

A Dynamic Threshold-Based Customer Redirection Rule for Minimizing Customer Transfer Costs in Bike-Sharing Systems

Simon Anthony D. Lorenzo, Lowell L. Lorenzo and Iris Ann G. Martinez
Department of Industrial Engineering and Operations Research, College of Engineering, University of the Philippines Diliman, Quezon City 1101, Philippines

Abstract — *This paper produces a new dynamic threshold-based customer redirection rule that minimizes customer extra effort in bike-sharing systems while maintaining a target service level. The proposed rule entails the redirection of both the customer's trip origin and destination following thresholds that are dynamically based on the stations' forecast supply and demand for bikes. A simulation model based on London's Santander Cycles bike-sharing system was used to evaluate the effectiveness of the proposed rule. Results show that the new rule was able to reduce the system average extra effort for customers by as much as 43.35% while maintaining the service level, over 100 replications of the simulation.*

Keywords — bike-sharing systems, customer redirection, extra effort minimization

I. INTRODUCTION

Bike-sharing is a new type of transportation system that is starting to develop around the world. It is a system where a customer rents a bike at a station, uses it to move to another station, then leaves the bike at a designated dock. It is a cheap, convenient and environment-friendly mode of transportation for commuters. Today, large-scale bike-sharing systems can be found major cities such as New York, Chicago, Montreal, Barcelona, London, Shanghai and Melbourne.

Since bike-sharing systems are services, it is central to maintain an acceptable service level across the system. In the context of these systems, service level is defined to be the ratio of number of events where services are fulfilled (i.e. a customer wanting to rent a bike and being able to) to the total number of events in the system (which includes instances where services are not fulfilled). Some bike-sharing systems (such as the Santander Cycles in London) are required by its country's government to maintain a specific service level. In general, maintaining a high service level is one of the main goals of a bike-sharing system operator.

However, the operation of bike-sharing systems is not without problems. It was found out by Fricker and Gast [1] that unregulated bike-sharing systems often experience low service levels, even in the ideal case where the system is balanced (the inflow of bikes is equal to its outflow, for each station). In order to address this problem, operators often implement control measures which aim to maintain a level of bikes in each station that would try to minimize the occurrence of both no-bike (a customer wanting to rent a bike but not being able to) and no-dock (a biking customer wanting to deposit a bike but not being able to) events. The common measures taken to address this problem are repositioning and customer redirection. With repositioning, an external entity (e.g. truck) manually moves bikes from one station to another. This is done when system activity is either negligible (called static repositioning) or at a high level (called dynamic repositioning).

An alternative to repositioning was introduced by Fricker and Gast [1]: customer redirection. This scheme, instead of utilizing trucks, involves asking customers to divert their routes and walk extra distances in order to avoid no-service events. In exchange, customers are usually offered incentives such as discounts on payment for future rentals. It was found out that service levels similar to those achieved through repositioning are attained even through customer redirection alone [1]. The implementation of customer redirection, however, brings difficulties. This scheme asks customers to divert their routes, causing them to walk distances they would normally not need to. This extra distance walked by the customers may lead to dissatisfaction, and may affect future interactions with the bike-sharing system. It should be emphasized that customers use bike-sharing systems with the intention of travelling intermediate distances; their destinations are often beyond walking distance resulting in the further need for bikes, so it is logical to reduce customer redirection as much as possible. This brings forth the focus of the study: minimizing the negative impact customer redirection has on bike-sharing system customers due to the extra effort being asked from them.

II. LITERATURE REVIEW

The following table summarizes the literature that deal with the optimization of measures taken to reduce bike-sharing system service level:

Table 1. Summary of Related Literature in bike-sharing systems.

Author	Description
Schuijbroek et al. [2] Rainer-Harbach & Papazek [3] Benchimol et al. [4] Hernández-Pérez et al.[5] Hernández-Pérez et al. [6] Huang et al. [7]	The earliest studies that dealt with the rebalancing problem of bike-sharing systems solved the static repositioning problem. These papers came up with heuristics to provide solutions to variants of the static repositioning bike-sharing model.

Author	Description
Ghosh et al. [8] Contardo et al. [9] Caggiani & Ottomanelli [10] Pfrommer et al. [11] Parhizkar et al. [12]	Following the wave of static repositioning papers are those that dealt with dynamic repositioning. These papers provided solutions for truck routing in dynamic repositioning.
Waserhole et al. [13] Chemla et al. [14] Bam et al. [15] Singla et al. [16] Aeschbach et al. [17]	These papers discuss the customer redirection solution to the bike-sharing system balancing problem. The papers focus on different incentive schemes to attract customers to conform to customer redirection.

The earliest studies that dealt with the rebalancing problem of bike-sharing systems solved the static repositioning problem. Hernandez-Perez et al. [6] was the first to tackel this, coining the term one-commodity pickup-and-delivery problem to describe this specific problem. They proved it to be NP hard, promoting studies that followed to develop heuristics. Numerous studies, such as Schuijbroek et al. [2], Rainer-Harbach & Papazek [3], and Benchimol et al. [4] developed heuristics for different variants of the NP-hard problem. This trend has been continuing up until the present, where studies such as Huang et al. [7] discussion solutions to the static repositioning model through a random forest model.

Following the wave of static repositioning papers are those that dealt with dynamic repositioning. These papers again solved for the optimal truck routes by developing heuristics. The papers that dealt with providing solutions to the dynamic repositioning problem are Ghosh et al. [8], Contardo et al. [9], Caggiani et al. [10], and Pfrommer et al. [11]. Pfrommer et al. [11] was the first paper to come up with a scheme that combines both repositioning and customer redirection. In their paper, the authors discussed the optimal dynamic bike repositioning scheme for the then Barclay's Cycle Hire in London, as well as a customer redirection scheme that focused on what incentives to give to customers in order to regulate the fill levels of the stations in the system. The study also made use of a utility plateau concept, which says that there is always an optimal fill level range which should be targeted for each station. This utility plateau was the basis of both the dynamic repositioning and customer redirection schemes proposed by the study.

The least explored bike-sharing system regulation measure among the three discussed is customer redirection. The first paper that dealt with the subject was done by Waserhole et al. [13], which tried to formulate a function that gives the probability that a customer will go along with redirection given varying degrees of incentives. It was found out that a higher incentive would increase the chance that a customer would go along with the redirection. Studies such as Ban et al. [15] and Singla et al. [16] incorporated the same procedure of investigating the relationship between the incentive and customer participation in redirection.

Aesbach [17] decouples the financial aspect of the operation of bike-sharing systems by assigning a customer cooperation factor, c, that gives the ratio of proportion of customers that are willing to cooperate with rerouting schemes, whatever pricing policy is used. The study, instead, focused on implementing four different redirection policies (i.e. what percentage of customers to redirect and under which specific system conditions) in order to maximize the service level of the system.

As discussed, most previous studies dealt with repositioning. As this requires the additional resource of an external entity to move bikes among different stations, the contribution of this paper is to not require external resource but on the voluntary movement of customers through customer redirection. Within customer redirection, the paper moves away from the exploration of incentive schemes (which many have already explored) and aims to address the extra effort incurred by customers due to redirection.

III. RESEARCH METHODOLOGY

To be able to address the illustrated gap, the study is set to, firstly, define clearly what the problem is. This is done through the illustration of the bike-sharing system as a network model, then through the mathematical formulation of the model. Given the formulation, the appropriate tool is used and an optimal solution is solved, if possible. If the model will not permit a globally optimal solution, a local optimal solution shall suffice (or at least, a solution better than the current best). Upon getting the solution, it is validated against a real-life bike-sharing system to ensure that the solution is both theoretical and actual applicable.

IV. PROBLEM DESCRIPTION AND MODEL FORMULATION

4.1 Problem Description

For the case of the bike-sharing system, the problem can be summarized as follows: as Figure 1 shows, given that the bike-sharing system has stations (nodes) that are have distances (d_{ij}) distances between them, arrival rate of customers or users of the bike (λ_i) for station i, and the station capacity (boxes) and the currently-available bike level (shaded boxes), to which station should the customer be redirected to strike a balance between the overall bicycle system service level and the amount of extra cost (effort) of the customer for transferring from one station to another.

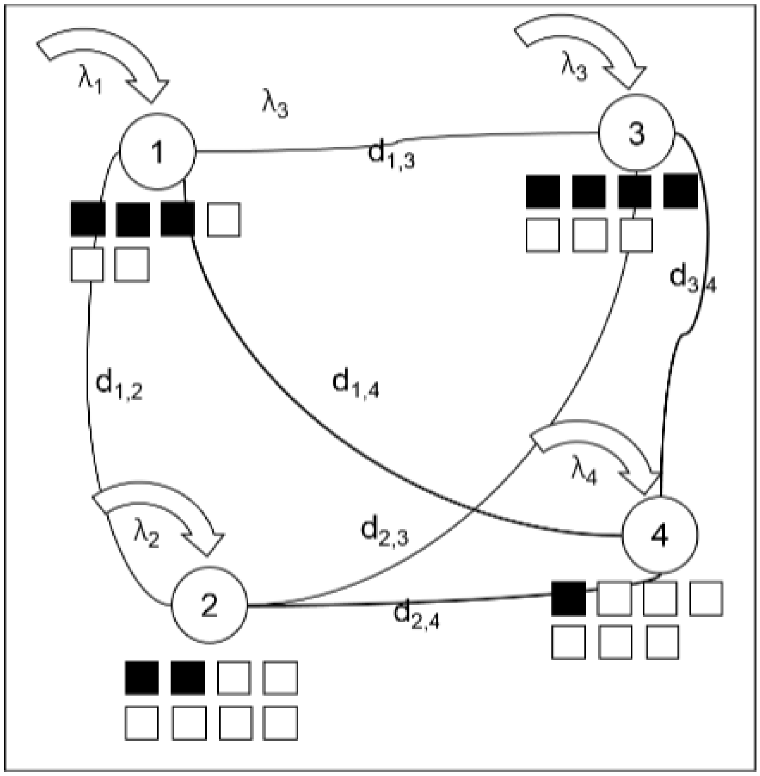

Figure 1. Illustration of the bike-sharing system model.

Increasing the overall system level is the metric that is used to evaluate the performance of the bike-sharing system. This is equal to the ratio of all the service events (incident) to all events. Mathematically,

$$Service\ level = \frac{(bike\ events + dock\ events)_t}{(bike\ events + dock\ events + no\ bike\ events + no\ dock\ events)_t} \qquad \text{(Eqn. 1)}$$

Furthermore, extra cost (effort) to the customer is defined as the distance the customer walks due to redirection. It can be incurred both at the redirection of the origin and the destination. The average extra effort across the system can then be computed using the following equation:

$$Average\ extra\ effort = \frac{origin\ redirection\ distance + destination\ redirection\ distance}{Total\ cooperative\ customers\ that\ entered\ the\ bike\text{-}sharing\ system} \qquad \text{(Eqn. 2)}$$

4.2 Mathematical Model

4.2.1 Overview

This study models the problem stated above using a transient analysis model to capture the bike-share queuing system before it reaches its steady state. Other studies that have focused on solving bike-share system problems, particularly those that look at the network of the bike-share system, have modeled the system as a closed queuing system. In modeling using the closed queuing system, each station is represented as a server to bike share customers who go around through the network. The limitation with modeling using closed queuing network systems is it assumes steady-state behavior. Specifically, the problem that this study solves concerns immediate queue length and waiting time of the customer and these are not captured by the steady-state characteristic of a network model. Hence, this study uses transient analysis to look at the queuing system before steady state of the bike-share system is reached.

4.2.2 Model

This study models the bike-share system that maximizes service level while retaining a certain service level as follows:

$$Min \sum_{i=1, i\neq j}^{m} \sum_{j=1}^{m} \sum_{k=1, k\neq l}^{m} \sum_{l=1}^{m} R_{ijkl}^{s}(D_{ik} + D_{lj})\lambda_i^s P_{ij}^s$$

s.t.

$$1 - \sum_{s=1}^{t}\sum_{i=1}^{m} \frac{\int_0^T \left(\pi_{i,b_i^s,0}^{s}(t)\lambda_i^s + \pi_{i,b_i^s,C_i}^{s}(t)\mu_i^s\right)dt}{\int_0^T \left(\pi_{i,b_i^s,0}^{s}(t)\lambda_i^s + \pi_{i,b_i^s,C_i}^{s}(t)\mu_i^s\right)dt + T(\lambda_i^s + \mu_i^s)} \geq SL_{target} \quad \text{(Constraint 1)}$$

$$b_i^s = min(max(b_i^{s-1} + \eta_i^s, C_i), 0), for\ s = 1, 2, 3, \dots, t \quad \text{(Constraint 2)}$$

$$\sum_{k=1, k\neq l}^{m} \sum_{l=1}^{m} R_{ijkl}^{s} = 1\ for\ s = 1,2,3, \dots, t; l = 1,2,3, \dots, m; k = 1,2,3, \dots, m, k \neq l \quad \text{(Constraint 3)}$$

$$R_{ijkl}^{s} \leq 1, for\ s = 1, 2, 3, \dots, t \quad \text{(Constraint 4)}$$

$$R_{ijkl}^{s} \geq 0, \forall\ i, j, k, l, s \quad \text{(Constraint 5)}$$

where:

Table 2. System parameters present in the mathematical formulation.

Variable	Description
R_{ijkl}^{s}	Proportion of customers intending to take a trip from i *(i = 1, 2, …, m)* to j *(j = 1, 2, …, m)* but was redirected to a trip from k *(k = 1, 2, …, m)* to l *(l = 1, 2, …, m)* during time slice s *(s = 1, 2, …, t)*
D_{ij}	Distance from station i *(i = 1, 2, …, m)*to station j *(j = 1, 2, …, m)*
P_{ij}^{s}	Probability of a customer originating from station i *(i = 1, 2, …, m)* to go to station j *(j = 1, 2, …, m)* during time slice s *(s = 1, 2, …, t)*
t	Number of time slices in the time horizon
m	Number of bike stations in the system
$\pi_{i,j,k}^{s}(t)$	The probability that station i *(i = 1, 2, …, m)* has a bike level of k *(k = 1, 2, …, m)* after t time units, given that it started with j bikes
λ_i^s	Mean bike arrival rate at station i *(i = 1, 2, …, m)* during time slice s *(s = 1, 2, …, t)*

Variable	Description
μ_i^s	Mean bike departure rate at station *i (i = 1, 2, …, m)* during time slice *s (s = 1, 2, …, t)*
T	Time horizon of the bike-sharing system operation
b_i^s	Bike level of station *i (i = 1, 2, …, m)* at the end of time slice *s (s = 1, 2, …, t)*
SL_{target}	Target service level of the system
b_i^0	Initial bike level of station *i (i = 1, 2, …, m)*
η_i^s	Net change in the bike level of station *i (i = 1, 2, …, m)* before and after time slice *s (s = 1, 2, …, t)*
C_i	Capacity of station *i (i = 1, 2, …, m)*

The lone decision variable R_{ijkl}^s gives the proportion of customers intending to take a trip from station *i* to station *j* but were redirected to a trip from station *k* to station *l* during time slice *s*. The objective function makes use of this decision variable to compute for the extra effort by multiplying the proportion of customers redirected to each possible trip (R_{ijkl}^s), the extra distance incurred due to redirection ($D_{ik} + D_{lj}$), and the expected number of customers taking the original trip from *i* to *j* $\lambda_i^s P_{ij}^s$, for each time slice s. The quadruple summation considers all possible combinations for the original origin and destination stations, as well as the new origin and destination stations. These summations are also defined such that the origin station is never equal to the destination station.

Constraint 1 expresses the total number of no-service events in the numerator for the fraction. This entire expression is directly taken from the mathematical formulation of Raviv and Kolka [13]. It is then divided by the total number of events to get the proportion of no-service events to the total. Finally, it is subtracted from 1 to get the service level, which is set to be greater than or equal to the target service level, SL_{target}.

Constraint 2 makes sure that the bike level of each station follows the value of both the bike arrival rate, λ_i^s, and the bike departure rate, μ_i^s for all time slices. This continuity constrain equates the bike level of a station at time slice s (b_i^s) and makes sure it is equal to the bike level of the previous time slice plus the change in bike level ($min(max(b_i^{s-1} + \eta_i^s)$. The constrain is formulated such that the bike level will not exceed the capacity of the station and will not be less than zero.

Constraint 3 ensures that the proportion of all redirected trips add up to 1, as proportions should by definition.

Constraints 4 and 5 ensure that the proportion for redirection cannot exceed 1, and cannot be negative at the same time.

Since Constraint 1 deals with transient analysis, this paper uses appropriate tools from other studies. Leguesdron et al. [14] provides equations for the probability of a future state given the present state. These equations are modified to fit the notation used in this study, thus producing the following equations to solve for the value of the probability that for a certain station *i*, the state (i.e. bike level) has a value of *k* after a time period *t* given that it started with a value of *j* all within a slice *s*, denoted by the variable $\pi^{s}_{i,j,k}(t)$:

$$\pi^{s}_{i,j,k}(\mathrm{t}) = e^{-(\lambda^{s}_{i}+\mu^{s}_{i})t}\left(\frac{\lambda^{s}_{i}}{\mu^{s}_{i}}\right)^{\frac{k-j}{2}}[I_{k-1}\left(2t\sqrt{\lambda^{s}_{i}\mu^{s}_{i}}\right) - I_{j+k+2}\left(2t\sqrt{\lambda^{s}_{i}\mu^{s}_{i}}\right) + \left(\frac{\mu^{s}_{i}}{\lambda^{s}_{i}}\right)^{i+1}\pi_{i,0,j+k+1}(t)$$

(Eqn. 3)

$$I_k(x) = \sum_{m=0}^{+\infty}\frac{(\frac{x}{2})^{2m+k}}{m!\,(k+m)!} \qquad \text{(Eqn. 4)}$$

$$p^{s}_{i} = \frac{\lambda^{s}_{i}}{\lambda^{s}_{i}+\mu^{s}_{i}} \qquad \text{(Eqn. 5)}$$

$$q^{s}_{i} = \frac{\mu^{s}_{i}}{\lambda^{s}_{i}+\mu^{s}_{i}} \qquad \text{(Eqn. 6)}$$

In Equation 4, $I_k(x)$ is defined to be a modified Bessel function of the first kind, which Leguesdron et al. [14] used to solve the Bessel differential equations that were involved in deriving an expression for the transient probabilities of the M/M/1 queue.

4.2.3 Model Assumptions

A simplifying assumption is that the quantities λ^{s}_{i} and μ^{s}_{i} are independent of each other. In actual bike-sharing systems, these two quantities may be related but for ease of computation, it is assumed that each one has its own value/distribution.

The variable η^{s}_{i} is defined to be the net change in bike level of station i at time slice s. To get the value of η^{s}_{i}, the following equation is used:

$$\eta^{s}_{i} = \lambda^{s}_{i} - \mu^{s}_{i} \qquad \text{(Eqn. 7)}$$

This equation is merely an approximation of the change in bikes, since this equation loses accuracy once no-service events are involved. However, one important characteristic validates this approximation: service levels of bike-sharing systems involving repositioning and redirection are generally high, meaning the occurrence of no-service events will likely be negligible.

V.PROPOSED SOLUTION

5.1 Weak Preemptive Control - Mobile App

Of all studies reviewed, none have attempted to minimize extra effort, and only Aeschbach et al. [10] attempted to record it. Consumer psychology is too broad a topic on its own, so tackling this would hinder discussion on extra effort minimization. Thus, Aeschbach et al. [10] decouples the psychological aspect of the operation of bike-sharing systems by assigning a customer cooperation factor, c, that gives the ratio of proportion of customers who are willing to cooperate with rerouting schemes, whatever pricing policy is used. This factor allows the study to accounting for varying levels of customer acceptance, giving way for customer extra effort analysis.

As previously mentioned, the study proposed four different redirection policies (i.e. what percentage of customers to redirect and under which specific system conditions). Each of these redirection policies has a different set of parameters upon implementation, and thus has a different effect on the customers of the system. Due to the complexity of the stochastic nature of the system, instead of attempting to solve the problem through linear/non-linear programming, this study opted to use simulation to test their proposed rule. The simulation conducted used data from the Santander Cycles (then Barclays Cycle Hire) bike-sharing system, including 5,110,650 bicycle trips made from January 5, 2014 up to July 19, 2014, with 745 bicycle stations and 9,109 bikes. Two new concepts were introduced: the previously mentioned customer cooperation factor and the concept of a neighborhood (all stations within a predefined neighborhood radius, r).

The study proposed four rules for customers to follow: No Control (NC), Minimal Intervention Control (MIC), Preemptive Control – Mobile App (PC) and Weak Preemptive Control – Mobile App (WPC). The description of each rule is as follows:

Table 3. Redirection rules presented by Aeschbach et al. (2015).

Name of Rule	Description
No Control (NC)	No redirection is implemented
Minimal Intervention Control (MIC)	If a no-bike event occurs, the customer is redirected to the station with the highest fill level in neighborhood. If a no-dock events occurs, the customer is redirected to the station with the lowest fill level within the neighborhood.
Preemptive Control – Mobile App (PC-MA)	Upon identification of the intended trip, origin is redirected to highest fill level in neighborhood and destination is redirected to lowest fill level in neighborhood
Weak Preemptive Control – Mobile App (WPC-MA)	Similar to PC-MA, but origin redirection occurs only if the origin has a fill level below a fixed threshold of 50%, and destination redirection occurs only if the destination has a fill level above a fixed threshold of 50%. The trip origin is redirected to the nearest station within the neighborhood with a fill-level above 50%, while the trip destination is redirected to the nearest station within the neighborhood with a fill-level below 50%.

For a specific value of c, both the WPC and PC rule had a service level of 98%, but the system average extra effort recorded by the PC rule was double that of the WPC rule. This means that some rules are strictly better than others with respect to service level and extra effort.

The WPC-MA rule involved redirecting the trip origin and trip destination of a customer when its fill level is below 50% and above 50%, respectively. Henceforth, these quantities shall be referred to as the Origin Redirection Limit (ORL) and Destination Redirection Limit (DRL) respectively. The origin is then redirected to the nearest station within a radius r that has a fill level that is greater than the ORL, and the destination is redirected to the nearest station within a radius r that has a fill level that is less than the DRL. The idea of this redirection is to avoid renting bikes at stations which have a low fill level, and avoid docking bikes at stations with high fill levels. The authors do not explicitly state the reason for setting both the ORT and DRT to 50%. However, it can be inferred that the rule tries to stay away from 0% and 100% fill levels, since those are the only sources of no-service events.

5.2 Utility of Changes

5.2.1 Definition of Utility

The current best rule for redirection uses a fixed value of 50% for both the ORL and DRL. However, Ruch et al. [15] explicitly states that setting a 50% target fill level for each station is not optimal, as it does not consider the arrival and departure patterns of each station in the near future. The latter's study proposes a procedure that considers the utility of any change done to a station's bike level in order to assess the optimal repositioning/redirection scheme to be implemented. The utility any given bike level of a station is defined to be the number of customers expected to be served over a finite time-horizon (called the look-ahead period). The optimal bike level for each station at any given point in time can be computed by maximizing the utility. However, this exhaustive method increases the complexity of the computations involved in the model significantly, bringing forth the need for a key simplifying assumption: arrivals are deterministic within a time slice, and are determined by historical arrivals. This allows for the solving of the utility of a finite number of system states, while still considering the expected demand for bikes/docks in the near future.

5.2.2 The Utility Plateau

Through the considerations in the previous section, it can be shown that for any set of values for the demand, there always exists two values, henceforth called the Utility Plateau Lower Limit (UPLL) and Utility Plateau Upper Limit (UPUL), for the station bike level b, where the utility inclusively between the two values is at its highest. The utility then decreases linearly when moving outside these two values. To illustrate this concept, a plot of the bike level (b, on the x-axis) versus its utility (U(b,i,s), on the y-axis) for any bike station i is given. The graph will always be a piecewise function illustrated by Figure 2, the slope of which is defined by the following equation:

$$\frac{d}{db}U(b,i,s) = \begin{cases} 1, & if\ b < UPLL \\ 0, & if\ UPLL \leq b \leq UPUL \\ -1, & if\ UPLL < b \end{cases} \qquad \text{(Eqn. 8)}$$

Where:
b is the bike level
i is the station number
s is the time slice

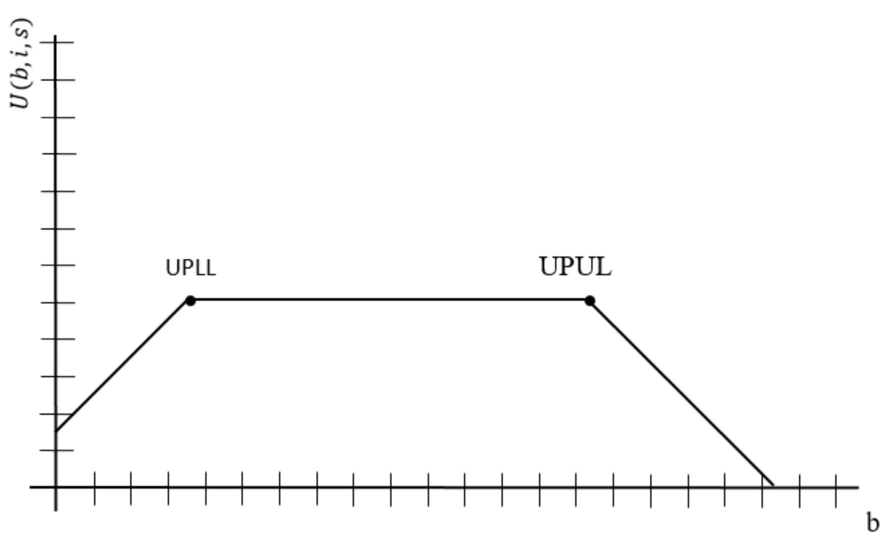

Figure 2. Plot of utility vs. bike level.

Since the concept of the utility plateau is central to the rule to be proposed in this study, it should be emphasized that it has already been proven by Ruch et al. [15] that the two values for UPUL and UPLL always exist, and that the maximum utility occurs for a bike level b, where UPLL ≤ b ≤ UPUL.

The concept of the utility plateau is used as the guide on how to redirect customers (or in the case of Pfrommer et al. [16], reposition bikes) in ways that will be favorable to the system's service level. For this study, the utility plateau will be used to determine a dynamic set of values for the bike level as targets for repositioning, for each station and for each time slice. This is seen as a more adaptive approach compared to the one used by Aeschbach [10] which uses a fixed ORL and DRL of 50%, since it is computed for each station and each time slice, and it considers the predicted demand for bikes/docks that the station will experience.

5.3 Proposed Rule

The proposed rule is adapted from the WPC-MA redirection rule of Aeschbach [10] but the fixed value of 50% for the ORL and DRL will be replaced with a value obtained through computations for the UPLL and UPUL. For the new rule, two quantities are computed to guide in redirection: the Effective Origin Threshold (EOT) and the Effective Destination Threshold (EDT). Since these two quantities are dependent on parameters whose values change throughout different time slices, there is a need to compute for the value of the EOT and EDT for each station and each time slice. Hence, the following variables that will be added to the simulation model are formally introduced:

Table 4. Variables to be used in defining the Effective Origin and Destination Thresholds.

Variable	Description
EOT_i^s	Effective origin threshold of station *i (i = 1, 2, …, m)* at time slice *s (s = 1, 2, …,t).*
EDT_i^s	Effective destination threshold of station *i (i = 1, 2, …, m)* at time slice *s (s = 1, 2, …, t).*
β	Bike level buffer
$UPLL_i^s$	Utility plateau lower limit of station *i (i = 1, 2, …, m)* at time slice *s (s = 1, 2, …, t).*
$UPUL_i^s$	Utility plateau upper limit of station *i (i = 1, 2, …, m)* at time slice *s (s = 1, 2, …, t)*
ORL	Origin redirection limit. This value is fixed at 50%.
DRL	Destination redirection limit. This value is fixed at 50%.

Six of the variables here have already been previously discussed. Four of these, namely the EOT_i^s, EDT_i^s, $UPLL_i^s$ and $UPUL_i^s$ are specified for each existing station per time slice. The ORL and DRL are both fixed at 50%, since these were the values of the best existing rule (WPC-MA). The only new variable to be used is β, the bike level buffer. This buffer makes sure that the target bike levels of the rule to be proposed will never be set to 0 or filled to the capacity. These bike levels are to be strongly avoided since these will be the only sources of no-service events. A more thorough discussion on the consideration of the buffer will be done later on.

The variables EOT_i^s and EDT_i^s are then computed through the following formulas:

$$EOT_i^s = max(min(UPLL, ORL), \beta) \quad \text{(Eqn. 8)}$$
$$EDT_i^s = min(max(UPUL, DRL), C_i - \beta) \quad \text{(Eqn. 9)}$$

The new hybrid rule for redirection is as follows:

> *A cooperative customer is to be redirected to the nearest station within the neighborhood radius from the origin station that has a fill level that is greater than or equal to the EOT, unless the origin station itself has a fill level that is greater than or equal to the EOT. The cooperative customer's destination is then redirected to the nearest station within the neighborhood radius from the destination station that has a fill level that is less than or equal to the EDT, unless the destination station itself has a fill level that is less than or equal to the EDT.*

The hybrid rule makes use of the original 50% *ORL* and *DRL* used by Fricker and Gast [1] in order to ensure a high service level for the system. In order to reduce the average customer extra effort of the system implementing the rule, the *UPLL* and *UPUL* of each system are considered, alongside the *ORL* and *DRL*.

For the origin redirection, WPC-MA checks for a minimum fill level of 50% since the system wants to avoid stations that are "too empty" and reduce the chance of a no-bike event occurring. While this may seem true, the utilization concept tells us that a 50% fill level is not always ideal when talking about service level. In fact, there is a possibility that the best fill level of a station is 0% (this occurs when the bike inflow of the station is very high). Thus, the 50% *ORL* is used in conjunction with the *UPLL*, in order to establish an effective lower limit, the *EOT*.

One last consideration for the *EOT* is β, the bike level buffer, for which the system accounts for improbable arrivals in the station that the *UPLL/UPUL* assumptions do not account for. Without the buffer, it is possible that the value for *EOT* is 0. A simple example would be a station that has a η_i^s that is very high for all values of the time slice *s*. This would lead to an *EOT* of 0 since the rule would like an empty station to accommodate the high inflow of bikes. This, however, opens up the station to a no-bike event should a customer go to the station and look for a bike. To avoid this, a lower limit for the UPLL equal to β is set. Likewise, it is also possible for *EDT* to be equal to the station capacity. The *EDT* of any station *i* will be capped at a value equal to $C_i - \beta$ in an effort to completely avoid no-dock events from full bike stations.

As the *UPLL* may recommend a very low fill levels for bike stations (possible 0%, as stated above), there is a chance that the next customer would look for a bike to ride, thus incurring a no-bike event. This is one of the weaknesses of the *UPLL* (and *UPUL*): it assumes that bike changes are deterministic. The buffer quantity makes sure that this type of occurrence is minimized. Considering all these, the previous expression for the *EOT* is obtained. A similar justification can be made for the *EDT*, which now considers the *DRL* and the *UPUL* in its computations.

This rule is expected to perform better than WPC-MA of Aeschbach et al. [10] since instead of setting a fixed threshold of 50% for redirection, the expected change of bikes in each station is now considered. This is seen to lead to less unnecessary customer redirection, particularly in stations that have an optimal fill level that is far from 50%.

Figure 3 gives an illustration on how to implement the new hybrid rule. Given a customer that wants to travel from station 2 to station 4. After hypothetically determining the EOT and EDT of stations 2 and 4 respectively, the customer's origin will likely be redirected to 1 and the customer's destination will likely be redirected to 4. This is to avoid station 2 being empty and station 5 being full, and at the same time this prevents station 1 from becoming full and station 5 from becoming empty, thus avoiding no-service events.

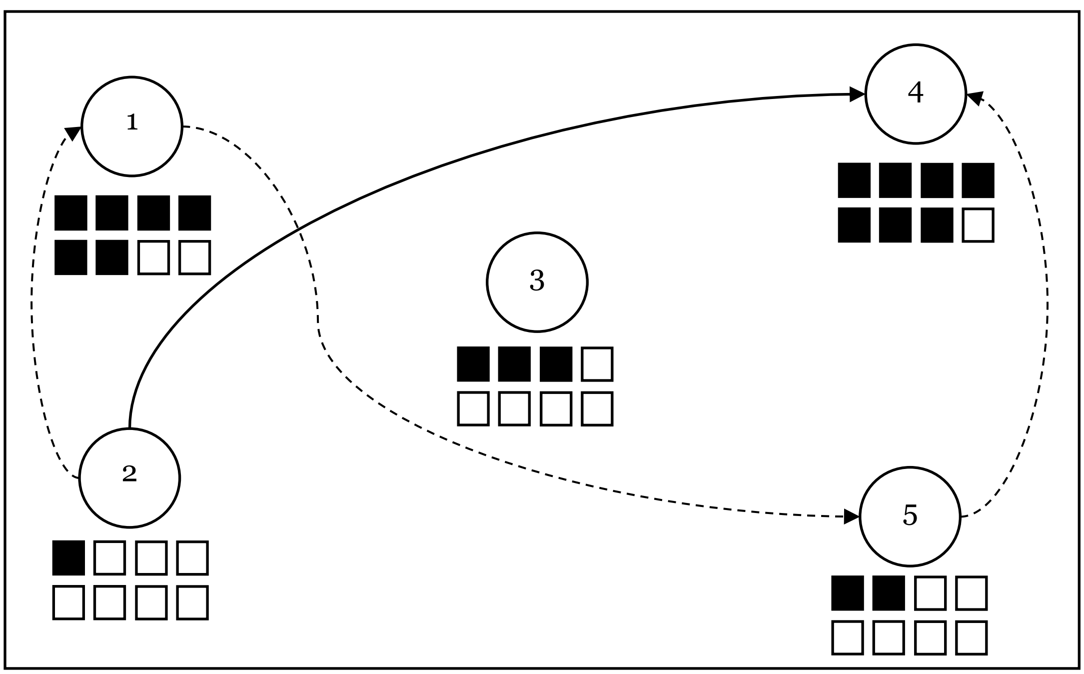

Figure 3. Illustration of new hybrid rule redirection.

VI. Validation

6.1 Simulation as Validation

Going back to the model formulation, the following expression is obtained from the first constraint of the mathematical model:

$$\int_{0}^{T} \left(\pi_{i,b_i^s,0}^{s}(t)\lambda_i^s + \pi_{i,b_i^s,C_i}^{s}(t)\mu_i^s\right) dt$$

As mentioned above, this expression in the first constraint was lifted from Raviv and Kolka [13]. Of this expression to evaluate the expected number of unsatisfied customers, Raviv and Kolka [13] says the following: *"The computational challenge in the evaluation of [the equation] is the calculation of the transition matrix π(t). There is no closed-form solution for the dynamics of this system." (Optimal inventory management of a bike-sharing station, Raviv, Kolka, 2013, p. 1081).*

Since the main constraint of the problem to be solved has a mathematically intractable solution, it would not be possible to solve for the global optimal solution of the NLP. This then calls for means such as simulation to be used for validation. Furthermore, simulation is used in majority of studies that deal with bike-sharing systems.

6.2 Simulation Model

The simulator used is OMNeT++ 5.2, an open source discrete event simulator that is based on the C++ programming language.

The simulation model is based on the Santander Cycles (prev. Barclays Cycle Hire) bike-sharing system in London. The model will make use of the publicly available data provided by the Transport for London (TfL), the government body that oversees Greater London's public transportation system. It should be made clear that the simulation does not intend to recreate the actual events that transpired within the specified dates of the raw data obtained. Rather, it would imitate the behavior of the system by adapting the statistical distributions of the different parameters of the system.

So as to be at par with Aeschbach et al. [10], the model will be using the exact same data set that was used in the said study, which claims to include 745 stations, and 5,110,650 trips of 9,109 bicycles dated from January 5, 2014 – July 19, 2014. The data retrieved is stored in seven *.csv* (comma separated values) files. Each file contains a table with nine headers: *Rental ID, Trip Duration, Bike ID, Trip End Time, Destination Station Number, Destination Station Name, Start Time, Origin Station Number and Origin Station Name*. The raw data was inspected and the actual data obtained was compared to the data described by Aeschbach [10], as summarized in the following table:

Table 5. Comparison of data retrieved from TfL to data used by Aeschbach et al. (2013).

Parameter	Value Retrieved from Transport for London website	Value used by Aesbach et al.	Deviation of Parameter
Data range of bicycle trips	January 5, 2014 – July 19, 2014	January 5, 2014 – July 19, 2014	None
Number of stations	755	745	1.34%
Number of Bicycle Trips	5,126,825	5,110,650	0.32%
Number of Bicycles	11,882	9,109	30.44%

The deviations between the data obtained are deemed to be negligible apart from the number of bicycles. Since the current best rule was developed using a model with 9,109 bikes, this value will be used for the simulation of the proposed rule as well.

Upon obtaining and inspecting the raw data, it was adjusted to match the assumptions of Aeschbach et al. [10] and be more illustrative of how a real-world bike-sharing system works. The following measures were done, in the given order, to clean the raw data obtained:

To clean the raw data, weekend trips were first excluded from the data set since there is a different behavior of bike and customer movement during weekends. Trips that ended at its origin were then removed from the data set, since these trips do not represent the intended purpose of bike-sharing systems. Finally, outliers (with respect to trip duration) were removed using the outlier removal functionality of Minitab 15 (points outside the Q_1-1.5*IQR to Q_3+1.5*IQR range were removed).

The data set to be used is now trimmed down from an original value of 5,126,825 trips to 3,660,304. From this processed data set, simulation parameters are computed.

Three types of parameters are identified: fixed, time-slice dependent and time-dependent parameters. Fixed parameters are set prior the simulation runs and do not change all throughout. Time-slice dependent parameters are identified to capture the periodicity of the system behavior. A 24-hour day is sliced into 8, with the first slice ranging from 12 MN – 3:00 AM. Lastly, time-dependent parameters freely change anytime during the simulation run. The following table summarizes all parameters used in the simulation run:

Table 6. Different parameters used in the simulation model.

Parameter	Description/Remarks	Type
Station Capacity, C_i	Obtained from a separate file from the same source; set to 749 throughout the entirety of the simulation to match the station capacity parameter of [17].	Fixed
Customer cooperation factor, c	The proportion of all customers that are willing to undergo redirection; to be set to three values (0.2, 0.5, 0.9) covering low, medium and high scenarios, respectively.	Fixed
Neighborhood radius, r	The distance by which the neighbors of a certain station are determined; to be set to three values (600m, 850, 1200m) covering low, medium and high scenarios, respectively.	Fixed
Inter-station bike times, β_{ij}^{s}	the mean travel time from station *i* to station *j* during time slice *s*.	Time-slice Dependent
Inter-station walk times, ω_{ij}^{s}	the mean walking time from station *i* to station *j* during time slice *s*.	Time-slice Dependent
Destination Probability, P_{ij}^{s}	the probability of station *j* being the destination for a given trip originating from station *i* during time slice *s*.	Time-slice Dependent
Bike inflow rate, λ_i^s	the number of bikes expected to arrive at station *i* (regardless of availability of docks) during time slice *s*.	Time-slice Dependent
Bike outflow rate, μ_i^s	the number of bikes expected to depart from station *i* (regardless of destination) during time slice *s*.	Time-slice Dependent
Net bike flow rate, η_i^s	the net change in bikes in station *i* during time slice *s*.	Time-slice Dependent
Station bike level, b_i^t	The bike level of station *s* at time *t*.	Time Dependent
Station fill level, f_i^t	The fill level of station *s* at time *t*.	Time Dependent

This entire set of parameters will be ran twice, once using the current best WPC-MA from Aeschbach et al. [10], and another time using the proposed rule. The results will then be compared.

To ensure the repeatability of the proposed rule, the results of each simulation run will be averaged over one hundred repetitions each. Taking into consideration the fact that there are three levels for *c* and *r*, there is a total of 3*3*100 = 900 simulation runs for each of the two redirection schemes. This is deemed to be enough since Aeschbach et al. [10] also used 100 replicates for each of their simulation runs.

To be able to manage the more complicated aspects of the system, some simplifying assumptions are made. Firstly, since the study focus solely on distances walked by the customer, the average customer extra effort does not include the variation in biking distance due to redirection. Regarding the bike levels of stations at the start of a simulation run, the initial distribution of bikes is assumed to be uniform throughout the stations of the system, since no distribution of bikes is given. For customer arrival distributions, all are assumed to be Poisson distributed. On time slices, model parameters would not change within a time slice (three hours). If a trip starts and ends at a different slice, the trip parameters will still follow the values from the time slice at the start of the trip. Lastly, the number of bikes in the system is constant throughout the simulation (no bike theft or failure is considered).

6.3 Results Discussion

As mentioned above, the study considered two simulation models (one for each rule) each with 9 different combinations of parameters, and each being run 100 times. The following table summarizes the results of all runs (both service level and extra effort are shown):

Table 7. Summary of results for the two tested redirection schemes.

c value	*r* value	WPC-MA Service Level	WPC-MA Average Customer Extra Effort	Hybrid Rule Service Level	Hybrid Rule Average Customer Extra Effort
0.2	600 m	0.741	371.77	0.710	267.25
0.2	850 m	0.707	487.03	0.697	381.42
0.2	1200 m	0.741	639.68	0.723	557.08
0.5	600 m	0.721	371.42	0.735	294.60
0.5	850 m	0.769	653.69	0.772	539.34
0.5	1200 m	0.742	656.68	0.798	568.16
0.9	600 m	0.785	423.38	0.821	413.11
0.9	850 m	0.882	750.27	0.832	652.85
0.9	1200 m	0.963	858.78	0.939	827.07

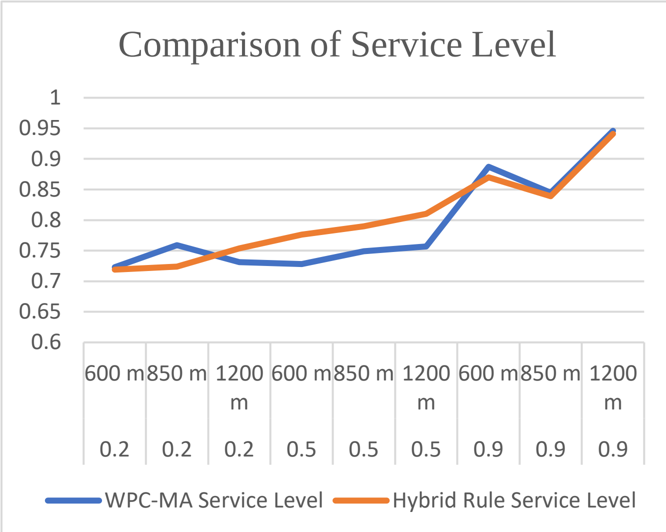

Figure 4. Comparison of Service Level.

Figure 5. Comparison of Extra Effort.

As a point of comparison, the average service level obtained when no rule is implemented is a significantly lower 0.186859. This shows that both these rules have a large contribution to the high service level of the system.

A set t-tests is done on the data collected through the multiple runs of each configuration. Two tests are run: one to ensure the retention (or improvement) of the service level and another to prove that the average customer extra effort was reduced.

The results of the t-tests validate that the simulation runs indeed prove that the new rule has an average customer extra effort that is significantly lower than that of the WPC-MA rule without worsening the service level, for all tested values of c (0.5, 0.75. 0.9) and r (600 m, 850 m, 1200 m). As mentioned above, this conclusion has been backed by statistical tests.

VII. SENSITIVITY ANALYSIS AND OTHER DISCUSSIONS

7.1 Sensitivity Analysis

One of the main assumptions of this study, as well as Fricker and Gast [1], where WPC-MA was used, is that customer inter-arrival times are exponentially distributed, leading to Poisson distributed arrivals. In order to test the robustness of the hybrid rule, a comparison is done on the service levels and extra efforts between the two rules for inter-arrival time distributions other than exponential. Simulation runs will be compared using a normally distributed IAT and an IAT following the Weibull distribution. For Weibull, two scenarios were simulated: $\beta = 0.5$ (decreasing arrival rate over time) and $\beta = 2$ (increasing arrival rate over time). The same statistical tests were done to validate the results.

Table 8. Summary of results for the two tested redirection schemes under normally distributed IATs.

c value	*r* value	WPC-MA Service Level	WPC-MA Average Customer Extra Effort	Hybrid Rule Service Level	Hybrid Rule Average Customer Extra Effort
0.2	600 m	0.723	424.91	0.719	247.52
0.2	850 m	0.759	634.71	0.724	437.90
0.2	1200 m	0.731	574.83	0.754	511.59
0.5	600 m	0.728	380.64	0.776	354.69
0.5	850 m	0.749	512.42	0.790	545.12
0.5	1200 m	0.757	643.46	0.810	618.01
0.9	600 m	0.887	537.44	0.870	484.68
0.9	850 m	0.845	718.29	0.839	648.75
0.9	1200 m	0.946	807.57	0.941	765.68

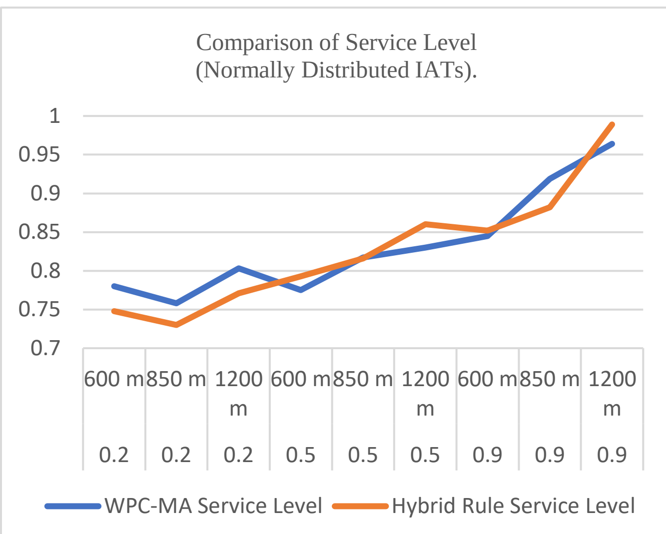

Figure 6. Comparison of Service Level (Normally Distributed IATs).

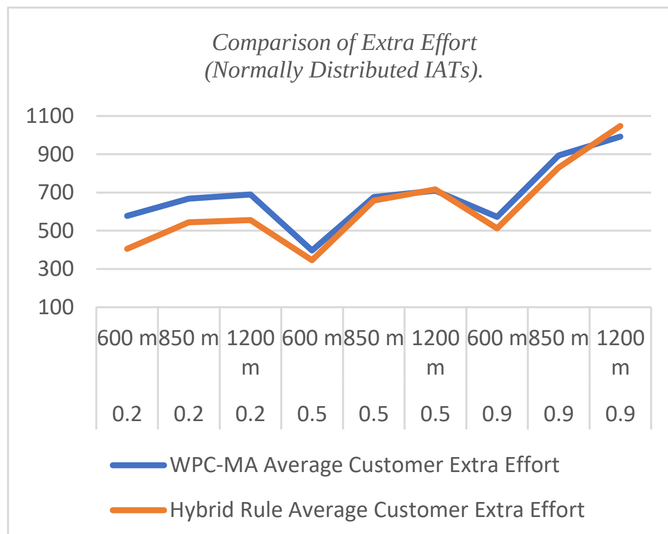

Figure 7. Comparison of Extra Effort (Normally Distributed IATs).

Table 9. Summary of results for the two tested redirection schemes under Weibull (β = 0.5) IATs.

c value	*r* value	WPC-MA Service Level	WPC-MA Average Customer Extra Effort	Hybrid Rule Service Level	Hybrid Rule Average Customer Extra Effort
0.2	600 m	0.780	577.38	0.748	405.09
0.2	850 m	0.758	667.15	0.730	544.56
0.2	1200 m	0.803	689.98	0.771	556.00
0.5	600 m	0.775	396.68	0.793	345.93
0.5	850 m	0.817	675.98	0.816	657.61
0.5	1200 m	0.83	709.22	0.860	716.09
0.9	600 m	0.845	572.96	0.852	512.64
0.9	850 m	0.919	893.80	0.882	829.58
0.9	1200 m	0.964	992.04	0.989	1047.82

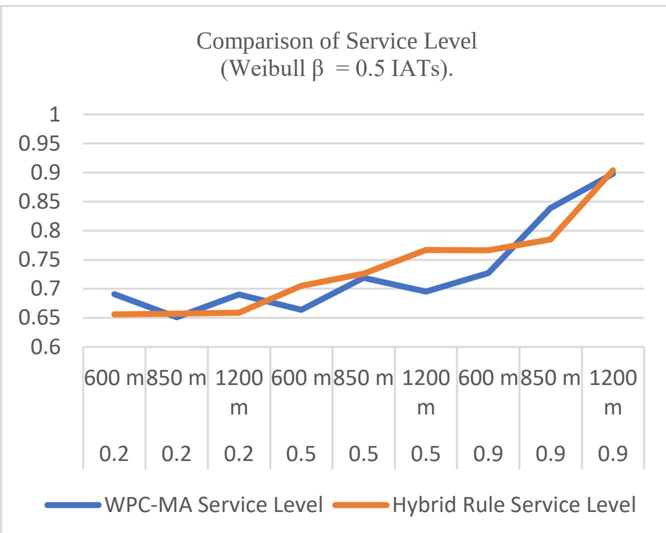

Figure 5. Comparison of Service Level (Weibull β = 0.5 IATs).

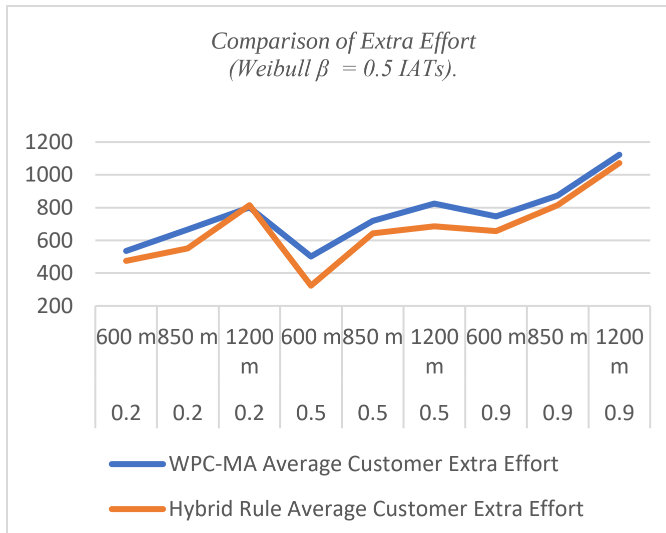

Figure 5. Comparison of Extra Effort (Weibull β = 0.5 IATs).

Table 10. Summary of results for the two tested redirection schemes under Weibull (β = 2) IATs.

c value	*r* value	WPC-MA Service Level	WPC-MA Average Customer Extra Effort	Hybrid Rule Service Level	Hybrid Rule Average Customer Extra Effort
0.2	600 m	0.691	534.85	0.656	474.38
0.2	850 m	0.651	665.90	0.657	551.80
0.2	1200 m	0.690	801.38	0.659	814.03
0.5	600 m	0.664	501.55	0.705	323.76
0.5	850 m	0.719	718.22	0.726	642.97
0.5	1200 m	0.695	824.17	0.767	685.40
0.9	600 m	0.727	745.49	0.766	656.66
0.9	850 m	0.839	872.70	0.785	815.52
0.9	1200 m	0.898	1122.92	0.904	1072.18

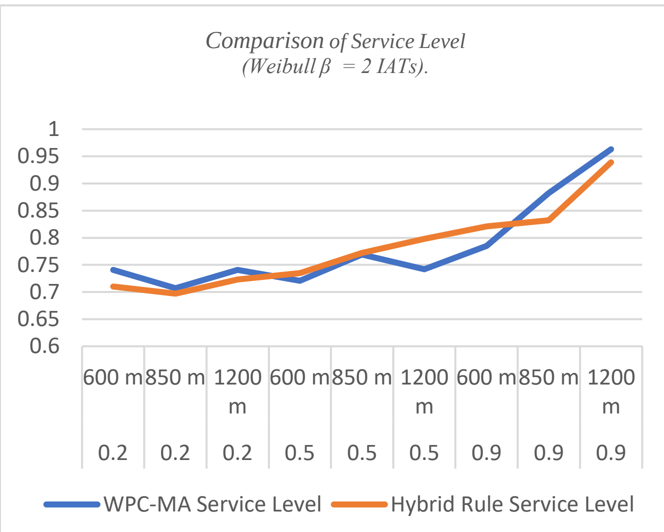

Figure 7. Comparison of Service Level (Weibull β = 2 IATs).

Figure 7. Comparison of Extra Effort (Weibull β = 2 IATs).

Through statistical tests, it has been proven that the average customer extra effort for the hybrid rule is significantly less than the average customer extra effort for the WPC-MA rule, but the service level was still retained, even when the distribution of the inter-arrival time distribution was changed from exponential to both normal and Weibull. This further solidifies the findings that the hybrid rule is better than the WPC-MA rule.

As a closing note on the simulation model, it is worth remembering that it was run using OMNeT++, thus the implemented algorithm of the bike-sharing system is limited by the programming language through which OMNeT++ runs: C++. Some of the adjustments that had to be made include rounding off quantities such as the inter-arrival time (in seconds), inter-station distances (in meters) and travel times (in seconds) to the nearest integer. The model would be more accurate another, more robust simulator is used.

7.2 Customer Cooperation and Extra Effort

The study dealt with the customer acceptance of redirection by allocating an adjustable variable, *c*, in the model. However, in the real world, not all customers are willing to be avail of redirection, even with incentives. We must remember that bike-sharing systems are meant for the traversing of intermediate distances; this would often include business districts and schools, where most of the customers would like to avoid lengthy trips so as not to be late for work or class. As such, these customers would likely not avail of redirection regardless of the incentive offered.

Singla et al. [16] conducted a survey among bike-sharing system customers in a city in Europe, and the results showed that roughly 20% were unwilling to walk/participate in redirection regardless of incentives given. Similarly, Ban et al. [15] conducted a survey in Korea on the willingness of bike-sharing customers to participate in redirection with incentives, and found out that more incentives are required the larger the extra effort of redirection is. Customer participation goes down from 94.3% to 47.4% when extra effort is increased from 200 m to 1000 m. Both studies show that customers are willing to participate in customer redirection, generally speaking. As such, the implementation of the proposed rule is expected to be adhered to by the customers of the bike-sharing system. In any case, extra effort has been shown to have been reduced outside very low values for the customer cooperation factor *c*.

7.3 Operationalization of the Solution

It was previously mentioned that Aeschbach et al. [17] made use of a mobile app that the bike-sharing system customers would use whenever they would rent a bike. If the proposed rule were to be operationalized, the same concept of a mobile app would be one of the easiest methods, due to mobile devices' wide usage and portability. The use of a mobile app would ensure that the information of the entire system can be relayed to all its customers, and any changes in the state of any of the entities present in the system (e.g. bike level of stations) would be shown in real time. All redirection parameters can also be immediately presented to the bike-sharing users (including neighborhood radius), so that customers can be more informed in their decision to whether or not avail of customer redirection. It should be noted that even if only a fraction of the bike-sharing system's customers use the mobile and engage in customer redirection, the benefits would be felt by everyone using the system.

VIII. SUMMARY, CONCLUSION AND FURTHER STUDIES

Bike-sharing systems experience imbalances in bikes throughout its stations. This is answered by implementing repositioning or customer redirection. With customer redirection, customers are asked to walk distances they would normally not have to. The study addresses the need for a rule that minimizes this extra distance walked (called extra effort) while maintaining a target service level for the system. A dynamic threshold-based rule was developed, which considers the historical net movement of bikes per station. The improvement brought about by the proposed rule was then validated through a simulation model based on the Santander Cycles bike-sharing system in London, which consisted of 749 bike stations, 9,109 bikes and 3,660,304 bike trips.

The study found out that the dynamic threshold-based rule gives a smaller value of system average extra effort over 100 simulation runs compared to the existing WPC-MA rule, while still holding the same service level. Both these conclusions (improvement in extra effort and retention of service level) have been proved statistically with a 5% level of significance. From here, it can be said that the hybrid rule is now the best exiting rule in terms of minimizing extra effort expended by redirected bike-sharing customers.

Future studies can try to find an even better rule in terms of both service level and extra effort, since the global optimal rule has not yet been formulated. Future studies can also investigate models with factors that were not considered in the study, such as incorporating customer redirection alongside static/dynamic repositioning.

REFERENCES

[1] Fricker C, Gast, N. 2016. Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity. Euro Journal on Transportation and Logistics. 5(3):261-291.

[2] Schuijbroek J, Hampshire RC, Van Hoeve WJ. 2017. Inventory rebalancing and vehicle routing in bike sharing systems. European Journal of Operational Research. 257(3):992-1004.

[3] Rainer-Harbach M, Papazek P, Hu B, Raidl GR. 2013, April. Balancing bicycle sharing systems: A variable neighborhood search approach. European conference on evolutionary computation in combinatorial optimization; Springer, Berlin, Heidelberg. p. 121-132.

[4] Benchimol M, Benchimol P, Chappert B, De La Taille A, Laroche F, Meunier F, Robinet L. 2011. Balancing the stations of a self service “bike hire” system. RAIRO-Operations Research. 45(1):37-61.

[5] Hernández-Pérez H, Rodríguez-Martín I, Salazar-González JJ. 2009. A hybrid GRASP/VND heuristic for the one-commodity pickup-and-delivery traveling salesman problem. Computers & Operations Research. 36(5):1639-1645.

[6] Hernández-Pérez H, Salazar-González JJ. 2004. Heuristics for the one-commodity pickup-and-delivery traveling salesman problem. Transportation Science. 38(2):245-255.

[7] Huang D, Chen X, Liu Z, Lyu C, Wang S, Chen X. 2020. A static bike repositioning model in a hub-and-spoke network framework. Transportation Research Part E: Logistics and Transportation Review. 141:102031.

[8] Ghosh S, Varakantham P, Adulyasak Y, Jaillet P. 2015, April. Dynamic redeployment to counter congestion or starvation in vehicle sharing systems. In Proceedings of the International Conference on Automated Planning and Scheduling. 25(1).

[9] Contardo C, Morency C, Rousseau LM. 2012. Balancing a dynamic public bike-sharing system. Montreal, Canada: Cirrelt. 4.

[10] Caggiani L, Ottomanelli M. 2013. A dynamic simulation based model for optimal fleet repositioning in bike-sharing systems. Procedia-Social and Behavioral Sciences. 87:203-210.

[11] Pfrommer J, Warrington J, Schildbach G, Morari M. 2014. Dynamic vehicle redistribution and online price incentives in shared mobility systems. IEEE Transactions on Intelligent Transportation Systems. 15(4):1567-1578.
[12] Parhizkar T, Hogenboom S, Vinnem JE, Utne IB. 2020. Data driven approach to risk management and decision support for dynamic positioning systems. Reliability Engineering & System Safety. 201:106964.
[13] Waserhole A, Jost V, Brauner N. 2013. Pricing techniques for self regulation in vehicle sharing systems. Electronic Notes in Discrete Mathematics. 41:149-156.
[14] Chemla D, Meunier F, Calvo RW. 2013. Bike sharing systems: Solving the static rebalancing problem. Discrete Optimization. 10(2):120-146.
[15] Ban S, Hyun KH. 2019. Designing a user participation-based bike rebalancing service. Sustainability. 11(8):2396.
[16] Singla A, Santoni M, Bartók G, Mukerji P, Meenen M, Krause A. 2015, February. Incentivizing users for balancing bike sharing systems. Proceedings of the AAAI Conference on Artificial Intelligence. 29(1).
[17] Aeschbach P, Zhang X, Georghiou A, Lygeros, J. 2015, December. Balancing bike sharing systems through customer cooperation-a case study on London's Barclays Cycle Hire. 54th IEEE Conference on Decision and Control (CDC). IEEE. p. 4722-4727.
[18] Raviv T, Kolka O. 2013. Optimal inventory management of a bike-sharing station. IIE Transactions. 45(10):1077-1093.
[19] Leguesdron P, Pellaumail J, Rubino G, Sericola B. 1993. Transient analysis of the M/M/1 queue. Advances in Applied Probability. p. 702-713.
[20] Ruch C, Warrington J, Morari M. 2014, June. Rule-based price control for bike sharing systems. 2014 European Control Conference (ECC). IEEE. p. 708-713.