may have a significant impact on the decision of the subordinate level. Since this work focuses on the integration of recorded trip data analysis and optimization for implemented systems, the planning levels are described from bottom to top. Although we present the classification using the example of BSS, the classification is applicable to other SMS.

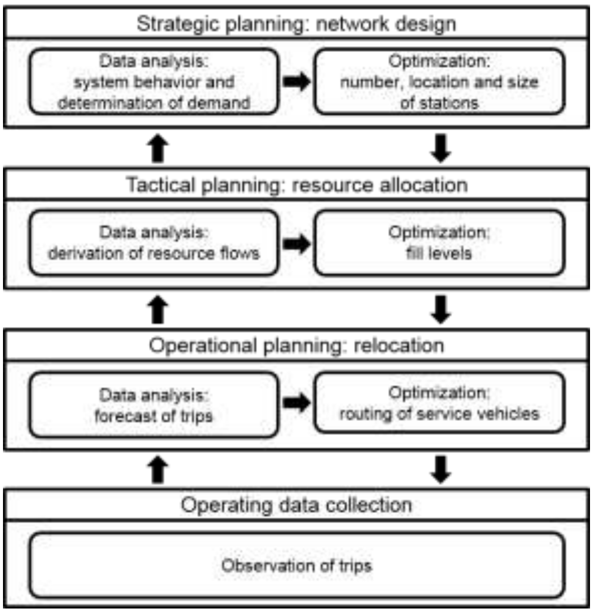

Fig. 1 Classification of planning tasks for shared mobility systems

- Within system operation, data about rental and return times at stations are continuously recorded for tracking and billing purposes. An exemplary trip data record comprises the specific rental station, the time of rental, the return station and the time of return.
- On the operational level, operators react to short-term variation of demand by relocating bikes. Maintaining the defined service level for a specific setup of the system requires detailed routing of the service fleet. Typical decisions are where and when to pick up and return how many bikes using which service vehicle. Due to the short-term planning horizon, routing relies on *detailed information* about current and expected bike demand, fill levels at stations, available service vehicles and staff. Based on historical trip data, short-term forecasts of trips may serve as input for the optimization of relocation operations. Forecasts have to incorporate short-term influences such as weather, events and traffic conditions.
- Tactical planning aims at rational and efficient management through allocation of system resources among a service network to improve the system's performance over medium-term horizons. Resource allocation targets a certain *service level* in order to avoid full or empty stations. Fill levels of bikes at stations are to be determined as input for operational planning. Thus, tactical relocation decisions need to take a wide-range of demand variation into account. To optimize resources for a given service level, tactical planning requires a suitable aggregation of input data and anticipation of operational decisions. Our tactical planning approach is detailed in Section 3.

- On the strategic level, network design specifies the number, location and (re-)size of stations ensuring sufficient overall coverage and capacity of the implemented system. The analysis of historical trip data as well as external data such as demographic and land-use data lead to insights on user behavior. Strategic planning also allows for the determination of overall demand for network design purposes.

In the following, a literature overview of the above planning tasks is provided, especially with regard to planning and optimization of BSS. We also discuss the role of observed data and its required level of aggregation for the particular planning tasks.

For *operational planning*, a system operator needs to determine the expected trips on a very detailed short-term level, supporting decisions on relocation tours. Statistical analysis of observed trip data may provide the required data input. Vogel and Mattfeld (2011) model and forecast bike rentals on an hourly basis while considering seasonal influences by detailed weather data. Borgnat et al. (2010) incorporate weather data and event data (holidays, strikes) for short-term forecasts of bike rentals. Kaltenbrunner et al. (2010) and Froehlich et al. (2009) refer to data about bike fill levels in order to forecast bike availability at stations. Many authors relate the optimization of relocation tours to the one commodity pickup and delivery problem (PDP) and the swapping problem (SP). In the PDP, a fleet of vehicles transports a commodity from pickup to delivery stations. In the SP, multiple commodities are considered and a station serves as both a pickup and delivery station. Relocations are studied as a static or dynamic problem. In the static problem, relocations are realized at night time only when no demand occurs. In the dynamic problem, demand variation and several decision points over time are considered. Benchimol et al. (2011) combine PDP and SP and present a static model and solution methods. Raviv et al. (2012) study the static relocation problem minimizing user dissatisfaction by means of penalty costs and operating costs for relocation. Ricker et al. (2012) introduce a simulation-based approach to determine the cost-efficient daily number of relocation operations. Weighted sums of transportation costs and costs for unserved customers are considered. Rainer-Harbach et al. (2013) propose a variable neighborhood search in combination with a greedy heuristic, maximum flow approach and LP to determine the routes and number of relocated bikes for the static relocation problem. Contardo et al. (2012) present an arc-flow optimization model for the dynamic routing of service vehicles minimizing “lost demand”. Lost demand is caused by customers who cannot rent or return rented vehicles at empty or full stations. Caggiani and Ottomanelli (2012) propose a decision support system for the dynamic relocation problem. Here, a neural network is used to forecast rentals and returns at stations.

For car sharing systems, Kek et al. (2009) present a mixed integer program (MIP) minimizing costs for service staff and relocation operations. Lost demand is considered in terms of penalty costs. Nair and Miller-Hooks (2011) propose a stochastic MIP with chance constraints to obtain a least-cost plan for the relocation of vehicles. Their cost function comprises fixed costs for the relocation of vehicles, relocation between stations, and penalty costs for the utilization of additional service vehicles.

For *tactical planning*, George and Xia (2011) model SMS by means of a closed queuing network. A profit maximizing optimization is applied in order to determine the optimal fleet size and allocation of rental vehicles. Raviv and Kolka (2013) also use queuing models. With the help of a user dissatisfaction function, the optimal fill level at a bike station is determined.

For *strategic planning*, information on typical system behavior is required. Based on data analysis of a large data set of customer trips, Vogel et al. (2011) determine temporal demand "activity clusters" describing typical rental and return activities at stations in the course of day. Cluster analysis reveals groups of stations with similar trip purposes represented by the activity. Borgnat et al. (2010) characterize interrelated stations by cluster analysis of bike flows between stations. Wang et al. (2012) apply linear regression to model the correlation of bike activity at stations and external factors like demography and transportation infrastructure.

Consideration of spatial relations between bike rentals at stations and location of stations may support *strategic decisions* on the number, location and size of stations. Lin and Yang (2011) present a hub-location model that determines the number and locations of bike stations as well as the network of bike paths. Here, customers' travel costs and setup costs for bike stations and bike paths are minimized. In an extended version of their optimization model, also decisions on the bike inventory at stations are taken into account (Lin et al. 2013). Martinez et al. (2012) propose an MIP to optimize the location of bike sharing stations and the size of the bike fleet. Nair and Miller-Hooks (2014) present an MIP for the optimal configuration of SMS by determining the station locations and sizes as well as vehicle inventories neglecting operational decisions, though.

The above work either focuses on a distinct planning level (strategic, tactical or operational) or discusses either data analysis or optimization methodology, respectively. Integrated approaches are rare. Especially for tasks of tactical planning, anticipation of operational decisions is crucial for the viability of SMS. Costly relocation could be alleviated by appropriate fill levels that compensate expected variation of demand. To the best of the authors' knowledge, only three integrated approaches exist, anticipating relocation operations in resource allocation:

- Correia and Antunes (2012) present multi-periodic MIP formulations to maximize the profit of a car sharing system considering the revenue of trips, costs of depot and vehicle maintenance as well as costs of vehicle relocation. They determine the number and the location of stations as well as the number of vehicles at each station in each period of daily operation. They consider static relocation at the end of the day where vehicles are relocated between stations to reset the initial fill level. The validity of the MIP approach is investigated by means of a simulation model (Jorge et al. 2012).

- Sayarshad et al. (2012) introduce a dynamic LP formulation to maximize profit in BSS. Relocation, maintenance, capital and holding costs of bikes as well as penalty costs for lost demand are deducted from revenue generated by trips. Unutilized bikes can be relocated in every period of daily operation.
- Schuijbroek et al. (2013) minimize the costs of relocation tours and incorporate service level requirements at stations. They consider the static case in which no varying user demand is considered. The service level is precalculated for each station without anticipation of the routing decisions. A cluster-first route-second heuristic is proposed to solve the problem.

In sum, recent approaches do not sufficiently reflect the interaction of fill levels and relocation operations within the scope of tactical optimization. A general methodology that benefits from usage and aggregation of detailed trip data for tactical planning is missing. Thus, we extend existing optimization approaches focusing on the adequate anticipation of relocation tours and present a new approach to aggregate trip data.

3 Tactical resource allocation in bike sharing systems

The tactical planning perspective requires anticipation of operational decisions and a suitable aggregation of trip data.

We present an optimization model as a linear programming formulation aiming at cost-efficient allocation of bikes to stations ensuring a predefined service level for different scenarios of bike demand. In particular, decisions regard target fill levels of bikes at stations, expected relocation regimes and the required number of bikes in the system. Target fill levels ensure the provision of service depending on the time of the day for a given scenario such as high demand for bikes on a working day in the main season. Relocation regimes yield the expected costs of relocation to compensate insufficient fill levels, anticipating costs of relocation operations. A relocation regime refers to a set of regular relocation requests for a given scenario.

An information model is proposed yielding scenarios of bike demand. Scenarios of bike demand are considered as realizations of the typical bike flows between stations in terms of time-dependent origin / destination matrices. We refer to an intelligent data analysis approach abstracting from individual trips recorded in BSS. In particular, intelligent data analysis carves out the spatio-temporal distribution of main trips. In combination with approaches from the field of traffic modeling, trip purposes allow for the prediction of different scenarios of bike flows serving as input for tactical planning.

The information model yielding trip purposes and demand scenarios is subject to Section 3.1. The optimization model considers the scenarios for determination of optimal target fill levels while anticipating relocation (cf. Section 3.2). For computational experiments, we exemplify the evaluation of the information and optimization model with the help of real bike sharing data in a subsequent case study (cf. Chapter 4).

3.1 An information model representing typical bike flows

Information systems for BSS automatically collect extensive amounts of detailed trip data. Trip data represent individual observations of customer behavior and are therefore not suited as input for tactical planning. Instead, aggregation of trip data is required in order to model trip purposes and generate bike flows appropriately. We refer to an intelligent data analysis approach providing a spatio-temporal information model and a generic representation of bike flows. Compared to simply averaging bike flows between stations, the information model comprises more information on system behavior. Furthermore, the information model yields expected values for bike flows allowing for the prediction of future demand.

To fully explore spatio-temporal characteristics of trips, intelligent data analysis is aligned along the Urban Transportation Planning Systems (UTPS) process (Johnston 2004). The UTPS is a common approach to model trips in urban areas. It comprises trip generation, trip distribution, mode choice and route selection, and provides an estimate of traffic flows for individual links of the considered transportation network. Input data for the UTPS is usually derived from costly surveys. Based on a rather small sample size, surveys provide a general picture on customer behavior, including background information on the purpose of a trip and the customer's attitude towards the usage of SMS. As opposed to this, we derive trip purposes from an extensive amount of detailed trip data, which are available at low costs. The challenge is to carve out the different trip purposes at stations from the observed rental and return activities. This allows for generalization of mobility behavior and generation of demand scenarios in terms of bike flows. Nevertheless, we cannot model the influence of varying resource allocation on future customer behavior. Thus, we assume that tactical planning decisions do not immediately change customer behavior.

Sophisticated aggregation of trips is necessary in order to model the distribution of trips. Depending on time of day and location of a station, a station may serve as trip generator, attractor or both. Due to commuter trips, for example, stations in residential areas would mainly show a high rental activity in the morning and a high return activity in the afternoon, whereas stations in working areas would mainly show the opposite activity. However, a station may not be used for one trip purpose only and therefore rental and return activity according to different trip purposes overlap. Simple inspection of activity is thus not sufficient to distinguish main trip purposes at stations. We employ intelligent data analysis to segment bike stations according to their temporal rental and return activity in the course of the day in order to determine the temporal distribution of trip purposes. With the temporal distribution at hand, the spatial distribution of trips between groups of stations with similar temporal activity can be determined.

The determination of temporal activity clusters requires the aggregation of trips according to a suitable timescale. The idea is to represent demand variation as accurately as possible without smoothing out relevant information, providing a compact representation of demand variation for information and optimization models. The segmentation occurs by cluster analysis yielding temporal activity clusters. A *temporal activity cluster* refers to a group of stations with similar rental and return activity at different times of the day. We implement cluster analysis by means of the Expectation-Maximization algorithm (Dempster et al. 1977). Results are evaluated by several internal validation measures as well as external validation by exploratory data analysis and expert interviews. For details on the cluster analysis, see Vogel et al. (2011). The result is a compact representation of the different temporal rental and return activity at stations.

When it comes to the determination of the *spatial distribution of trips*, the direct consideration of trips between individual stations would not be appropriate. Even in BSS with a high usage, the number of trips between most pairs of stations will be rather low. Due to the small sample size, the influence of non-typical behavior in trips, e.g. starting rain, events and full or empty stations, at individual stations could superimpose the activity and thus main trip purposes. Hence, the distribution of trips of between stations is derived from the associated temporal activity clusters in two steps:

- The *inter-cluster distribution* describes trip distribution patterns between stations of individual activity clusters. They are specified by the proportion of trips from a particular activity cluster to a particular cluster for a given time of the day. For instance, in the morning, the majority of trips is directed from residential to working clusters, whereas the opposite is true for afternoon hours.
- The *intra-cluster distribution* specifies how trips are distributed from a particular station to the stations of a cluster. We determine destinations according to the distance between stations and the trip duration. The distribution of trip durations can be empirically derived from observed trip data.

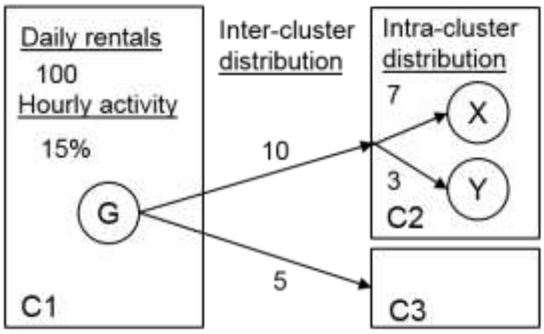

Fig. 2 Flow generation according to the temporal activity and spatial distribution

Let us exemplify the generation of typical bike flows as follows (cf. Fig. 2). In a specific hour of the day, we consider a generator station *G* in cluster *C1* with a particular rental activity and two destination clusters *C2* and *C3*. *C2* consists of two stations *X* and *Y*. *X* is closer to *G* then *Y*. For the sake of simplicity, the stations of *C3* are not considered. Combining the temporal and spatial distribution of trips, the flows from station *G* to *X* and *Y* are then derived as follows:

- The expected number of daily rentals for station *G* is 100.
- The number of daily rentals is then temporally distributed according to the rental activity cluster of this station, which reflects that 15% of the daily rentals account for the specific hour of the day.
- Next, rentals are spatially distributed to activity clusters according to the time-dependent inter-cluster distribution. According to the inter-cluster distribution, 2/3 of rentals are distributed from *C1* to cluster *C2* and 1/3 from *C1* to *C3*. Hence, 10 rentals are distributed to *C2* and 5 to *C3*.
- Rentals are further distributed within each activity cluster according to the intra-cluster distribution, contributing to a particular flow between a given OD pair. Assume that 70% of the trips have a short duration and 30% have a long duration. Thus, 70% of the rentals from *G* are assigned to station *X* and 30% to station *Y*, because *X* is closer to *G* then *Y*.

In sum, for a specific hour of the day, the derived flow from station *G* to *X* is set to 7 bikes and from *G* to *Y* is set to 3 bikes.

Since each station also serves as an attractor of trips and the number of daily rentals and returns may differ, the procedure is executed again to determine expected return activities at each station. Assume, for example, that the derived flow to station *G* from *X* is 4 bikes and to *G* from *Y* is 1 bike. In the end, the average of the generator and attractor bike flows is computed, providing time-dependent OD matrices. Here, the average flow from station *G* to *X* is 5.5 and from *G* to *Y* 2.

The presented approach has the drawbacks of smoothing demand variation and generating real valued bike flows. The information model distributes the total number of trips in the BSS to each spatio-temporal relation of the network. However, the derived flows can be interpreted as expected values for flows on the specific relations and allow for prediction of different demand scenarios. A method to scale and transform the real-valued bike flows into integer bike flows is presented in the following. The approach is based on the derived typical flows and assumes no changes in the general system behavior due to increasing demand, for example.
The two steps of scaling and transforming bike flows are as follows:

- In the scaling step, the real valued flows $f_{ij,t}$ between stations i and j in period t are multiplied such that they equal a desired number of flows d in relation to the total number of observed flows o: $f'_{ij,t} = mult \cdot f_{ij,t} \forall i,j \in N, t \in T$ with $mult = {}^{d}/_{o}$.
- In the transformation step, the flows are rounded by determining a threshold for rounding up and down such that the total number of rounded flows amounts to the desired number of flows: $Round\left(threshold, f'_{ij,t}\right) = d \ \forall i,j,t$
  with $Round\left(threshold, f'_{ij,t}\right)$: $If\ f'_{ij,t} - \lfloor f'_{ij,t} \rfloor < threshold\ then \lfloor f'_{ij,t} \rfloor\ else\ \lceil f'_{ij,t} \rceil$.

A binary search is applied in order to determine the threshold yielding the desired number of bike flows.
In order to evaluate the derived flows and adjust the parameters of the information model if necessary, flows are generated and compared to the observed flows. We present the evaluation of generated flows on the basis of real bike sharing data in Section 4.
The presented information model aims at representing dynamic system behavior as required by tactical resource allocation. Note that it could also be used to depict typical bike flows for strategic planning and individual trips for operational planning if parameterized accordingly. For strategic planning, spatio-temporal distributions of to be implemented stations can be forecasted from existing stations, e.g., one could assume that a new station in a residential area will likely show the same main trip purposes as an existing station in a similar area. For operational planning, the expected values have to be incorporated into probability functions such as the Poisson distribution. In combination with the distribution of trip durations, this allows for generating individual trips between stations. The information model is also applicable to other SMS recording a large amount of operational data.

3.2 An optimization model for tactical resource allocation in bike sharing systems

The following optimization model follows the work of Crainic (2000) on tactical service network design in freight transportation. Decisions on the tactical level aim at the optimal allocation and utilization of resources to fulfill customer service and economic goals. Total costs comprise fixed costs for offering a regular transportation service request between two locations in a network and variable costs that arise for a particular set of transported goods. In the case of BSS, the service operator transports bikes in capacitated trucks from full to empty stations to maintain a given service level. Thus, fixed transportation costs arise for a relocation request and variable costs for the particular handling of transported bikes.

We propose an LP formulation determining optimal target fill levels at stations in the course of day ensuring the fulfillment of demand scenarios at a predefined service level, e.g. avoid full or empty stations. The objective is to obtain fill levels at minimal total expected costs of system operation arising from relocation. We refrain from modeling the resource allocation as a computational challenging inventory routing problem (Campbell et al. 1998). Instead, relocation is anticipated by means of a dynamic transportation problem (Bookbinder and Sethi 1980) yielding the required relocation demand. We express relocation demand in terms of relocation regimes. A relocation regime refers to a set of relocation requests, i.e. pickup and return station, time-period, number of relocated bikes, for a given scenario of customer demand. The determined target fill levels serve as input for the optimization of relocation tours on the operational level. The relocation regimes aid the operator in constructing relocation tours on the operational level. However, tours have to be adjusted depending on the actual demand realization of the particular day.

Let N be a set of rental stations and T the set of periods in a day. The total number of bikes in the system is represented by b. The typical demand for bikes and bike racks is depicted by bike flows $f_{ij,t}$ between stations i and j in period t. Fulfillment of demand at stations depends on the given design and configuration of system infrastructure, i.e., the size of a station s_i for returns and the number of allocated bikes at each station and period $B_{i,t}$ for rentals. Relocation requests enable the compensation of missing bikes or bike racks by relocation of bikes $R_{ij,t}$ in each period at each day of system operation.

The objective is to minimize the total costs for relocation requests while ensuring the availability of rental and return resources for time-dependent "safety buffers" of bikes $sb_{i,t}$ and bikes racks $sbr_{i,t}$. Based on time-dependent OD matrices, the information model provides a scenario of resource flows $f_{ij,t}$, that serve as input for optimization. Note that the number of bikes in the system can be adjusted, e.g., when transition into another season or when bikes have to be taken out due to maintenance.

The resource allocation model reads as follows:

Sets

- $N = \{1, \dots, n\}$: set of stations
- $T = \{0, .., t_{max}\}$: set of periods, e.g., hours of the day. For resetting the number of allocated bikes at the end of the day, t_{max} includes the first period of the next day

Decision variables

- $B_{i,t} \in \mathbb{R}$: number of bikes at station i in time period t
- $R_{ij,t} \in \mathbb{R}$: relocated bikes between stations i and j in time period t

Parameters

- s_i : size of stations in terms of bike racks at station i
- b : total number of bikes in the system
- $f_{ij,t}$: bike flow between stations i and j in time period t
- ch_t : average handling costs of one relocated bike in time period t
- ct_{ij} : average transportation costs of one relocated bike between stations i and j
- $sb_{i,t}$: bike safety buffer at station i in time period t
- $sbr_{i,t}$: bike rack safety buffer at station i in time period t

With this notation the optimization model reads:

$$\text{Minimize} \quad \sum_{t=0}^{t_{max}} \sum_{i=1}^{n} \sum_{j=1}^{n} (ch_t + ct_{ij}) \cdot R_{ij,t} \quad (1)$$

subject to

$$B_{i,t+1} = B_{i,t} + \sum_{j=1}^{n} (f_{ji,t} - f_{ij,t} + R_{ji,t} - R_{ij,t}) \ \forall i \in N, t \in T \backslash t_{max} \quad (2)$$

$$B_{i,t} - \sum_{j=1}^{n} f_{ij,t} + \sum_{j=1}^{n} f_{ji,t} - \sum_{j=1}^{n} R_{ij,t} \geq \mathrm{sb}_{\mathrm{i,t}} \ \forall i \in N, t \in T \quad (3)$$

$$s_i - B_{i,t} - \sum_{j=1}^{n} f_{ji,t} + \sum_{j=1}^{n} f_{ij,t} - \sum_{j=1}^{n} R_{ji,t} \geq \mathrm{sbr}_{\mathrm{i,t}} \ \forall i \in N, t \in T \quad (4)$$

$$R_{ij,1} = 0 \ \forall i, j \in N \quad (5)$$

$$B_{i,1} = B_{i,t_{max}} \ \forall i \in N \quad (6)$$

$$\sum_{i=1}^{n} B_{i,t} = \mathrm{b} \ \forall \mathrm{t} \in \mathrm{T} \quad (7)$$

$$B_{i,t}, R_{ij,t} \geq 0 \ \forall i, j \in N, t \in T \quad (8)$$

In the objective function (1), the costs for relocation requests are minimized, comprising handling and transportation costs. Depending on the given infrastructure configuration, potentially missing bikes or bike racks are compensated by relocation of bikes $R_{ij,t}$ between stations for each period of the day. Equation (2) ensures flow conservation, i.e., the number of bikes at a station in the next period depends on the current number of bikes plus returns from customers (f) and relocation activities(R), minus customer rentals and relocation pickups.

Here, we assume that a particular relocation request is realized within one period. If relocation requests take longer, (2) has to be adjusted by setting $R_{ji,t-1}$ as well as the range of index t. The availability of resources is ensured by constraints (3) and (4). On the one hand, it is guaranteed that a sufficient number of bikes (3) is present at every station and period, i.e., the number of bikes minus customer rentals plus costumer returns and relocation pickups is always larger than a given bike safety buffer $sb_{i,t}$. On the other hand, the number of free bike racks (bikes racks minus allocated bikes, customer and relocation returns plus customer rentals) is always larger than the bike rack safety buffer $sbr_{i,t}$ (4). These two constraints also ensure that rented bikes and bike racks are not available for relocation in the particular period and all demand is satisfied. This is a rather optimistic approach, since customer rentals and returns are interchanged simultaneously. An alternative approach is to handle bikes and bike racks as separate resources, but this would result in a too pessimistic modeling since recently returned bikes could not be used by the next customer immediately. The above constraints enable particular buffers for bike and bike racks depending on the time of the day. For instance, in periods with a high rental activity and a low return activity at a station, the bike safety buffer can be set to a high value while the bike racks safety buffer can be kept low. Relocation requests are not allowed in the first period (5), and the initial fill level is restored at the end of the day (6). Equation (7) ensures that all existing bikes need to be allocated. Decision variables must be non-negative (8).

The above optimization model is based on an LP relaxation of relocation requests, assuming that there is no restriction on the frequency and capacity of relocations. We thus model relocation requests as real-valued variables. Note that this is a very rough approximation of relocation requests, which cannot replace detailed optimization for relocation from an operational perspective by means of subsequent vehicle routing. Our LP formulation handles fixed costs and variable costs for relocation in terms of an average handling and transportation cost per bike, ignoring that relocation of several bikes, for example, would decrease the cost of relocation for a particular bike. An alternative to the LP model introduced above would be a MIP formulation. Here, an integer variable would allow constraints on the frequency and capacity of relocation requests by pooling relocation operations. However, MIP service network design models are usually hard to solve (Crainic 2000). To keep our model computationally tractable, we decided against implementing an MIP formulation, assuming that the LP relaxation yields a sufficient approximation for the tactical planning perspective.

4 Decision support for tactical resource allocation of "Citybike Wien"

In the following study, the presented approach is applied to real bike sharing data in order to demonstrate the interplay of information and optimization model. The information model is parameterized based on extensive trip data arising from operation of Vienna's BSS "Citybike Wien" (CW) and two demand scenarios are generated (Section 4.1). Furthermore, target fill levels in the course of day for each station depending on the scenarios are determined (Section 4.2). In addition, characteristics of relocation regimes are investigated.

4.1 Modeling trip purposes and predicting typical bike flows

CW provided trip data for the years 2008 and 2009. The dataset comprises approx. 750'000 data records for 59 bike stations with a total of 1253 bike racks and 627 bikes. In order to reflect the typical usage of the system, we focus on summer trips only (April to October) accounting for 72% of all trips. At average, 1569 trips a day (~2.5 trips per bike) occur at that time of the year. The data analysis tool RAPIDMINER (http://rapid-i.com/) was used for generation, documentation and implementation of the information model. The transformation and scaling of flows was implemented in JAVA.

4.1.1 Temporal distribution of trips

For temporal modeling, we look for an appropriate timescale in order to aggregate individual trip data for resource allocation. In particular, the duration of trips is calculated according to rental and return times, and then aggregated in one minute buckets. Trip durations follow a Poisson-like distribution (Fig. 3a). The average trip lasts approx. 27 minutes with a median of 16 minutes. About 92% of trips are shorter than 60 minutes. Moreover, almost 70% of the trips end within the same hour, e.g., a trip that starts at 3:xx will end at 3:xx. Hourly aggregation thus seems sufficient to reflect temporal variation of bike rentals.

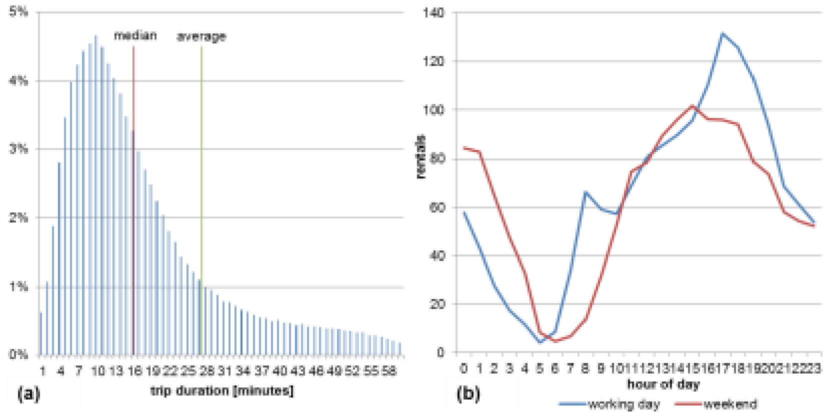

Fig. 3 Distribution of trip duration **(a)** and rental patterns on working and weekend days **(b)**

Next, temporal rental patterns on working days and weekend days are considered. The aggregation of trips results in 24*7 values representing the average rental activity in every hour of each day. Working days and weekend days show distinct rental patterns (Fig. 3b). Working days have three peaks: (1) a night peak probably resulting from the cessation of subway service, (2) morning commutes, and (3) the overall daily peak in the afternoon hours due to overlapping commuter and leisure usage. Weekend days clearly indicate a leisure-dominated activity by a distinct night peak, whereas the morning peak is missing, accompanied by a major usage in the early afternoon. In the following, we focus on the analysis of working days only, since relocation is usually not carried out on weekends. Rental and return activities for every station and each hour of the day are considered with “activity” referring to the hourly fraction of daily rentals or returns, respectively.

For the above data set, cluster analysis yields five activity clusters by means of cluster centroids representing the main trip purposes at stations that were assigned to the cluster. Fig. 4 shows the obtained rental and return activities as well as the geographical distribution of clusters in the city of Vienna:

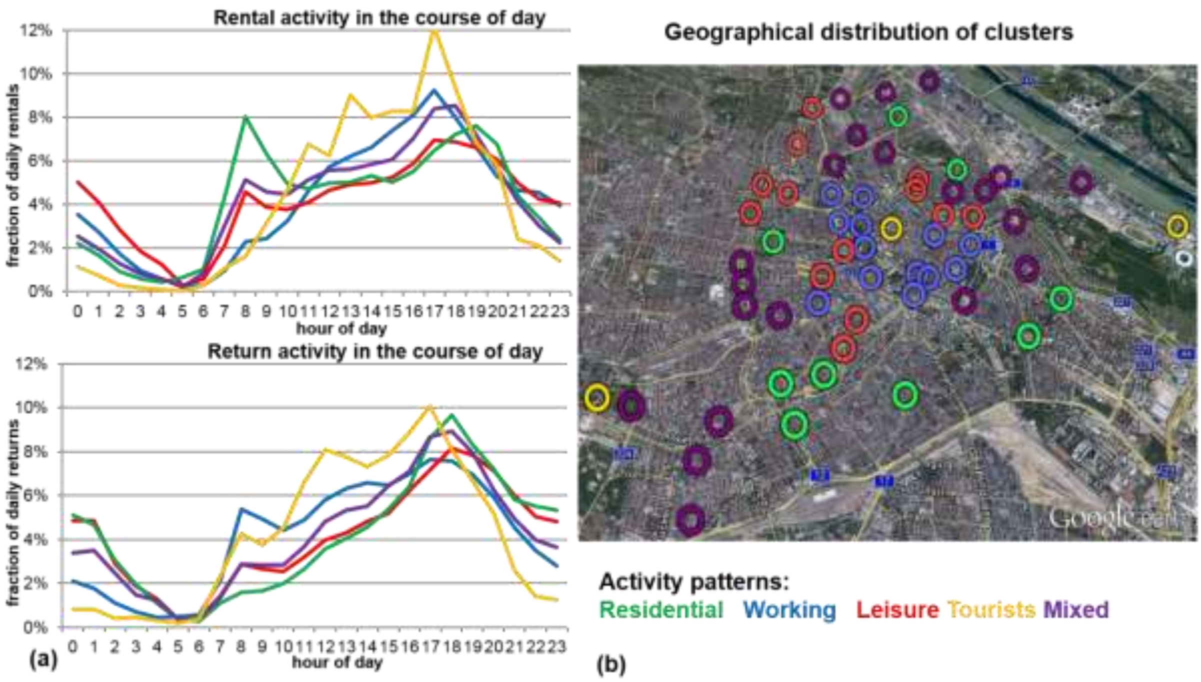

Fig. 4 Rental and return activity clusters **(a)** and geographical distribution of clusters **(b)**

- Stations within the *working* cluster are characterized by commuter trips showing a return activity peak in the morning and a rental activity peak in the late afternoon. These stations are located in the city center, having a high number of working places and points of interest as well as a low proportion of residents.
- The *residential* cluster shows the opposite activity of commuter trips: dominating rental activity in the morning and return activity in the afternoon. These stations are located at the periphery, which has more residential buildings.

- The *leisure* cluster shows activities similar to the residential cluster, but stands out due to different nighttime activities due to leisure trips. These activities are probably caused by popular nightlife districts.
- The *tourist* cluster is distinguished by high daytime rental and return activity, but almost no nighttime activity. Stations are close to popular tourist attractions in the west (castle Schoenbrunn), east (Prater carnival) and the city center (St. Stephan's Cathedral). Note that CW's "tourist card" is also handed out next to the city center station, which may explain the distinguished activity of this station.
- The *mixed* cluster represents stations that cannot be distinguished according to their main trip purposes and thus reflects the average rental and return activity on working days (cf. Fig. 3b). This observation is also underlined by the location of mixed stations, which is usually between stations of other clusters.

4.1.2 Spatial distribution of trips

Based on the activity clusters, the spatial distribution of trips between activity clusters is determined considering the different purposes. Results for the time-dependent inter-cluster distribution are exemplified by stations of the residential cluster (cf. Fig. 5). The figure shows the distribution of trips starting at stations of the residential cluster and ending at other clusters.

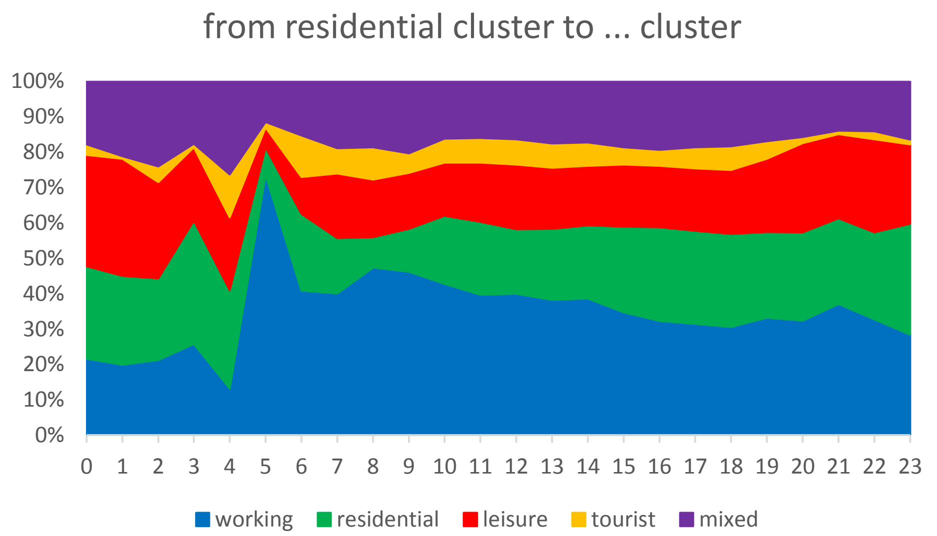

Fig. 5 Inter-cluster distribution between the residential cluster and other clusters

In the morning hours, more than 40% of trips starting at the residential cluster end at the working cluster reflecting commuter trips. Striking is the peak hour at 5 am showing 70% commuter trips. Since this is the hour with the lowest overall usage, commuter trips might be overrepresented. During the rest of the day, commuter trips decline with small peaks at 21pm and 3 am. Trips from the residential cluster to the residential cluster show a nearly opposite behavior: less than 10% in the morning hours and up to 20% in the late afternoon and night hours. The latter increase might be caused by trips between residential stations without a direct public transport connection. The distribution of trips to the leisure cluster shows a similar run of the curve like the residential cluster. However, the percentage at night is higher presumably caused by pub goers. The distribution of trips to the tourist cluster is near zero during the night and only 3% - 8% during the hours from 6 am to 6 pm. Trip distribution to the mixed cluster shows no specific distinctions and is almost stable around 20% in the course of day.

4.1.3 Generation and validation of flows

The CW serves as an example for the generation of flows. Bike flows are derived according the methodology presented in Section 3.1, providing 24 time-dependent OD matrices for 59 stations. The data set contains the derived typical flows comprising 1569 daily trips performed with 627 bikes. Thus, the information model distributes the 1569 trips to 83544 OD pairs (59 stations times 59 stations times 24 time-periods). An extract from the data set showing one period and three stations is depicted below. Rows present origin stations, columns destination stations and data entries the absolute real-valued derived flows:

$$\begin{bmatrix} 0.1591 & 0.0596 & 0.2153 & \dots \\ 0.0659 & 0.6611 & 0.1594 & \dots \\ 0.2212 & 0.1860 & 0.3703 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

We compare derived and observed flows in order to validate the quality of the information model. Therefore, the normalized rental and return activity for each station is calculated from the derived flows.

Fig. 6 shows the average difference of observed and derived activity aggregated for each cluster. The derived cluster activity patterns almost match the observed activity patterns. In general, the difference of derived and observed activity is very low, ranging from -0.5% to 0.5%. Note that the hourly sum of the absolute difference for each cluster amounts to only 5% on average. Only the return activity for the tourist cluster shows outliers in the morning and evening. The Prater station is responsible for this deviation. Due to the fringe location and touristic usage, the fraction of pendulum trips and trip durations are higher than at other stations. On the station level, the derived activity is less accurate. The hourly sum of the absolute difference for each cluster amounts to 14.5% on average with a standard deviation of 4.7%.

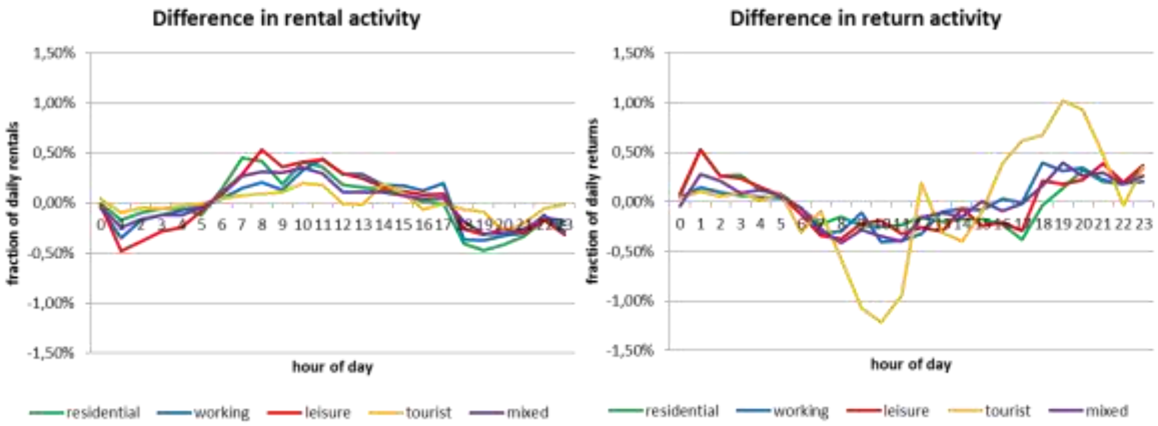

Fig. 6 Observed minus derived cluster activity

It can be concluded that the information model reasonably reproduces the CW behavior for tactical planning purposes. Incorporating cluster-dependent trip durations might improve the quality of the information model, though.

The generated flows have the flaw to be real-valued. Therefore, flows are transformed into integers. First, the data set is normalized such that the total number of trips equals the number of bikes in the system by dividing each data object by 1569 / 627 = 2.5… (below left). Second, the threshold (0.0884) yielding 627 trips after rounding is determined. Applying this threshold generates integer flows (below right):

$$\begin{bmatrix} 0.0636 & 0.0238 & 0.0860 & \dots \\ 0.0263 & 0.2642 & 0.0637 & \dots \\ 0.0884 & 0.0743 & 0.1480 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix} => \begin{bmatrix} 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 1 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

In order to predict different demand scenarios, the procedure is executed with different scaling factors. For the following computational experiment, integer flows for the typical main season working day with an observed usage rate of 2.5 trips per bike a day are used. In order asses increasing demand on the existing infrastructure, integer flows with a usage rate of 5 trips per bike a day are applied.

4.2 Supporting decisions on fill levels and relocation regimes

The following computational experiment aims to determine optimal target fill levels while minimizing expected relocation costs using two scenarios of demand. As a result, we can identify fill levels at stations in the course of day in order to ensure a given service level. The evaluation of relocation regimes yields the extent that stations need relocation request and shows main flows of relocation.

The setup of the experiment is as follows. We consider the BSS network of CW which comprises $n = 59$ bike stations with a given number of bike racks and a total of $b = 627$ bikes (~50% fill level). Time is discretized in terms of $t_{max} = 24$ (hour) periods, and we assume that relocation requests take one hour on average (approx. 15-20 minutes for loading and unloading plus travel times between stations). According to the system operator, handling costs depend on the time of the day. Daytime handling costs are set to $ch_{day} = 4$ Euro (in effect for periods 8 to 17), while night time handling costs are more expensive ($ch_{night} = 7$ Euro). Transportation costs are assumed independent of the time of the day and amount to $ct_{ij} = 0.5$ Euro per kilometer. Bike and bike rack safety buffers are set to 10% of the station size for each station and period. The total number of initially allocated bikes obtained from the optimization is rounded and adjusted to $b = 627$. Two scenarios serve as input: typical flows for a usage rate of 2.5 (1569 trips) and 5 (3135 trips). The LP model was implemented in IBM ILOG OPL and solved with CPLEX 12.5 on an INTEL Core i5 processor at 3.2 GHz and 8 GB RAM running Windows 7 64 Bit. All instances were optimally solved within 5 seconds of run time.

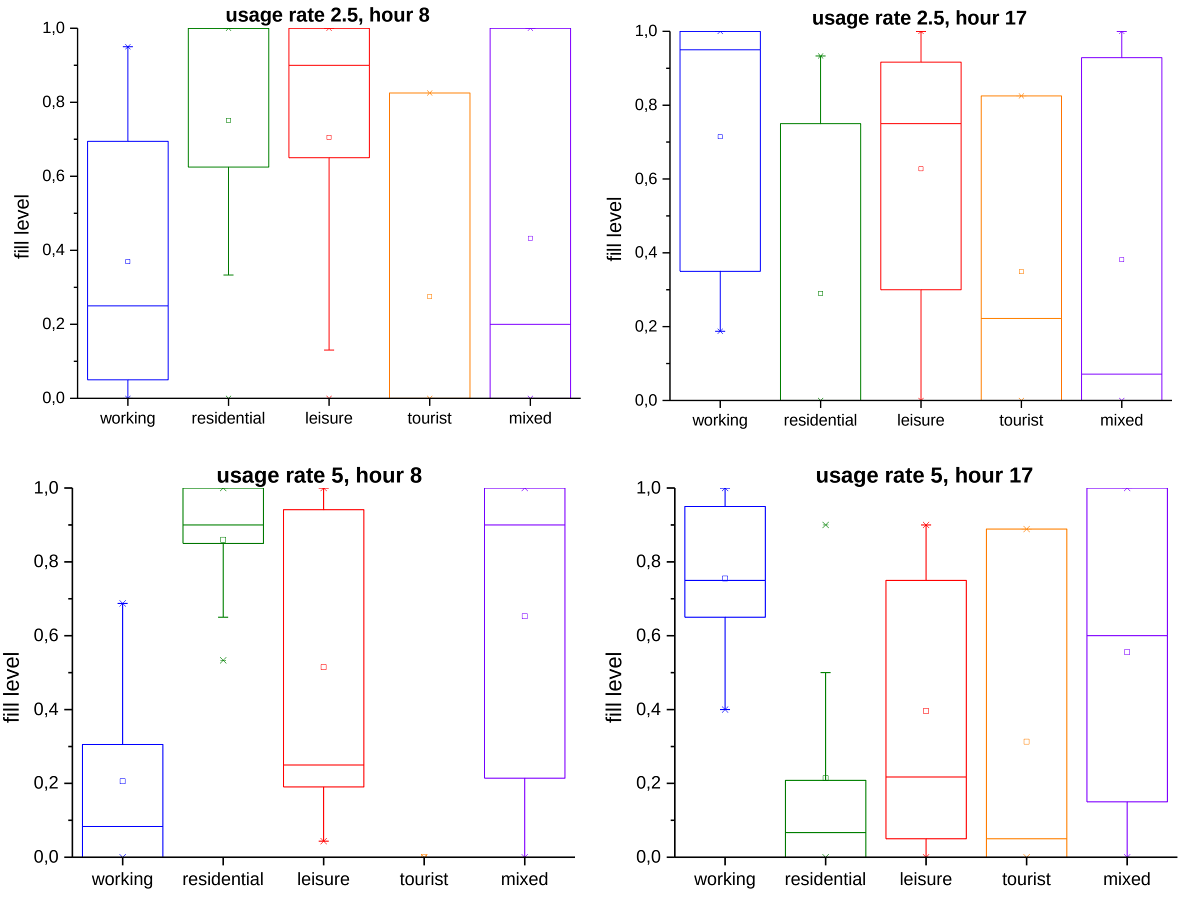

Fig. 7 *Boxplots of fill levels per cluster for peak hours in two demand scenarios*

Evaluation of determined fill levels occurs by boxplots. Fill levels aggregated by cluster for the morning and afternoon peak hours are depicted in Fig. 7. The scenario with a usage rate of 2.5 rentals per bike a day reflects the observed demand for a typical working day. In the morning peak hour, stations belonging to the working cluster require a low fill level whereas residential and leisure cluster stations require a high fill level. Commutes from residential and leisure cluster stations to working cluster stations are responsible for the high fill levels. However, the variance of fill levels within the working cluster is significantly higher than in the residential and leisure clusters. Working cluster stations do not only serve as return stations in the morning. They are also used as rental stations to bridge non-existing subway connections. In addition, bikes are needed in the afternoon. The average fill levels at stations belonging to the tourist and mixed clusters is almost 50% and fill levels feature a high variance. This observation is especially obvious for the mixed cluster since trip purposes are diverse at these stations. Tourist stations are either (almost) full or (almost) empty. Demand at tourist stations shows no typical behavior but high variation which cannot be reflected by the tactical information model. Thus, tourist stations play a minor role within tactical planning.

In the afternoon peak hour, working and residential cluster stations require opposite fill levels compared to the morning hour. Striking is the high variance of the residential cluster. The reason for this behavior is similar to the reason for working cluster stations in the morning: in the afternoon residential cluster stations also serve as rental stations to bridge subway connections. In addition, unused bikes simply remain in the residential cluster stations for the morning peak of the upcoming day. Fill levels at tourist and mix cluster stations are similar to the morning hour. In sum, fill levels reflect the rental and return activity of the clusters.

Within the scope of tactical planning, it is of interest how fill levels change if demand increases. Here, we evaluate the effect on fill levels for a usage rate of 5 rentals per bike a day. Focus lies on the working, residential, leisure and tourist clusters since the mixed cluster shows similar fill levels compared to the lower usage rate. In the morning peak hour, the higher fill levels at residential cluster stations and lower fill levels at working cluster stations are striking. In order to fulfill the increased demand of commuter trips in the morning, the working cluster requires lower fill levels and the residential cluster higher fill levels. However, bikes are still needed for trips within the working cluster. Thus, additional bikes are withdrawn from the leisure cluster to the residential cluster. The tourist cluster stations are empty in the morning, since the morning activity is zero. Commuter trips superimpose touristic trips in the information model and thus bikes are not required at tourist stations. In the afternoon peak hour, working cluster stations show higher fill levels due to increased rentals. However, the residential cluster also needs bikes for rentals. Thus, bikes are withdrawn from the leisure cluster. The activity at tourist stations is higher, indicated by higher fill levels.

In sum, the demand increase especially affects working and residential cluster stations. Stations belonging to the leisure serve as compensation stations for the increased demand for rentals and returns at working and residential cluster stations.

Also of interest is the effect of increased demand on expected relocation regimes. Doubling demand does not result in doubled relocation demand and costs. The number of relocated bikes increases from 119 (usage rate 2.5) by 1.8 to 215 bikes (usage rate 5). The expected relocation costs increase by a factor of 1.87 (546 to 1023 Euro). Reason for this is that bikes are relocated from or to further away stations. Nevertheless, the increase in relocation demand and costs is not as big as the demand increase. Due fill levels adapted to the higher demand, relocation effort and costs are saved. In order to get a better impression of relocation demand and regimes, the total relocation demand at stations is studied. Furthermore, characteristics of relocation regimes are investigated.

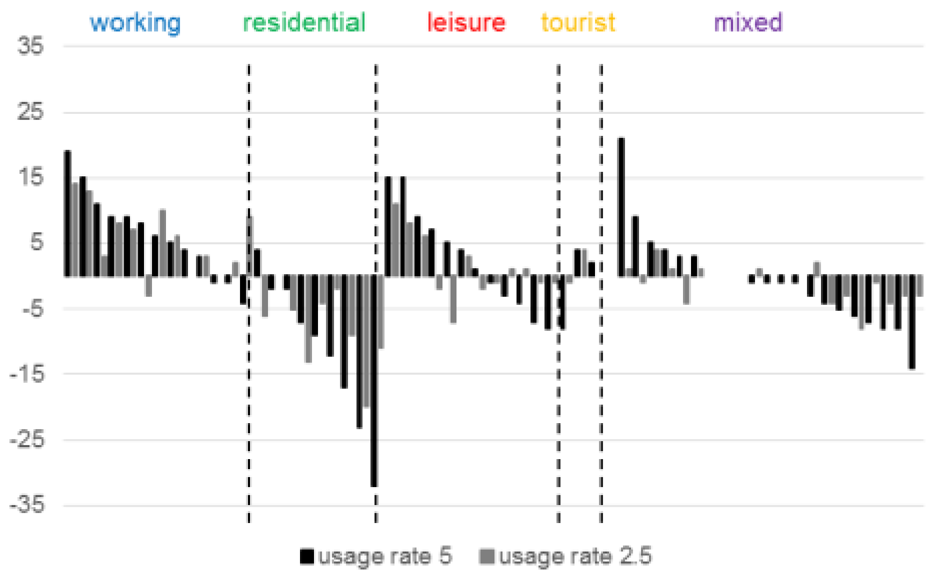

Fig. 8 *Total number of returned (positive) and picked up (negative) bikes by relocation*

Fig. 8 shows the total number of returned and picked up bikes by relocation at each station, sorted by cluster assignment for the two scenarios. We can clearly identify stations that require relocation returns, pickups or stay balanced. Striking is that either relocation pickups or returns occur at a station. In particular, stations that require relocation returns mainly belong to the working and leisure cluster. Stations requiring relocation pickups belong to the residential and tourist cluster. The mixed cluster stations need both relocation pickups and returns, but they are often able to (almost) balance pickups and returns properly. Comparison of the two scenarios shows that the tendency of a station requiring either relocation returns or pickups remains except for a few stations.

Based on these observations, we can derive the urgency for relocation at stations. Regarding the usage rate of 2.5, 35% of the stations require one bike or less to be relocated. Hence, the fill levels properly compensate the demand variation and only little relocation is needed in the course of the day. In contrast, almost 12% of the stations need more than ten bikes to be relocated and are thus of high priority for relocation, e.g. regular relocation tours each day. Regarding the scenario with a usage rate of 5, 20% of the stations require one bike or less to be relocated. Instead, almost 20% of the stations require more than ten bikes to be relocated. Hence, implications for the operational level are that these stations have to be served on a daily basis whereas stations with lower priority may be served only once a week, for example. Note that these numbers are probably a lower bound due to the discussed limitations of the LP model and tactical modeling of demand.

Table 1 *Proportion of relocation flows between clusters*

Usage rate 2.5

origin\destination	working	residential	leisure	tourist	mixed
working	2,52%	--	--	--	--
residential	38,66%	--	15,97%	--	4,20%
leisure	5,88%	--	3,36%	3,36%	
tourist	--	--	--	--	--
mixed	15,97%	--	5,88%	--	4,20%

Usage rate 5

origin\destination	working	residential	leisure	tourist	mixed
working	4,19%	--	--	--	1,86%
residential	20,47%	3,72%	8,84%	1,86%	17,21%
leisure	9,30%	1,86%	3,26%	--	--
tourist	--	--	--	--	--
mixed	10,70%	--	13,95%	0,93%	1,86%

Visualizing the direction of relocation flows helps to understand characteristics of relocation regimes. Table 1 shows the proportion of relocation flows between clusters for the two scenarios. Striking in both scenarios is the high proportion of relocation flows to the working cluster. Although main bike flows are directed from the peripheral clusters to the working cluster in the morning, it seems that there are still bikes needed at the working cluster in the afternoon, because at that time, the rental activity is higher than the return activity (cf. Fig. 4). Furthermore, most of the bike flows in the afternoon are directed from the working cluster to the peripheral clusters. Hence, working cluster stations have to be filled before the afternoon peak. The situation for stations of the leisure cluster is similar: they have to be filled in the morning (after the early morning rental peak) and in the afternoon (before the rental peak at night).The characteristics of relocation regimes support BSS operators in the planning and implementation of relocation tours.

5 Conclusions and future research

In this paper, we have proposed an integrated approach of intelligent data analysis and mathematical optimization for tactical resource allocation in BSS. An information model has been presented allowing for the generic derivation of spatio-temporal customer demand in terms of time-dependent OD matrices. Derived OD matrices have served as input for an LP based resource allocation model. The optimization model determines the optimal fill level at stations minimizing the expected costs of relocation while ensuring a predefined service level. The computational study has shown that the approach yields reasonable fill levels at stations and identifies the extent that stations need relocation operations as well as typical flows of relocation regimes.

Future research could investigate improved ways of modeling relocation requests for BSS and the adaption of the resource allocation model to related SMS. Pooling of relocations with the help of a MIP approach would better reflect actual relocation operations than captured by an LP model. Furthermore, a more robust resource allocation model would be desirable. Generated demand scenarios reflect differences in demand variation. These scenarios could be applied to a stochastic optimization model. In addition, the identified relocation flows can support optimization of relocation tours on the operational level. Here, optimization models can be decomposed according to the clusters to reduce the complexity. Regarding the evaluation of tactical decisions, it might be interesting to study the influence of fill levels on resulting operational relocation tours. Another interesting future path is the determination of optimal safety buffers for each station and period to improve target fill levels. Whilst the above information model is generally applicable to SMS, the resource allocation model has to be modified for other means of transportation, system designs and business models.

Acknowledgements

We thank „Gewista Werbegesellschaft m.b.H.“ for providing data from their project “Citybike Wien” and in particular Dr. Hans-Erich Dechant for the helpful discussion of related issues and results.

References

Benchimol M, Benchimol P, Chappert B, et al. (2011) Balancing the stations of a self-service bike hire system. RAIRO-Operations Research 45:37–61.

Bookbinder JH, Sethi SP (1980) The dynamic transportation problem: A survey. Naval Research Logistics Quarterly 27:65–87.

Borgnat P, Robardet C, Rouquier JB, et al. (2010) Shared bicycles in a city: a signal processing and data analysis perspective. Advances in Complex Systems (ACS) 14:415–438.

Büttner J, Petersen T (2011) Optimising Bike Sharing in European Cities - A Handbook.

Caggiani L, Ottomanelli M (2012) A modular soft computing based method for vehicles repositioning in bike-sharing systems. Procedia-Social and Behavioral Sciences 54:675–684.

Campbell A, Clarke L, Kleywegt A, Savelsbergh M (1998) The inventory routing problem. Fleet management and logistics. Springer, pp 95–113

Contardo C, Morency C, Rousseau LM (2012) Balancing a dynamic public bike-sharing system.

Correia GHA, Antunes AP (2012) Optimization approach to depot location and trip selection in one-way carsharing systems. Transportation Research Part E: Logistics and Transportation Review 48:233 – 247.

Crainic TG (2000) Service network design in freight transportation. European Journal of Operational Research 122:272–288.

DeMaio P (2009) Bike-sharing: History, impacts, models of provision, and future. Journal of Public Transportation 12:41–56.

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society Series B (Methodological) 1–38.

Froehlich J, Neumann J, Oliver N (2009) Sensing and predicting the pulse of the city through shared bicycling. International Joint Conference on Artificial Intelligence. pp 20–26

George DK, Xia CH (2011) Fleet-sizing and service availability for a vehicle rental system via closed queueing networks. European Journal of Operational Research 211:198–207.

Johnston RA (2004) The Urban Transportation Planning Process. The geography of urban transportation

Jorge D, Correia G, Barnhart C (2012) Testing the validity of the MIP approach for locating carsharing stations in one-way systems. Procedia-Social and Behavioral Sciences 54:138–148.

Kaltenbrunner A, Meza R, Grivolla J, et al. (2010) Urban cycles and mobility patterns: Exploring and predicting trends in a bicycle-based public transport system. Pervasive and Mobile Computing 6:455–466.

Kek AGH, Cheu RL, Meng Q, Fung CH (2009) A decision support system for vehicle relocation operations in carsharing systems. Transportation Research Part E: Logistics and Transportation Review 45:149–158.

Lin JR, Yang TH (2011) Strategic design of public bicycle sharing systems with service level constraints. Transportation Research Part E: Logistics and Transportation Review 47:284–294.

Lin J-R, Yang T-H, Chang Y-C (2013) A hub location inventory model for bicycle sharing system design: Formulation and solution. Computers & Industrial Engineering 65:77–86.

Martinez LM, Caetano L, Eiró T, Cruz F (2012) An optimisation algorithm to establish the location of stations of a mixed fleet biking system: an application to the city of Lisbon. Procedia-Social and Behavioral Sciences 54:513–524.

Midgley P (2011) Bicycle-sharing schemes: enhancing sustainable mobility in urban areas.

Nair R, Miller-Hooks E (2011) Fleet management for vehicle sharing operations. Transportation Science 45:524–540.

Nair R, Miller-Hooks E (2014) Equilibrium network design of shared-vehicle systems. European Journal of Operational Research 235:47–61.

Rainer-Harbach M, Papazek P, Hu B, Raidl GR (2013) Balancing bicycle sharing systems: A variable neighborhood search approach. Evolutionary Computation in Combinatorial Optimization. Springer, pp 121–132

Raviv T, Kolka O (2013) Optimal inventory management of a bike-sharing station. IIE Transactions

Raviv T, Tzur M, Forma IA (2012) Static repositioning in a bike-sharing system: models and solution approaches. EURO Journal on Transportation and Logistics 1–43.

Ricker V, Meisel S, Mattfeld DC (2012) Optimierung von stationsbasierten Bike-Sharing-Systemen. Proceedings of MKWI 2012. pp 215 – 226

Sayarshad H, Tavassoli S, Zhao F (2012) A multi-periodic optimization formulation for bike planning and bike utilization. Applied Mathematical Modelling 36:4944 – 4951.

Schuijbroek J, Hampshire R, van Hoeve W-J (2013) Inventory Rebalancing and Vehicle Routing in Bike Sharing Systems.

Shaheen SA, Guzman S, Zhang H (2010) Bikesharing in Europe, the Americas, and Asia. Transportation Research Record: Journal of the Transportation Research Board 2143:159–167.

Vogel P, Greiser T, Mattfeld DC (2011) Understanding Bike-Sharing Systems using Data Mining: Exploring Activity Patterns. Procedia-Social and Behavioral Sciences 20:514–523.

Vogel P, Mattfeld DC (2011) Strategic and Operational Planning of Bike-Sharing Systems by Data Mining – A Case Study. In: Böse, J and Hu, H and Jahn, C and Shi, X and Stahlbock, R and Voß, S (ed) Computational Logistics. Springer Berlin / Heidelberg, Germany, pp 127–141

Wang X, Lindsey G, Schoner JE, Harrison A (2012) Modeling bike share station activity: the effects of nearby business and jobs on trips to and from stations. Transportation Research Record 43:45.

Zahory MN (2009) Request for Proposals for the operation of the Arlington Bike-sharing Program. http://www.metrobike.net/index.php?s=file_download&id=18. Accessed 30 Jun 2012