Simulating a Bluebikes network

Adam Dusenbery

July 22, 2022

Abstract

To examine problems facing the Bluebikes sharing service and provide descriptive solutions backed up by simulated networks using real world relevant data provided by the company itself. This project focuses mainly on issues regarding bike traffic buildup and provides solutions catering towards optimizing customer satisfaction and reducing unexpected travel time.

1 Introduction

Bluebikes, a popular bike sharing service in the Boston, MA region, offers customers the chance to rent bicycles at stations and store them in any other serviceable station they choose. A common issue facing the company as of recently is the buildup of traffic that causes stations to fill capacity, thus preventing new customers from docking at their preferred location. This in turn causes customers to move off of their normal route and dock the bike at a faraway destination. The result being unexpected and inconvenient travel time at the expense of the customer.

2 Project Overview

2.1 Goals and Focus

With customer satisfaction being the outlined loss, the goal of this project was to reduce buildup at any given station and consider any waiting time for a station to become available as an absolute worst case scenario. Furthermore, it is necessary to consider customers who wish to have bikes available to move from any given station. With these two goals in mind, the project of customer satisfaction becomes a balancing act of ensuring availability at all times.

2.2 Data Collection

The data used in this research was gathered both partly from real observations as well as directly from Bluebikes website, in which trip statistics are recorded and available in CSV format. Basic spreadsheet tuning was used in order to parse information into relevant figures. Primary metrics that were sought include arrival rates, (specifically arrival rates for certain stations) average daily use, and weekly variance.

2.3 Data Described

Our primary metric to use will be the arrival rates of customers at two separate stations. For the purposes of simplicity, we will consider only two stations. The first station, Davis (the central hub of the network and by far the most popular arrival station) will represent its real life counterpart. While the second station, Linear Park, will represent a collection of identical stations in the real world, all with comparable arrival rates and significantly less traffic. The arrival rates of each station is given below (See Table 1) and will be used for our simulation purposes.

Station Name	Arrivals
Davis	45
Linear Park	16

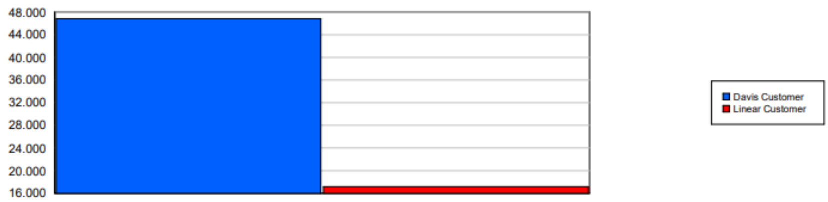
Table 1: Station Daily Arrivals

2.4 Process

With Arena as our chosen software we begin building the simulation with two separate entity arrivals, one for customers arriving at Davis, another for customer arrivals in Linear Park. From here we implement a decision block which routes customers to either exiting the simulation (if the station they are in is their preferred location) or beginning the process of acquiring a bike and subsequently begin their route to their preferred station. Either way, our customers end up always exiting the system so long as there are available bikes.

Chart 1: Process Flow

Due of the nature of this service, it is available all times of the day, however almost no observations were recorded during the night hours of the day. With this in mind, for the purposes of more realistic simulation, we set our daily clock to have only 12 hours per day and adjust our arrival rates accordingly. This will ensure the same number of arrivals, while simulating more precise queue time measurements.



Number Out	Average	Half Width	Minimum Average	Maximum Average
Davis Customer	45.1700	1.44	27.0000	64.0000
Linear Customer	16.6400	0.83	7.0000	25.0000

WIP	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Davis Customer	1.5935	0.05	1.0337	2.2264	0.00	9.0000
Linear Customer	0.5855	0.03	0.2501	0.8892	0.00	5.0000

3 Results and Observations

3.1 Results

With a working process and accurate data, arriving at an optimized solution involves a matter of simple trial and error. The goal of the project was to optimize availability to ensure customer satisfaction. Thus any queue time should be avoided at all costs. This can be achieved, quite realistically. Running the simulation at 100 repetitions gives an accurate count of queue time as well as entity in/out numbers. Most importantly, the findings of this study show queue times can be almost entirely avoided with a resource rate of 5 bikes in our starting location on the Linear Park station and 9 bikes in our Davis station.

Waiting Time	Average	Half Width
Grab Davis Bike.Queue	0.00	0.00
Grab Linear Bike.Queue	0.00	0.00

Other

Number Waiting	Average	Half Width
Grab Davis Bike.Queue	0.00	0.00
Grab Linear Bike.Queue	0.00	0.00

3.2 Observations and Concluding Remarks

One particular note is that the arrival rates fall more precisely on a non-homogeneous Poisson distribution, as specific intervals throughout the day experience much higher traffic than others. For instance, rush hour experiences approximately 16 customers arriving in the Davis station every half hour. This rush is short lived, and due to the extended time periods where almost no traffic is observed, it was adequate to consider the distribution mean on a full day schedule, allowing us to work with a simple Poisson arrival rate.

With all this being said, the observations found in the simulation reflect well with what follows the real life network. Bikes are a 'one size fits all' resource and thus customers are not overly picky with which they select. Thus transferring customers is a speedy process which eliminates lengthy process time and ensures availability with only a few resources required in spite of a large number of daily customers.