

Optimal Inventory Management of a Bike-Sharing Station

Tal Raviv and Ofer Kolka

September 2011

Bike-sharing systems allow people to rent a bicycle at one of many automatic rental stations scattered around the city, use them for a short journey and return them at any other station in the city. A crucial factor for the success of such system is its ability to meet the fluctuating demand for bicycles and for vacant lockers at each station. In order to meet the demand, the inventory of each station should be regularly reviewed. The paper introduces an inventory model suited for the management of bike rental stations and a numerical solution method to solve it. Extensive numerical study, based on real-life data is presented to demonstrate the effectiveness and efficiency of this method.

Keywords: Shared mobility systems, Inventory Management

1. Introduction and Literature Review

Bike-sharing systems (BSS) allow individuals to rent a bicycle at automatic rental stations scattered around a city; use them for a short journey; and finally return them to any other station in the city. As of December 2010, some 238 cities around the world deployed such systems and currently 53 are in the planning stage, see MetroBike LLC (2011). For a review of the history of BSS and on current trends, see DeMaio (2009).

Bike-sharing programs encourage residents to use bicycles as an environmentally sustainable and socially equitable mode of transportation. Such programs also serve to complement other modes of mass transit systems by mode sharing. In addition to these transportation functions, a municipal BSS may yield revenue for a city in the framework of a compliance carbon offset market, see, e.g., Capoor and Ambrosi (2009).

Modern BSS are supported by information systems that provide data about the state of the system (i.e., the number of bicycle and lockers available at each station). This information is accessible on-line via the World Wide Web and in data kiosks at the stations. In addition, since the automatic rental stations identify the bicycles using radio frequency identification technology (RFID), operators can easily obtain detailed statistics

regarding trips undertaken by users. This information can be used to support operational decisions and long-term planning.

A crucial factor for the success of a BSS is its ability to reliably cope with fluctuating demand. Indeed, the main complaints voiced by BSS users relate to the unavailability of bicycles at their point of origin and (even worse) of lockers at their destinations. For example, in order to improve Brussels' bike-sharing program (Villo), a voluntary group of users created a web service that gathers inventory data from the city's BSS website in order to monitor the service level and create public pressure on the operator to improve the system. According to the group's web site (<http://www.wheresmyvillo.be/>), they want to make "JCDecaux [the operator of Villo] drastically improve the availability of bikes and parking spaces through better reallocation of bikes".

Meeting the demand for bicycles and vacant lockers is a particularly challenging problem due to inherent imbalances in the renting and return rates at the various stations. While the flow of commuters is approximately balanced over the course of a day, this is not the case for the flow of bicycles. This is due to the fact that a BSS may be used as a partial substitute for other modes of transportation. For example, users may choose to ride a bicycle in one direction and take a bus on their way back for a variety of reasons: weather or traffic conditions; availability and frequency of the bus service at the time of their journey; topography of their route, etc. In some cases, the imbalance is persistent, e.g., relatively low return rate at stations located on the top of hills. In other cases, the imbalance is temporary, e.g., at suburban train stations there is apt to be a high return rate in the morning, as commuters into the city drop off their bikes and in the afternoon a high rental rate as commuters exit the train and begin to make their way home. Satisfying user demand subject to such imbalances requires a dedicated fleet of light trucks to regularly transfer bicycles among stations. We refer to this activity as *repositioning* bicycles.

There are two repositioning modes. Static repositioning is performed during the night, when the system is nearly idle; dynamic repositioning takes place during the day in order to cope with looming shortages. In practice, many operators work in both modes.

If one ignores the stochastic nature of the demand, the static bicycle-repositioning problem can be classified as a variation of the pickup and delivery problem (PDP), a type

of problem that has attracted considerable attention in recent years. Berbeglia et al. (2007) surveyed the literature on static PDP and classified these problems according to various parameters. Benchimol et al. (2010) and Chemla et al. (2011) studied a single vehicle, single commodity pickup and delivery problem without time constraints with the goal of minimizing the total travel distance of the vehicle as it completed a prescribed repositioning task. These studies are motivated by BSS. Chemla et al. (2011) present a branch-and-cut algorithm for solving a relaxation of the problem, from which a solution is obtained through a Tabu Search. Both of these papers assume a known target inventory level for each station in the system.

Nair and Miller-Hooks (2011) used a stochastic programming approach to handle dynamic repositioning planning in shared mobility systems. Their model assumes that the cost of moving a single vehicle between two given stations is known and fixed. This assumption is realistic for the one-way car-sharing systems that motivated their work. Vogel and Mattfeld (2010) presented a stylized model to assess the effect of dynamic repositioning efforts on service levels. Their model is useful for strategic planning but is not detailed enough to support repositioning operations.

Some authors consider strategic decisions regarding the capacity and locations of bike rental stations. Shu et al. (2010) proposed a stochastic network flow model to support these decisions. They used their model to design a BSS in Singapore based on demand forecast derived from current usage of the mass transit system. Lin and Ta-Hui (2011) considered a similar problem but formulated it as a deterministic mathematical model. Their model is aware of the bike path network and mode sharing with other means of public transportation.

The static repositioning operation poses a new and challenging inventory routing problem. In this paper, we focus on the inventory management part of this problem. Accordingly, we define a user dissatisfaction measure; model user behavior when facing shortages; and devise an efficient method to estimate the expected value of this measure given an initial inventory, station capacity, length of the replenishment cycle and the (stochastic) demand patterns.

The inventory problem associated with a bike sharing rental station can be classified as a closed-loop inventory problem. (There is considerable literature on this

topic. For a comprehensive review, see Dekker et. al. (2004). However, the available models assume that the replenishment operations by new and used items occur periodically. These models do not capture the minute-to-minute dynamics of a bike rental station where replenishment by users who return bicycles (the equivalent of replenishment by used items) occurs randomly at much higher rates but in smaller quantities than replenishment by repositioning vehicles (equivalent of replenishment by new items). The latter may occur every several hours or even every night. Moreover, in a BSS, users may choose to either abandon the station when bicycles or vacant lockers are not available or wait until one becomes available.

The main contributions of this paper are: introducing a user dissatisfaction function to measure the performance of a station; presenting an inventory model of a bike sharing rental station; establishing the convexity of two important special cases of this function; and devising an efficient approximation method to estimate them.

The inventory model we developed is useful for supporting operational tasks such as inventory routing of bicycles among stations, see Raviv, Tzur and Forma (2011). It can be also used for strategic decisions regarding the system-wide inventory of bicycles and for station size decisions (discussed in greater detail in the conclusion section). In addition we believe that our approach may be applicable for other closed-loop inventory systems.

Notation and a formal definition of the model are presented in Section 2. The convexity of the user dissatisfaction measure for two important special cases is proved in Section 3. In Section 4, a method to approximately solve the model under the assumption that the arrival processes are Poisson is devised. The results of numerical experiments that demonstrate the effectiveness and robustness of the approximation are presented in Section 5. We conclude with a discussion in Section 6 about possible straightforward applications of the results presented in the paper and some ideas for future research.

2. Notation and Problem Formulation

We consider a single bike-sharing station over a finite horizon $[0, T]$ with the following settings. At time 0, the inventory level (number of bicycles) in the station is set. During the horizon, users who wish to rent or return bicycles arrive at the station according to an

arbitrary stochastic process. If the desired service can be provided right away, the bicycle is rented or returned and the inventory level is updated. If the service cannot be provided (i.e., empty station for a renter or full one for a returner) then the user decides whether to queue at the station or to abandon the system, i.e., to forgo renting a bike on this particular occasion. The abandoner may decide to seek the service at another station but this is not within the scope of our single-station model. If the user decides to join the queue, we assume that she will wait until the service can be provided¹. The user's decision may be affected by her prior knowledge of the demand patterns at the station at the time of her arrival and on the length of the queue, i.e., the amount of time she will have to wait.

There are four sources of user dissatisfaction in this model. The system is penalized for each one of them.

p^A Penalty charged for each potential renter who abandons due to shortage of bicycles.

h^A Penalty charged for each returner who abandons due to shortage of vacant lockers.

p^Q Penalty charged per time unit of a renter waiting for an available bicycle.

h^Q Penalty charged per time unit of a returner waiting for a vacant locker.

The number of lockers in a station is denoted by C . We refer to this value as the *station capacity*. Next we present a notation that approximately describes a particular realization of the demand. The description is based on discretization of the horizon to Θ short period. These periods are indexed by $\theta \in \{1, 2, \dots, \Theta\}$. The state of the station is given by an integer number i . A state $i \in \{0, \dots, C\}$ refers to a situation where the number of bicycles in the station is i and no user is waiting for a bicycle or a locker. A state $i > C$ refers to a situation where the station is full and $i - C$ users are waiting for a vacant locker and a state $i < 0$ refers to a situation in which the station has no bicycles and $-i$, are waiting for one.

¹ While this is not necessarily a rational behavior, especially when the demand is non-homogenous over time, we believe that it approximately represents rational behavior because in many cases the expected remaining time is non-increasing over time.

$b_\theta(i)$ The number of users that arrived at the station at period θ in order to rent a bicycle and were ready to join the queue assuming the state of the system at the beginning of the period is i .

$s_\theta(i)$ The number of users that arrived at the station at period θ in order to return a bicycle and were ready to join the queue assuming the state of the system at the beginning of the period is i .

The random variables $b_\theta(i)$ and $s_\theta(i)$ are determined for each realization. Both are defined for $i = 0, 1, 2, \dots$ and must satisfy $b_\theta(i) \geq b_\theta(i - 1)$ and $s_\theta(i) \geq s_\theta(i + 1)$. That is, any user who is willing to wait in a long queue will be willing to wait in a shorter one. Next, we define the following set of state variables for $\theta \in \{1, \dots, \Theta\}$:

I_θ The state of the station at the end of period θ . If $0 \leq I_\theta \leq C$, then I_θ is the number of bicycles in the system. If $I_\theta \leq 0$, then the station is empty and there are $-I_\theta$ users waiting in the queue for the bicycle. If $I_\theta > C$, then the station is full and there are $I_\theta - C$ users waiting in the queue for a locker.

Finally, I_0 is a design variable that represents the initial inventory level of bicycles. Using this notation the inventory problem of a bike rental station can be formulated as follows:

$$\min_{I_0 \in \{0, \dots, C\}} E \left\{ \sum_{\theta=1}^{\Theta} \left(h^Q \max(0, I_\theta - C) + p^Q \max(0, -I_\theta) + h^A (s_\theta(C - 1) - s_\theta(I_{\theta-1})) + p^A (b_\theta(1) - b_\theta(I_{\theta-1})) \right) \right\} \quad (1)$$

Subject to

$$I_\theta = I_{\theta-1} - b_\theta(I_{\theta-1}) + s_\theta(I_{\theta-1}) \quad (2)$$

The objective function 0(1) to be minimized is the expected user dissatisfaction over all possible realizations expressed as a weighted sum of the total number of abandonments and the waiting time of users. The function in 0(1) is the *User Dissatisfaction Function (UDF)*. Constraint (2) is inventory balance equality.

It is possible to estimate a solution of (1) and (2) using numerical methods, such as a Monte Carlo simulation. Since the program consists of a single design variable that can only obtain values in the finite set $\{0, \dots, C\}$ it is possible to use such method in order to

solve the problem by evaluating all candidate solutions. However, our analysis in Section 3 reveals some useful structural properties of the UDF.

We point out that the value of the UDF is of interest for all values of initial inventory I_0 and not only for the ones that minimize the function. This is due to the fact that in many cases it is impossible or too costly to set the initial inventory levels at all stations to their ideal values. In order to understand the trade-off between the number of bicycles in the system, the cost of transferring bicycles among stations and user dissatisfaction, a method to efficiently calculate the UDF is needed.

3. Convexity of the User Dissatisfaction Function

Recall that the objective function (1) expresses our measure for the expected dissatisfaction caused by a shortage of bicycles and lockers in a station. In this section we prove that in some settings this function is convex with respect to the initial inventory, I_0 .

The convexity of the UDF is of interest for the following reasons. First, it allows a solution of the stochastic program (1)-(2) in $\log_2(C + 1)$ evaluation of the function. Without such knowledge, $C + 1$ evaluations are needed. Note that depending on the nature of the underlying stochastic process, the evaluation may be computationally demanding. This would be the case if, for example, the evaluation is done by discrete event simulation. Other merits of the complexity property are discussed in Section 6.

In this section we assume, without loss of generality, that the horizon is discretized to infinitesimally small periods and thus $P[b_\theta(i) + s_\theta(i) > 1] = 0$ for all θ . Let $F(x)$ denote expected user dissatisfaction as expressed by (1) for initial inventory level $I_0 = x$. Let us define the marginal dissatisfaction function from additional units of initial inventory as

$$F'(x) \equiv F(x) - F(x - 1).$$

Clearly, this marginal value function is a discrete counterpart of the differential in a continuous function. Equivalently, if $F'(x)$ is a non-decreasing function, then $F(x)$ is said to be convex.

We start by considering the case of impatient users, i.e., users who abandon the station if they cannot get service right away. While this is probably not the reality of

most BSS, it may very likely be a good approximation since users may abandon the station and seek service at neighboring stations if the rental station network is dense.

Theorem 1: If for all possible realizations for all $\theta \in 1, \dots, \Theta$, $b_\theta(i) = 0$ for all $i \leq 0$ and $s_\theta(i) = 0$ for all $i \geq C$, then the UDF, $F(x)$, is a convex function of the initial inventory I_0 .

Proof: The outline of our proof is as follows. First note that since the expected user dissatisfaction is obtained as a weighted sum of the user dissatisfaction over all possible realizations, it is enough to show that the total user dissatisfaction is convex for any particular realization. Next note that the total dissatisfaction for a given realization is a weighted sum of four components, namely:

1. number of potential renters that abandon the station
2. number of potential returners that abandon the station
3. total waiting time of renters
4. total waiting time of returners

However, since the users are all impatient, the last two components disappear. Hence, it is enough to show that the first two are convex functions of the initial inventory in any given realization. For a fixed realization, let us denote these quantities by $f_1(x), f_2(x)$ and let $I_\theta(x)$ denote the station's state at time period θ where $I_0 = x$.

Consider the state of two hypothetical systems under the same demand realization that differs by their initial inventory level by one unit. Let us follow the process $I_\theta(x - 1)$ and $I_\theta(x)$. If $I_\theta(x) > 0$ and $I_\theta(x - 1) < C$ for all $\theta \in \{0, \dots, \Theta\}$ then, there are no abandonees in either system and thus $I_\theta(x) - I_\theta(x - 1) = 1$ for all θ . However, if at some period abandonment occurs in one of the processes then, from this period on, the two processes coincide. Consequently, the difference between the numbers of abandonees in both systems is at most one. For a given scenario there are three alternatives:

1. If $I_\theta(x)$ hits 0 before $I_\theta(x - 1)$ hits C , then $f_1'(x) = -1$ and $f_2'(x) = 0$.
2. If $I_\theta(x - 1)$ hits C before $I_\theta(x)$ hits 0, then $f_1'(x) = 0$ and $f_2'(x) = 1$.
3. If $I_\theta(x)$ never hits 0 and $I_\theta(x - 1)$ never hits C , then $f_1'(x) = 0$ and $f_2'(x) = 0$.

Next, observe that $f_1'(x) = -1 \rightarrow f_1'(x - 1) = -1$. This is due to the fact that $I_\theta(x - 2)$ always hits 0 before $I_\theta(x - 1)$ hits it. Note that it also possible that $f_1'(x) = 0$

and $f_1'(x-1) = -1$ or $f_1'(x) = f_1'(x-1) = 0$. That is, $f_1'(x)$ is a non-decreasing function and so $f_1(x)$ is convex. Similarly $f_2'(x-1) = 1 \rightarrow f_2'(x) = 1$ and thus $f_2(x)$ is convex. **Q.E.D.**

Next we consider other extreme case where the patience of the users is infinite. That is, the users always wait at the station until they get service.

Theorem 2: If all renters (respectively, returners) always wait at the station until a bicycle (respectively, vacant lockers) is available, then the UDF, $F(x)$, is a convex function of the initial inventory I_0 .

Proof: Note that under the conditions of the theorem, $b_\theta(i)$ and $s_\theta(i)$ are fixed in any given realization for all i and $\theta \in 1, \dots, \Theta$. As in the proof of Theorem 1, we prove that the statement holds for each possible realization and thus for the expected value as well. In this case, since there are no abandonments, the dissatisfaction consists of a weighted sum of the total waiting time of renters. For a fixed realization, let us denote the total waiting time of renters by $f_3(x)$ and the total waiting time of returners by $f_4(x)$. Again, it is enough to show that these functions are convex.

Note that in this setting $I_t(x) - I_t(x-1) = 1$ for all possible values of initial inventory $x, x-1$ and for all $t \in [0, T]$.

$$f_3(x) = \int_0^T \max(0, -I_t(x)) dt$$

$$f_4(x) = \int_0^T \max(0, C - I_t(x)) dt.$$

Note that

$$f_3'(x) = f_3(x) - f_3(x-1) = \int_0^T [\max(0, -I_t(x)) - \max(0, -I_t(x-1))] dt.$$

Observe that $f_3'(x)$ is equal to minus the total time over the period $[0, T]$ in which $I_t(x) \leq 0$. That is,

$$f_3'(x) = - \int_0^T \mathbb{I}_{\{I_t(x) \leq 0\}} dt,$$

when $\mathbb{I}_{\{\cdot\}}$ is an indicator function.

Equivalently, $f'_3(x - 1)$ is equal to minus the duration of time in which $I_t(x - 1) \leq 0$. Now, since the latter is clearly smaller (larger in absolute value), it follows that $f'_3(x - 1) \leq f'_3(x)$ and so the convexity of $f_3(x)$ is proved. Similarly, $f'_4(x)$ [respectively, $f'_4(x - 1)$] equal to the duration of time in which $I_t(x) > C$ [respectively, $I_t(x - 1) > C$]. Hence $f'_4(x - 1) \leq f'_4(x)$ and the convexity of $f_4(x)$ is also proved.

Q.E.D.

We saw that the two extreme characterizations of user behavior, i.e., when all users are impatient and when all users have unlimited patience, result in a convex UDF. It is interesting to point out that this is not necessarily the case for a general user behavior model, i.e., when some users abandon and others choose to queue at the station. Consider for example the following (deterministic) arrival process with only two arrivals of renters and no arrivals of returners. At $t = 1$ an impatient renter arrives and at $t = 2$ a renter with unlimited patience arrives. It easy to observe that:

- with $I_0 = 2$ the station will not be penalized for the waiting time of users and for abandonments
- with $I_0 = 1$ the station will accumulate a total waiting time of $T - 2$ over the planning horizon and there will be no abandonments
- with $I_0 = 0$ the station will accumulate a total waiting time of $T - 2$ over the planning horizon and one renter will abandon.

Now $F'(1) = -p^A$ and $F'(2) = -p^Q(T - 2)$. Clearly it is possible to choose parameter values such that $F'(1) > F'(2)$ which implies that $F(\cdot)$ is not convex. Nevertheless, we conjecture that UDF is quasi-convex. This conjecture is supported by the 80 numerical experiments reported in Section 5.

4. Approximation of the User Dissatisfaction Function

In this section we present a method for calculating the UDF under some sound modeling assumptions on the nature of the stochastic process that governs the demand at each rental station. In particular, we assume that arrival the processes of renter and returners are non-homogenous Poisson processes with rates denoted by μ_t and λ_t respectively.

A renter who arrives at the station at time t decides with probability $\beta_t(I_t)$ to join the queue at the station and wait for service and to abandon the station with probability $1 - \beta_t(I_t)$. $\beta_t(I_t)$ is assumed to be a non-decreasing function of the current system state, I_t , and $\beta_t(I_t) = 1$ for $I_t > 0$. This reflects the fact that the tendency of a renter to join the line is determined by her expected waiting time, which is affected by the particular time of her arrival and the current length of the queue. If a renter arrives at the station when bicycles are available, she will certainly join the (empty) queue.

Similarly, a returner who arrives at the station at time t decides with probability $\sigma_t(I_t)$ to join the queue at the station, $\sigma_t(I_t)$ is non-increasing in I_t . For $I_t < C$, we have $\sigma_t(I_t) = 1$ because if vacant lockers are available, the returner will use one. Otherwise, he may either queue at the station or search for a vacant locker at a neighboring one.

The state of the station can be viewed as a non-time-homogenous birth and death process depicted as a Markov chain in Figure 1. The birth rate is λ_t for all states, I_t , in which vacant lockers are available at the stations (hence $\sigma_t(I_t) = 1$) and $\lambda_t\sigma_t(I_t)$ for $I_t \geq C$. Similarly, the death rate is $\mu_t\beta_t(I_t)$ for $I_t \leq 0$ and μ_t otherwise.

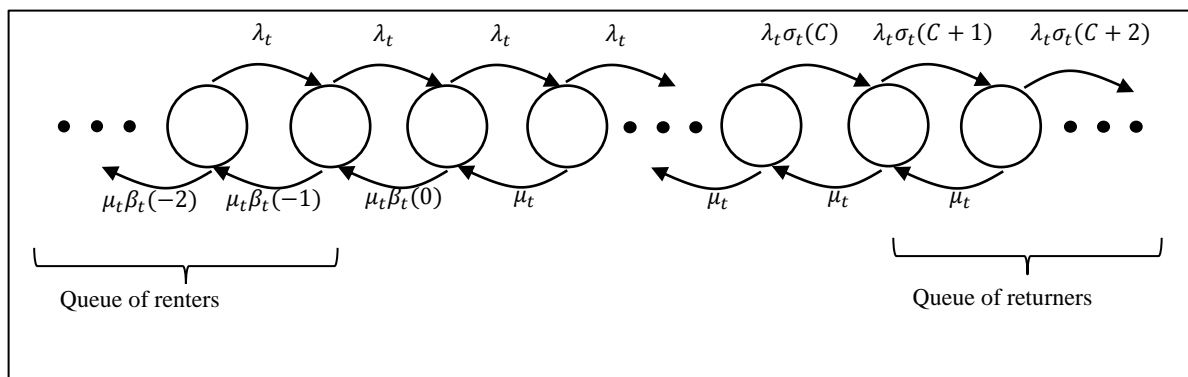


Figure 1: Continuous time Markov chain that represents the dynamics of the bicycles inventory Level

Note that this process admits no steady state and that we are interested in its dynamics. Let $\pi_{ij}(t)$ denote the probability of the station being at state j at time t providing that its initial state at time 0 was i . We use the notation $\boldsymbol{\pi}(t)$ to refer to the whole transition probability matrix. It is possible to state the UDF in terms of the transition probabilities as follows:

$$\begin{aligned}
F(x) = \int_0^T & \left(\sum_{j=-\infty}^0 \pi_{xj}(t) [p^A \mu_t(1 - \beta_t(j)) - jp^Q] \right. \\
& \left. + \sum_{j=C}^{\infty} \pi_{xj}(t) [h^A \lambda_t(1 - \sigma_t(j)) + (C - j)h^Q] \right) dt.
\end{aligned} \tag{3}$$

The first summation term in the integral represents the expected user dissatisfaction accumulated when the station is empty. During such a period, abandonments occurred at a rate of $\mu_t(1 - \beta_t(j))$ and each bears a penalty of p^A . In addition, the total waiting time of renters is accumulated at a rate that is equal to the queue length, $-j$. The second summation term represents the expected user dissatisfaction that accumulates when there are no vacant lockers at the station. During such a period of time, abandonments occurred at a rate of $\lambda_t(1 - \sigma_t(j))$ and each bears a penalty of h^A . In addition the total waiting time of returners is accumulated at a rate that is equal to the queue length, $(C - j)$.

The computationally challenging part of evaluating (3) is the calculation the transition matrix $\pi(t)$. There is no closed form solution for the dynamics of this system. In addition, it is important to note that no insight can be gained from a steady state analysis of some simplified version of the system, such as one with a homogenous arrival process. This is mainly because when the system is properly operated, it never approaches a steady state. Instead, its operator regulates the inventory level and periodically sets it at levels that are in general very far from the steady state mean ones to which it will deteriorate if left untouched. Next, we present a procedure to estimate (3) using discretization and truncation of the Markov chain in **Error! Reference source not found.** In the next section we validate this procedure by comparing it to a simulation that requires numerous replications and is thus much more demanding computationally.

We observe that the transition probabilities to extreme states, i.e., ones that are much below 0 or above C, are very small. There are two reasons for this: (1) the tendency of users to join the queue decreases as the length of the line increases. Indeed, since the user may have other alternatives, such as seeking a bicycle or vacant locker at another station the length of queues must be bounded; (2) renters are unlikely to form long queues at that time of day when the arrival rate of returners at the station is low relative to

the arrival rate of renters because at such time they expect that the waiting time will be very long. In addition, when the arrival rate of returners is relatively high, a queue of renters is unlikely to develop even if the renters are willing to queue. A similar argument for the unlikeliness of a long returner queue also holds. These observations allow us to truncate extreme states from the Markov chain, which enables us to base our calculations on a finite number of states.

Next, we claim that while the arrival rates of renters and returners is non-homogenous over time, it is reasonable to assume that these rates change in a finite number of steps over a planning horizon of, let us say, every 15 or 30 minutes. This modeling assumption can be justified merely by the fact that it is typically impossible to estimate the rates reliably for shorter periods.

Now our strategy to estimate $F(I_0)$ is as follows. We discretize the planning horizon into short periods of one minute each. Let δ denote the length of such a period. For each period θ , we evaluate the transition probability matrix from the beginning of the period to its end. We denote this transition probability by P_θ . Now the transition probability matrix from time 0 to time $t = \tau \cdot \delta$ is given by

$$\boldsymbol{\pi}(t) = \prod_{\theta=1}^{\tau} P_\theta.$$

Moreover, using the recursive relation $\boldsymbol{\pi}(t) = \boldsymbol{\pi}(t - \delta) \cdot P_{t/\delta}$ one can calculate all the transition matrices $\boldsymbol{\pi}(t)$ for $t = \delta, 2\delta, \dots, \tau\delta$ in only τ matrix multiplication operations once the values of P_θ are estimated. Recall that since the arrival rates are constant during each period θ , the $P_\theta = e^{Rt}$ where R is the transition rate matrix of the Markov chain in Figure 1 with transition rates that correspond to time t . An approximation of e^{Rt} can be obtained by the identity

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n.$$

This limit can now be well approximated by calculating $\left(1 + \frac{x}{n}\right)^n$ [see, for example Ross (2010), Section 6.8].

Based on the approximate value of $\boldsymbol{\pi}(t)$, we can obtain a reliable approximation of the user dissatisfaction function using the following discretization procedure

$$F(x) \approx \delta \sum_{\theta=0}^{\frac{T}{\delta}-1} \left(\sum_{j=-L}^0 \pi_{x_j}(t + 0.5\delta) [p^A \mu_\theta (1 - \beta_\theta(j)) - jp^Q] + \sum_{j=C}^U \pi_{x_j}(t + 0.5\delta) [h^A \lambda_\theta (1 - \sigma_\theta(j)) + (j - C)h^Q] \right). \quad (4)$$

where L (respectively, U) denotes an upper bound on the length of the queue of renters (respectively, returners). The approximation procedure is parameterized by L, U and δ . Clearly, as $L, U \rightarrow \infty$ and $\delta \rightarrow 0$ the approximate value approaches $F(x)$ but the computational effort increases. Note that we use the transition probability to the midpoint of each discretized period as a measure of continuity correction.

In the special case, when users are assumed to have no patience at all and thus abandon the system if they cannot obtain a bicycle or a locker right away, the infinite number of states that represents users waiting in the queue can be safely truncated and a more accurate approximation is obtained by a simplified version of (4)

$$F(x) \approx \delta \sum_{\theta=0}^{\frac{T}{\delta}-1} (\pi_{x_0}(t + 0.5\delta) p^A \mu_\theta + \pi_{x_C}(t + 0.5\delta) h^A \lambda_\theta). \quad (5)$$

We note that while the no-patience assumption is not realistic in such a service system, we believe that such a modeling assumption is useful in deriving a nearly optimal initial inventory. Indeed, the UDF that is calculated in (5) ignores the potential waiting time of users but at the same time increases the number of abandonments. The merits of using this model are clear:

- one can use this model without estimating the tendency of users to abandon (β and σ)
- the model can be calculated much more accurately
- a lower and upper bound on the estimation error can be constructed as demonstrated below

$$F(x) \geq \delta \sum_{\theta=0}^{\frac{T}{\delta}-1} (\min(\pi_{x0}(t), \pi_{x0}(t+1)) p^A \mu_{\theta} + \min(\pi_{xC}(t), \pi_{xC}(t+1)) h^A \lambda_{\theta}). \quad (6)$$

$$F(x) \leq \delta \sum_{\theta=0}^{\frac{T}{\delta}-1} (\max(\pi_{x0}(t), \pi_{x0}(t+1)) p^A \mu_{\theta} + \max(\pi_{xC}(t), \pi_{xC}(t+1)) h^A \lambda_{\theta}). \quad (7)$$

Inequalities (6) and (7) are valid due the fact that during each period θ , the rate of both renters and returners abandonments accumulation is monotonous, either non-increasing or non-decreasing depending on the relation between μ_{θ} and λ_{θ} . For example, if the renter arrival rate is greater than the returner arrival rate, $\mu_{\theta} > \lambda_{\theta}$, the probability of the station being empty, $\pi_{x0}(t)$, increases during the period while the probability of the station being empty, $\pi_{xC}(t)$ decreases. Calculating the number of abandonments based on the lower probability between the results in a lower bound while using the higher probabilities results in an upper bound.

In Figure 2 below, we illustrate the effectiveness of the bounds described above. The UDF presented in the figure is for a 20-hour period (6am-2am) at one of the busiest stations of Capital Bikesharing in Washington DC, a station with 15 lockers. The arrival rate of users was estimated based on data collected in the station during weekdays in the winter of 2010-2011. The average number of arrivals and returners per day during this period was 60.88 and 65.12 respectively. The computation was based on time discretizations of 15 and 5 minutes. Since the cost of abandonment of types was set to one ($p^A = h^A = 1$), the y-axis represents the expected number of user abandonments per day.

One interesting observation from the graphs is that although the station faces, on average, a similar number of renters and returners, it is best to set the initial inventory of the station at almost empty. This is due the the fact that most of the returners arrive in the morning while most of the renters arrive in the afternoon, as can be observed in Figure 3. This is a typical situation for many stations that are located in city centers.

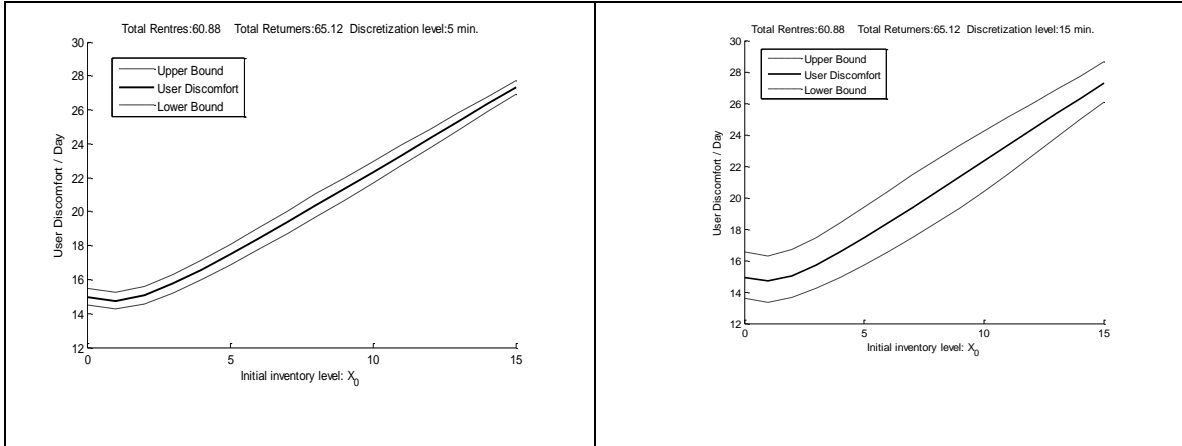


Figure 2: UDF and bounds, assuming users with no patience with discretization levels of 5 and 15 minutes

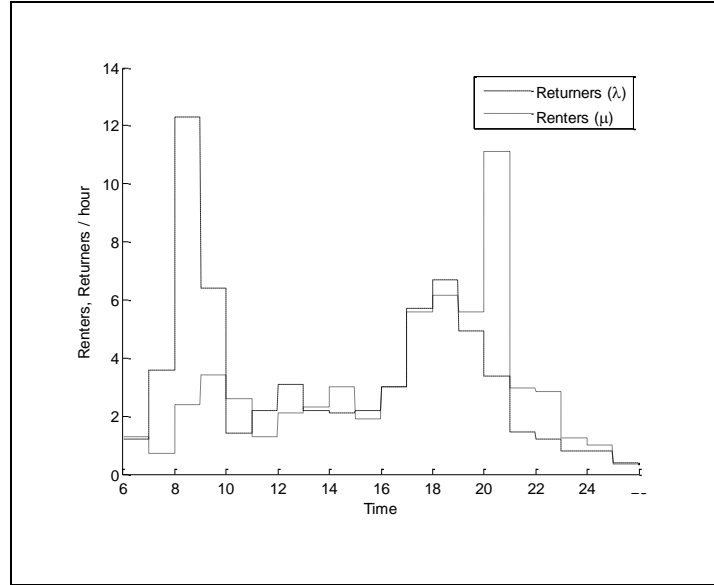


Figure 3: Renters and returners arrival rate at a station

It is possible to bound (4) in similar way to (6) and (7). That is

$$\begin{aligned}
 F(x) \geq & \delta \sum_{\theta=0}^{\frac{T}{\delta}-1} \left(\sum_{j=-\infty}^0 \min(\pi_{x_j}(t), \pi_{x_j}(t+1)) [p^A \mu_{\theta}(1 - \beta_{\theta}(j)) - jp^Q] \right. \\
 & \left. + \sum_{j=C}^{\infty} \min(\pi_{x_C}(t), \pi_{x_C}(t+1)) [h^A \lambda_{\theta}(1 - \sigma_{\theta}(j)) + (j - C)h^Q] \right) \quad (8)
 \end{aligned}$$

$$\begin{aligned}
F(x) \leq & \delta \sum_{\theta=0}^{\frac{T}{\delta}-1} \left(\sum_{j=-\infty}^0 \max(\pi_{xj}(t), \pi_{xj}(t+1)) [p^A \mu_\theta (1 - \beta_\theta(j)) - jp^Q] \right. \\
& \left. + \sum_{j=C}^{\infty} \max(\pi_{xC}(t), \pi_{xC}(t+1)) [h^A \lambda_\theta (1 - \sigma_\theta(j)) + (j - C)h^Q] \right) \tag{9}
\end{aligned}$$

But, since the approximation of these expressions is affected also by the state truncation, it is impossible to use them for explicit calculation of valid bounds.

The computationally demanding part of (4)-(7) is the approximation of $\pi(t)$. This calculation requires numerous matrix multiplications that can be implemented very efficiently and effortlessly using specialized mathematical packages such as Matlab™ or the open-source package Octave.

As discussed above, if users are patient, we had to truncate the Markov chain and remove states that represent long queues, which are unlikely to occur during the period $[0, T]$. Observe that for state ≤ 0 , the probability that the next state will be $i - 1$ is approximately $\mu_\theta \beta_t(i) / (\lambda_\theta + \mu_\theta \beta_t(i))$ [this statement is exact if arrival rates are constant]. Now if this probability is small enough, the process is unlikely to arrive at state -1 . Furthermore, since $\beta(i)$ is non-increasing, if it ever arrives at this state it is likely to return to state i in a short time. Hence we truncate from the left side of the chain (queue for bicycles) all the states that admit $\mu_\theta \beta_t(i) / (\lambda_\theta + \mu_\theta \beta_t(i)) \leq \epsilon$ for some small positive number ϵ . Similarly from the right hand at the states that admit $\lambda_\theta \sigma_t(i) / (\mu_\theta + \lambda_\theta \sigma_t(i)) \leq \epsilon$ are truncated.

5. Numerical Results

The goal of the numerical study in this section is to demonstrate the efficiency and effectiveness of the approximation method presented in Section 4. This goal is accomplished by comparison to the presented bounds and to the results obtained by simulation.

Our experiments are based on data collected from ten of the busiest stations in the ‘‘Capital Bike-Share’’ program which operates in Washington, DC and Arlington, VA.

The capacity of these stations ranges from 11 to 27 and the total daily demand for bicycles ranges from 30 to 60. Since less than 1% of the demand in these stations occurs between 2 am and 6pm, we decided to ignore this demand and allotted these hours to the replenishment tasks. The UDF is calculated for 6am to 2am ($T = 20$ hours).

We estimated the demand rate for bicycles and lockers at each hour during the day in regular weekdays (Monday-Thursday, excluding holidays) based on demand data collected by the system during a three-month period starting on November 1, 2010. We note that the data is censored because the demand for bicycles (respectively, lockers) is not recorded during periods in which the station is empty (respectively, full). We employed standard statistical methods to overcome this difficulty. The data on the estimated demand rate for each station and for each hour of the day is available from the authors upon request. We assumed the arrival process is a Poisson one.

Demand during this period was particularly low since the Capital Bikeshare system was in its early stages and weather conditions were harsh due to the Washington winter. Consequently, we created 10 additional instances with the same demand patterns and station capacities multiplied by three. The capacities of the inflated stations range from 33 to 81 and the daily demand for bicycles ranges from 90 to 180. These demand rates and station sizes are similar to those in the busiest stations of mature BSSs such as the Paris-based Velib, a large-scale public bicycle sharing system (from whom we could not obtain detailed demand data). The ten original stations are denoted below by serial numbers 1-10 and the ten inflated stations are denoted by 1*-10* respectively.

In all the experiments, the penalty cost associated with abandonment was set to $p^A = h^A = 1\$$ per occurrence with a penalty of $p^Q = h^Q = 2\$$ per waiting hour. The likelihood that a user will join the queue is modeled as follows. Each user has her own level of patience expressed in terms of the expected time that she is willing to wait. If, on arriving at the station, the expected waiting time is shorter than the user's patience, she will join the queue and wait to be served. Otherwise, she will immediately abandon the station. The patience of the user is assumed to be identically distributed with known parameters. For our numerical experiment, we checked three patience distributions: exponential distributions with a mean of ten minutes (long patience); two minutes (shorter patience); and impatient users who abandon the station if the service cannot be

provided immediately. The values of $\beta_t(I_t)$ and $\sigma_t(I_t)$ were calculated accordingly. We could not verify this user behavior model because the “Capital Bike-Share” system does not collect information on waiting times and abandonments. The probability of a renter with mean patience γ who arrives at the station at period θ to join the queue when its state is $i \leq 0$ is given by

$$\beta_\theta(i) = 1 - \mathcal{F}_\gamma\left(\frac{-i + 1}{\mu_\theta}\right)$$

where \mathcal{F} is the cumulative distribution function of the exponential distribution with γ . Similarly, for $i \geq C$,

$$\sigma_\theta(i) = 1 - \mathcal{F}_\gamma\left(\frac{i - C + 1}{\lambda_\theta}\right).$$

In fact, more accurate calculation should be based on the rate of the returns and rentals not only during the current period, θ , but also during future periods but such a complication is not needed for our proof of concept here.

A truncation parameter $\epsilon = 0.0001$ was selected and our simulation study confirmed that this choice is appropriate.

In total, 120 problem instances were tested: 20 stations (ten original and ten inflated) \times 2 discretization levels ($\delta = 1$ min., 5 min.) \times 3 users’ mean patience ($\gamma = 0, \gamma = 2$ min., $\gamma = 10$ min.). Each of these 120 instances was solved by our method. The 40 instances that correspond to the impatient case ($\gamma = 0$) were compared to the lower and upper bounds, (6) and (7). The 80 instances that correspond to patient users ($\gamma = 2$ min., $\gamma = 10$ min.) were compared to the results obtained by simulation. The calculation of (4) and the simulation was programmed in MathWorks Matlab™ the code is available from the author by request. All the experiments were run on an Intel Core i7™ desktop under Windows 7™ 64 bits.

In Table 1 we present the results of our experiments for the impatient case. The expected dissatisfaction function was approximated for each of the 20 stations with both discretization levels. Each station is described by its capacity and expected daily demand for bicycles / lockers. For each run we reported on:

- the computation time (in seconds) including the calculation of the bounds (*CPU time*)

- the maximum relative difference between either the lower or upper bounds over all possible initial inventory levels (*Max Error*)
- the average relative difference between the bounds divided by two (*Average Error*)
- the maximum relative difference between the values approximated by the two levels of time discretization (*Max Diff.*)

Let us denote an approximated UDF using a discretization of level of k minutes by $F^k(x)$. Its lower bound is obtained with the same discretization level by $LB^k(x)$ and the upper bound by $UB^k(x)$. The above quantities can be formally defined as follow:

$$Max\ Error = \max \left\{ \max_{i \in \{0, \dots, C\}} \left(\frac{UB^k(x) - F^k(x)}{F^k(x)} \right), \max_{i \in \{0, \dots, C\}} \left(\frac{F^k(x) - LB^k(x)}{F^k(x)} \right) \right\}$$

$$Average\ Error = \frac{\sum_{i=0}^C [UB^k(x) - LB^k(x)] / F^k(x)}{2(C + 1)}$$

$$Max\ Diff. = \max_{i \in \{0, \dots, C\}} \left| \frac{F^5(x) - F^1(x)}{F^1(x)} \right|$$

Station			5 Minutes Discretization			1 Minute Discretization			Max Diff.
ID	Capacity	Daily demand	CPU time (sec.)	Max Error	Average Error	CPU time (Sec.)	Max Error	Average Error	
1	15	31/26	0.047	2.80%	2.32%	0.199	0.56%	0.46%	0.01%
2	11	32/44	0.048	2.77%	1.98%	0.196	0.55%	0.40%	0.01%
3	15	32/36	0.047	2.42%	2.06%	0.195	0.48%	0.41%	0.01%
4	15	36/33	0.046	2.21%	1.93%	0.199	0.44%	0.39%	0.01%
5	27	37/39	0.058	2.66%	2.17%	0.239	0.53%	0.43%	0.01%
6	15	38/41	0.043	2.20%	2.00%	0.191	0.44%	0.40%	0.01%
7	11	40/38	0.041	1.75%	1.25%	0.19	0.35%	0.25%	0.00%
8	19	45/42	0.049	2.77%	2.17%	0.211	0.55%	0.43%	0.01%
9	15	52/42	0.048	2.78%	2.34%	0.204	0.54%	0.47%	0.01%
10	15	61/65	0.045	3.59%	2.92%	0.2	0.70%	0.58%	0.01%
1*	45	92/78	0.087	5.00%	3.49%	0.3	0.97%	0.70%	0.02%
2*	33	95/131	0.081	3.68%	2.61%	0.264	0.72%	0.52%	0.02%
3*	45	95/109	0.094	4.33%	3.43%	0.312	0.85%	0.69%	0.04%
4*	45	108/99	0.092	3.88%	3.13%	0.314	0.76%	0.63%	0.02%
5*	81	110/116	0.23	4.49%	3.24%	0.669	0.86%	0.65%	0.02%
6*	45	114/122	0.133	3.49%	2.94%	0.311	0.68%	0.59%	0.03%
7*	33	120/115	0.067	2.92%	2.03%	0.349	0.58%	0.41%	0.02%
8*	57	134/126	0.097	4.63%	2.91%	0.365	0.89%	0.58%	0.03%
9*	45	157/127	0.084	4.13%	3.17%	0.311	0.79%	0.63%	0.04%
10*	45	183/195	0.081	5.53%	4.08%	0.289	1.05%	0.81%	0.06%
Avg.			0.076	3.4%	2.61%	0.275	0.66%	0.52%	0.02%

Table 1: Results for the model with no patience

It is apparent from Table 1 that instances of the impatient case could be approximated very quickly by the proposed method. Although the solution time increased with discretization level and station capacity, even the largest instance, corresponding to a station with 81 lockers and time discretization of one minute, could be solved in 0.669 seconds.

The approximation accuracy seems to improve significantly as the discretization level increased. For example, the average Max Error for one minute discretization was on average about fifth of that obtained by a five minute discretization. However, this improvement is mainly due to the tightening of the bounds rather than actual changes in the approximated values. Indeed the maximum relative difference between the two methods is, on average, merely 0.02%. We concluded that the method is fairly accurate, even with a time discretization of five minutes. Thus, if there is a need to solve numerous instances of the problem quickly. For example, in an on-line setting, it is possible to use rougher time discretization in order to save time. Furthermore, it is unlikely that further refinement of the discretization will lead to a significantly more accurate approximation.

Next, the performance of the approximation method for a model that assumes users with limited patience was tested. For each instance and each possible initial inventory $X_0 = 0, 1, \dots, C$ we also ran a simulation with 10000 replications in order to create the whole UDF. For each of the 20 tested stations Table 2 presents:

- the computation time in seconds required for the approximation procedure (*cpu time*)
- the maximal and average relative difference between the mean results of the simulation (*Max Diff.* and *Avg. Diff*)
- the percentage of the results obtained by the approximation results that are outside a 99% confidence interval created based on the simulation (*Outside C.I*)

The above results are reported for both one and five minute discretization level. Formally, let us denote the mean value of $F(x)$ obtained by the simulation by $\bar{F}(x)$ and the half with of the confidence interval corresponds to this point by $h(x)$ then

$$Max\ Diff. \equiv \max_{x \in \{0, \dots, C\}} \frac{|\bar{F}(x) - F(x)|}{F(x)}$$

$$Avg. Diff. \equiv \frac{1}{C+1} \sum_{x=0}^C \frac{|\bar{F}(x) - F(x)|}{F(x)}$$

$$Outside C.I. \equiv \frac{\sum_{x=0}^C \mathbb{I}_{\{F(x) \notin [\bar{F}(x)-h(x), \bar{F}(x)+h(x)]\}}}{C+1}$$

where $\mathbb{I}_{\{\cdot\}}$ is an indicator function.

Station			5 Minutes Discretization				1 Minute Discretization			
ID	Capacity	Daily demand	CPU time (sec.)	Max Diff.	Avg. Diff.	Outside C.I	CPU time (sec.)	Max Diff.	Avg. Diff.	Outside C.I
1	15	31/26	0.064	0.98%	0.32%	0.00%	0.292	0.98%	0.31%	0.00%
2	11	32/44	0.144	0.63%	0.37%	0.00%	0.683	0.62%	0.37%	0.00%
3	15	32/36	0.063	1.80%	0.54%	0.00%	0.304	1.79%	0.53%	0.00%
4	15	36/33	0.071	1.41%	0.55%	0.00%	0.329	1.39%	0.55%	0.00%
5	27	37/39	0.094	3.80%	1.00%	0.00%	0.462	3.79%	1.00%	0.00%
6	15	38/41	0.089	1.77%	0.52%	0.00%	0.422	1.78%	0.53%	0.00%
7	11	40/38	0.06	0.95%	0.46%	0.00%	0.286	0.94%	0.47%	0.00%
8	19	45/42	0.062	1.11%	0.38%	0.00%	0.299	1.09%	0.38%	0.00%
9	15	52/42	0.113	0.83%	0.33%	0.00%	0.55	0.85%	0.33%	0.00%
10	15	61/65	0.195	0.92%	0.36%	0.00%	0.935	0.89%	0.35%	0.00%
1*	45	92/78	0.136	3.60%	0.78%	2.17%	0.644	3.65%	0.79%	2.17%
2*	33	95/131	0.488	0.54%	0.17%	0.00%	2.392	0.52%	0.16%	0.00%
3*	45	95/109	0.285	3.11%	0.56%	2.17%	1.327	3.14%	0.56%	2.17%
4*	45	108/99	0.193	2.81%	0.79%	0.00%	0.912	2.77%	0.79%	0.00%
5*	81	110/116	0.651	12.73%	1.63%	2.44%	4.679	12.79%	1.63%	2.44%
6*	45	114/122	0.339	1.60%	0.50%	0.00%	1.614	1.57%	0.49%	0.00%
7*	33	120/115	0.213	2.32%	0.51%	2.94%	1.027	2.34%	0.51%	2.94%
8*	57	134/126	0.303	2.65%	0.47%	5.17%	1.41	2.70%	0.46%	5.17%
9*	45	157/127	0.399	0.81%	0.26%	0.00%	1.924	0.74%	0.24%	0.00%
10*	45	183/195	0.635	0.85%	0.27%	0.00%	3.095	0.80%	0.27%	0.00%
Avg.			0.23	2.26%	0.54%	1.23%	1.18	2.26%	0.54%	1.23%

Table 2: Results for the model with $\gamma = 2$ min.

In Table 3, the same information is presented assuming the mean patience of users is ten rather than two minutes, i.e., the tendency of users to wait for service in the station is higher. The structure of this table is identical to the structure of Table 2.

Station			5 Minutes Discretization				1 Minute Discretization			
ID	Capacity	Daily demand	CPU time (sec.)	Max Diff.	Avg. Diff.	Outside C.I	CPU time (sec.)	Max Diff.	Avg. Diff.	Outside C.I
1	15	31/26	0.214	1.54%	0.58%	0.00%	1.05	1.55%	0.58%	0.00%
2	11	32/44	0.777	0.93%	0.26%	0.00%	3.679	0.94%	0.26%	0.00%
3	15	32/36	0.449	1.95%	0.49%	0.00%	2.089	1.96%	0.49%	0.00%
4	15	36/33	0.321	1.32%	0.56%	0.00%	1.578	1.32%	0.56%	0.00%
5	27	37/39	1.351	2.66%	0.81%	0.00%	3.03	2.64%	0.81%	0.00%

6	15	38/41	0.56	1.48%	0.59%	0.00%	2.733	1.47%	0.59%	0.00%
7	11	40/38	0.316	1.07%	0.53%	0.00%	1.544	1.07%	0.53%	0.00%
8	19	45/42	0.575	1.69%	0.42%	0.00%	4.242	1.71%	0.42%	0.00%
9	15	52/42	0.588	1.51%	0.36%	6.25%	2.86	1.53%	0.37%	6.25%
10	15	61/65	1.037	1.00%	0.37%	0.00%	5.084	1.02%	0.37%	0.00%
1*	45	92/78	1.058	2.73%	0.67%	2.17%	3.435	2.74%	0.67%	2.17%
2*	33	95/131	3.365	0.74%	0.16%	2.94%	15.368	0.74%	0.15%	2.94%
3*	45	95/109	1.739	2.01%	0.73%	2.17%	8.149	2.04%	0.73%	2.17%
4*	45	108/99	1.213	2.94%	0.81%	2.17%	5.516	2.92%	0.81%	2.17%
5*	81	110/116	4.332	8.86%	1.69%	2.44%	20.325	8.92%	1.69%	2.44%
6*	45	114/122	2.309	1.52%	0.41%	0.00%	10.68	1.48%	0.40%	0.00%
7*	33	120/115	1.138	1.31%	0.43%	0.00%	7.567	1.32%	0.43%	0.00%
8*	57	134/126	1.787	1.74%	0.42%	0.00%	8.373	1.70%	0.41%	0.00%
9*	45	157/127	2.79	0.96%	0.22%	0.00%	1.05	1.55%	0.58%	2.17%
10*	45	183/195	7.786	1.18%	0.24%	0.00%	3.679	0.94%	0.26%	0.00%
Avg.			1.69	1.96%	0.54%	1.07%	7.38	1.95%	0.54%	1.23%

Table 3: Results for the model with $\gamma = 10$ min.

Several observations are derived from Table 2 and Table 3. Most important, it seems that the results for all the 80 instances and for all the possible initial inventory levels nicely agree with the simulation results with mean relative deviations of merely 0.54%. Moreover, the results obtained from the approximation procedures fall in all but 1.19% of the cases within a 99% confidence interval constructed by the simulation results. Note that these cases can be attributed to an estimation error of the simulation rather than an approximation error of the algorithm. Indeed, in 1% of the cases, the true value is expected to fall outside a 99% confidence interval.

The strongest inconsistency between the approximation algorithm and the simulation occurred at station 5* in regard to both patience and discretization levels. We drilled down into the results of these instances in order to obtain a better understanding of what was taking place. **Error! Reference source not found.** presents the UDF, as approximated by (4), and the lower and upper bounds of the 99% confidence interval created by the simulation. One can observe that while the graph of the approximated function crosses the confidence interval twice, its overall trend is very similar to the estimation obtained by the simulation and is very likely to lead to the same decision regarding the optimal initial inventory. Indeed, the greatest difference between the optimal initial inventory prescribed by the simulation and the one prescribed by the approximation procedure, is, in all the 80 experiments, one bicycle. Moreover, even in these cases the relative difference between the expected value of the solution based on

these two methods, when measured according to the approximation procedure, never exceeds 1.24%.

It is also apparent from all our experiments that the results obtained with five-minutes discretization is very similar to those obtained by a one-minute discretization while the running time increased proportionally with the discretization level (the number of periods per hour). This implies that if computation time is a consideration, it is better to use rough rather than fine discretization.

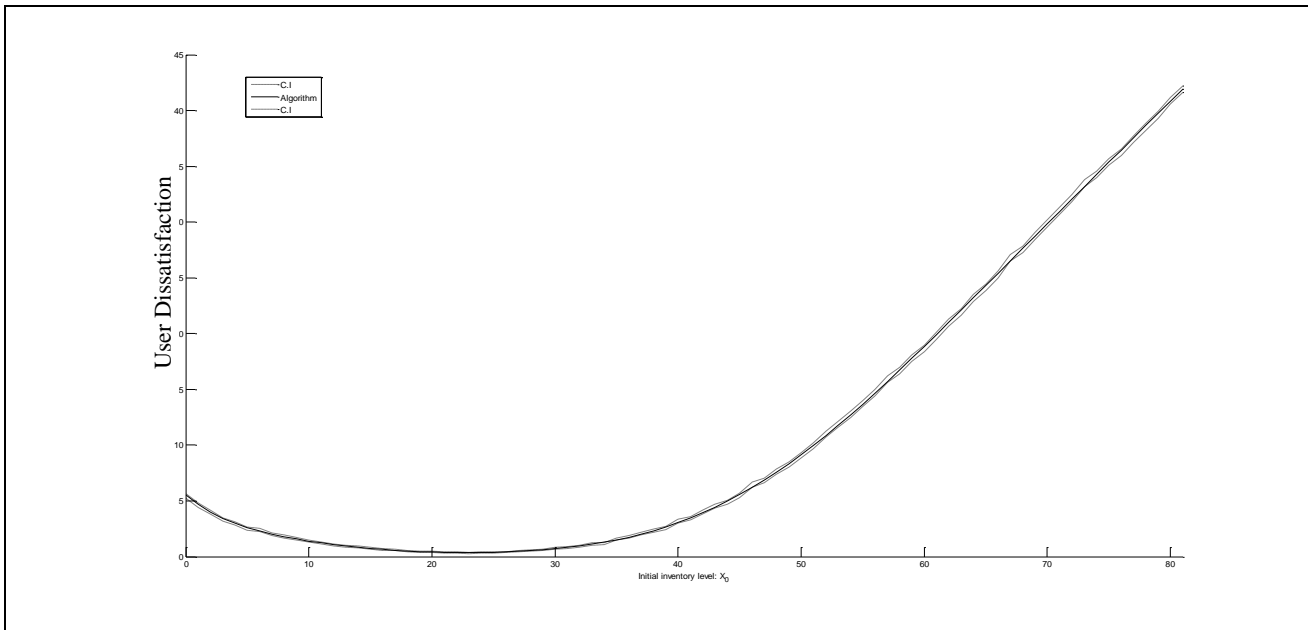


Figure 4: Approximation algorithm vs. simulation results. Station 5*, $\gamma = 10$, discretization level 1 minute

Finally, we compared the results obtained by the model with $\gamma = 0$ and $\gamma = 10$ in terms of optimal initial inventory levels. It seems that the two models prescribe very similar optimal initial inventory levels. Indeed, in 18 out of 20 cases, the same solution was obtained while in two cases (station 6* and 10*) there was a difference of one bicycle. In both of these cases, the relative difference in the value of the UDF (in terms of $\gamma = 10$ model) was less than 0.2%. We conclude that the impatient ($\gamma = 0$) model can be safely used as a surrogate for the more detailed model that allows users to wait at the station.

6. Discussion and Conclusions

This paper presents an inventory model for bike rental stations. This model may be used by operators of BSS in their daily operations while rebalancing their system. Moreover, the convexity property together with the efficient approximation method and the observation that the impatient user model can be used as a surrogate for the more general (but not necessarily convex) model make the results of this study very useful for various operational and strategic decisions that may be made by the BSS operators. In particular, the inventory model may serve as a building block for modeling the following problems: optimal station capacity; optimal global bicycle inventory; and the static repositioning problem. These problems and their relation to the inventory model are briefly discussed below.

Optimal Station Capacity: Consider a BSS operator with a repositioning fleet that allows replenishing stations every fixed interval, T , say every night or twice a day. Let us redefine the UDF (3) as a function of the station capacity, C , and initial inventory, x , $F(C, x)$. Next, let $g(C)$ denote the cost of deploying a rental station with C lockers. If user dissatisfaction can be expressed in monetary value and $g(C)$ is amortized, in terms of the inter replenishment period T , then $G(C, x) = F(C, x) - g(C)$ is the total cost associated with installing C lockers and set the initial inventory level to x . Since we now possess an efficient approximation method for $F(C, x)$, it is possible to find the optimal combination of (C, x) , at least within a bounded range of C , merely by enumerating all possible values. It is interesting to study the properties of $G(C, x)$. We note that it is reasonable to assume that $g(C)$ is concave due to economies of scale. In such a case, convexity of $F(C, x)$ implies convexity of $G(C, x)$.

If it is not possible to express user dissatisfaction and/or station capacity costs in common monetary units, one can still plot an efficiency frontier with station size on one axis and UDF, assuming optimal initial inventory, on the other. This would enable the decision-maker to pick one Pareto optimal solution.

Optimal Global Bicycle inventory: The decision about the optimal number of bicycles in the system is a crucial tactical decision that should be reviewed by the operator whenever demand patterns or station capacity change. Clearly, if the global

inventory is too low the renter is more likely to face bicycle shortages; if it is too high, the returners are apt to face a shortage of lockers.

Let us denote the cost associated with purchasing and maintaining, y , bicycles in the system, amortized to the replenishment period T by $v(y)$. Let x_i (resp C_i) denote the optimal initial inventory (respectively capacity) of station i . Now the following mathematical model can be used to determine the optimal system wide bicycle inventory.

$$\min \sum_{i=1}^n [F_i(C_i, x_i) - g_i(C_i)] - v(y) \quad (10)$$

s.t

$$\sum_{i=1}^n x_i = y \quad (11)$$

$$x_i \in \mathbb{N}^n, \quad y \in \mathbb{N} \quad (12)$$

If the capacity of the stations can be changed simultaneously with the global inventory level, C_i can be regarded as decision variables rather than parameters. While a solution method for the non-linear integer program (10)-(12) is a topic for future research, the proposed efficient approximation method for the UDF may be instrumental in any such method. Since the demand patterns change seasonally, it is recommended that the operator will have a storage capacity to allow addition and removal of bicycles to/from the system as needed.

Repositioning: Preliminary results of this study were used by Raviv, Tzur and Forma (2011) to formulate a mixed integer linear model for routing inventory in order to rebalance stations using a fleet of vehicles subject to capacity and time constraints. The sum of the UDFs of all stations is incorporated into the objective function of the repositioning model rather than used to prescribe target inventory levels. This approach takes advantage of the fact that the UDF is typically flat around the optimal level and thus in many cases it is not worthwhile to visit a station that is off but near its optimal inventory level. A future challenge is to solve the dynamic repositioning problem with the objective of minimizing the total user dissatisfaction. However, in the dynamic case, the exact time of the next visit at the station is unknown because it depends on the future

state of the system. Hence, some adaptation may be needed to allow the function to handle T as a random variable that depends on the initial inventory.

We believe that the single station inventory model presented in this paper may be very useful for BSS operators. However, like any mathematical model, it is based on some simplified assumptions. In particular, the model presented here abstracts out complex interdependencies among the different stations. In fact, we see two types of such interdependencies. First, the demand process at each station may be affected by the state of its neighboring stations since unsatisfied renters and returners are likely to go to other stations. Second, the arrival processes of returners at destination stations are affected by the state of other origin stations. Indeed, if a renter decides to abandon Station A because it is empty, she will never be a returner at her presumed destination B . Moreover, in the long run, the service level experienced by users affects their demand in future periods since disappointed users are less likely to use the system again. In conclusion, we believe that future researchers should create detailed and empirically supported user behavior models. Such a model is unlikely to be efficiently computable but studying it by simulation may help to validate and fine tune the model presented here.

BSS operators generally sign service level agreements with the municipal authorities. In some of these agreements the service level is defined as the percentage of the time in which the station is allowed to be completely empty or completely full. For example, the operator may be committed to ensuring that none of the station will be at such states more than 5% of the time in each day. The parties agree upon a penalty for each breach of this commitment. We believe that this is an inaccurate measure for service level. The service level is affected not only by the length of the periods in which the station cannot provide the service but also and more importantly by the timing of these periods. For example, an empty station next to a high school just before the beginning of the school day is desirable because such a station can accept the expected wave of returners. However, if the same station is empty in the afternoon, just before the end of school day, it is likely to provide a poor service for the expected wave of renters. Hence, we argue that the number of unsatisfied renters and returners is a more appropriate service level measure and that it should be included in service level agreements between cities and BSS operators. Indeed, this measure cannot be calculated accurately but it can

be easily estimated from the system log. However, If the operator does interested in minimizing the length of the period in which the station is empty or full, the definition of the UDF in (3) should be modified to

$$F(x) = \int_0^T \left(\sum_{j=-\infty}^0 \pi_{xj}(t) + \sum_{j=c}^{\infty} \pi_{xj}(t) \right) dt.$$

The adaptation of the approximation formulas, (4) and (5), is straightforward.

Finally, we would like to point out that the *sharing economy* is not limited to shared mobility systems. While the concept of renting equipment instead of buying it is not new, modern information technology makes it a very attractive alternative. For example, construction firms may prefer renting power tools and machinery at the time and location where they need them instead of owning and transporting them. Similarly, some individuals may prefer renting camping equipment in order to save the need to store it at their apartments, etc. Managing the supply chain of a multi-branch renting organization is a challenging task compared to a retail organization due to the need to cope with the dynamics of at least two incoming streams of goods: returned items and items acquired vertically or horizontally from other echelons. The literature on reverse logistics is extensive but primarily focused on systems with items that must be reprocessed in some way before being delivered again. However, in the rental business, returned items can be used again immediately after they are returned. Modeling such systems requires integrating inventory and queuing models. Hence, we believe that the modeling concepts and solution methods presented in this paper may be relevant for many problems originating from the growing sector of renting/sharing organizations.

Acknowledgment: the authors wish to thank Mr. Brodie Hylton and Mr. Danny Quarrel from Alta Bicycle Share Ltd. for providing data on the Capital Bikeshare; Professor Michal Tzur, from Tel Aviv University, for helpful advice; and to Mr. Avraham Edison for help in processing the demand data.

References

Benchimol, M., Benchimol, M., Chappert, B., Taille, A.D.L., Laroche F., Meunier, F. and Robinet, L. (2010). Balancing the Stations of a Self Service “Bike Hire” System, to appear in *RAIRO Operations Research*.

Berbeglia, G., Cordeau, J.-F., Gribkovskaia I., and Laporte, G. (2007). Static pickup and delivery problems: a classification scheme and survey. *TOP* **15**, 1-31.

Capoor K. and Ambrosi, P. (2009). State and Trends of the Carbon Market 2009, The World Bank, Washington, DC.

Chemla, D., Meunier, F. and Wolfler Calvo, R. (2010) Bike hiring system: solving the rebalancing problem in the static case. Working paper.

Dekker, R. (2004). Reverse logistics: quantitative models for closed-loop supply chains. Berlin ; New York, Springer.

DeMaio P. (2009), Bike-sharing: History, Impacts, Models of Provision, and Future, *Journal of Public Transportation*, Vol. 12, No. 4, 2009.

Lin, J-R. and Ta-Hui, Y. (2011) Strategic design of public bicycle sharing systems with service level constraints. *Transportation Research Part E*, 47, 284–294.

MetroBike LLC (2011). observed in July 21, 2011, http://bike-sharing.blogspot.com/2010_12_01_archive.html

Nair R. and Miller-Hooks E. (2011), Fleet Management for Vehicle Sharing Operations, to be appear in *Transportation Science*.

Raviv, T., Tzur, M., and Forma, I.,A., (2011), Static Repositioning in a Bike-Sharing System: Models and Solution Approaches, working paper

Shu J., M. Chou, Q. Liu, C-P. Teo, I-L. 2010. Bicycle-Sharing System: Deployment, Utilization and the Value of Re-distribution, working paper

Ross, S. M. (2010). Introduction to probability models. Amsterdam ; Boston, Academic Press.

Vogel, P. and Mattfeld, D.C (2010), Modeling of Repositioning Activities in Bike-Sharing Systems, Proceeding of the 12th World Conference on Transport Research, July 11-15, 2010, Lisbon, Portugal.