Generating Rental Data for Car Sharing Relocation Simulations on the Example of Station-Based One-Way Car Sharing

Alfred Benedikt Brendel
University of Goettingen
abrende1@uni-goettingen.de

Christian Rockenkamm
University of Goettingen
c.rockenkamm@stud.uni-goettingen.de

Lutz M. Kolbe
University of Goettingen
lkolbe@uni-goettingen.de

Abstract

Developing sophisticated car sharing simulations is a major task to improve car sharing as a sustainable means of transportation, because new algorithms for enhancing car sharing efficiency are formulated using them.

Simulations rely on input data, which is often gathered in car sharing systems or artificially generated. Real-world data is often incomplete and biased while artificial data is mostly generated based on initial assumptions. Therefore, developing new ways for generating testing data is an important task for future research.

In this paper, we propose a new approach for generating car sharing data for relocation simulations by utilizing machine learning. Based on real-world data, we could show that a combined methods approach consisting of a Gaussian Mixture Model and two classification trees can generate appropriate artificial testing data.

1. Introduction

The demand for new forms of mobility is increasing, driven by the trend of more people living in urban areas. The UN predicts that by 2030, around 60% of the total world population will reside in urban areas [1], suggesting that the demand for urban mobility will rise immensely, along with the growing need for additional roads and parking space [2]. Hence, the available space will decrease. Therefore, future (passenger) transportation systems have to develop alternatives to privately owned cars in form of flexible, dynamic and sustainable mobility services [3]. Under this conditions, understanding mobility as a service presents a welcome development [4]. Particularly young people begin to use a mixture of various mobility services, instead of a privately owned car [5]. In this context, the (sustainable) potential of new services like car sharing have been the object of interest and focus of past research [6].

Notably, car sharing has been reported as a flexible and sustainable mobility service. It provides several vantages for its users and the environment. Most notable is, that it is convenient and cost effective for users [7], [8], and additionally, by being a flexible (short-term) transportation service, it can complement classical means of transportation by intertwining individual mobility with existing public transportation options [9], [10]. Furthermore, earlier studies have exhibited that compared to trips with a privately owned car, car sharing has the capability to decrease the individual car-bound mobility by up to 30% [2], [11]. Beyond the reduced car-bound mobility and the derived lower emissions, using car sharing results in a reduced parking demand and less noise, which makes it an environmentally sustainable form of transportation, for urban areas [2], [11]. Therefore, car sharing can be viewed as a mean to increase sustainability and counter urban transportation problems, e.g., the lack of space and harmful emissions [11], [12]. Hence, it is favorable that car sharing continues to grow [13] to live up to its expectations.

However, to lower operating costs and to increase flexibility remains a key factor for the success of car sharing services [14]. One of the major cost factors is balancing vehicle supply and demand [14]. Since the available amount of vehicles can vary throughout the day, there may be a situation where the supplies cannot meet the demand in some operation areas [14]–[16]. Many existing car sharing provider practice operator-based relocation, which is often more costly than user-based relocation [17], [18]. Considering the financial cost difference, it would be more sustainable and cost-efficient if car sharing provider could substitute operator-based relocation with user-based relocation [18], [19].

HICSS

For car sharing are information systems (IS) a key factor [9], [20], especially to apply specific relocation methods (e.g. user-based relocation) [14], [21]. The development of new and efficient relocation algorithms is a major task of current research [9], [14], [22]. By using advanced simulations, relocation algorithms can be tested in an artificial car sharing environment and compared with other algorithms [1], [7], [14], [17], [23]–[25]. One major problem besides the development of improved algorithms remains: gathering enough representative rental data. Using real-world rental data [21], [26] or artificially generated data [24], [25] are often the only, suboptimal options. Using real-world data leads to biased results, caused by real-world data being biased. Brendel et al. [27] used the terms "incomplete data" and "relocation bias" to describe the problem of real-world data. Hence, we will elaborate on these terms in the following to get a better understanding of the problem. We argue that both issues are not independent and therefore will describe them interrelated in the following.

Incomplete data is caused by having no record of aborted booking processes and denied requests, caused by a lack of vehicle supply. This is partly due to the implemented relocation method incapability to supply vehicles for every rental request (called relocation bias). Furthermore, the relocation bias is inherent in real-world datasets by having the real-world relocation method as there optimal solution. Hence, each simulated relocation method can only be as good as the real-world approach by default.

Therefore, it is necessary to generate reliable and unbiased data [27] to compare relocation methods objectively via simulation. Artificially generated data used in current research is often unreliable, hypothetical and based on predefined rules, assumptions and/or demand estimations [24], [25], [28]. Hence, it is important to find a way to generate reliable data based on real-world data [27].

Accordingly, this paper aims to answer the following research question:

How can car sharing rental data for relocation simulations be generated?

To answer this question, we used a combination of two machine learning algorithms: Gaussian Mixture Model (GMM) and Classification Tree. Based on gathered real-world rental data, this combined methods approach can identify patterns and generate data correspondingly. Data generated this way is scalable and is not based on assumptions. This way, data can be generated, which is without relocation bias and can additionally be completed to mimic hidden demands.

2. Related Work

The following will describe the status-quo of car sharing research, especially regarding vehicle relocation and data for car sharing simulations.

2.1 Relocation Research

Relocation is the measure of re-distribution a vehicle with the intention to solve the issue of demand and supply imbalances within a car sharing system [19], [26], [29], [30].

Current research distinguishes car sharing systems into three types [14]: station-based two-way car sharing, station-based one-way car sharing, and free-floating car sharing. The conditions of station-based two-way car sharing let customers rent a vehicle from a station, but customer have to return the vehicle to the same station at the rentals end, thus preventing the need for vehicle relocation between stations [15], [19]. Station-based one-way car sharing enhances station-based two-way car sharing by giving its customers the advantage to return their vehicle to any available station [15], [16]. An akin form of car sharing is free-floating car sharing. This version has no static stations, and customers can (ideally) pick up and return rental vehicles wherever they want within the operation area of the car sharing provider [7], [15], [18], [21].

Station-based one-way and free-floating car sharing have a similar relocation problem. Free-floating car sharing can be understood as a station-based one-way car sharing system with an infinite number of stations. Even though their relocation methods cannot simply be applied to both of them, they can be compared [27], [21]. In the following sections, we will present an examination of the current relocation research regarding both car sharing versions.

In car sharing, the staff member of car sharing provider are commonly the ones who rearrange the vehicles by driving, towing or ride-sharing them to the desired station [30], [18]. This practice is termed "operator-based relocation".

Jorge et al. [31] were able to show, that operator-based relocations can lead to a more balanced station-based one-way car sharing system. Extending this concept, Kek et al. [30] used an optimization-trend simulation to substantially decrease the number of relocations, and therefore, reducing the total costs of operator-based relocation. Even though, the costs of operator-based relocation can still be considered a major issue. User-based relocation bases on the idea to motivate car sharing users to return their vehicle at stations in demand for vehicles, instead of stations

2

with a high supply of vehicles [21], [32]. Hence, Clemente et al. [17] have pointed out that operator-based relocation is more expensive and inadequate when related to user-based relocation. As an implication, user-based relocation should be favored.

Weikl and Bogenberger [18] researched operator-based and user-based relocation as an opportunity to solve the imbalance in free-floating car sharing. They concentrated on demand prediction methods and algorithms for computing the optimal vehicle distribution within the operation area of the car sharing provider. Nonetheless, they did not evaluate their concept empirically or in a case-study. Recently, Wagner et al. [21] developed a relocation framework for user-based relocation support systems in a free-floating car sharing context. Their goal was to minimize the idle time of the vehicles and in doing so they demonstrated that applying their approach of relocation would decrease the idle time and increase the rentals per car.

2.2 Test Data for Car Sharing Simulations

To evaluate the status quo of research regarding car sharing simulation and test data usage, we conducted a literature review. Expanding on the literature review of Jorge and Correia. [14], we identified 13 publications regarding car sharing relocation simulations from the past 15 years (see Table 1). We analyzed the used data (artificial data and real-world data), as well as if it addressed the problems regarding artificial or real-world data.

Research regarding the simulation of car sharing systems to develop and evaluate new relocation algorithms has gained more attention in the recent years, but is still an underresearched field.

Some publications used surveys regarding the mobility demand or historical data to generate artificial data (e.g.; [31]). The assumption is often, that car sharing can substitute the conducted demand completely.

Carlier et al [24] are using a self-developed data generator instead of real-world data, because they could not gather enough real-world data. The data generator generates data based on random demand variations over the course of a day.

Kek et al. [35] used real-world data to evaluate their relocation approach. They could replicate the performance observed in the original system, and also reduce the number of car parking lots and relocation employees needed.

Jorge et al. [31] used fragmented real-world data to accumulate patterns for their simulation. This way they could simulate their car sharing system using partly real world-data and partly artificial data. While doing so, they identified the problems of their approach and explain in which way it affects their results.

Wagner et al. [37] gathered rental data of 250,000 rentals to verify their user-based relocation approach. They addressed the problem of incomplete data briefly by stating that insufficient historical real-world car sharing data could lead to undetected hidden demands in some areas.

Brendel et al. [27] used real-world data from a different form of car sharing for their simulation. They altered the data to fit the context and pointed out the discovered problems regarding the usage of real-world data within the context of car sharing simulations. The two major problems are: incomplete data and relocation bias.

Despite the potential of machine learning, none of

Article	Using artificial data	Using real-world data	Addressing the problems of artificial data	Addressing the problems of real-world data
Alfian et al. [33]		x		
Barth et al. [34]	x	x		
Brendel et al. [27].	(x)	(x)		x
Carlier et al. [25]	x			(x)
Carlier, Aur´elien [24]	x			
Clemente et al. [17]	x			
Cucu et al. [23]	x	x		
Jorge et al. [31]	x	(x)	(x)	
Kek et al [35]		x		
Lopes et al. [28]	x			
Repoux et al. [26]		x		
Wagner et al. [21]		x		
Wang et al. [36]	x	x		(x)

Table 1: Literature overview - car sharing relocation simulations and test data

the gathered research publications used it to generate proper rental data. Instead, they use real-world data or data generators based on demand estimations or assumptions. Often the problems of using artificial data and/or real-world data were not addressed.

3. Data Set

A rental data record includes many variables, i.e. customer ID, start time, end time, origin station, destination station, etc. (e.g., [21], [27].). In the context of the relocation simulations, a rental r is defined as the following tuple:

$$r = (origin, destination, \tau_s, \tau_t)$$

Where $origin$ stands for the station the rental starts from; $destination$ is the station the rental ends at; τ_s start time describes the time-point the rental starts; and τ_t is the travel time of the rental.

We collected rental data from a station-based one-way car sharing system of a city in Germany (around 100,000 inhabitants) over the course of 107 days (November to March). In total we gathered 2062 rentals.

For data preparation we transformed τ_s into values ranging from 0 (representing 00:00:00 on Monday) up to 604799 (representing 23:59:59 on Sunday) for every second of the week.

4. Method

In the following, we will present the process for generating the data with the machine learning algorithms used and how we evaluate the generated data.

For the implementation of the described method we used scikit-learn and GridSearch to obtain the configurations of the used models [38].

4.1 Data Generation Method

To generate appropriate rental data, the initial input data has to be multiplied without duplicating it.

We used the following generative process for each week with N data points: (1) Choose N tuple (τ_s, τ_t). (2) For each of N tuple do. (2a) Choose an $origin$. (2b) Choose a $destination$.

This procedure is based on the following equation:

$$P(\tau_s \cap \tau_t \cap origin \cap destination) = $$
$$P(\tau_s \cap \tau_t) \cdot P(origin \mid \tau_s \cap \tau_t)$$
$$\cdot P(destination \mid \tau_s \cap \tau_t \cap origin)$$

In the first step we are using a trained GMM to choose the tuple (τ_s, τ_t). The GMM was used because it can predict values not inherent in the training dataset and also generates multiple different tuples, both are abilities needed in this case. We will describe the GMM briefly in the following.

A GMM is the weighted sum of M component densities given by [39], [40]:

$$p(x|\lambda_n) = \sum_{i=1}^{M} w_i g(x|\mu_i, \Sigma_i)$$

Where M is the number of mixtures of Gaussian components; x is the D-dimensional data vector; w_i, $i = 1, ..., M$ are the mixture weight; $w_i g(x|\mu_i, \Sigma_i)$, $i = 1, ..., M$ are the component Gaussian densities. The density of each component is a D-variate Gaussian function of the form:

$$g(x|\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2}(x - \mu_i)'(c_i^m)^{-1}(x - \mu_i)\right\}$$

μ_i is the mean vector and Σ_i the covariance matrix. The constraint $\sum_{i=1}^{M} w_i = 1$ is satisfied by the mixture weights. The following notation can collectively represent the parameters:

$$\lambda = \{w_i, \mu_i, \Sigma_i\}\, i = 1, ..., M$$

The goal for a GMM-based system is to train the parameters λ, so that the Gaussian mixture density matches the distribution of the data vectors.

After training the GMM, starting time and travel time tuples (τ_s, τ_t) can be generated. The generated tuples have to be checked for negative travel times or out of range starting times, since a GMM will also generate them. For each improper tuple a new one is generated until the desired amount of acceptable tuples are generated. Other filters may also be feasibly, like for to short rental durations (e.g. one minute to travel between two far apart stations).

In the second step, the $origin$ and the $destination$ are generated for each of the N tuples. In steps 2a and 2b the $origin$ and $destination$ is computed by using a classification tree for each of them. We used the classification tree because it delivers the possibility of generating multiple different classifications, and can therefore be used to generate data including values of low possibility. In the following we will briefly describe the classification tree.

Our implementation uses the optimized version of the CART (Classification and Regression Tree) classification tree [38].

Based on the given training vector $x_i \in R^n, i = 1, ..., I\ y \in R^i$ the decision tree recursively splits the space to group the samples with the same labels. The data node m is represented by Q. For each split $\theta = (j, t_m)$ consisting of a feature j and threshold t_m, the data is split into $Q_{left}(\theta)$ and $Q_{right}(\theta)$ subsets:

$$Q_{left}(\theta) = (x, y)|x_j <= t_m$$

4

$$Q_{right}(\theta) = Q \backslash Q_{left}(\theta)$$

The impurity of *m* is computed based on *H( )*:

$$H(X_m) \sum_k p_{mk}(1 - p_{mk})$$

Using *H( )* the following function has to be minimized:

$$G(Q,\theta) = \frac{n_{left}}{N_m} H\left(Q_{left}(\theta)\right)$$
$$+ \frac{n_{right}}{N_m} H\left(Q_{right}(\theta)\right)$$

The parameters are selected to minimize the impurity:

$$\theta^* = argmin_\theta G(Q,\theta).$$

This procedure is repeated until the maximum depth is reached, $N_m < min_{samples}$ or $N_m = 1$.

After training the first classification tree (2a), it was used to generate an *origin* for each given tuple (τ_s, τ_t). For this, the probabilities are cumulated and an origin is selected using a unified distribution from 0 to 1. The second classification tree (2b) is trained after 2a and the *destination* is generated using the same procedure for each tuple $(\tau_s, \tau_t, origin)$.

4.2 Evaluation

For evaluating the quality of the generated data quantitatively we used the BIC (Bayes Information Criterion) and the AIC (Akaike Information Criterion) [41].

The AIC is a measurement for the relative quality of a model for a given set of data and described as the following:

$$AIC = -2\log\left(\mathcal{L}(\hat{\theta})\right) - K$$

Where $\mathcal{L}$ is the maximized value of the likelihood function of the model; θ stands for the parameters of the model; K describes the number of free parameters to be estimated.

The BIC is used for selecting a model. Models with a lower BIC have to be preferred. Therefore, a model with a lower BIC has a better fit on the given data and predicts it more precisely. The BIC is defined as:

$$BIC = -2\ln(\mathcal{L}) + K\log(n)$$

n is the number of data points in the observed data set.

5. Results

We generated 7200 rentals data points, divided into 600 data points per week for 12 weeks. By this process we more than tripled the original data.

To evaluate the quality of the generated data we used two methods: Comparing the descriptive patterns and comparing the AIC and BIC.

By comparing the patterns of each variable from the real-world data set and the generated data set it is possible to see the relation between the two data sets. The patterns are depicted in the Figures 1 through 4. The aim was to generate data that is similar to the original data, but not just a duplication of it.

The start time distribution of both data sets (see Figure 1) can be described as similar, but not identical. The starting times in the original data are spiking around 12 o'clock (the marking for each weekday are at 12 o'clock) and nearly no rentals are starting in the nights between two weekdays. The same applies for the generated data.

By comparing the travel time probability (see Figure 2) patterns of the original data and the generated data, it is visible that both patterns spike around 20 minutes and are very similar with only some differences.

The distribution of the origin station (see Figure 3) has its highest value at station 8 and very low percentages at station 1 and 9, both in the original and in the generated data.

Furthermore, in the destination distribution (see Figure 4), the station distribution is also quite similar.

To compare the data, we trained a GMM on the real-world data set and computed the AIC and BIC. To be able to compare the data quantitatively, we used the trained GMM to compute the AIC and BIC for the generated data set. This way we could compare how similar the data sets are e.g., how well they can be predicted by the same model.

	AIC	BIC
Original	0.1035	113630
Generated	1.4236	436102

Table 3: Fit comparison of input data and generated data

According to the AIC and BIC, the GMM can predict the original data better than the generated data. It has to be noted, that the high values for both of the BICs are interfering with this conclusion. Meaning, the model is not able to predict both data sets sufficiently. Nonetheless, it is an indicator for the difference of the data sets. Hence, the generated data is not similar enough to the original data to have a similar predictability by the same model.

Combining the pattern analysis and the fit comparison, we can derivate the result that the generated data is similar to the original data regarding the start time, travel time, origin and destinations, but is not a simple duplication.
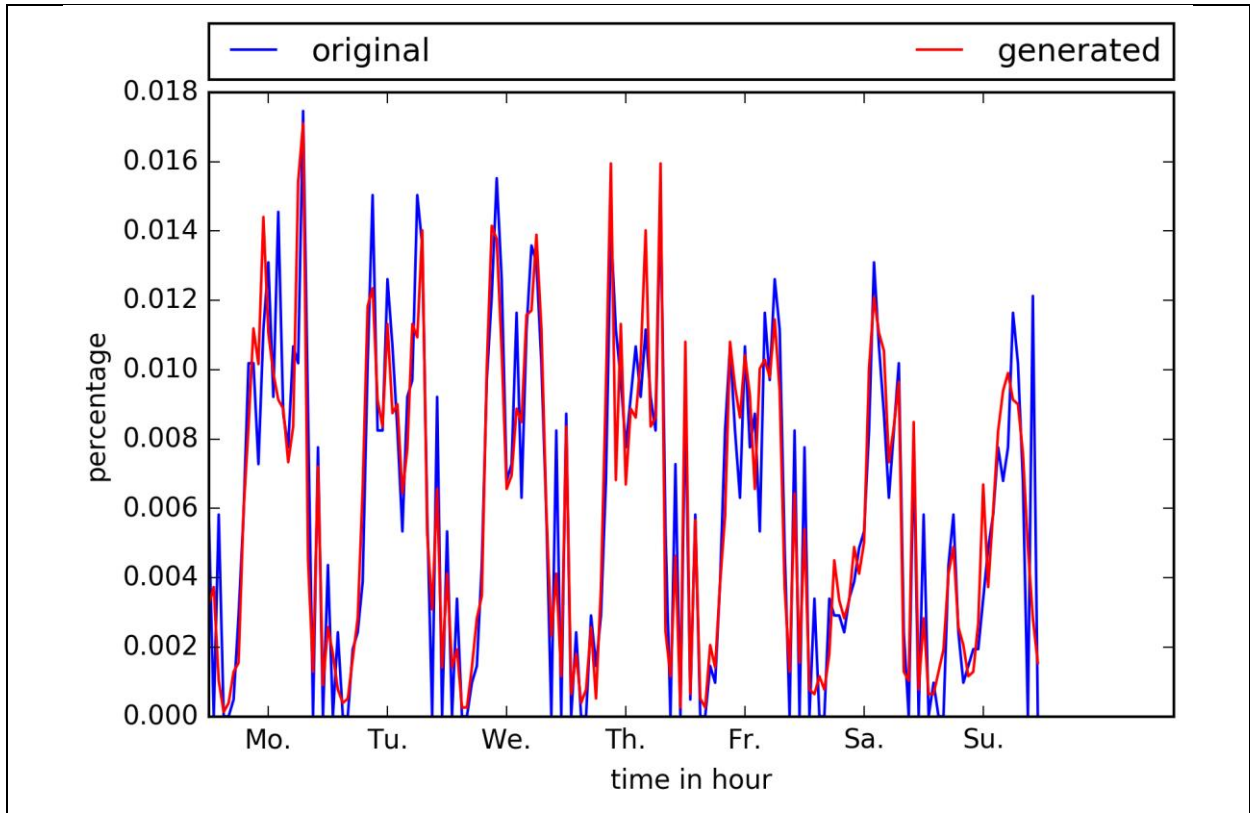
5

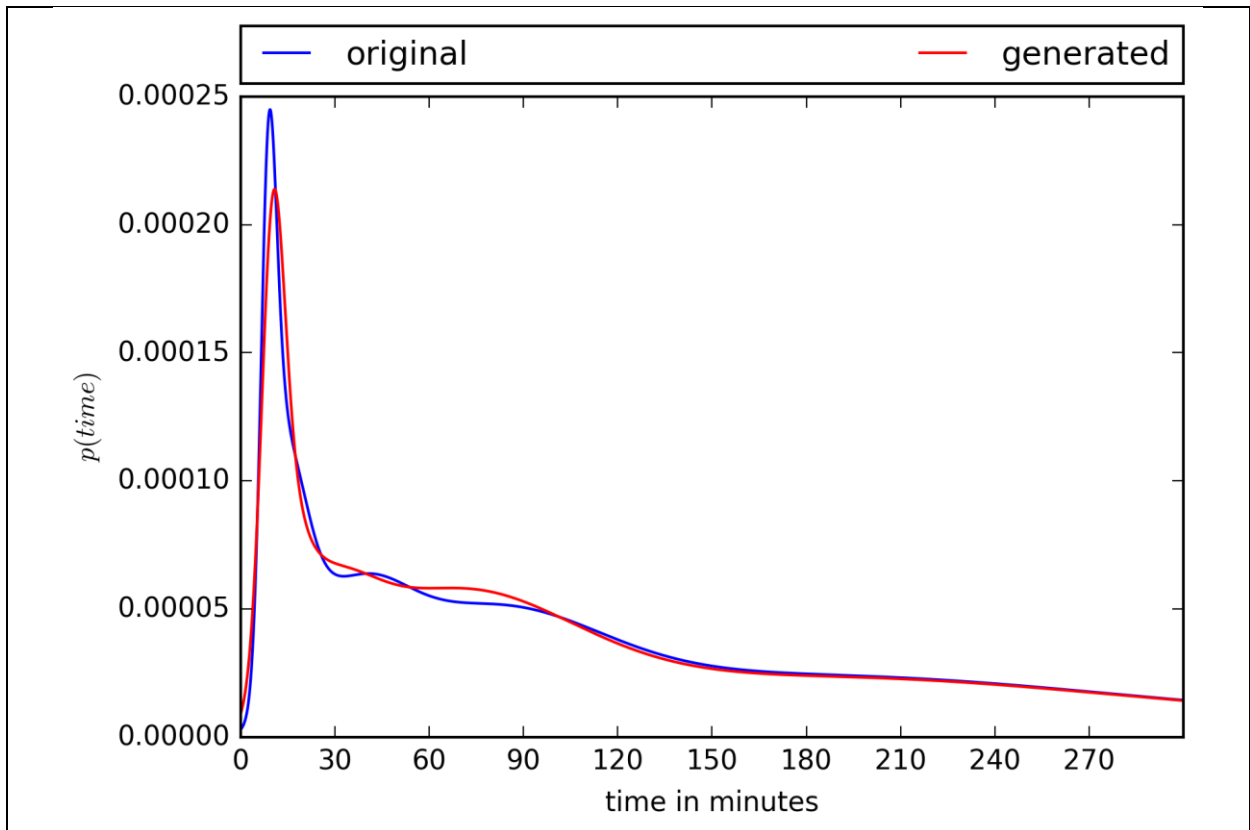**Figure 1: Start time distribution (7200 generated data points)**

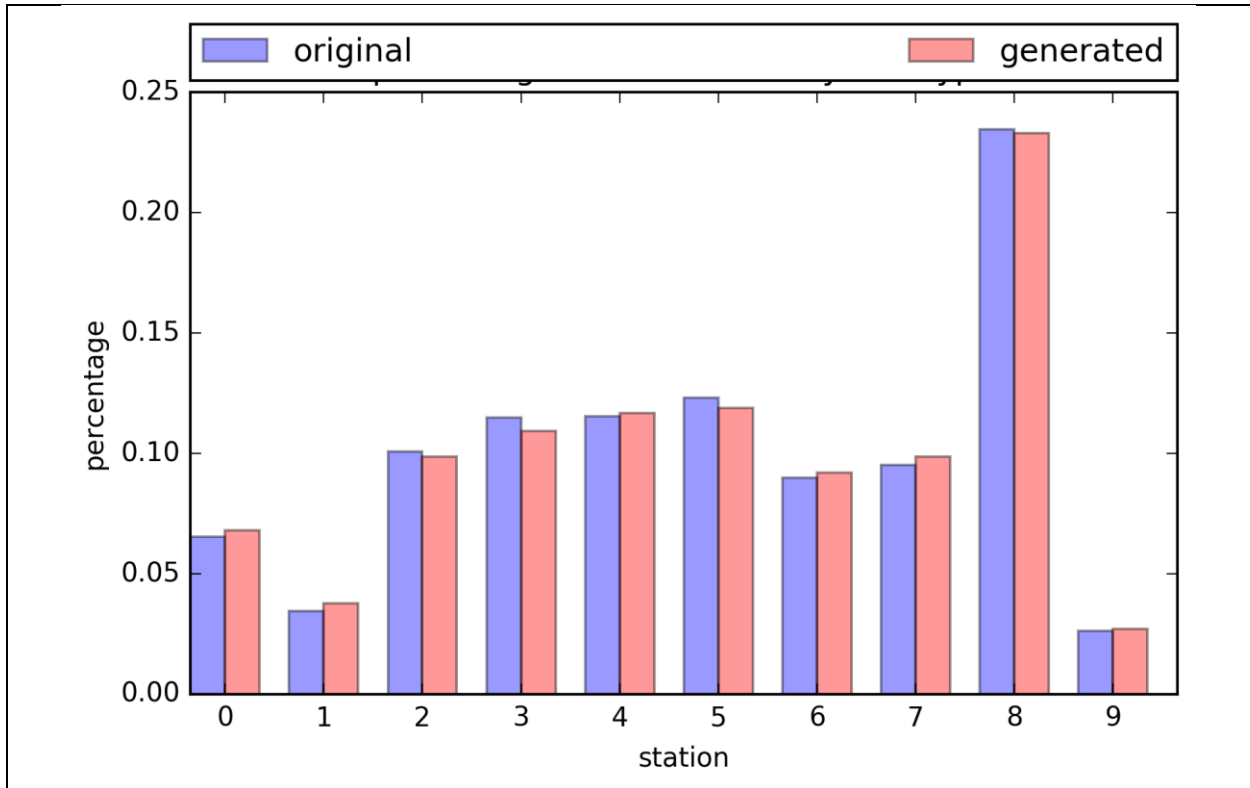**Figure 2: Travel time probabilities (7200 generated data points)**

6

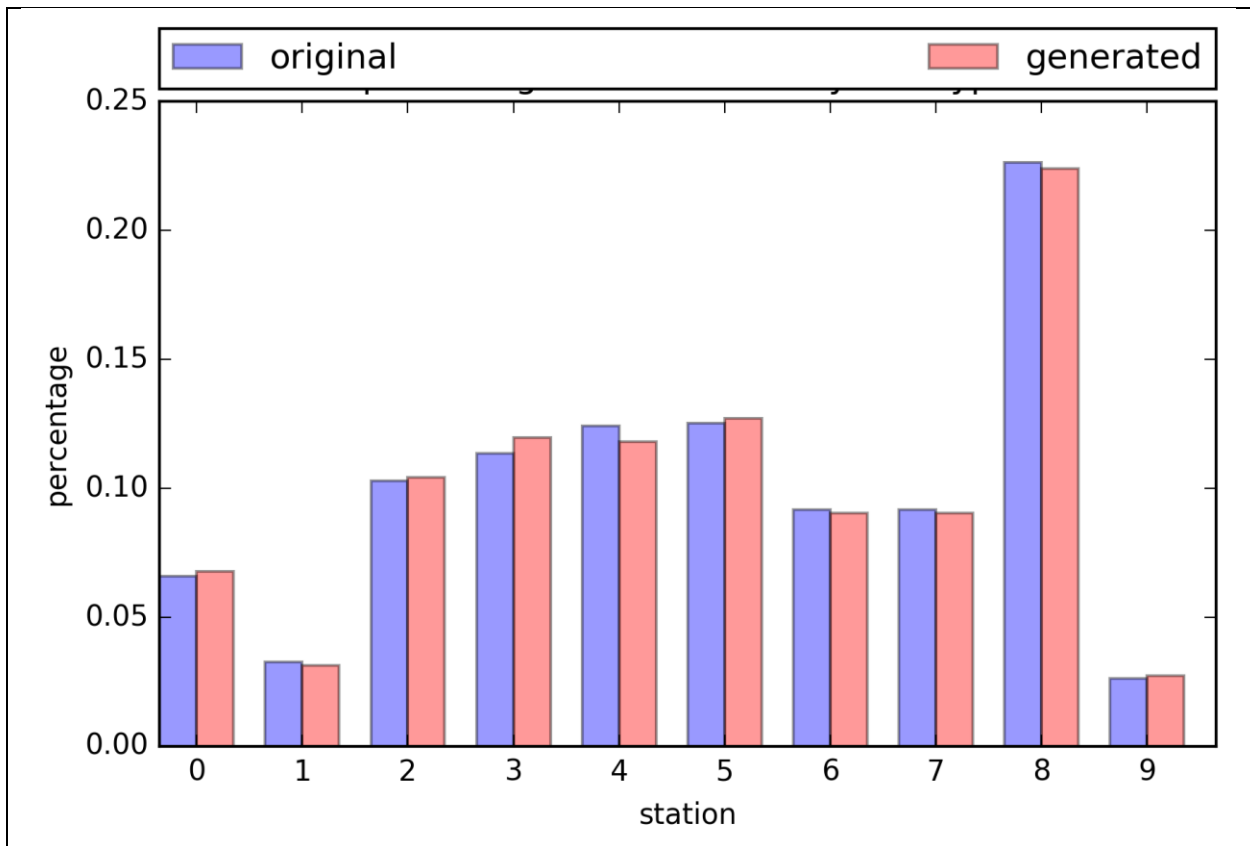**Figure 3: Origin distribution (7200 generated data points)**

Figure 4: Destination distribution (7200 generated data points)

7

6. Discussion

The results demonstrate that appropriate data can be generated with the presented approach. Models have to be trained for each car sharing system, but the configuration and the results presented in this paper show, that the approach is feasible.

In contrast to currently used methods for data generation (e.g., [24]), the proposed methods is only based on the original input data and is not involving initial assumptions on patterns or anything similar. This can be beneficial in particular for testing new relocation algorithms via simulations when more data is needed.

To simulate existing car sharing environments, generating more data can help to get an established system to its limits. For example, scenarios involving the same number of cars and stations, but many more rental requests, are possible to test. This can give indications for the need of more cars and/or stations in a system.

Like mentioned before, the relocation bias is inherent in the real-world dataset by having the real-world implemented relocation method as its optimal solution. Each simulated relocation method can only be as good as the real-world approach. By generating a data set with the presented method, there is no optimal solution which could limit the results of the relocation method. Relocations methods can be compared objectively via simulations with the same data set.

Furthermore, generating data for suspected hidden demands is possible by generating one data set and additionally generating a data set filtered for hidden demands (e.g. for one specific station as the origin). By combining both data sets a data set can be constructed, which includes additional data to mimic hidden demands. Thus, our approach counters incomplete data by offering the option to add filtered data to complete a data set.

The findings of this study have to be interpreted with caution due to the following limitations.

Firstly, the proposed method was only applied on one data set from a single car sharing system. Hence, the results have to be applied and verified for other systems and cases e.g., free-floating car sharing and e-car sharing. For free-floating car sharing often areas are used to describe and simplify the relocation problem [37], [42], thus this method can be applied by swapping stations for areas.

Secondly, the two used methods deliver satisfying results, but for other car sharing systems and other circumstances different methods could deliver better results. Therefore, future research should explore which machine learning algorithm should be used for which car sharing context.

Thirdly, other methods for comparing the original and the generated data set could conclude differently. Hence, more studies on how to compare original and generated car sharing rental data sets is needed.

Fourthly, the optimal configurations were obtained by using GridSearch [38]. It could be beneficial to set some parameters manually to reach other desired outcomes (e.g. even lower similarity to the original data).

7. Conclusion

In this paper we addressed the problem of generating rental data for car sharing simulations. We developed a data generation method to deal with the problems of incomplete data and relocation bias.

This study contributes to the field of IS research as follows. Following Watson et al.'s ([43]) call for more IS research on environmental sustainability, we proposed a data generation method for car sharing simulations that can help to improve station-based one-way car sharing. Furthermore, by processing real-world data, we were able to assess the capabilities of this method. In conclusion, this study is a step further to improve the sustainability of car sharing as an environmental friendly transportation service, thus contributing to sustainability within our society.

Moreover, our study provides a valuable method for practitioners. They can generate more data to "stress-test" their system via simulation. This could be helpful to prepare for events (e.g., trade fairs), where a higher vehicle demand is anticipated.

8. References

[1]   UN, "World Urbanization Prospects The 2007 Revision Highlights," *New York*, vol. ESA/P/WP/2, no. 4, p. 883, 2007.

[2]   H. Nijland, J. Van Meerkerk, and A. Hoen, "Impact of car sharing on mobility abd co2 emissions," *PBL Note - PBL Publ. number 1842*, no. July, 2015.

[3]   B. Nykvist and L. Whitmarsh, "A multi-level analysis of sustainable mobility transitions: Niche development in the UK and Sweden," *Technol. Forecast. Soc. Change*, vol. 75, no. 9, pp. 1373–1387, 2008.

[4]   KPMG, "Which companies will survive the digital revolution?," 2014. [Online]. Available: http://www.kpmg.com/BR/en/Estudos_Anali ses/artigosepublicacoes/Documents/Wich-

Companies-Will-Survive-The-Digital-Revolution.pdf.

[5]   A. Millard-Ball and L. Schipper, "Are we Reaching Peak Travel? Trend in Passenger Transport in eight Industrialized Countries," *Transp. Rev.*, vol. 31, no. 3, pp. 1–26, 2011.

[6]   A. B. Brendel and M. Mandrella, "Information Systems in the Context of Sustainable Mobility Services : A Literature Review and Directions for Future Research," in *AMCIS*, 2016, pp. 1–10.

[7]   S. Wagner, T. Brandt, M. Kleinknecht, and D. Neumann, "In Free-Float: How Decision Analytics Paves the Way for the Carsharing Revolution," *ICIS 2014 Proc.*, pp. 1–17, 2014.

[8]   C. Wagner and S. Shaheen, "Car Sharing and Mobility Management: Facing New Challenges with Technology and Innovative Business Planning," *J. World Transp. Policy Pract.*, vol. 4, no. 2, pp. 39–43, 1998.

[9]   B. Hildebrandt, A. Hanelt, T. Nierobisch, E. Piccinini, L. Kolbe, and T. Nierobisch, "The Value of IS in Business Model Innovation for Sustainable Mobility Services - The Case of Carsharing," in *WI 2015*, 2015, pp. 1008–1022.

[10]   C. Nawangpalupi and O. Demirbilek, "Investigation of the drivers and the barriers for travel behaviour changes and analysis of the impact: a case study of car sharing in Australia Catharina Nawangpalupi Oya Demirbilek," *Int. J. Environ. Cult. Econ. Soc. Sustain.*, vol. 4, no. 4, pp. 1–12, 2008.

[11]   E. Martin and S. Shaheen, "The Impact of Carsharing in Household Vehicle Ownership," *Access Mag.*, pp. 22–27, 2011.

[12]   J. Firnkorn and M. Müller, "What will be the environmental effects of new free-floating car-sharing systems? The case of car2go in Ulm," *Ecol. Econ.*, vol. 70, no. 8, pp. 1519–1528, 2011.

[13]   S. Shaheen and A. Cohen, "Innovative Mobility Carsharing Outlook - Carsharing Market Overview, Analysis, and Trends Summer 2015," *Transp. Sustain. Res. Cent.*, pp. 1–4, 2015.

[14]   D. Jorge and G. Correia, "Carsharing systems demand estimation and defined operations: A literature review," *Eur. J. Transp. Infrastruct. Res.*, vol. 13, no. 3, pp. 201–220, 2013.

[15]   M. Balac and F. Ciari, "Modeling station-based carsharing in Switzerland," in *14th Swiss Transport Research Conference*, 2014, no. May.

[16]   F. Ciari, C. Dobler, and K. W. Axhausen, "Modeling one-way shared vehicle systems: an agent-based approach," *13th Int. Conf. Travel Behav. Res.*, no. July, pp. 1–14, 2012.

[17]   M. Clemente, M. P. Fanti, A. M. Mangini, and W. Ukovich, "The vehicle relocation problem in car sharing systems: Modeling and simulation in a Petri net framework," in *Application and Theory of Petri Nets and Concurrency*, Berlin, Springer Berlin Heidelberg: D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, J.-M. Colom and J. Desel, 2013, pp. 250–269.

[18]   S. Weikl and K. Bogenberger, "Relocation strategies and algorithms for free-floating car sharing systems," *IEEE Intell. Transp. Syst. Mag.*, vol. 5, no. 4, pp. 100–111, 2013.

[19]   A. Di Febbraro, N. Sacco, and M. Saeednia, "One-Way Carsharing," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2319, pp. 113–120, 2012.

[20]   K. Degirmenci and H. Breitner, "Carsharing: A Literature Review and a Perspective for Information Systems Research," *Tagungsband MKWI 2014*, pp. 962–979, 2014.

[21]   S. Wagner, T. Brandt, and D. Neumann, "Data analytics in free-floating carsharing: Evidence from the city of Berlin," *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, vol. 2015-March, pp. 897–907, 2015.

[22]   A. Hanelt, I. Nastjuk, H. Krüp, M. Eisel, and C. Ebermann, "Disruption on the Way ? The Role of Mobile Applications for Electric Vehicle Diffusion," in *WI Proceedings*, 2015, pp. 1023–1037.

[23]   T. Cucu, L. Ion, and S. M. Boussier, "Management of Public Transportation Service: Car Sharing Service," *6th Int. Conf. theory Pract. Perform. Meas. Manag. Dunedin, New Zealand.*, pp. 1–15, 2009.

[24]   A. Carlier, A. Munier-Kordon, and W. Klaudel, "Optimization of a one-way carsharing system with relocation operations," in *Proceedings of 10th Internation Conference on Modeling, Optimization and Simulation*, 2014.

[25]   A. Carlier, A. Munier-Kordon, and W. Klaudel, "Mathematical Model for the Study of Relocation Strategies in One-way Carsharing Systems," *Transp. Res. Procedia*,

9

vol. 10, no. July, pp. 374–383, 2015.

[26] M. Repoux, B. Boyaci, and N. Geroliminis, "Simulation and optimization of one-way car-sharing systems with variant relocation policies," in *94th Annual Meeting of the Transportation Research Board*, 2014.

[27] A. B. Brendel, B. Brauer, and B. Hildebrandt, "Toward User - Based Relocation Information Systems in Station-Based One-Way Car Sharing," in *AMCIS*, 2016, pp. 1–10.

[28] M. M. Lopes, L. M. Martinez, and G. H. de A. Correia, "Simulating Carsharing Operations through Agent-based Modelling: An Application to the City of Lisbon, Portugal," *Transp. Res. Procedia*, vol. 3, no. July, pp. 828–837, 2014.

[29] B. Boyaci, K. G. Zografos, and N. Geroliminis, "An optimization framework for the development of efficient one-way car-sharing systems," *Eur. J. Oper. Res.*, vol. 240, pp. 718–733, 2014.

[30] A. G. H. Kek, R. L. Cheu, Q. Meng, and C. H. Fung, "A decision support system for vehicle relocation operations in carsharing systems," *Transp. Res. Part E Logist. Transp. Rev.*, vol. 45, no. 1, pp. 149–158, 2009.

[31] D. Jorge, G. H. A. Correia, and C. Barnhart, "Comparing optimal relocation operations with simulated relocation policies in one-way carsharing systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 4, pp. 1667–1675, 2014.

[32] F. Schulte and S. Voß, "Decision Support for Environmental-friendly Vehicle Relocations in Free- Floating Car Sharing Systems: The Case of Car2go," *Procedia CIRP*, vol. 30, pp. 275–280, 2015.

[33] G. Alfian, J. Rhee, Y. S. Kang, and B. Yoon, "Performance comparison of reservation based and instant access one-way car sharing service through discrete event simulation," *Sustain.*, vol. 7, no. 9, pp. 12465–12489, 2015.

[34] M. Barth, M. Todd, and L. Xue, "User-Based Vehicle Relocation Techniques for Multiple-Station Shared-Use Vehicle Systems," *Transp. Res. Board 83th Annu. Meet.*, no. 04, 2004.

[35] A. Kek, R. Cheu, and M. Chor, "Relocation Simulation Model for Multiple-Station Shared-Use Vehicle Systems," *Transp. Res. Rec.*, vol. 1986, no. 1, pp. 81–88, 2006.

[36] H. Wang, R. Cheu, and D. H. Lee, "Dynamic relocating vehicle resources using a microscopic traffic simulation model for carsharing services," in *3rd International Joint Conference on Computational Sciences and Optimization, CSO 2010: Theoretical Development and Engineering Practice*, 2010, vol. 1, pp. 108–111.

[37] S. Wagner, C. Willing, T. Brandt, and D. Neumann, "Data Analytics for Location-Based Services: Enabling User-Based Relocation of Carsharing Vehicles," *ICIS*, vol. 3, no. 5, pp. 279–287, 2015.

[38] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

[39] D. Reynolds, "Gaussian mixture models," *Encycl. Biometrics*, no. 2, pp. 659–663, 2009.

[40] Y. Huang, K. B. Englehart, B. Hudgins, and A. D. C. Chan, "A Gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 11, pp. 1801–1811, 2005.

[41] K. P. Burnham and R. P. Anderson, "Multimodel Inference: Understanding AIC and BIC in Model Selection," *Sociol. Methods Res.*, vol. 33, no. 2, pp. 261–304, 2004.

[42] S. Weikl and K. Bogenberger, "A practice-ready relocation model for free-floating carsharing systems with electric vehicles – Mesoscopic approach and field trial results," *Transp. Res. Part C Emerg. Technol.*, vol. 57, 2015.

[43] R. T. Watson, M.-C. Boudreau, and A. J. Chen, "Information systems and environmentally sustainable development: energy informatics and new directions for the is community," *MIS Q.*, vol. 34, no. 1, pp. 23–38, 2010.

10