Bike commuters contribution to balance shared bike systems during peak load

Dmitry Gimon
College of Management
Park University
Parkville, Missouri
dmitry.gimon@park.edu

Abstract— **Bike sharing systems are a new and rapidly growing feature of modern cities. The most common problem in these systems is the capacity limitation. Demand depends on the station's location and the time of the day. In this study, we analyze how user behavior incentivized or disincentivized changes depending on the station's load, and as a result allows the stations to remain functional. We analyzed a dataset of NYC Citi Bike system rides and station states. The study shows that users tend to choose other stations as the number of bikes available decreases, and they are more likely to do so when an incentive is introduced. This paper contributes to the studies of collective use of city infrastructure.**

Keywords—bike sharing, behavior analysis, collective systems

I. Introduction

Bike sharing systems have become the norm in cities around the world. Bike sharing systems are an example of the collective use of city infrastructure. Many bike sharing systems use docking stations with limited capacities. The vast majority of bike sharing literature studies the fleet rebalancing problem of these systems to better serve rush hours and inconsistent ride flow over the course of the day. The proposed solutions to this problem include rebalancing using trucks and optimal station allocation [1][2]. The goal of this study is to understand how user behavior helps keep shared systems functioning during peak loads. Users may contribute to the system performance by choosing a less busy station despite their individual preferences, or by choosing other modes of transportation.

A user's contribution to the system performance may be unconscious, when the system benefit is not known to the user, or conscious when user knows that her action will benefit the system. The latter may be either solely altruistic, when the user does not benefit from the action, or it may be incentivized. For example, the Citi Bike system in New York City has a program called Bike Angels, where stations are dynamically being assigned certain points, to encourage users to pick up bikes at stations with a high number of bikes and return them to ones with high number of docks available.

In this study, we provide a generalized model of users' choice using utility theory [3]. Then we studied the observed contribution of bike commuters to decrease the system load during the morning commute by alternating their choices of bike stations. We limited our study to those users who repetitively use bike sharing systems (regular commuters), since their alternations of the station choice can be tracked; To minimize the effect of possible differences in preference models during outgoing and return commutes, we studied morning commutes only.

We compared two types of rides: one when a commuting users decides on a pickup station for her regular commute from a station with no incentives, and when she can earn incentives by alternating the pickup station. We hypothesize that a load of the user's preferred station has an effect on the choice of the station, i.e. if a preferred pickup station has fewer bicycles, then the user more likely decides to use another station. If no bikes or docks are available for pickup or drop off respectively, then the user has no choice.

II. Model

Let S be a set of stations in a bike sharing system. Let each commuter i have a set of stations $P_i \in S$ in close proximity to her home or the transfer point where she transfers to a bike from another mode of transportation, and $R_i \in S$ be a set of stations in proximity to her work or transfer point she transfers to from a bike. For each $s_j \in P_i$, $U^P(s_j)$ represents a numerical value of the preference of picking up a bike at this station. For each $s_j R_i$, $U^R(s_j)$ represents her preference to return the bike at this location.

If we apply the utility theory assuming the rational behavior of commuters, she will pick up the bike at the most preferred station for her pick up $s_i^{P*} = \max_{S_j \in P_i} U^P(s_j)$, and return it to the most preferred station for drop off $s_i^{R*} = \max_{S_j \in R_i} U^R(s_j)$.

To model those preference functions, we need to list the factors that may change her preferences.

Bikes may not be available at her preferred pickup station, or at her drop off station; or there may be very few bikes or docks available, so by making a decision on a station some time before pick up or drop off, the risk that the station becomes unavailable (no bikes or no docks) is high.

Let us assume that each station s_j has a static capacity C_j and dynamic parameters of the number of bikes parked $b_j(t)$, and the number of docks available $d_j(t)$, where $C_j(t)=b_j(t)+d_j(t)$. Then $U^P(s_j, t)$ should be a function of $b_j(t)$ and $d_j(t)$.

There are other factors outside of a network's state that influence a user's decision, such as the weather (if the weather is bad, then the walking distance to a station may become more critical). Let us combine them into a coefficient k_j $(t)=[0..1]$, k_j $(t)=1$, when there are no other factors making a station less/more desirable, if k_j $(t)=0$, other factors make the station unacceptable. Let $I_j^P(t)=[0..1]$ be incentive offered to pick up a bike from station i. Then the user's utility can be described as:

$$U^P(s_j,\ t)=k_j\ (t)*f^P\left(b_j(t),d_j(t),I_j^P(t)\right) \qquad (1)$$

$U^R(s_j,\ t)$ can be defined the same way. Then, depending on the station load and incentives, a rational user would choose their station.

Hypothesis I. A station j becomes less attractive for commuters when there are fewer bikes or docks available. As a result, commuters will alternate to use other stations.

Hypothesis II. If an incentives is offered to use another station, the degree of unattractiveness of stations increases.

III. Data analysis

We used an anonymous dataset of Citi Bike morning rides in New York City for September 2016 [4]; during that month we have also recorded the states of the system (station loads, and incentives offered with the 15 minutes interval). Note that due to the data privacy, we were not able to obtain data on whether a user actually earned points; instead we studied the potential for earning points. Then, we applied a custom clustering algorithm to identify clusters of regular commuters, by clustering rides made by annual subscribers taking a similar route during a similar time of the day during the workweek, no more than one ride per day. We also used provided data on gender and age to verify clusters. Our dataset includes 1350 commuters with the median value of six rides per month. For each cluster we computed the most common pickup and drop off locations for each commuter.

We selected the cases when no incentives were offered at the station where a user picked up a bike, and the cases where a station a user picked up a bike from offered incentives.

Figure 1. shows the likelihood of alternating from the most preferred station depending on its load.

The results demonstrate that commuters tend to choose alternative stations more often when their most preferred station is not available. When incentives are not offered, only extremely low bike availability (between 0 and 10%) results in growth of the probability to alternate.

The alternations are more common when commuters may earn provided incentives (note, our data does not allow us to tell if they actually earned it).
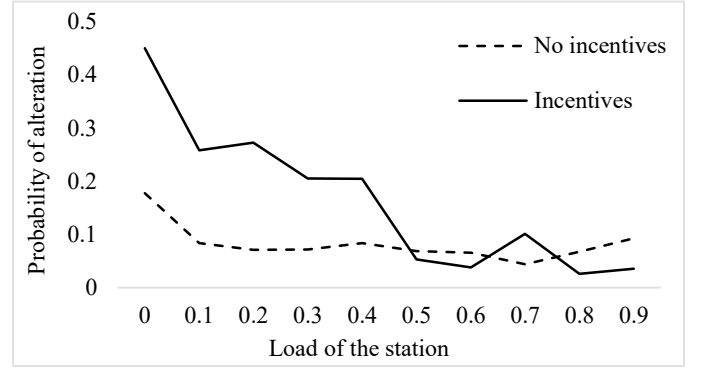

Fig. 1. Probability of station alternation.

The results shows that once incentives are introduced, commuters are more sensitive to the load of their preferred stations, i.e., they are more willing to alternate even when more bikes are available, when compared to the case with no incentives. Therefore, the preliminary results support H1 and H2; however, more formal data analysis should be conducted.

IV. Discussion

The results of our study demonstrate that even when bikes are available at the commuters' preferred stations, they tend to use another station as less bikes are available. This observation allows us to conclude that commuters help to keep stations open for new users by choosing other stations. When incentives are introduced, commuters are more likely choose to alternate between stations when compared to cases with no incentives. That means that the commuters accept incentives, and incentives help commuters satisfy the needs of the shared system.

The limitation of our study is that we did not look into other factors outside of the load of the most preferred station to pick the bike (e.g. weather, or altruism). Future studies should also focus on the station and system perspective, which would allow accessing the applicability of the proposed models to study an individual's roles in the well-being of other shared infrastructure of modern cities.

References

[1] D. Freund, S. G. Henderson, and D. B. Shmoys, "Minimizing Multimodular Functions and Allocating Capacity in Bike-Sharing Systems," Springer, Cham, 2017, pp. 186–198.

[2] J. Schuijbroek, R. C. Hampshire, and W.-J. van Hoeve, "Inventory rebalancing and vehicle routing in bike sharing systems," *Eur. J. Oper. Res.*, vol. 257, no. 3, pp. 992–1004, Mar. 2017.

[3] P. C. Fishburn, "Utility theory for decision making". McLean VA: Research analysis corp, 1970.

[4] "Citi Bike System Data | Citi Bike NYC." [Online]. Available: https://www.citibikenyc.com/system-data. [Accessed: 20-Jun-2018].