Proceedings of 7th Transport Research Arena TRA 2018, April 16-19, 2018, Vienna, Austria

Semi-Automated Location Planning for Urban Bike-Sharing Systems

Markus Straub[a,*], Christian Rudloff[a], Anita Graser[a], Christian Kloimüllner[b,c], Günther R. Raidl[b], Markus Pajones[d], Felix Beyer[e]

[a]*AIT Austrian Institute of Technology GmbH, Giefinggasse 2, 1210 Vienna, Austria*
[b]*TU Wien, Institute of Computer Graphics and Algorithms, Favoritenstraße 9-11, 1040 Vienna, Austria*
[c]*Research Industrial Systems Engineering, Concorde Business Park F, 2320 Schwechat, Austria*
[d]*University of Applied Sciences Upper Austria, Department of Logistics, Wehrgrabengasse 1-3, 4400 Steyr, Austria*
[e]*Rosinak & Partner ZT GmbH, Schloßgasse 11, 1050 Vienna, Austria*

Abstract

Bike-sharing has developed into an established part of many urban transportation systems. However, new bike-sharing systems (BSS) are still built and existing ones are extended. Particularly for large BSS, location planning is complex since factors determining potential usage are manifold. We propose a semi-automatic approach for creating or extending real-world sized BSS during general planning. Our approach optimizes locations such that the number of trips is maximized for a given budget respecting construction as well as operation costs. The approach consists of four steps: (1) collecting and preprocessing required data, (2) estimating a demand model, (3) calculating optimized locations considering estimated redistribution costs, and (4) presenting the solution to the planner in a visualization and planning front end. The full approach was implemented and evaluated positively with BSS and planning experts.

Keywords: Bike-Sharing; Active Mobility (Cycling / Walking); Spatial Planning / Last Mile; Intermodality; Transport on Demand; Mobility As a Service

[*] Corresponding author. Tel: +43 50550-6289; fax: +43 50550-6439
Email address: markus.straub@ait.ac.at

1. Introduction

Bike-sharing has developed into an established part of many urban transportation systems as it provides an ecofriendly and healthy way of travelling through (large) cities (DeMaio (2009)). Particularly for large bike-sharing systems (BSS), location planning is complex since factors determining potential usage are manifold and usage of stations is influenced by neighbouring stations. However, choosing good station locations from the start is an important success factor directly impacting potential usage. Furthermore, high costs may arise from station relocations.

The state of the art of **location planning** is described in ITDP (2013) and Schroeder (2014) as mostly manual process consisting of two steps. First a list of potential station locations is drafted by following guidelines such as choosing well-connected places with high demand (general planning). Second the candidate selection and exact station placement is finalized through site visits and stakeholder engagement (detailed planning). General planning can be supported through geographic information systems (GIS), such as the analysis of area and population served by planned stations. For detailed planning, an automatic approach is not feasible since it is influenced by a multitude of intangible factors such as land ownership, existing infrastructure (e.g. gas pipes), or lines of sight. While there is plenty of research centring on BSS (for a good overview see Fishman (2016)), currently, little research is available on an automated support to the location planning of BSS. One exception is García-Palomares et al. (2012), where a GIS approach is applied to plan a BSS. While the expected demand is estimated using a simple approach based on the number of home and work locations in an area the main part of the paper applies GIS methodologies.

A reliable **demand model** is an essential ingredient for location planning. In the area of demand estimation for bikes in a BSS several papers exist both for system wide demand as well as for station by station demand. Most of the existing work is based on count models. Examples are Rudloff and Lackner (2014), where demand for bikes and return slots for each station is modelled with special consideration of neighbouring stations, Noland et al. (2016), where trip generation was modelled based on variables like land use and public transport accessibility from BSS stations, and Faghih-Imani and Eluru (2016), where spatio-temporal interactions are included in the demand model. However, here a **cell based approach** based on splitting the planning area into cells around one hectare in size is used. Each cell is a station candidate. This makes modelling of the unobserved demand between cells necessary. While count models would be the natural fit for the cell based demand model, due to combinatorial complexity these models are not applicable here. Thus, a simpler, linear approach is needed to estimate the demand of a new station.

Nearly all work in combinatorial optimization on the **Bike-Sharing Station Planning Problem (BSSPP)** uses different problem variations, i.e. different constraints, problem statements and optimization goals. Most apply mixed integer programming (MIP) approaches to solve these problems. Some solve the problem exactly by applying their mathematical models to a commercial MIP solver, see Lin and Yang (2011), Saharidis et al. (2014), Chen and Sun (2015), Hu and Liu (2014), or Frade and Ribeiro (2015). Others utilize a hybrid of a (meta)heuristic and a MIP based approach like, e.g., Martinez et al. (2012) and few works also apply pure (meta)heuristics to the problem, see Yang et al. (2010) or Lin et al. (2013). A detailed description of the algorithmic approach used within this work but applied on a slightly simplified problem variant was published in Kloimüllner and Raidl (2017). Last but not least, see Gavalas et al. (2016) for a survey on the design of vehicle sharing systems. However, all these approaches are only tested on small instances with the largest ones containing just a few hundred cells. No well scalable algorithms exist in literature that allow to solve problem instances with thousands of cells as required for solving the problem for large cities in real-world scenarios.

In this work we propose a tool for **semi-automatic location planning** for real-world sized BSS, from here on referred to as **SALP-tool**. The tool is designed to only rely on commonly available data and trip data from an existing BSS in order to make it applicable to the same or a comparable city. Therefore, the method is applicable to most planning scenarios worldwide. The target audience for such a tool are (traffic) planners, who can use it as follows: The planning process starts with the collection and preprocessing of required data (1). Then a demand model is estimated using the collected data (2). An algorithm calculates optimal station locations using the demand model as input (3). Finally the planner can fine-tune the suggested station locations in a planning front end (4). Planners can also define planning scenarios in step three, which can differ in many aspects such as the planning region or the available budget.

In the remainder of the paper we present the requirements and the four planning steps of the SALP tool in detail. The paper concludes with an evaluation and an outlook to future work.

2. Requirements

To define requirements for the SALP-tool an interview guide for semi-structured expert interviews was prepared based on literature review. In total 16 experts from the fields of bike sharing (three international, one national), planning (six national), and cycling in general (six national) were interviewed on the phone. Together with the experts the optimization goal was defined and the SALP-tools' potential inputs, optimization parameters, and outputs were prioritized.

2.1. Optimization Goal

The optimization goal was decided early on since method development depended on it: given a planning area disaggregated into smaller areas where a station can be placed (cells), a maximum budget for construction, a maximum budget for operation, and a number of possible station configurations (see parameter *costs* in Table 2), a BSS shall be planned by defining target cells and configurations of stations, such that the expected number of trips is maximized, i.e. as much mobility demand as possible is fulfilled via bike-share trips. According to our interviews a fixed budget or a fixed number of stations are common starting points for planners. We decided to use the former as it seems to be the more sensible regarding the quality of the BSS.

2.2. Inputs, Parameters, and Outputs

The proposed and chosen **inputs** are shown in Table 1. In addition to these inputs the proposed tool relies on (1) a complete road graph, (2) historical trip data and station fill levels for the reference BSS, and (3) historical weather data (temperature and precipitation) for the same period as the historical trip data. For this work we collected data for Vienna, Austria and the Viennese BSS Citybike Wien.

Table 1: Potential optimization inputs as ranked by experts in decreasing order of importance. Some inputs such as population and job data are commonly provided on the level of city districts or registration districts and available as (Open) Government Data (GD/OGD) from local governments or statistics offices. All other inputs can be extracted from the free world map OpenStreetMap (OSM). All entries with a check-mark are actually required as input by our proposed SALP-tool.

	input	data source	comment
✓	topography	OGD	
✓	public transport stops (count)	OSM	four categories: bus stops, tram stops, entries to underground and suburban train stations, entries to main train stations (where intercity trains stop)
✓	(quality of) bicycle routes	OSM	part of the road complete graph
✓	demographic data	OGD	limited to population data in our case
✓	jobs (count)	OGD	
✓	education POIs (count)	OSM	two categories: universities, schools
✓	population (count)	OGD	
✓	leisure POIs (count)	OSM	e.g. bars, cinemas, parks,...
-	cycling modal split	OGD	omitted: not useful for integration into demand model or optimization
-	timeseries for motorized traffic	GD	omitted: not commonly/freely available and not ranked high enough
✓	shopping POIs (count)	OSM	
-	traffic counts	GD	omitted: not commonly/freely available and not ranked high enough
✓	tourism POIs (count)	OSM	addition: suggested by experts; e.g. attractions, hotels, viewpoints,...

The **optimization parameters** (see Table 2) can be used by planners to influence the optimization process of the SALP-tool. This is especially useful for interactive and iterative planning scenarios where the whole optimization process is run multiple times. Some desired parameters can not be supported because they are not suitable regarding the defined optimization goal, i.e. the optimization actually optimizes the parameter.

The desired **outputs** ranked by the experts are listed in Table 3. Except for the revenue, which can not be calculated since we did not implement tariff systems due to the high complexity, everything is covered by the output of the SALP-tool, which is an assignment of the predefined station configurations to the planning cells

Table 2: Optimization parameters as ranked by experts in decreasing order of importance. All entries with a check-mark are actually supported by our proposed SALP-tool.

	parameter	comment
✓	spatial accessibility	can be adjusted by weighing e.g. metro stops higher
-	station density	omitted: automatically determined by the optimization
-	goal of the BSS	only one optimization goal was within the scope of this work
✓	costs	the planner must define one or more station configurations such as a station with 10, 20, or 40 boxes for bicycles and assign both construction and operation costs. Construction costs include all costs arising during the construction phase such as building stations or buying terminals and bicycles. Operation costs are defined for a certain period such as three years and include e.g. maintenance, repair, and redistribution logistics.
-	tariff system	omitted: complexity too high, out of scope
✓	station sizes	see costs
-	number of stations	omitted: automatically determined by the optimization; can be implemented when using a different overall optimization goal
✓	planning area	a polygon of the planning area (including holes for barriers or large areas where cycling is not possible such as a railway station or a river)
✓	types of bicycles	not relevant for conventional bicycles / ranked high enough (implicitly included through costs), e-bike-sharing data not available in this work
✓	types of stations	not relevant / ranked high enough (implicitly included through costs)

and the estimated demand between the planned stations. For more details about the requirements the reader is referred to Pfoser and Pajones (2016).

Table 3: Optimization outputs as ranked by experts in decreasing order of importance. All entries with a check-mark are actually outputs of our proposed SALP-tool or possible to do with the output.

	output	comment
✓	station locations	
✓	costs	
✓	optimum number of stations in planning area	implicit output
✓	estimated usage	
-	estimated revenue	omitted: complexity too high, out of scope
✓	station size	implicit output: the optimization chooses one of the predefined station configurations
✓	estimated redistribution effort	
✓	manual editing of the planning result	

3. Data Collection and Data Preprocessing

Step one defines a commonly available dataset that is required for the demand model. The demand model requires (1) potential locations to choose from and (2) matrices containing the altitude difference and travel time between these locations for the modes of transport walking (when going to and from a station) and cycling (for cycling between stations). To define these potential locations, a novel tessellation approach (Graser, 2017) is used which generates cells that cover the whole planning area seamlessly using a road graph. These cells are then used to map all relevant input variables for the demand model.

3.1. Cell Generation

Our cell generation approach aims at creating a tessellation that is suited for planning BSS. In general, tessellations are subdivisions of space without overlaps or gaps. Common tessellations of urban areas are city blocks or districts and regular square or hexagonal grids. However, these common tessellations are not well suited for BSS planning. City blocks or districts focus on the area between streets rather than the street space itself. When a location recommendation refers to a city block, it remains unclear for the BSS planner which streets enclosing the block should be preferred. Regular grids, on the other hand, are sensitive to the selected cell size. Tessellation with small cells are likely to contain cells that are disconnected from the street network. Large cells, on the other hand, may contain areas that are disconnected due to barriers, such as rivers or rail infrastructure without crossings. Furthermore, large cells are of limited use for BSS planning since most pedestrians will only walk a limited distance to get to a station. In our tessellation, cells are centred around street intersections. Our approach extracts suitable tessellation seeds at intersections from the underlying road graph and then constructs Voronoi cells (that is, the area consisting of all points closer to that seed than to any other) which account for unsuitable areas and barriers. For a detailed description of the algorithm, the reader is referred to Graser (2017).

3.2. Mapping Input Data to Cells

To map **population and job data** to our cells, we determine population and job density per square meter of the input area. Then we compute intersections of input area geometries and our cell geometries. Finally, we can determine the population and jobs for each cell by summing up the products of density and intersection area. If land-use data is available, population or job data can be additionally refined using dasymetric mapping (Eicher and Brewer, 2001). In short, if an area's land-use is mostly "park", we assign fewer inhabitants than if the land-use is "residential".

POI data and public transport stations are mapped to cells by counting the number of points per type per cell. In addition points that do not lie within a cell but are within 100 meters of the planning region are handled as if they would lie within the nearest cell. This approach ensures that our model does not miss the influence of points which lie slightly outside of the planning area but are relevant since they attract bike-sharing trips.

4. Demand Modelling

Step two estimates a statistical demand model for the potential bicycle trip demand between cells defined in the previous step. The model is estimated for a city with an existing BSS since historical trip data is needed for the estimation. Similarly to mode-choice models (see e.g. Gunn (2001)), it is expected that the demand models are transferable to comparable cities, e.g. cities in the same geographical region. However, this transferability needs to be tested in future projects.
The main challenges for the model are the potentially large number of cells and the fact that demand between cells can only be observed indirectly via demands for trips between stations of existing BSS. An illustration (Figure 1) of the modelled demand $D_{\alpha\beta}$ between two cells α and β and the observed demand between two Citybike Wien stations D_{AB}.

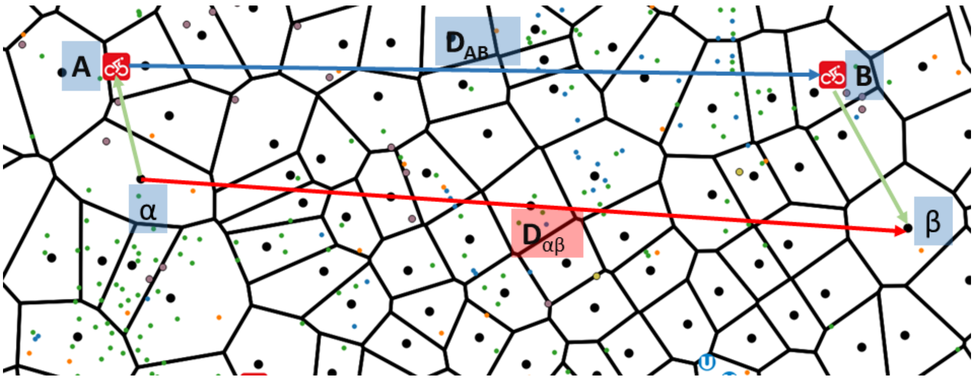

Figure 1: Observed demand and modelled demand for bikes between cells for the example of Vienna. The coloured dots show POIs in the different cells.

The demand between cells is modelled as an exponential of a weighted sum of the influencing factors $D_{\alpha\beta} = exp(\beta_{start}X_\alpha + \beta_{end}X_\beta + \beta_{trip}X_{trip_{\alpha\beta}})$ of the start and end stations and the trip between them (here X_* denote the data matrices and β_* the estimated parameter vectors. To estimate weights, the trips between stations A and B are modelled as weighted sum of all trips between the cells closest to stations, weighted by a function $f(d_{\alpha A}, d_{\beta B})$ of the walking distances $d_{\alpha A}$ and $d_{\beta B}$ to and from cells to BSS stations A and B, respectively, i.e. $D_{AB} = \sum_{\alpha \in C_A, \beta \in C_B} D_{\alpha\beta} f(d_{\alpha A}, d_{\beta B})$. The demand is estimated for four distinct time periods per workday: the morning and evening peaks, off-peak time in between, and nights, since theses show quite distinct usage patterns. Weekends were omitted for modelling simplicity and because most trips happen on workdays.

The variables applied in the demand model are the cell variables for population, jobs, POIs and public transport (see Table 1) in each start and end cell, distance and square distance, as well as altitude gained and lost on the cycle route between the cell. In addition, average rain and temperature were included as well as the time that there were no bikes at the starting or no return slots at the end BSS station closest to start and end cell respectively. The last variables were added to account for unobserved lost demand due to empty or full stations. The parameter values and signs are consistent with expectations. In the morning people leave from home or arrive at main train stations from their commute and cycle to work. In the afternoon shopping locations or underground stations attract more trips. The reason for the small attraction of main train stations may be that in Vienna these lie uphill and on the borders of the Citybike Wien system and shared bikes tend to be used more on downhill stretches. For a detailed account of the modelling and estimation approach as well as more detailed results, the reader is referred to Rudloff et al. (in preparation).

5. Calculation of Optimal Station Locations

Step three applies an optimization algorithm on the demand model created in the previous step. The algorithm selects a (near-)optimal set of cells representing the new stations to be built together with suitable station configurations. Furthermore, satisfiable demand as well as costs are obtained. In order to achieve the necessary scalability we propose an approach based on hierarchical clustering of the input data. The optimization problem is formulated as a mixed-integer linear program (MIP) which is then approximately solved following the multilevel-refinement paradigm. Experiments indicate that meaningful results to large-scale instances with up to 7,000 cells are obtained in reasonable calculation times. To solve our large scale problem, for our approach MIP models are suitably embedded in a metaheuristic framework to enable scalability to substantially large real-world instances.

The underlying optimization follows the optimization goal specified in Section 2.1 and searches for a subset of optimal station configurations such that the user demand can be fulfilled as much as possible and all budget constraints are respected. As real-world instances tend to be very large, the input of the algorithmic approach is given as hierarchical clustering which is defined as rooted tree where leaf nodes represent stations and all other nodes in between represent cluster nodes. In practice, a tree height of $log_2 n$ or $log_4 n$ is usually meaningful. Demands are aggregated such that all demands below a certain threshold are summed up and multiple demand arcs are combined into a single demand arc in the hierarchically clustered input data. This input structure is later on also exploited by the proposed algorithm.

By using $x_{s,k}$ as the binary decision variables whether configuration $k \in K_s$ can be built at station $s \in S$, we define the underlying optimization problem as follows.

$$\max \sum_{t \in T} D(\mathbf{x}, t) \tag{1}$$

$$\sum_{s \in S} \sum_{k \in K_s} \left(x_{s,k} \cdot \left(b^{\text{conc}}_{s,k} + b^{\text{opc}}_{s,k} \right) + b^{\text{reb}} \cdot Q_x(s) \right) \leq B^{\text{tot}}_{\text{max}} \tag{2}$$

$$\sum_{s \in S} \sum_{k \in K_s} b^{\text{conc}}_{s,k} \cdot x_{s,k} \leq B^{\text{conc}}_{\text{max}} \tag{3}$$

$$\sum_{k \in K_s} x_{s,k} = 1 \qquad s \in S \tag{4}$$

$$x_{s,k} \text{ binary} \qquad s \in S, k \in K_s \tag{5}$$

Every element of the solution vector $\mathbf{x} = \{x_1, \ldots, x_n\}$ denotes a particular station configuration, i.e., $x_s = x_{s,k} \mid x_{s,k} = 1$. Each configuration $k \in K_s, \forall s \in S$ consists of particular construction costs $b^{\text{conc}}_{s,k}$ and operation costs $b^{\text{opc}}_{s,k}$ as well as a specific number of slots. The function $D(\mathbf{x}, t)$ is used to compute the satisfiable demand for a given solution vector $\mathbf{x}$ and a particular time period $t \in T$. Moreover, the function $Q(x_s, s)$ estimates the number of bikes which need to be rebalanced for station $s \in S$. The objective function (1) maximizes the satisfiable demand. Inequalities (2) and (3) are used to restrict the costs of the planned system to a given total budget and construction budget respectively. Inequalities (4) is used such that for each prospective station exactly one configuration is chosen. All decision variables $x_{s,k}$ are binary (5).

This model defines the optimization problem. Functions $D(\mathbf{x}, t)$ and $Q_(\mathbf{x}, s)$ are modelled via linear programs (LP), for details we refer to Kloimüllner and Raidl (2017). By putting these elements together in a MIP model (the overall MIP model and the two LP models), the problem can be solved to optimality using a commercial MIP solver. However, this is only possible for smaller instances up to about 300 cells according to our experiments in Kloimüllner and Raidl (2017).

As real-world instances for large cities may consist of thousands of cells, we have to increase scalability, which we do by means of a multilevel-refinement approach. The idea is to reduce input size by iteratively merging cells following a hierarchical clustering until the instance reaches a size that can be solved reasonably well with the MIP model.

Based on a hierarchical clustering an algorithm based on the **multilevel refinement** methodology was chosen. This general approach was first proposed by Walshaw (2002, 2004). The original algorithm as proposed by Walshaw is to **coarsen** the problem until it is small enough so that it can be solved easily by a MIP formulation, like in our case, but one can also use a (meta)heuristic for the so called **initialization** procedure in which the initial solution is computed. Coarsening is done by aggregating the demand upwards the hierarchical clustering tree. For the station candidates it is possible to compute a minimum and maximum number of slots which can be built in a clustered cell on a higher level of the rooted tree. On higher levels of the clustering tree we simplify the problem such that we allow an arbitrary number of slots to be chosen for each station candidate instead of choosing an exact station configuration. On the one hand due to aggregation it is not possible to chose exact station configurations and on the other hand it is also not necessary. Only at the lowest level station configuration have to respected. The problem is iteratively **extended**, i.e., the MIP is solved with the relaxation to arbitrary slot numbers, and optionally also **refined**. The extension step is used to propagate the incumbent solution downwards the hierarchical clustering tree. The optional refinement step was not used/implemented in this work but could be implemented in future work to improve the incumbent solution after each extension step. Many possible refinement techniques are possible also including, e.g., some large neighbourhood search. When the problem is extended to the lowest level, the solution to the overall problem is retrieved, i.e., the assignment of station configurations to the particular stations and the total satisfiable demand as well as the according total and construction costs for building the system.

6. Visualization and Planning Front End

Step four is a web-based planning front end, which serves both as decision support system for detailed planning by the planner as well as visual debugging aid for development of the tool. It supports both with the following features, which correspond to the planning steps:

1. **Data collection and data preprocessing**: visualization of the planning area split into cells as well as visualization and inspection of all variables such as the population density per cell
2. **Demand model**: visualization of estimated incoming and outgoing demand for each cell
3. **Calculation of optimal station locations**: visualization of planned stations, the estimated demand in these stations, and the (largest) flows between them
4. **Visualization and interactive planning**: visualization of the coverage area and estimated demand of the already placed stations and interactive planning with free placement of stations, that allows planners to e.g. further adapt an optimization result

A set of **performance indicators** is calculated for the stations of the existing system (optionally), the result of the optimization, and for the results of interactive planning. The calculation is based on the outputs defined in Table 3. The selected indicators are the covered demand (of the total demand possible), costs for construction, maintenance, and redistribution logistics, system coverage area and population reached in this area, and station density. With these indicators different solutions can be easily compared to each other, which is especially helpful for interactive planning. An example of the cell visualization for steps two and three can be seen in Figure 2. An overview of the features supporting the planner in the selection of stations during interactive planning is given in Figure 3. A solution calculated with the optimization is given in Figure 4.

7. Evaluation

The implemented prototype of the proposed SALP-tool was evaluated to identify its strong points and weaknesses where future work should provide further improvements. For this purpose we conducted qualitative expert interviews with six experts from the fields of traffic planning and BSS. The interviews were conducted in form of an interactive demonstration of all planning steps in the SALP-tool from visualizing and explaining the relevance of different input data sets to showcasing the interactive planning mode (see Section 6).
From these interviews we identified three strong points the SALP-tool is offering:

1. **Tangible outputs relevant for political decision makers**: planning new BSS in cities is a process, where political decision makers are strongly involved. In this context it is important to present the potential impacts of a future BSS in a way, which supports politicians in their decisions and strategic considerations.

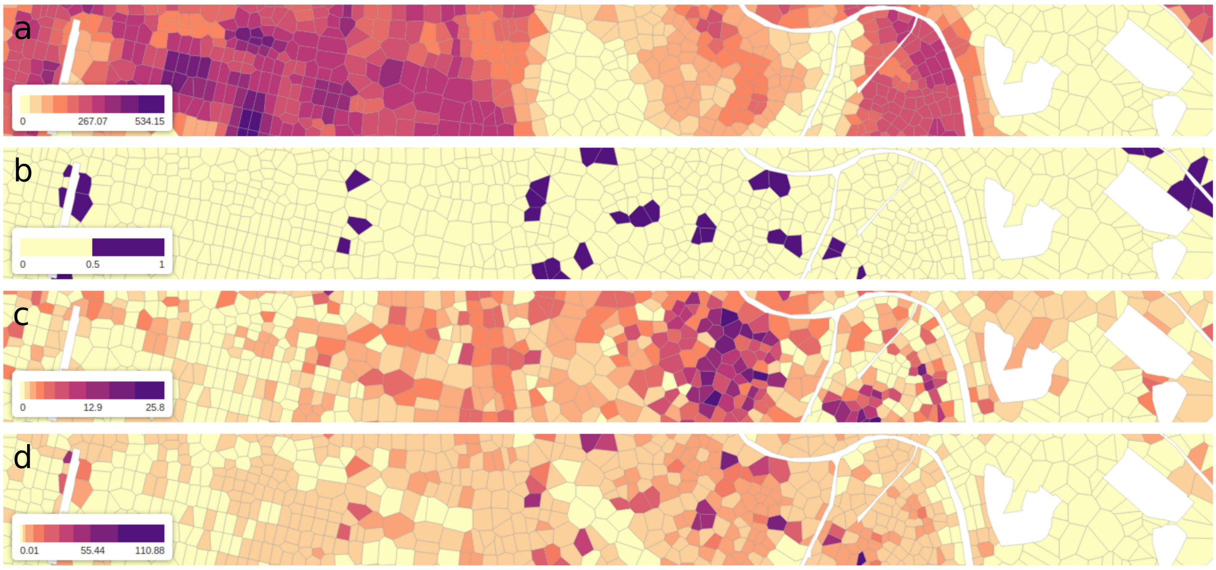

Figure 2: Visualization of different variables for cells: (a) population, (b) number of entrances to underground (and suburban train) stations, (c) number of POIs (all types combined), (d) demand of bikes going to the cell, i.e. demand of free slots to return a bike

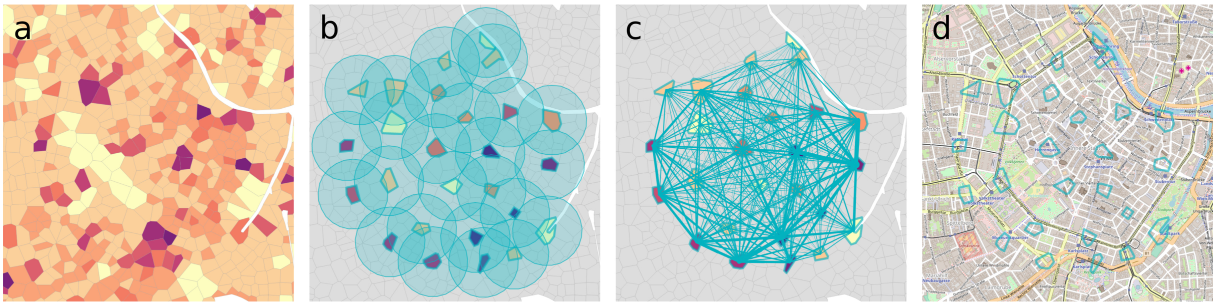

Figure 3: Exemplary planning process of a BSS: (a) the visualization, e.g. of the demand of bikes going to the cell for each cell helps the planner choose the most suitable station locations, (b) for each planned station a 300m radius helps seeing uncovered areas; the cell colour represents the predicted demand of bikes going to each planned station, (c) predicted trips between the stations are shown as lines where the thickness represents the relative amount of traffic, (d) the stations on a map (© OpenStreetMap and contributors). The planned system consists of 26 stations and its performance indicators are as follows: it covers an area of $4.3km^2$, has a station density of 6.0 stations per km^2, an average distance between stations of $299m$ and covers 5.8% of the daily demand (of the complete planning area).

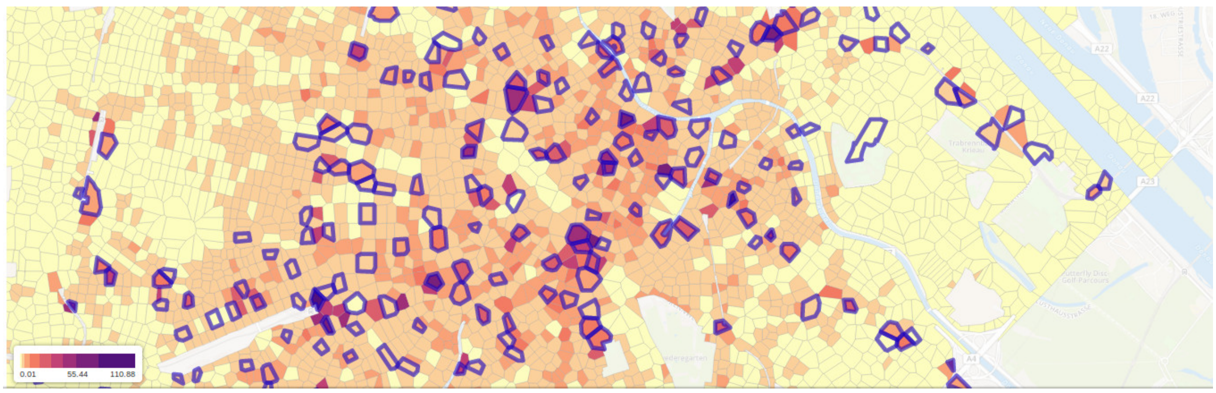

Figure 4: Exemplary solution obtained by the optimization algorithm consisting of 280 stations and the following performance indicators: it covers an area of $39.8km^2$, has a station density of 7.0 stations per km^2, and an average distance between stations of $234m$. The background map shows the demand going to the cell as in Figure 3a. The total available budget was set to 10 million Euros (construction: 6.5 million, operation: 3.5 million). The problem was coarsened until a maximum of 32 customer nodes has been reached. Calculation took 1.5 hours on an Intel Xeon E5540 2.53GHz Quad Core processor with 40GB of RAM. The algorithms have been written in C++ and have been compiled with gcc-6.2.0. For solving the LPs and MIPs Gurobi 7.0 was used (background map © Mapbox, OpenStreetMap and contributors).

For the interviewed experts the provision of outputs to support support political decision makers is one of the most useful applications of the planning front end.

2. **Decision support through advanced visualization**: the visualization of used input data and BSS demands is helpful for a better understanding and support of the planning process of new BSS as well as the extension of existing systems. The planning front end depicts outcomes of complex data models and algorithms behind them in a transparent way.
3. **Interactive planning**: the tool supports local planning experts in their planning process and for validating their own thoughts when planning new or expanding existing BSS.

With these interviews potential issues and extensions were identified as well:

- **More detailed input data**: if available the integration of more detailed input data could improve the quality of the demand model: e.g. passenger frequency at transit stops, hospitals as POIs, public services such as town halls as POIs.
- **Transparency**: the model should not behave as black box but it should be clear which parameters of the input data have which effect on the estimated demand.
- **Manual tweaking**: possibility for manual changes in the optimization process such as adapting the influence of certain types of POIs.
- **Temporal resolution**: using a higher temporal resolution for the demand model, e.g. hourly intervals, and also including data for weekends could further improve the model quality.
- **Transferability**: the transferability of the demand model to other cities must be proven.
- **Visualization**: a detailed visualization of the topography is important for interactive planning, even if the topography is already used by the optimization.
- **Cost-benefit analysis**: the assessment of costs and benefits is critical for political decision-making. An additional performance indicator to support this are costs per BSS trip, which are an indicator for the efficiency of the system.

8. Conclusions and Future Work

The developed SALP-tool can support BSS planners in their work by providing a clearly arranged graphical overview of a great variety of basic input data and by producing a first solution of a possible BSS station distribution throughout a given area. The input data are disaggregated into an optimized set of cells, that helps to narrow down the possible location for the BSS stations. The exact location of each station must always be manually determined in a next step considering the given local circumstances. The proposal for the distribution of the BSS stations is based on the optimization algorithm that aims for maximization of demand satisfaction given a defined budget. The result is presented in an interactive map giving the optimal location (cell) and size for each station. The first solution can subsequently be manually adjusted by moving, deleting or adding stations. Alternatively the BSS can be planned by hand from scratch using only the knowledge of the planner supported by the layers of the basic data. The SALP-tool is a powerful instrument in order to help BSS planners in optimizing the layout of a BSS and preparing maps and charts to support political decisions.

8.1. Future Work

Within this work it was not possible to quantitatively compare the effort for planning a BSS with and without the SALP-tool. It is presumed that after a simplification of the data entry into the system and a validation of the algorithms there will be a clear benefit of the SALP-tool in the planning process.

Due to the complexity of estimating **demand models** for unobserved demand between cells via the observed demand for rides between bike sharing stations, only a linear approach to model mean demand was feasible in this application. To improve the cell based demand models further the input data needs to be improved. As mentioned above, the number of passengers frequenting PT stops would be an important addition to the data. Other improvements could be a better differentiation to the POI data, e.g. splitting leisure POIs into more uniform subgroups like parks, bars etc. However theses improvements would make a proper variable selection process even more important to avoid overfitting of the models. Due to the high complexity of the estimation process of unobserved demand between cells at the current stage variable selection, even with relatively fast forward procedures would be too time consuming. Hence, faster estimation techniques need to be developed to get more reliable models for forecasting demand. It would then also be feasible to estimate separate models for weekends and more fine-grained day categories to further improve the model accuracy. Furthermore, the increasing availability of e-bike-sharing systems would be an interesting extension to the planning system since e-bikes allow for longer travel distances and reduce the impact of topography.

As noted by experts during the evaluation in Section 7 proving the transferability of the demand model is important future work. It is especially interesting which factors such as population size, transit system, or modal shares, influence the transferability in a negative way. Further research can contribute to a more accurate calibration of the model or may lead to a differentiation between certain types of cities that have to be determined.

Regarding the **calculation of optimized locations** there are still many open research questions and interesting research directions to explore. First, it would be nice to implement a refinement strategy into the multilevel refinement to improve the solution after each extension step. As this was only a first approach for solving the problem, it would be interesting to extend the approach to an iterative one. There are two options. There could be applied either an iterative multilevel refinement or a POPMUSIC, see Taillard and Voss (2002), based approach could also be used. It would also be interesting to compare both of them. Furthermore, it would also be interesting to implement a (meta)heuristic based extension procedure instead of using a MIP model and to compare solution quality as well time which was needed to extend the problem instance.

Acknowledgements

The authors gratefully acknowledge support from the Austrian Federal Ministry for Transport, Innovation and Technology (BMVIT) within the strategic program *Mobilität der Zukunft* under grant 849028 (PlanBiSS).

References

Chen, Q., Sun, T., 2015. A model for the layout of bike stations in public bike-sharing systems. J. Adv. Transport. 49 (8), 884–900.
DeMaio, P., 2009. Bike-sharing: History, impacts, models of provision, and future. Public Transportation 12 (4), 41–56.
Eicher, C. L., Brewer, C. A., 2001. Dasymetric mapping and areal interpolation: Implementation and evaluation. Cartography and Geographic Information Science 28 (2), 125–138.
Faghih-Imani, A., Eluru, N., Jun. 2016. Incorporating the impact of spatio-temporal interactions on bicycle sharing system demand: A case study of New York CitiBike system. Journal of Transport Geography 54, 218–227.
Fishman, E., Jan. 2016. Bikeshare: A Review of Recent Literature. Transport Reviews 36 (1), 92–113.
Frade, I., Ribeiro, A., 2015. Bike-sharing stations: A maximal covering location approach. Transport. Res. A-Pol. 82, 216–227.
García-Palomares, J. C., Gutiérrez, J., Latorre, M., 2012. Optimizing the location of stations in bike-sharing programs: A GIS approach. Applied Geography 35 (1-2), 235–246.
Gavalas, D., Konstantopoulos, C., Pantziou, G., 2016. Design & management of vehicle sharing systems: A survey of algorithmic approaches. In: Obaidat, M. S., Nicopolitidis, P. (Eds.), Smart Cities and Homes: Key Enabling Technologies. Elsevier Science, Ch. 13, pp. 261–289.
Graser, A., 2017. Tessellating urban space based on street intersections and barriers to movement. GI_Forum 2017 1, 114–125.
Gunn, H., 2001. Spatial and temporal transferability of relationships between travel demand, trip cost and travel time. Transportation Research Part E: Logistics and Transportation Review 37 (2), 163 – 189, advances in the Valuation of Travel Time Savings. URL http://www.sciencedirect.com/science/article/pii/S1366554500000235
Hu, S.-R., Liu, C.-T., 2014. An optimal location model for a bicycle sharing program with truck dispatching consideration. In: IEEE 17th International Conference on Intelligent Transportation Systems (ITSC). IEEE, pp. 1775–1780.
ITDP, 2013. The bike-share planning guide. Tech. rep., Institute for Transportation and Development Policy.
Kloimüllner, C., Raidl, G. R., 2017. Hierarchical clustering and multilevel refinement for the bike-sharing station planning problem. In: Conference Proceedings of Learning and Intelligent Optimization Conference (LION 11). LNCS. Springer.
Lin, J.-R., Yang, T.-H., 2011. Strategic design of public bicycle sharing systems with service level constraints. Transport. Res. E-Log. 47 (2), 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. Comput. Ind. Eng. 65 (1), 77–86.
Martinez, L. M., 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 Soc. Behav. Sci. 54, 513–524.
Noland, R. B., Smart, M. J., Guo, Z., Dec. 2016. Bikeshare trip generation in New York City. Transportation Research Part A: Policy and Practice 94, 164–181.
Pfoser, S., Pajones, M., 2016. Specification of input, output and setting options for a bike sharing planning system. International Journal of Transport Development and Integration 1 (1), 84–91.
Rudloff, C., Graser, A., Straub, M., in preparation. A cell based flow model for bike sharing as a basis for location planning 2430 (1).
Rudloff, C., Lackner, B., 2014. Modeling Demand for Bikesharing Systems. Transportation Research Record: Journal of the Transportation Research Board 2430 (1), 1–11.
Saharidis, G., Fragkogios, A., Zygouri, E., 2014. A multi-periodic optimization modeling approach for the establishment of a bike sharing network: A case study of the city of Athens. In: Proceedings of the International MultiConference of Engineers and Computer Scientists 2014 Vol II. No. 2210 in LNECS. Newswood Limited, pp. 1226–1231.
Schroeder, B., 2014. Bicycle Sharing 101: Getting the Wheels Turning. Moonshine Media.
Taillard, É. D., Voss, S., 2002. Popmusic—partial optimization metaheuristic under special intensification conditions. In: Essays and surveys in metaheuristics. Springer, pp. 613–629.
Walshaw, C., 2002. A multilevel approach to the travelling salesman problem. Oper. Res. 50 (5), 862–877.
Walshaw, C., 2004. Multilevel refinement for combinatorial optimisation problems. Ann. Oper. Res. 131 (1), 325–372.
Yang, T.-H., Lin, J.-R., Chang, Y.-C., 2010. Strategic design of public bicycle sharing systems incorporating with bicycle stocks considerations. In: 40th International Conference on Computers and Industrial Engineering (CIE). IEEE, pp. 1–6.