Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

REBALANCING AND FLEET SIZING OF MOBILITY-ON-DEMAND NETWORKS WITH COMBINED SIMULATION, OPTIMIZATION AND QUEUEING NETWORK ANALYSIS

Peter Bazan
Anatoli Djanatliev
Marco Pruckner
Reinhard German

Computer Science 7
Friedrich-Alexander University Erlangen-Nürnberg
Martensstr. 3
Erlangen, 91058, GERMANY

Christoph Lauer

Development Autonomous Driving
AUDI AG
I/EF-12
Ingolstadt, 85045, GERMANY

ABSTRACT

Car rental companies have to operate the fleet of vehicles considering a cost effective fleet sizing and rebalancing of the vehicles. Additionally, they have to ensure a good vehicle availability at the rental stations. We present a simulation model combined with optimization and queueing network analysis for mobility-on-demand networks with one-way car sharing. The system is modeled as a closed queueing network. On the one hand, this allows an estimation of the vehicle availability; on the other hand, constraints for the optimization model can be derived from the mathematical equations of the queuing network model. To optimize the total revenue, the optimization model considers price incentives for customer trips and cost of empty trips resulting in a mixed-integer linear programming problem. The simulation uses the optimization component for an optimized operation of the mobility-on-demand networks and computes the optimal fleet size, the cost of empty runs, and the vehicle availability.

1 INTRODUCTION

All over the world big cities are faced with increasing problems in the transport sector. Car drivers are increasingly stuck in traffic and they are confronted with driving restrictions because of a high level of fine particulate matter pollution. In order to take vehicles off the road, alternative transport concepts such as car sharing or ride sharing become more and more attractive. According to (Shaheen et al. 2014) car sharing can help to take 9 to 13 vehicles per shared vehicle off the road and to cut the number of vehicle miles traveled by 27 %. Car sharing means that third-party organizations maintain, manage, and insure a vehicle fleet which can be shared among a group of members. Only recently, the German original equipment manufacturer BMW and Daimler agreed to combine their car sharing services DriveNow and Car2Go (Daimler AG 2018). Both services operate in 31 major international cities a total of over 20,000 vehicles. Due to the combination of both services, a better vehicle availability can be ensured.

In the future, car sharing becomes even more interesting when it will be combined with autonomous driving (Shaheen et al. 2014). According to (Lang et al. 2016), so-called robo-taxi fleets could enable cities to meet improving traffic efficiency, reducing congestion, and lowering air pollution caused by auto engine emissions. Moreover, they can also provide elderly and children with better access to transportation (Lang et al. 2016). Nevertheless, alternative transport concepts, such as robo-taxis, can only prevail if they are economical feasible for the providers and cost-attractive as well as easily accessible for the customers. For an easy access to transportation, a good vehicle availability is inevitable.

In this paper we present a simulation model combined with optimization and queuing network analysis for mobility-on-demand networks with one-way car sharing. The formal model of the mobility-on-demand network is first mapped to a closed queueing network model. Then we consider price incentives for rebalancing the system by the help of customers to reduce the number of empty trips. For the optimization, a mixed integer linear problem is constructed. The queueing model and the optimization model is then combined with the simulation for improving the fleet size.

The remainder of the paper is organized as follows. In Section 2 we discuss related work. The considered mobility-on-demand network is presented in Section 3. A description of the associated closed queuing network model is given in Section 4. Section 5 presents the mathematical formulation and solution of the optimization model. A general description of our simulation model is given in Section 6. Section 7 shows simulation results and a comparison with results provided by the optimization model. Finally, Section 8 concludes the paper and gives an outlook on future work.

2 RELATED WORK

The analysis of mobility-on-demand networks with queueing network approaches is commonly used for the rebalancing and optimization of empty trips. The work published by (Briest and Raupach 2011) models the network as a closed queueing network with single-server queueing systems. The cost for rebalancing the system is proportional to the distance a vehicle has to travel without a customer. The model does not consider price incentives for customer trips and therefore the optimization problem can be formulated as a linear program.

The queueing network approach published by (Zhang and Pavone 2015) uses single-server nodes at points where the vehicles are waiting for customers and infinite-server nodes for the connection of these points. The mobility-on-demand network relies on a team of drivers to rebalance the vehicles and is modeled as two coupled qeueing networks, which can be solved approximately by two decoupled linear programs, whereas an exact solution is based on a nonlinear optimization technique. This approach is extended to self-driving vehicles and road networks with restricted capacities in (Iglesias et al. 2018). Both approaches don't use incentives for the rebalancing of the vehicles in the network.

Queueing network approaches are also used for the fleet-sizing of car sharing systems. An iterative solution and an easier to solve approximate solution is presented in (George and Xia 2011). The system is modeled as closed queueing network with single-server and infinite-server queueing systems. The fleet sizing for an electric car sharing system is published by (Fanti et al. 2014). Because of the charging of vehicles, the system is now modeled as a closed queueing network with additional multiple-server queueing systems.

Fleet sizing and rebalancing is investigated by (Köchel et al. 2003). At first they model the system with a queueing network. Then they combine a simulation with an heuristic optimizer based on genetic algorithms. With the queueing network and a steady-state analysis, the outcome of the combined simulation and optimization can be checked for bottlenecks.

3 MOBILITY-ON-DEMAND NETWORK

The considered mobility-on-demand network with one-way car sharing connects a number of cities with each other as shown in Figure 1. It provides a number of vehicles in each city that customers can use to drive to another city. For the formal description of the system (Figure 2), the following values are required in addition to the number of vehicles. The customers rent a vehicle with the mobility rates $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ at each city. The probability of driving from city i to city j is given by $\mathbf{P} = [p_{i,j}]$ and the corresponding travel time is notated as $\mathbf{T} = [t_{i,j}]$. For the example with 3 cities, the formal description in matrix notation is therefore given by (1), (2), and (3).

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3) \tag{1}$$

$$\mathbf{P} = \begin{bmatrix} 0 & p_{12} & p_{13} \\ p_{21} & 0 & p_{23} \\ p_{31} & p_{32} & 0 \end{bmatrix} \tag{2}$$

$$\mathbf{T} = \begin{bmatrix} 0 & t_{12} & t_{13} \\ t_{21} & 0 & t_{23} \\ t_{31} & t_{32} & 0 \end{bmatrix} \tag{3}$$

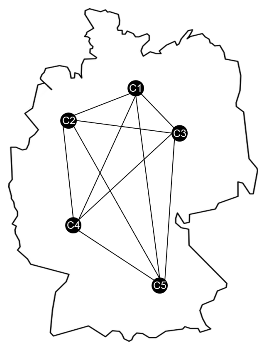

Figure 1: A mobility-on-demand network with five cities and ten connections.

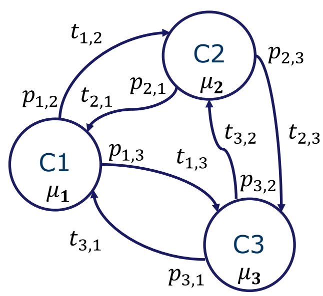

Figure 2: Formal model of a mobility-on-demand network with three cities.

Since the trips from city to city are one-way trips, empty trips must be carried out so that sufficient vehicles are available in every city. Of course, the number of these trips should be reduced as much as possible. Furthermore, the fleet size should be low due to the cost of acquisition and maintenance, but at the same time not too many customers should be rejected because no vehicles are currently available. To achieve these goals, techniques from queueing network analysis, optimization, and simulation are used in the following.

4 QUEUEING NETWORK MODEL

The described mobility-on-demand network can be modeled as a closed queueing network (Bolch et al. 2006) where the number of jobs in the network is fixed. The vehicles correspond to the jobs in the queueing network. Vehicles in a city wait in the corresponding queue of the queueing network for arriving customers. It is assumed that the inter-arrival times between the arrivals of customers are exponentially distributed with rates $\boldsymbol{\mu}$ and that they correspond to the service-times of a queueing system. Therefore, a city with cars and customer is modeled as M/M/1-FCFS-queue according to Kendall's notation (Bolch et al. 2006). The first letter M denotes the exponentially distributed inter-arrival times of the cars at a city, whereas the second letter M indicates the exponentially distributed service-times of the cars. The cars are served in the order in which they arrive (First Come First Served (FCFS)) and only one car is served at a time by the arrival of a customer. The number of cars at this queueing system is not limited.

The connections between the cities are modeled as M/G/∞ (Infinite Server) queueing systems. The second letter G indicates that every vehicle on the connection can be served simultaneously with general distributed service-times. This corresponds to the arrival of a car at the destination city after the respective travel time **T**. The destination city is selected according to the branching probability **P**.

The steady-state performance measures of the queueing network model can now be computed with the Mean Value Analysis (MVA) or any other appropriate algorithm (Bolch et al. 2006). The MVA of the queueing network model calculates for a given number of cars (jobs) the mean number $\overline{\mathbf{k}}$ of vehicles and the mean availability $\overline{\boldsymbol{\rho}}$ of vehicles per city for $t \to \infty$. With the mean availability $\overline{\boldsymbol{\rho}}$ the requested mean percentage of rejected customers per city and over all cities can be determined as:

$$\overline{\boldsymbol{c}}_{rej} = (\mathbf{1} - \overline{\boldsymbol{\rho}})100\% \qquad \overline{c}_{rej} = \boldsymbol{\mu}^\top \overline{\boldsymbol{c}}_{rej} / \boldsymbol{\mu}^\top \mathbf{e}\,.$$

Figure 3 shows an example of a mobility-on-demand network with three cities modeled as a closed queueing network. The nodes *Node01*, *Node02*, and *Node03* model the cars in each city waiting for customers. The other six nodes (e.g., node *Node12*) represent the cars traveling to their destinations.

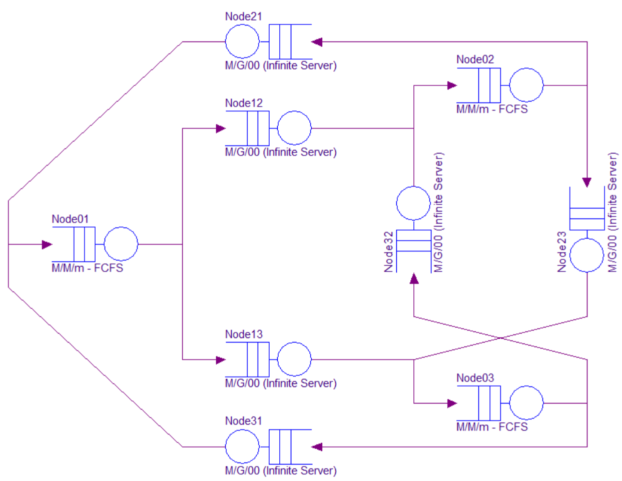

Figure 3: Queueing network model of the mobility-on-demand network with three cities (*Node01*, *Node02*, and *Node03*).

With 100 vehicles in the system and the service rates $\boldsymbol{\mu}$, the branching probabilities **P**, and the service times **T** as follows

$$\boldsymbol{\mu} = (10, 10, 10) \qquad \mathbf{P} = \begin{bmatrix} 0 & 0.1 & 0.9 \\ 0.8 & 0 & 0.2 \\ 0.7 & 0.3 & 0 \end{bmatrix} \qquad \mathbf{T} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \tag{4}$$

the MVA computes the mean distribution of vehicles per city as $\overline{\mathbf{k}} = (48.0, 0.6, 27.7)$ and the mean availability of vehicles as $\overline{\boldsymbol{\rho}} = (0.99, 0.39, 0.97)$. The overall number of vehicles minus the sum of the vehicles in the cities is the mean number of vehicles traveling from city to city. It can be observed that the number of rejected customers $\overline{\boldsymbol{c}}_{rej}$ in *Node02* is particularly high at 61 %.

In the following, instead of using service rates and branching probabilities, we use rates for the formulation of the model. For each connection the rate is calculated as $r_{ij} = \mu_i \cdot p_{ij}$. This leads to the rate matrix $\mathbf{R} = \text{diag}(\boldsymbol{\mu}) \cdot \mathbf{P}$, where $\text{diag}(\boldsymbol{\mu})$ is a matrix which diagonal elements contain the vector $\boldsymbol{\mu}$; all other elements of the matrix are zero (every ith row of $\mathbf{P}$ is multiplied with μ_i). The rate matrix $\mathbf{R}$ contains the same information as $(\boldsymbol{\mu}, \mathbf{P})$, as can be seen in the following example.

$$\boldsymbol{\mu} = (10,10,10) \quad \mathbf{P} = \begin{bmatrix} 0 & 0.1 & 0.9 \\ 0.8 & 0 & 0.2 \\ 0.7 & 0.3 & 0 \end{bmatrix} \quad \mathbf{R} = \begin{bmatrix} 0 & 1 & 9 \\ 8 & 0 & 2 \\ 7 & 3 & 0 \end{bmatrix}$$

After defining the vector $\mathbf{e} = (1,\ldots,1)$ and the matrix $\text{diag}^{-1}(\boldsymbol{\mu})$ containing the reciprocal values of $\boldsymbol{\mu}$ as diagonal entries, the reconversion of $\mathbf{R}$ to $(\boldsymbol{\mu}, \mathbf{P})$ can be computed. For the calculation of $\boldsymbol{\mu}$ the multiplication of $\mathbf{R}$ with $\mathbf{e}$ gives the row sums of $\mathbf{R}$. For the calculation of $\mathbf{P}$ every ith row of $\mathbf{R}$ is divided by μ_i.

$$\boldsymbol{\mu} = \mathbf{R} \cdot \mathbf{e} \quad \mathbf{P} = \text{diag}^{-1}(\boldsymbol{\mu}) \cdot \mathbf{R} \tag{5}$$

With the new formulation of the model, empty trips $\mathbf{R}^+$ from selected sources to targets can be added to achieve a more balanced mean availability of vehicles per city. This results in a rate matrix $\mathbf{R}^*$ whose new service rates $\boldsymbol{\mu}^*$ and branching probabilities $\mathbf{P}^*$ can be calculated from (5).

$$\mathbf{R}^* = \mathbf{R} + \mathbf{R}^+ \quad \boldsymbol{\mu}^* = \mathbf{R}^* \cdot \mathbf{e} \quad \mathbf{P}^* = \text{diag}^{-1}(\boldsymbol{\mu}^*) \cdot \mathbf{R}^*$$

The idea is now to select the empty trips such that a rate equilibrium in each city is established. This means that for each city the incoming rate equals the outgoing rate. Under the assumption that the queues of a city are never empty (i.e., the probability for this event is negligibly low), we get the rate equilibrium for city i as:

$$\sum_{j=1,\ldots,n; j \neq i} r^*_{ji} = \sum_{j=1,\ldots,n; j \neq i} r^*_{ij}$$

or in matrix notation with the meaning that the column sums are equal to the row sums:

$$\mathbf{e}^\top \cdot \mathbf{R}^* = \mathbf{R}^* \cdot \mathbf{e} \,. \tag{6}$$

5 OPTIMIZATION MODEL

Finding the empty trips with minimum costs under the condition of the rate balance (6) can be solved with linear programming. For the cost calculation, the distances $\mathbf{D} = [d_{ij}]$ between source city i and target city j are needed in addition to the previously described formal model resulting in the new formal model $(\boldsymbol{\mu}, \mathbf{P}, \mathbf{T}, \mathbf{D})$.

The optimization has to find an optimal rate matrix with empty trips $\mathbf{R}^+$. The new rate matrix of the model is then given by $\mathbf{R}^* = \mathbf{R} + \mathbf{R}^+$. The linear program has to consider the constraints $\mathbf{R}^+ \geq \mathbf{0}$ and $\mathbf{e}^\top \cdot \mathbf{R}^* = \mathbf{R}^* \cdot \mathbf{e}$. The target function for cost minimization of empty trips is

$$\min \Big[\sum_{i=1}^{n} \sum_{j=1}^{n} r^+_{ij} \cdot d_{ij} \Big] \,.$$

The solution of this linear optimization problem (called Optimization I) provides a new mobility grid $(\boldsymbol{\mu}^*, \mathbf{P}^*, \mathbf{T}, \mathbf{D})$ with rate equilibrium through minimum empty trips.

For the example with three cities from (4) and $\mathbf{T} = \mathbf{D}$ the optimization yields the new rate matrix

$$\mathbf{R}^* = \begin{bmatrix} 0 & \mathbf{6} & 9 \\ 8 & 0 & 2 \\ 7 & \mathbf{4} & 0 \end{bmatrix}$$

and with (5) the new mobility grid $(\boldsymbol{\mu}^*, \mathbf{P}^*, \mathbf{T}, \mathbf{D})$ including empty trips.

$$\boldsymbol{\mu}^* = (\mathbf{15}, 10, \mathbf{11}) \qquad \mathbf{P}^* = \begin{bmatrix} 0 & \mathbf{0.4} & \mathbf{0.6} \\ 0.8 & 0 & 0.2 \\ \mathbf{0.6364} & \mathbf{0.3636} & 0 \end{bmatrix}$$

With the optimized values for empty trips, the MVA computes a significant improvement regarding the balance of the mean distribution $\overline{\mathbf{k}} = (22, 22, 22)$ and the mean availability $\overline{\boldsymbol{\rho}} = (0.97, 0.97, 0.97)$ of vehicles.

Instead of relying only on empty trips, it can also be attempted to use the customer behavior for balancing the network by changing prices of trips on certain routes. A more expensive route causes a decrease in customers, while a cheaper route increases the demand compared to the same route at the normal price.

To describe the system, some additional constants are introduced. The yield per km for a normal trip is y, for a discounted trip y_{red}, and for a more expensive trip y_{inc}. A discounted trip increases the customer arrival rate by the factor s_{inc} and a more expensive trip decreases the customer arrival rate by the factor s_{red}. The cost per km empty trip is denoted as c.

As before, the rate of an empty trip between i and j is r^+_{ij} and r^*_{ij} is the total rate between i and j. Furthermore, some additional indicator variables which can only take the value 0 or the value 1 are needed for the formulation of the optimization problem. These indicator variables are $\gamma_{ij} \in \{0, 1\}$ for trips with ordinary price, $\delta_{ij} \in \{0, 1\}$ for trips with increased price, and $\eta_{ij} \in \{0, 1\}$ for trips with decreased price. The matrix notation of the constants and variables are given in Table 1.

Table 1: Matrix notation of the prices, rates, and indicators .

	Price	Rate	Indicator
Ordinary trip	$y\mathbf{D}$	$\mathbf{R}$	$\boldsymbol{\Gamma}$
Discounted trip	$y_{red}\mathbf{D}$	$s_{inc}\mathbf{R}$	$\boldsymbol{\Delta}$
More expensive trip	$y_{inc}\mathbf{D}$	$s_{dec}\mathbf{R}$	$\mathbf{H}$
Empty trip	$c\mathbf{D}$	$\mathbf{R}^+$	

Because of the indicator variables, the optimization model (called Optimization II) is now a Mixed Integer Linear Problem (MILP). A total rate is the sum of the rate for an empty trip and a normal, a discounted, or an expensive trip. The sum of the indicator variables for a route is equal to 1 and assures that a trip is only of the type indicated. All variables are non-negative and the rate equilibrium (6) must be maintained. The additional constraints are

$$r^*_{ij} = r_{ij}\left[\gamma_{ij} + s_{inc}\delta_{ij} + s_{red}\eta_{ij}\right] + r^+_{ij}$$

$$\gamma_{ij} + \delta_{ij} + \eta_{ij} = 1 \text{ for } i \neq j \text{ and } \gamma_{ii} + \delta_{ii} + \eta_{ii} = 0 \; .$$

The target function for the formal mobility network model $(\boldsymbol{\mu}, \mathbf{P}, \mathbf{T}, \mathbf{D})$ and the derived rate matrix $\mathbf{R}$ is given as

$$\max \left[\sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij} d_{ij} \left[y\gamma_{ij} + y_{red} s_{inc} \delta_{ij} + y_{inc} s_{red} \eta_{ij} \right] - \sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij}^{+} d_{ij} c \right] .$$

The solution of this MILP problem yields a new mobility grid $(\boldsymbol{\mu}^*, \mathbf{P}^*, \mathbf{T}, \mathbf{D})$ with rate equilibrium based on incentives and empty trips, and with maximum profit. For the example of a mobility-on-demand network with seven cities and the parameters of Tables 2 to 5, the optimized values for the rates (normal, discounted, expensive, and empty trips) are given in Tables 6 and 7. The analysis is done with a total number of 7916 vehicles and the price incentives

$$y = 1.0, \quad y_{red} = 0.8, \quad s_{inc} = 1.1, \quad y_{inc} = 1.2, \quad s_{red} = 0.9, \quad c = 2.0 .$$

Table 2: Customer arrival rates $\boldsymbol{\mu}$ for a mobility grid with seven cities in customers/min.

	City1	City2	City3	City4	City5	City6	City7
$\boldsymbol{\mu}$	245	40	48	121	72	97	42

Table 3: Transition probabilities **P** for a mobility grid with seven cities.

	City1	City2	City3	City4	City5	City6	City7
City1	0	0.12	0.22	0.12	0.16	0.17	0.22
City2	0.12	0	0.20	0.11	0.17	0.24	0.17
City3	0.17	0.17	0	0.15	0.26	0.10	0.15
City4	0.13	0.12	0.22	0	0.18	0.16	0.18
City5	0.10	0.13	0.28	0.12	0	0.20	0.17
City6	0.16	0.25	0.12	0.14	0.22	0	0.11
City7	0.21	0.15	0.17	0.14	0.20	0.12	0

Table 4: Travel times **T** for a mobility grid with seven cities in min.

	City1	City2	City3	City4	City5	City6	City7
City1	0	339	375	182	349	338	373
City2	339	0	179	237	84	378	284
City3	375	179	0	306	131	257	159
City4	182	237	306	0	265	452	399
City5	349	84	131	265	0	348	242
City6	338	378	257	452	348	0	155
City7	373	284	159	399	242	155	0

The verification of the results with the MVA shows that on the average approximately 100 % of the requests in all cities can be served (Table 8). Table 9 shows a comparison of the profit per month between Optimization I (LP) and Optimization II (MILP) of the mobility grid with seven cities. The costs for the empty trips could be reduced significantly, with at the same time increased profit.

Table 5: Distances between cities **D** for a mobility grid with seven cities in km.

	City1	City2	City3	City4	City5	City6	City7
City1	0	532	551	288	572	585	632
City2	532	0	251	362	69	637	426
City3	551	251	0	497	192	393	206
City4	288	362	497	0	424	775	657
City5	572	69	192	424	0	575	369
City6	585	637	393	775	575	0	233
City7	632	426	206	655	369	233	0

Table 6: Total service rates $\mathbf{R}^*$.

	City1	City2	City3	City4	City5	City6	City7	R^*e
City1	0	27	47	26	35	37	48	**221**
City2	8	0	7	49	6	8	6	**85**
City3	77	7	0	8	11	4	6	**114**
City4	16	14	24	0	20	18	20	**111**
City5	38	9	18	10	0	13	11	**98**
City6	17	22	11	12	19	0	10	**91**
City7	65	6	7	7	7	10	0	**101**
$e^\top R^*$	**221**	**85**	**114**	**111**	**98**	**91**	**101**	**820**

Table 7: Rates $\mathbf{R}^+$ of empty trips.

	City1	City2	City3	City4	City5	City6	City7
City1	0	0	0	0	0	0	0
City2	3	0	0	44	0	0	0
City3	68	0	0	0	0	0	0
City4	0	0	0	0	0	0	0
City5	30	0	0	0	0	0	0
City6	0	0	0	0	0	0	0
City7	55	0	0	0	0	5	0

6 SIMULATION MODEL

The presented queueing network model and the optimization model can be combined with a discrete-event simulation (DES) model for a further improvement of the mobility-on-demand network's operation. Instead of using the MVA for the steady-state analysis, the DES can take the dynamic behavior of the queueing network model of Section 4 into account.

At first, the simulation model uses the Optimization II of Section 5 to determine the optimal rates for empty, normal, discounted, and expensive trips. With these rates inserted into the queueing network model and repeated application of the MVA for a different number of vehicles in the mobility-on-demand network, the simulation determines a range of initial number of vehicle values around a predetermined acceptable value of rejected customers.

With the rates for empty trips $\mathbf{R}^+$ from the optimization and the formula presented below, the simulation is able to flexibly respond to the number of available cars in a city and adjusts the empty trips. The aim is to reduce the number of rejected customers with the smallest possible fleet size, since this can reduce the costs of acquisition and maintenance.

Table 8: Mean value analysis with price incentives.

	City1	City2	City3	City4	City5	City6	City7
$\bar{\mathbf{k}}$	524	524	524	524	524	524	524
$\bar{\boldsymbol{\rho}}$	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table 9: Profit for seven cities in million €/month.

	Optimization I	Optimization II
Ordinary trips	234	4
Discounted trips	-	25
More expensive trips	-	217
Empty trips	-184	-155
In total	**50**	**91**

To achieve this goal, the simulation calculates adjusted rates $\widehat{r}_{ij}^{+}$ according to (7). The average number of waiting vehicles per city $\bar{k}_i$ is computed from the queueing network model using the MVA and the adjustment is based on the deviation of the actual number of waiting vehicles in a city k_i from the average number of waiting vehicles per city $\bar{k}_i$. The rates are adjusted at every time-step of the simulation.

$$\widehat{r}_{ij}^{+} = r_{ij}^{+}\left(1 - \frac{\bar{k}_i - k_i}{2\bar{k}_i}(100\bar{\rho}_i - 99) + \frac{\bar{k}_j - k_j}{2\bar{k}_j}(100\bar{\rho}_j - 99)\right) \tag{7}$$

Figure 4 shows the improvement by approximately 3800 % of the simulation with continuously adjusted rates compared to the Optimization II regarding the mean percentage of lost customers $\bar{c}_{rej}$ for the example with seven cities and a fleet size of 7700 vehicles. This improvement, calculated as $(\bar{c}_{rej}^{opt}/\bar{c}_{rej}^{sim} - 1)100\%$, is achieved because the optimization for the steady-state model has no knowledge about the time dependent behavior and therefore can not adjust the rates. The calculated overall profit of the simulation compared to the optimization is only slightly reduced by 0.37 %.

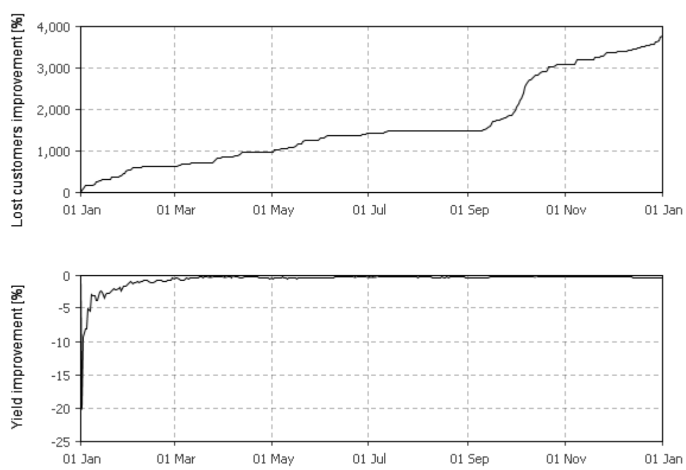

Figure 4: Lost customers improvement and yield improvement of the simulation compared to the Optimization II for the example with seven cities and a fleet size of 7700 vehicles.

7 RESULTS

The DES model of Section 6 is now used to improve the results of the mobility-on-demand network example with seven cities of Section 5. At first, the simulation calculates the mean number of vehicles in a city and the mean availability of vehicles in a city for a range of total vehicles in the network. Because it uses the optimized queueing network, the values are independent from the considered city. If the acceptable level of rejected customers $\overline{c}_{rej}$ is less than 0.2 %, then at least about 7700 cars are needed (Table 10).

Table 10: Mean value analysis of the optimized queueing network with different number of vehicles.

vehicles	4900	5600	6300	7000	7700
$\overline{k}_i$	98	195	294	394	493
$\overline{\rho}_i$	0.9912	0.9956	0.9971	0.9978	0.9983
$\overline{c}_{rej}$	0.88	0.44	0.29	0.22	0.17

With the mean number of vehicles in a city $\overline{k}_i$ from the MVA, the DES model can adjust the rates for empty trips of the optimized queueing network according to (7). Instead of a fleet size of 7700 vehicles, the rate adaptive simulation model can achieve a mean level of rejected customers $\overline{c}_{rej}$ of less than 0.2 % with a fleet size of only 5600 vehicles (Table 11).

Table 11: DES of the optimized queueing network with different number of vehicles and continuously adjusted rates.

vehicles	4900	5600	6300	7000	7700
$\overline{c}_{rej}$	0.47	0.06	0.01	0.01	0.00

Figure 5 shows the improvement by approximately 800 % of the simulation with continuously adjusted rates compared to the Optimization II regarding the mean number of lost customers $\overline{c}_{rej}$ for the scenario with a fleet size of only 5600 vehicles. This improvement is not achieved at the expense of profit, which increases by a factor of 0.42 %.

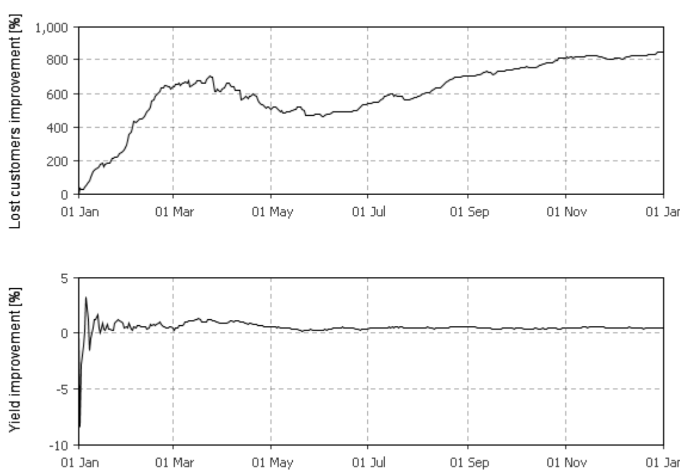

Figure 5: Lost customers improvement and yield improvement of the simulation compared to the Optimization II for the scenario with a fleet size of 5600 vehicles.

8 CONCLUSION

In this paper, we have combined a queueing network model, a MILP model, and a simulation model to address the optimal rebalancing problem and fleet-sizing problem of mobility-on-demand networks. In order to optimize the number of empty trips, price incentives for customers were also taken into account in order to minimize the number of such costly trips.

A closed queueing network with single-server nodes and infinite-server nodes is used as model for the mobility-on-demand network. This model allows the computation of the vehicle availability used by the simulation model for a dynamic adjustment of the rates of empty trips. This dynamic adjustment reduces the number of rejected customers and therefore, the fleet-size for operating the mobility-on-demand networks can be reduced.

To optimize the total revenue, the optimization model considers price incentives for customer trips and cost of empty trips resulting in a mixed-integer linear programming problem. The simulation uses the optimization component for an optimized operation of the mobility-on-demand networks and computes the optimal fleet size, the cost of empty runs, and the vehicle availability. The effectiveness of this combined approach is then shown by means of examples.

For the adjustment of the empty trips, the presented approach uses a formula that depends on the actual number of vehicles available in a city. Instead, more sophisticated methods could be applied, such as neural networks. With such a network, the correction factor could adapt more flexibly to the respective situation.

REFERENCES

Bolch, G., S. Greiner, H. De Meer, and K. S. Trivedi. 2006. *Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications*. 2nd ed. Hoboken, New Jersey: John Wiley & Sons.

Briest, P., and C. Raupach. 2011. "The Car Sharing Problem". In *Proceedings of the 23rd Annual ACM Symposium on Parallelism in Algorithms and Architectures*, 167–176. New York: ACM.

Daimler AG 2018. "BMW Group and Daimler AG Agree to Combine Mobility Services". http://media.daimler.com/marsMediaSite/ko/en/34636751. Accessed July 25nd, 2018.

Fanti, M. P., A. M. Mangini, G. Pedroncelli, and W. Ukovich. 2014. "Fleet Sizing for Electric Car Sharing System via Closed Queueing Networks". In *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1324–1329. Piscataway, New Jersey: IEEE.

George, D. K., and C. H. Xia. 2011. "Fleet-Sizing and Service Availability for a Vehicle Rental System via Closed Queueing Networks". *European Journal of Operational Research* 211(1):198 – 207.

Iglesias, R., F. Rossi, R. Zhang, and M. Pavone. 2018. "A BCMP Network Approach to Modeling and Controlling Autonomous Mobility-on-Demand Systems". *The International Journal of Robotics Research*.

Köchel, P., S. Kunze, and U. Nieländer. 2003. "Optimal Control of a Distributed Service System with Moving Resources: Application to the Fleet Sizing and Allocation Problem". *International Journal of Production Economics* 81-82:443 – 459.

Lang, N., M. Rüßmann, A. Mei-Pochtler, T. Dauner, S. Komiya, X. Mosquet, and X. Doubara. 2016. "Self-Driving Vehicles, Robo-Taxis, and the Urban Mobility Revolution". Technical report, The Boston Consulting Group, Boston, Massachusetts.

Shaheen, S., M. Galczynski, and R. M. Juster. 2014. "Autonomous Carsharing/Taxi Pathways". In *Envisioning Automated Vehicles within the Built Environment - Ancillary Workshop to the TRB Automated Vehicle Symposium 2014*. San Francisco, California.

Zhang, R., and M. Pavone. 2015. "A Queueing-Network Approach to the Analysis and Control of Mobility-On-Demand Systems". In *Proceedings of the 2015 American Control Conference (ACC)*, 4702–4709. Piscataway, New Jersey: IEEE.

AUTHOR BIOGRAPHIES

PETER BAZAN is a Research Assistant at the Chair for Computer Networks and Communication Systems at the Friedrich-Alexander University Erlangen-Nürnberg (FAU), Germany. He studied Computer Science

at FAU and obtained his master's degree (Dipl.-Inf. Univ.) with a thesis on the area of performance modeling. He worked a couple of years as a consultant for industry projects before he went into business for himself. In addition to his business he worked for the Performance Modelling and Process Control research group at FAU. Since 2001 he worked in various positions at the Department of Computer Science, where he also received his Ph.D. degree (Dr.-Ing.). He is mainly interested in the modeling and analysis of complex systems. This includes the performance evaluation of computer networks, communication and production systems, as well as the hybrid simulation of renewable energy systems and smart grids. His e-mail address is peter.bazan@fau.de.

ANATOLI DJANATLIEV is an Assistant Professor (AkadR) at the Chair for Computer Networks and Communication Systems at the Friedrich-Alexander University Erlangen-Nürnberg (FAU). He received his PhD degree (Dr.-Ing.) and his diploma in computer science (Dipl.-Inf. Univ.) from FAU. His research interests include various topics on simulation and modeling using system dynamics, discrete-event simulation, agent-based simulation, hybrid simulation, and methods for healthcare decision-support. More recently, his research activities focus on simulation of vehicular networks and other topics from the area of connected mobility and smart cities. His e-mail address is anatoli.djanatliev@fau.de.

MARCO PRUCKNER is an Assistant Professor of Energy Informatics at the Department of Computer Science at the Friedrich-Alexander University Erlangen-Nürnberg (FAU). He received his Ph.D. degree in engineering (Dr.-Ing.) and his M.Sc. degree in mathematics (Dipl. Math. Univ.) from FAU in 2015 and 2011, respectively. He is especially interested in the analysis of future electrical power systems with a high share of renewable energy resources based on hybrid simulation techniques. His further research interests include modeling and simulation of electro mobility issues and future mobility solutions within smart cities. His e-mail address is marco.pruckner@fau.de

REINHARD GERMAN received his diploma in computer science in 1991, the PhD degree in 1994, and the habilitation degree in 2000 from the Computer Science Department, Technical University of Berlin. Thereafter, he joined the Department of Computer Science at the Friedrich-Alexander University Erlangen-Nürnberg (FAU). First, he was an associate professor (system simulation), then he became a full professor in 2004 (computer networks and communication systems), and served as head of the department and as dean of the faculty of engineering. He is also contact professor for INI.FAU, a research cooperation with the AUDI AG and he also conducts simulation projects at the Energie Campus Nürnberg (EnCN). His research interests include performance and dependability analysis of networked systems based on numerical analysis, network calculus, discrete-event simulation, measurements, and testing. Vehicular communications, smart energy systems and healthcare constitute major application domains. His e-mail address is reinhard.german@fau.de.

CHRISTOPH LAUER studied Computer Sciences at the Friedrich-Alexander University Erlangen-Nürnberg (FAU) and Computer and Information Sciences at the Auckland University of Technology. He graduated 2008 and 2007 respectively and received his PhD from FAU in 2011 for his work on model-driven development of automotive safety systems. Since 2011 he is working for AUDI AG in the field of safety and self-driving car systems. His e-mail address is christoph.lauer@audi.de.