The Bike Share Program

Marko Mitrovic, Alborz Namazi, Sarah Lin, Miguel Valdez

Department of Mathematics, Simon Fraser University, Burnaby, BC, Canada

Abstract

This paper proposes and solves a modified allocation problem that aims to implement the Bike Sharing Program. Here, we seek to maximize the total benefit of each bike stations while minimizing the total operational cost. Through this two-step process, various optimization methods and techniques, including both linear and non-linear programming, are used to: a) determine the most optimal bike stations among all the candidate locations, and b) decide on the most optimal number of bikes needed for each station. Considering the constraints imposed by the software used in this project, Microsoft Excel, the problem is limited to selecting four stations, while operating over a 24-hour cycle. In part a, the solver has selected stations B, G, H, and N among 14 different candidates to maximize the total benefit. In part b, the optimal number of bikes during the opening hour at each of these four stations is calculated as: 297 bikes at station B, 106 bikes at station N, and no bikes at stations G and H.

1 Introduction

There are numerous ecological issues that environmentalists have been dealing with over the past few decades; one of which is pollution. For instance, in China, the amount of carbon emissions, brought out by cars and factories, have been so great that the government is having a hard time controlling air quality. In addition diminishing oil resources, and rising gas prices are a few other factors that may persuade the general public more than ever to seek other means of transportation.

As a solution to this problem, this paper proposes the bike-share program must be supported and implemented by various municipalities. This program develops a green and sustainable mode of transportation. Of course, now the question to answer is where to place these bike stations and how many bikes are required to operate.

2 Background

The bike-share system enables users to rent a bike without having to pay a fee, or just pay an affordable fee. Many cities have already implemented a bike-share system to reduce private vehicles use, traffic congestion, noise and air pollution and also to reduce the number of car theft and vandalism. One of the most famous bike-share systems is the “Bixi”. Bixi is the largest public bike-share system in Canada, where it was started in Montreal, in 2007. Each bike station includes a pay station, bike docks and bikes. Bixi uses ANAT technology, which allows each pay station to deploy a single dock. The bike station is flexible; it can be easily moved and expanded.

The Bixi system is very user-friendly, and those who would like to rent a bike, may go to the Bixi website to find the nearest bike station. From the website, one may also find out which bike station is available for the bikes. Furthermore, there are two payment systems. The *shorter* one allows for a 24 hours access code for $5, and the *longer* version allows for 72 hour for $12. Users can get a subscription key from the “bixi” website; a one year subscription is $97, and 30 days is $41. After selecting a bike, the user will get a code, which is used to unlock the bikes. After a 30-minute trip, the rider may return the bike at any bike station of their choice. The Bixi system was implemented successfully in Montreal, and resulted in one million rides by the end of its first few months.

The “Bixi” system has flourished in Ottawa, Toronto, Boston, London and Washington, and will be launched in downtown Vancouver and Kitsilano area in summer of 2013. Compared to Montreal, Vancouver has a much tougher Helmet By-law, making it difficult for the program to be implemented. To resolve this issue the City officials are contemplating the idea of asking the Richmond-based company, Sandvault Group Global, to build reusable, helmet dispensing vending machines at the cost of $50,000 each. Similar to the return of the bike, you can return your helmets at any bike station. It is estimated that bringing Bixi to Vancouver will cost $1.9 million per year.

3 Problem Statement

There are two main questions to be answered when implementing the bike-share program: a) where to place the bike stations, and b) how many bikes to allocate to each station. Therefore, this research is divided into two parts, to first select the most optimal stations and next, to select the most optimal number of bikes in each station.

In order to characterize each potential location, the following attributes are defined (please note that, depending on each client's specific needs these attributes can easily be changed and modified):

i. Proximity to a skytrain station

ii. Proximity to major roads

iii. Proximity to bike roads

iv. Population Density

v. Construction Costs

vi. Potential demand

These characteristics have to be integrated into the objective function, using appropriate coefficients. In reality these coefficients are provided by the client, however for the purpose of testing the model, coefficients are defined by arbitrary figures. Under part a, possible constraints would be the maximum number of stations, having at least one station in each region or its neighbouring region, and also, binary and integer decision variables.

The second part of the research focuses on the number of bikes per station. It would be unrealistic to assume that the demand and supply of bikes maintain the same level over time; therefore, each station has assigned a Poisson Distribution with a different λ for its supply and demand. The inconsistent supply and demand gives birth to the 'availability problem'. This problem arises in two areas: a) when a popular location captures all the demand, shifting the balance, creating excess supply and demand for bikes b) at the closing hour, all or most of the bikes are accumulated in a few stations.

To resolve these two issues, a penalty system is defined in order to capture the impact of the excess demand and supply. Also it is assumed that bikes can be transferred from one station to another at a cost.

Under part b of this problem, objective function minimizes operational costs in order to determine the number of bikes per station. Total operational costs consists of the sum of total cost of purchasing bikes, total penalty costs and total transportation costs.

Penalty cost is implemented when there is a shortage of bikes at a station (excess demand), or when a customer has to wait to return a bicycle due to lack of space (excess supply). Transportation costs involves transporting bikes and redistributing them to each station.

The constraints under this part must consider the available capacity, and the availability of bikes to be transferred, in addition to binary and integer constraints. These constraints will be discussed in the future sections of this report.

4 Model Formulation - Part A

Under part a, assume there are *n* regions neighbourhoods, or potential stations, where *k* stations must be selected in a way that each region, or its neighbouring region has at least one bike station. To test the model it is assumed that there are fourteen potential regions n $\epsilon\{A, B, C, \ldots, N\}$ and four stations are to be selected.

4.1 Basic Variables and Decision Variables

It is realistic to assume that one of the requirements put forward by the client is that there must be a minimum number of bike stations in each jurisdiction or its neighbouring jurisdictions (in this case, the minimum number is assumed to be one). In order to capture this constraint, a binary variable, y_{ij}, must be defined to show that whether neighbourhood i is next to neighbourhood j or not. So if region A is neighbouring region M, then $Y_{AM} = 1$, 0 otherwise.

Also, there must be binary variables to reflect the attributes associated with each station i:

SS_i	Proximity to Sky-Train Stations
BR_i	Proximity to Bike Roads
MR_i	Proximity to Major Roads
PD_i	Above average Population Density
CC_i	Construction Cost
ED_i	Excess Demand

The first four variables above have binary values of 0 or 1; meaning either they meet the criteria or not.

Also, in order to reflect the impact of each characteristic in the objective function, each variable has a coefficient associated to it. As stated previously, these coefficients are provided by the client, however for the purpose of testing the model, coefficients are defined as arbitrary figures:

s	Proximity to Sky-Train Stations
b	Proximity to Bike Roads
r	Proximity to Major Roads
d	Above average Population Density
c	Construction Cost
e	Excess Demand

Unlike the first four variables, the last two are not binary; rather they are represent by linear functions.

Construction costs, CC_i is directly correlated to the capacity of the station, i.e. the larger the station the higher the construction costs.

$$CC_i = c * C_i$$

C_i is the capacity at station *i* and *c* is the coefficient connecting the capacity to the cost. For example, $10,000 for each 100 bike capacity.
Excess Demand, ED_i is also a linear function based on the demand and capacity:

$$ED_i = C_i - DD_i$$

DD_i is the daily demand at station *i* which is based on the Poisson Distribution with the arrival rate of λ_i. Therefore, stations where the daily demand exceeds the capacity may not be the best option.
In part a, the decision variables would take a binary form. Defined by, X_i, the decision variables will have values 1 if selected and 0, otherwise.

4.2 Objective Function

The objective function is to maximize the total benefit by choosing the most optimal stations:

$$\sum_{i=1}^{n}[X_i][(s)(SS_i) + (b)(BR_i) + (r)(MR_i) + (d)(PD_i) - (c)(CM_i) - (e)(ED_i)]$$

The output of this function is the sum of the total net effects of the selected stations.

4.3 Constraints

The maximum number of stations to build, must be equal to the client's requirement (k):

$$\sum_{i=1}^{n}(X_i) \leq k$$

Also the number stations covering each region and its neighbouring region must be greater than or equal to the client's requirement (q):

$$\sum_{j=1}^{n}\sum_{i=1}^{n}(X_i)(Y_{ij}) \geq q$$

5 Model Formulation - Part B

In developing the second model, it was also assumed that each station has its own supply and demand following a Poisson Distribution. Therefore, over time it highly probable that all the bikes may accumulate at one location.

In order to resolve this issue, a service is put in place to transport bikes from one station to another, at a cost.

Also, as discussed previously, the constant changes in supply and demand may create situations where, there is no bike to borrow, or the station is full and the bike cannot be returned.

5.1 Basic Variables and Decision Variables

To accommodate the programming of this part the following variables are designed:

D_{it}	'Demand' at time t at station i
S_{it}	'Supply' at time t at station i
A_{it}	'Available Bikes' at time t at station i

The decision variables are the optimal number of bikes at different stations at time zero (T_0), BT_{i0}

Available Bikes at time *t* at station *i* or A_{it}, is a linear function of:

$$A_{it} = A_{t-1} + S_{it} - D_{it}$$
$$A_{i0} = BT_{i0}$$

The penalty, *p*, must be applied when A_{it} takes a negative value, i.e. the demand exceeds availability:

IF ($A_{it} \leq 0$) THEN p.ABS(A_{it}), otherwise 0.

The IF, THEN constraint must be converted into a non-linear constraint:

$$\text{Penalty} = PN_{it} = -p * A_{it}$$
$$-PN_{it} \leq p * A_{it} + My$$
$$-PN_{it} \geq p * A_{it} + My$$
$$A_{it} > M(1-y)$$

By setting M $= -p * A_{it}$, the penalty will be only applied when there is an excess demand. The same procedure can be applied to the excess supply.
As noted before, the transportation cost of bikes must be considered as well. TC_{ij} is the Transportation cost, the cost of moving bikes from station *i* to station *j*.
To calculate the transportation costs, a new set of variables are required; that is the number of bikes moved from one station to another, TB_{ij}, the number of bikes transferred from station *i* to station *j*.

5.2 Objective Function

The objective function aims to minimize the Total Operating Expense which is a combination of the following:

i. Penalty cost function:

$$\sum_{i=1}^{n}\sum_{t=1}^{24}(PN_{it}) = \sum_{i=1}^{n}\sum_{t=1}^{24}(-p.A_{it}), \forall A_{it} \leq 0$$

ii. Cost of Purchasing Bikes:

$$\sum_{i=1}^{n}(cb * BT_{i0})$$

cb is the cost to purchase a bike.
iii. Transportation cost:

$$\sum_{j=1}^{n}\sum_{i=1}^{n}(TB_{ij} * TC_{ij})$$

The objective function would minimize the sum of total penalty cost, total cost of purchasing bikes and total transportation costs.

5.3 Constraints

Similar to part a, in addition to the required constraints to obtain a solution, there are a few restrictions imposed by the client. For instance, the available budget to purchase bikes. However, due to limitations imposed by the solver in MS Excel, these constraints are kept to minimum:

- the number of available bikes at time t cannot exceed the capacity of the station:

$$A_{it} \leq C_i$$

- The number of bikes transported out from a station to other locations at the end of the day cannot exceed the number of bikes available:

$$\sum_{j=1}^{n}\sum_{i=1}^{n} TB_{ij} \leq A_{i\ closing-time}$$

- The number of bikes transported into a station must satisfy the opening hour demand:

$$\sum_{j=1}^{n}\sum_{i=1}^{n} TB_{ij} \geq BT_{i0}$$

6 Solution

6.1 Part A

In order to test this model, the following attributes and coefficients were fed to the solver, using MS Excel:

The Coefficients of Each Attribute	
Near a Skytrain Station	$ 50,000
Near a Major Road	$ 30,000
Near a Bike Road	$ 40,000
Population Density Above Average	$ 20,000
Basic Construction Cost	$ 10,000
Penalty for Excess Demand	$ 50

Potential Stations	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Near a Skytrain Station	1	0	0	1	0	1	1	1	0	0	1	0	0	1
Near a Major Road	1	1	0	0	0	0	1	1	1	1	1	0	0	1
Near a Bike Road	0	1	1	0	0	1	0	1	1	0	0	0	1	1
Population Density Above Average	1	1	0	1	1	0	1	0	0	1	0	0	1	0
Construction Cost Multiplier	5	7	5	7	6	7	7	5	6	7	7	8	7	5
Capacity	500	700	500	700	600	700	700	500	600	700	700	800	700	500
Rate of Daily Demand For Each Station (Lambda)	500	650	400	600	650	550	600	750	700	800	500	800	750	600
Daily Demand For Each Station (Poisson)	456	626	424	616	678	567	565	734	718	747	495	802	762	596
Excess Demand	0	0	0	0	78	0	0	234	118	47	0	2	62	96

Also, a hypothetical zoning map was developed to determine the neighbouring jurisdictions:

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Total
Neighbouring Jurisdictions Matrix	A	1	1	0	0	0	0	0	0	0	0	0	0	1	0	3
	B	1	1	1	0	1	1	1	0	0	1	0	1	1	1	10
	C	0	1	1	1	1	0	0	0	0	0	0	0	0	0	4
	D	0	0	1	1	1	0	1	0	0	0	0	0	0	0	4
	E	0	1	1	1	1	0	1	0	0	0	0	0	0	0	5
	F	0	1	0	0	0	1	1	1	0	1	0	0	0	0	5
	G	0	1	0	1	1	1	1	1	0	0	0	0	0	0	6
	H	0	0	0	0	0	1	1	1	1	1	1	0	0	0	6
	I	0	0	0	0	0	0	0	1	1	1	1	0	0	1	5
	J	0	1	0	0	0	1	0	0	1	1	1	0	0	1	6
	K	0	0	0	0	0	0	0	1	1	1	1	0	0	0	4
	L	0	1	0	0	0	0	0	0	0	0	0	1	1	1	4
	M	1	1	0	0	0	0	0	0	0	0	0	1	1	1	5
	N	0	1	0	0	0	0	0	0	1	1	0	1	1	1	6

Based on the input provided, stations B, G, H, and N are selected to maximize the total benefit. The Net Effect of each station, NE_i, are as follows:

NE_B = \$20,000

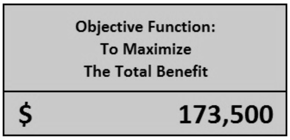
NE_G = \$30,000

NE_H = \$58,300

NE_N = \$65,200

As a result,

Jurisdiction	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Total	≤	Const.
Bike Station?	0	1	0	0	0	0	1	1	0	0	0	0	0	1	4		4

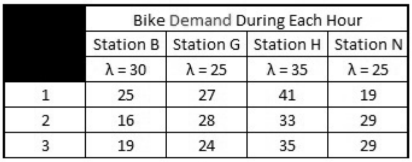


Objective Function:
To Maximize
The Total Benefit

\$ 173,500

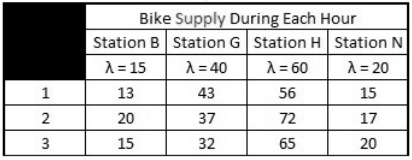
6.2 Part B

For Part B, the solver will seek the optimal number of bikes per station by minimizing the total operational costs, based on the following input:



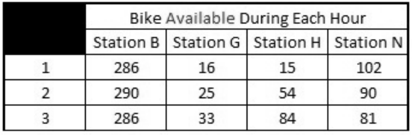
	Bike Demand During Each Hour			
	Station B	Station G	Station H	Station N
	λ = 30	λ = 25	λ = 35	λ = 25
1	25	27	41	19
2	16	28	33	29
3	19	24	35	29

⋮



	Bike Supply During Each Hour			
	Station B	Station G	Station H	Station N
	λ = 15	λ = 40	λ = 60	λ = 20
1	13	43	56	15
2	20	37	72	17
3	15	32	65	20

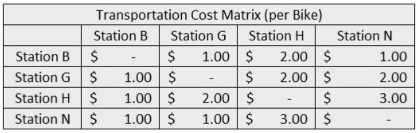
⋮



	Bike Available During Each Hour			
	Station B	Station G	Station H	Station N
1	286	16	15	102
2	290	25	54	90
3	286	33	84	81

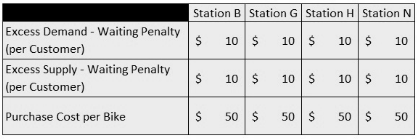
⋮

Transportation cost:



Transportation Cost Matrix (per Bike)				
	Station B	Station G	Station H	Station N
Station B	$ -	$ 1.00	$ 2.00	$ 1.00
Station G	$ 1.00	$ -	$ 2.00	$ 2.00
Station H	$ 1.00	$ 2.00	$ -	$ 3.00
Station N	$ 1.00	$ 1.00	$ 3.00	$ -

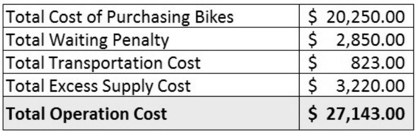
The penalty system, and the cost of purchasing a bike:



	Station B	Station G	Station H	Station N
Excess Demand - Waiting Penalty (per Customer)	$ 10	$ 10	$ 10	$ 10
Excess Supply - Waiting Penalty (per Customer)	$ 10	$ 10	$ 10	$ 10
Purchase Cost per Bike	$ 50	$ 50	$ 50	$ 50

Based on the information provided to the solver, an optimal solution by setting: BT_{B0} = 298, BT_{G0} = 0, BT_{H0} = 0, and BT_{N0} = 106

Generating the following operating cost:



Total Cost of Purchasing Bikes	$ 20,250.00
Total Waiting Penalty	$ 2,850.00
Total Transportation Cost	$ 823.00
Total Excess Supply Cost	$ 3,220.00
Total Operation Cost	**$ 27,143.00**

7 Improvements

In spite of the fact that we put our best efforts towards creating the most realistic model we could, we were inevitably constrained by the amount of time, information, and experience we had. As a result, we compiled a list of improvements that we considered both during the process of creating the model and the reflection procedure after we had finished the model.

The first and most obvious improvement we considered was collecting real data. One way to gather some of the data (for example construction costs, bike costs, and etc.) would be to contact the appropriate representatives of various municipalities, government officials, or private consulting firms who could in turn provide realistic figures.

However, the market data may not necessarily be available for all the variables, and there is a possibility that for certain variables there is no previous data collected. For example, it is unlikely that the City has information regarding the rate of daily demand at a particular location considering there has never been a bike station installed at that location. It would still be possible to collect a preliminary sample pool regarding the daily demand by surveying/interviewing the public; however, collecting such data would be a gargantuan task with many challenges. For example, such a survey would not only have to interview a large number of people, but it would also have to ensure that the right demographic is targeted. Conducting such surveys is on the level that private companies are usually contracted to do the job. Perhaps it would be possible to use and extrapolate data from cities with established bike-share programs. Although this is not the ideal solution, it would provide some insight into the possible supply and demand curves.

An excellent example of where realistic data is almost necessary for an optimal model is seen in the task of "zoning" the city. Figure 1 shows the neighbourhood plan for the City of Vancouver.

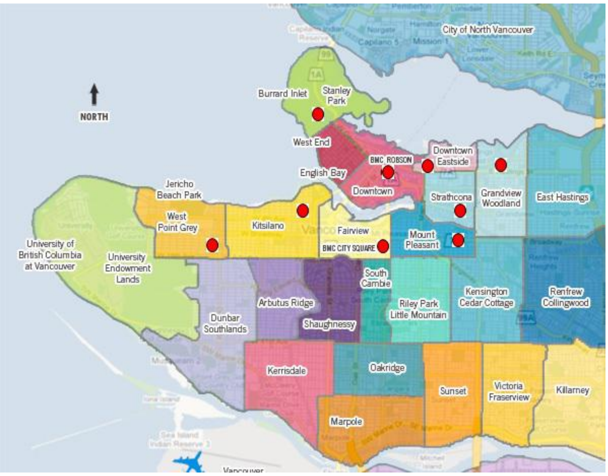


Figure 1: Figure 1: City of Vancouver - Neighbourhood Plan

In reality, one would have to examine exactly how to partition different regions, and whether or not more constraints are necessary for selecting an ideal location, which will lead to a far more complex problem.

Another improvement would be to upgrade the software used in solving this problem. Due to MS Excel limitations, many variables and constraints are left out of the equation to ease the load on the software. Perhaps, in case of expanding the problem, over a longer period of time with many more potential stations and attributes, it would be more appropriate to use other software, such as CPLEX Optimizer.

8 Extensions and Alternate Approaches

This problem of changing demands leads us into the second major area we reflected on; that is, the various possible approaches and expansions of the problem. For example, as mentioned earlier, we could vary parts of the model not only by hour, but also by day of the week and/or month of the year. Perhaps most importantly, we could vary the demands for the bikes at each station based on the day of the week and month of the year. It would seem intuitive that the bike demands at a station in Stanley Park, for example, would be much higher during the weekend and over the summer than they would be over weekdays and the winter.

We also considered another, quite different, approach to the entire problem. In our model, we assumed that the government is attempting to provide a public service with this bike-share program (perhaps as an extension to its public transportation system). Thus, our model is essentially trying to maximize benefit to the citizens at the lowest possible cost. An alternate idea would be to implement a "for-profit" model in which a private corporation was attempting to maximize its profit by renting out these bikes. In many ways the two approaches would be quite similar as they would both likely include hourly demands for the bikes at each station and other such features. However, the "for-profit" model would probably be much more focused on the queuing theory aspect since the company would want to balance the wait time at each station with the maximum amount of time a customer would be willing to wait and still rent a bike. Another exemplar of a difference between the two models is that "for-profit" model would have a section where we would try to optimize the number of bikes rented out by raising/lowering the price of rental (this new price variable could also benefit from the aforementioned variation by month of the year).

9 Conclusion

This project allowed us to see both the usefulness and difficulties associated with applying mathematical methods to "real-world" problems. In the end, we believe that with the time, resources and course material we had available, we produced an adequate model for a complicated problem that was particularly relevant to our community.

References

[1] Bixi Montreal. Ride with BIXI. 6 April, 2013. Retrieved from website: https://montreal.bixi.com/ride-with-bixi/functioning

[2] City of Vancouver (2011, Aug 8). VanMap: Public edition. [Interactive Media]. Retrieved from website: http://vanmapp.vancouver.ca/pubvanmap_net/default.aspx