Two-stage stochastic optimization for bike sharing systems

Ruijian An[*] **Tianshu Shen**[†]

Abstract

In this paper, we formulate the bike sharing problem as a two stage stochastic optimization model which determines the optimal number of bikes per station and optimal transshipment of bikes. We present a Benders Cut and an extensive method to solve the problem. An extension to multistage problem is made where a bike is allowed to be rented multiple times a day. Computational result is presented with various number of scenarios and different demand distributions.

1 Introduction

In recent years, public bike sharing systems (BSS) are deployed in hundreds of cities around the world. In the bike-sharing problem, a service provider need to manage a fleet of bikes over a set of bike stations, with given capacities. A customer can rent a bike at one station and return the bike to another station. At the end of the day, bikes are transshipped between stations so that each station has optimal number of bikes during the next cycle of service. The utility that a bike sharing system can provide is heavily dependent on the stochastic demand of customers. Ideally, we would like to avoid two cases: i) no bike is available when a customer wants to rent, ii) no slot is available when a customer wants to return a bike. However, the number of bikes per station is decided before the stochastic demand is realized. This paper presents a two-stage stochastic optimization based on *Stochastic optimization models for a bike-sharing problem with transshipment*. The first stage decision is how many bikes should be assigned to each station. The second stage decision is the optimal number of bikes transshipped from one station to another. The objective is to minimize the cost caused by procurement cost of bikes, transshipment cost and cost generated by case i) & ii). In this paper, we introduce a new optimization model, Benders Cut, to compute the problem and implement the multistage formulation using node-based extensive form.

2 Problem Description

The bike sharing system being modelled by this project involves several actions induced by the users including realizing bike rental demands (i.e. arrival at the bike stations), actual rental action, redirection of bikes in case of station overflow. A demonstration of the overall sequence of operations flow could be found in Figure 1.

The formulation of the two-stage stochastic program problem is written as follows:

First stage variables:

x_i : the number of bikes to assign to bike-station i $\in$ B at the beginning of the service

Second stage variables:

β_{ijs} : Number of rented bikes from bike-station i to bike-station j in scenario s

I_{is}^{+} : Realized surplus of bikes at bike-station i in scenario s

I_{ijs}^{-} : Realized shortage of bikes at origin-destination pair i, j in scenario s

ρ_{ijs} : Number of redirected bikes from bike-station i to bike-station j in scenario s (in case of overflow)

O_{is}^{+} : Residual capacity at bike-station i in scenario s

O_{is}^{-} : Overflow at bike-station i in scenario s

τ_{ijs} : Number of transshipped bikes from bike-station i to bike-station j in scenario s

T_{is}^{+} : Excess of bikes at bike-station i in scenario s

$T_i s^{-}$: Lack of bikes at bike-station i in scenario s

[*]Department of Computer Science, University of Toronto; ruijian.an@mail.utoronto.ca

[†]Department of Mechanical and Industrial Engineering, University of Toronto; tina.shen@mail.utoronto.ca

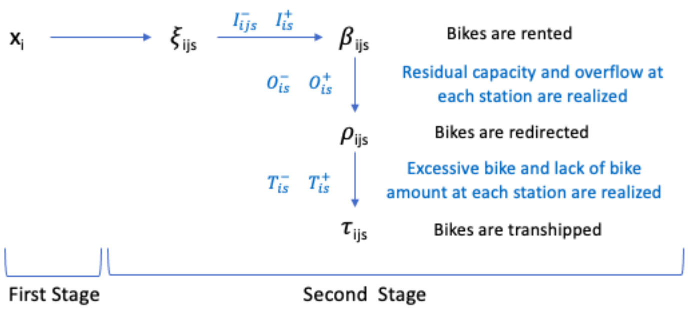


Figure 1: Sequence of operations

3 Stochastic optimization models

This section lists the three models being implemented and further analysed for the project. This includes the two-stage problem constructed in its extensive form, the benders decomposition model, and finally, the multi-stage model.

3.1 Extensive Model

$$\text{minimize: } c\sum_{i}^{B} x_i + \sum_{s=1}^{S} p_s \sum_{i=1}^{B} [v_i \sum_{j=1}^{B} I^-_{ijs} + w_i O^-_{is} + \sum_{j=1}^{B} t_{ij}\tau_{ijs}] \tag{1}$$

$$\text{subject to: } x_i \leq k_i, \forall i \in B \tag{2}$$

$$\beta_{ijs} = \xi_{ijs} - I^-_{ijs}, \forall i,j \in B, s \in S \tag{3}$$

$$I^+_{is} - \sum_{j=1}^{B} I^-_{ijs} = x_i - \sum_{j=1}^{B} \xi_{ijs}, \forall i \in B, s \in S \tag{4}$$

$$O^+_{is} - O^-_{is} = k_i - x_i + \sum_{j=1}^{B} \beta_{ijs} - \sum_{j=1}^{B} \beta_{ijs}, \forall i \in B, s \in S \tag{5}$$

$$\sum_{j=1}^{B} \rho_{ijs} = O^-_{is}, \forall i \in B, s \in S \tag{6}$$

$$\sum_{j=1}^{B} \rho_{jis} \leq O^+_{is}, \forall i \in B, s \in S \tag{7}$$

$$T^+_{is} - T^-_{is} = k_i - O^+_{is} + \sum_{j=1}^{B} \rho_{jis} - x_i, \forall i \in B, s \in S \tag{8}$$

$$\sum_{j=1}^{B} \tau_{ijs} = T^+_{is}, \forall i \in B, s \in S \tag{9}$$

$$\sum_{j=1}^{B} \tau_{jis} = T^-_{is}, \forall i \in B, s \in S \tag{10}$$

$$x_i, I^+_{is}, O^+_{is}, O^-_{is}, T^+_{is}, T^-_{is} \in \mathbb{Z}^+, \forall i \in B, s \in S \tag{11}$$

$$\tau_{ijs}, \beta_{ijs}, \rho_{ijs}, I^-_{ijs} \in \mathbb{Z}^+, \forall i,j \in B, s \in S \tag{12}$$

Where the objective function (1) is minimizing the expected total cost, this is the total sum of the procurement cost of initial bike assignments, expected stock-out cost in case of shortage, the expected time-waste cost in case of overflow and the expected transshipment cost for re-positioning bikes. Constraint (2) sets the upper bound of bike assignments to be each station's capacity k_i. Constraint (3) sets the equality of rented bike to be the difference between the demand and the realized shortage. For this problem, the realized shortage is set to be a decision variable, the optimal solution is therefore an optimistic value. Constraint (4) demonstrates the calculation of the realized surplus and shortage and constraint (5) ensures the balance between the residual capacity and overflow after bikes are being rented or redirected. Constraint (6) and (7) ensures the residual overflow at each station are being redirected and the receival of all redirected bikes at each station are within the residual capacity. After the action of rental and redirection, constraint (8) demonstrates the balance between the shortage and overflow bike amount at each station. Constraint (9) and (10) demonstrates the relation between the second stage decision variables and the amount of excessive and lack of bike amount for each station. Constraints (11) and (12) define the integrality and non-negativity of the first-stage and the second-stage variables.

3.2 Benders Cut

The following section lists out the Benders cut formulation for the two-stage problem. This includes both the benders multi-cut and single-cut model.

For the Benders multi-cut problem, decisions variables $\eta_s, \forall s \in S$ is introduced for the approximation task. Similarly, for the Benders Single-cut problem, decision variable θ is introduced.

Benders Multi Cut Master Problem

$$\text{minimize: } c\sum_{i}^{B} x_i + \sum_{s=1}^{S} p_s \eta_s \tag{13}$$

$$\text{subject to: } x_i \leq k_i, \forall i, j \in B \tag{14}$$

$$\eta_s \geq 0, \forall s \in S \tag{15}$$

Benders Single Cut Master Problem

$$\text{minimize: } c\sum_{i}^{B} x_i + \theta \tag{16}$$

$$\text{subject to: } x_i \leq k_i, \forall i, j \in B \tag{17}$$

$$\theta \geq 0 \tag{18}$$

Subproblem

$$\text{minimize: } Q_s(x) = \sum_{i=1}^{B}[v_i \sum_{j=1}^{B} I_{ij}^- + w_i O_i^- + \sum_{j=1}^{B} t_{ij}\tau_{ij}] \tag{19}$$

$$\text{subject to: } \beta_{ij} = \xi_{ijs} - I_{ij}^-, \forall i, j \in B \tag{20}$$

$$I_i^+ - \sum_{j=1}^{B} I_{ij}^- = x_i - \sum_{j=1}^{B} \xi_{ij}, \forall i \in B \tag{21}$$

$$O_i^+ - O_i^- = k_i - x_i + \sum_{j=1}^{B} \beta_{ij} - \sum_{j=1}^{B} \beta_{ij}, \forall i \in B \tag{22}$$

$$\sum_{j=1}^{B} \rho_{ij} = O_i^-, \forall i \in B \tag{23}$$

$$\sum_{j=1}^{B} \rho_{ji} \leq O_i^+, \forall i \in B \tag{24}$$

$$T_i^+ - T_i^- = k_i - O_i^+ + \sum_{j=1}^{B} \rho_{ji} - x_i, \forall i \in B \tag{25}$$

$$\sum_{j=1}^{B} \tau_{ij} = T_i^+, \forall i \in B \tag{26}$$

$$\sum_{j=1}^{B} \tau_{ji} = T_i^-, \forall i \in B \tag{27}$$

$$x_i, I_i^+, O_i^+, O_i^-, T_i^+, T_i^- \in \mathbb{Z}^+, \forall i \in B \tag{28}$$

$$\tau_{ij}, \beta_{ij}, \rho_{ij}, I_{ij}^- \in \mathbb{Z}^+, \forall i, j \in B \tag{29}$$

4 Multistage Formulation

This section lists out the formulation of a multi-stage BSS problem. While the previous two-stage problem assumed the bike rental is only once per day, the multi-stage formulation splits the rental demand of each origin-destination pair into three rental periods in morning, afternoon, and evening. The bikes are now allowed to be rented in all these three periods and the transshipment are performed over night to re-balance the number of bikes at each station. The available bikes at each station in each period is a function of the initial assignment, rentals in the periods before, including the re-directions and transshipment operations.

$$\text{minimize: } c\sum_{i}^{B} x_{i0} + \sum_{n=1}^{N} p_n\Big[\sum_{i=1}^{B}\Big[v_i\sum_{j=1}^{B} I_{ijn}^{-} + w_i O_{in}^{-}\Big]\Big] + \sum_{n=N-F+1}^{N} p_n\Big[\sum_{i=1}^{B}\sum_{j=1}^{B} t_{ij}\tau_{ijn}\Big] \tag{30}$$

$$\text{subject to: } x_{i0} \leq k_i, \forall i \in B \tag{31}$$

$$\beta_{ijn} = \xi_{ijn} - I_{ijn}^{-}, \forall i,j \in B, n \in \mathcal{N}\backslash\{0\} \tag{32}$$

$$I_{in}^{+} - \sum_{j=1}^{B} I_{ijn}^{-} = x_i - \sum_{j=1}^{B}\xi_{ijn}, \forall i \in B, n \in \mathcal{N}\backslash\{0\} \tag{33}$$

$$O_{in}^{+} - O_{in}^{-} = k_i - x_i + \sum_{j=1}^{B}\beta_{ijn} - \sum_{j=1}^{B}\beta_{ijn}, \forall i \in B, n \in \mathcal{N}\backslash\{0\} \tag{34}$$

$$\sum_{j=1}^{B}\rho_{ijn} = O_{in}^{-}, \forall i \in B, n \in \mathcal{N}\backslash\{0\} \tag{35}$$

$$\sum_{j=1}^{B}\rho_{jin} \leq O_{in}^{+}, \forall i \in B, n \in \mathcal{N}\backslash\{0\} \tag{36}$$

$$T_{in}^{+} - T_{in}^{-} = k_i - O_{in}^{+} + \sum_{j=1}^{B}\rho_{jin} - x_i, \forall i \in B, n \in F \tag{37}$$

$$\sum_{j=1}^{B}\tau_{ijn} = T_{in}^{+}, \forall i \in B, n \in F \tag{38}$$

$$\sum_{j=1}^{B}\tau_{jis} = T_{is}^{-}, \forall i \in B, n \in F \tag{39}$$

$$x_{i0} \in \mathbb{Z}^{+}, \forall i \in B \tag{40}$$

$$\tau_{ijn}, T_{in}^{-}, T_{in}^{+} \in \mathbb{Z}^{+}, \forall i,j \in B, \forall n \in F \tag{41}$$

$$\beta_{ijn}, \rho_{ijn}, I_{ijn}^{-}, I_{in}^{+}, O_{in}^{+}, O_{in}^{-} \in \mathbb{Z}^{+}, \forall i,j \in B, n \in \mathcal{N}\backslash\{0\} \tag{42}$$

In the multistage formulation, the paper introduces the following notations:

$\mathcal{N} = \{n : n = 0, 1, ..., N\}$: ordered set of nodes of the scenario tree
$0 \in \mathcal{N}$: root node
$\mathcal{F} = \{n : n = N - F + 1, ..., N\}$: leaf nodes of the scenario tree
$\text{pa}(n)$: parent node of n
x_{i0} : total number of bikes to assign to bike-station $i \in B$ at the beginning of the service
p_n : probability of node n defined by $\begin{cases} \frac{1}{\|\mathcal{F}\|} & n \in F \\ \sum_{m \in \mathcal{N}\backslash\{0\}, pa(m)=n} p_m & n \in \mathcal{N}\backslash\mathcal{F} \end{cases}$

5 Computational Results

All the source code for data geneneration, computational results, and qualitative results can be found on Git repository [3]

[3] `https://github.com/TinaBBB/Stochastic_optimization_modles_BSP`

5.1 Data Generation

While the value of the deterministic parameters were provided by the paper, the stochastic demand data $\xi_{ijs} \in \Xi \subset \mathbb{Z}^+$ for each bike-station (i, j) pair needs to be generated following a probabilistic distribution.

The dataset being used for the original paper were generated from 4 different selected distributions based off a limited number of historical data provided by the bike-sharing service "LaBiGi" in Bergamo, Italy (mobile app: `https://www.atb.bergamo.it/it/mobilita/bike-sharing`) . Since this historical data is not publicly accessible, the parameters used for the distribution of the stochastic demands were randomly generated within a range that produces reasonable demands and feasible optimal solutions.

5.1.1 Probability distributions of stochastic demands

This project proposed four types of possible probability distributions:

- Uniform distribution $\mathcal{U}$
- Exponential distribution $\mathcal{E}$
- Normal distribution $\mathcal{N}$
- Log-normal distribution $\mathcal{L}$

For simplicity, all four distributions share the same mean $\bar{\xi}_{ij}$ and standard deviation σ_{ij} parameters.

5.1.2 Realistic stations capacity and fixed stations capacity

Although the original paper has provided the system's capacities as presented in Table 1, this capacity level led to the problem of forcing the optimal assignment x_{i0} to each station's available capacity. In order to better understand how the different probability distribution for the stochastic demand affects the optimal solutions, the deterministic value of capacity for each station were set to 50 as well. Its computational results were compared with the realistic station capacity in Table 2 and Table 3.

Table 1: Capacity k_i of each bike-station

	$Station_i$											
	A	B	C	D	E	F	G	H	I	J	K	
k_i	22	10	10	17	19	12	20	8	8	20	10	
	$Station_i$											
	L	M	N	O	P	Q	R	S	T	U	V	Total
k_i	19	8	10	10	10	18	10	8	12	12	10	283

5.1.3 Number of scenarios

Number of scenarios S = 100, 300, 500, 700, 900, 1000 has been considered and used to produce stochastic demand data for further analysis. The probability of each individual scenario is drawn from the Dirichlet distribution so that $\sum p_s = 1$.

5.2 Two-Stage problem in Extensive Form

After the data generation process, the defined two-stage problem was solved in its extensive form under different scenario numbers and using different probabilistic distribution for the stochastic demand. The optimal objective values (i.e. system's total cost) is shown in Table 2 and Table 3.

The computational results by using different station capacities suggested that given the settings in this project, the number of scenarios does not affect the optimal solution significantly. When comparing between different probabilistic distributions, utilizing the stochastic demands generated by Normal distribution yields the best optimal results (the least objective value). Using log-normal distributed stochastic demands yields the "worst" results among the 4 potential distributions. This is due to the shape of the exponential distribution that higher standard deviation leads to higher probability of generating a large demand. Therefore the optimal assignment under a log-normal distributed stochastic demands was forced to be the capacity of each station as shown in Table 4.

As mentioned in the previous section, in order to avoid the limitation caused by the initial station capacity settings, the team fixed each station's capacity to be 50 and investigated the computational results under the fixed capacity.

As shown in table 3, without the limitation from the bike station capacity, the overall optimal objective value has decreased. This objective value decrement indicates a prevention of shortage cost and redirection cost enabled by the incremented bike station capacity and initial assignment.

The results from this section matched with our initial expectation of having a normally distribution stochastic demand helps to analyse the problem in a more conservative and consistent way, while using a log-normal distribution gives a different perspective of having more unexpected demands than the average expectations. This observation is slightly different from the referred paper that this project is based on. The original paper had the exponential distribution as the "outlier" for producing pessimistic results. This difference is due to the different dataset being used, using a different value of standard deviation would lead to opposite behaviours between the log-normal and exponential distributions.

Based on the observations from the extensive form results, further model implementation is going to be using Normally distribution stochastic demands with 100 scenarios and the realistic station capacities.

Table 2: Optimal Solution with given capacity

	Scenario numbers					
	100	300	500	700	900	1000
$\mathcal{U}$	1283.99971	1283.999856	1284.00036	1283.99964	1284.000432	1283.99993
$\mathcal{E}$	726.3494365	726.6937227	723.7350438	726.1084282	727.5106984	723.75143
$\mathcal{N}$	563.9447413	562.3745437	562.3240139	562.7263149	562.5395563	564.332514
$\mathcal{L}$	3374.611781	3375.915331	3377.054484	3379.593556	3378.074512	3374.76947

Table 3: Optimal Solution with fixed capacity (k_i = 50)

	Scenario numbers					
	100	300	500	700	900	1000
$\mathcal{U}$	924	924	924	924	924	924
$\mathcal{E}$	706.3695815	708.4676332	706.2223105	708.4844494	710.4845903	706.6015822
$\mathcal{N}$	555.3210648	553.5348993	553.5485166	554.2538413	554.0197485	555.5529113
$\mathcal{L}$	2039.383342	2039.084215	2038.313809	2039.702939	2039.317024	2037.070869

Table 4: Optimal bike station assignment (100 scenarios)

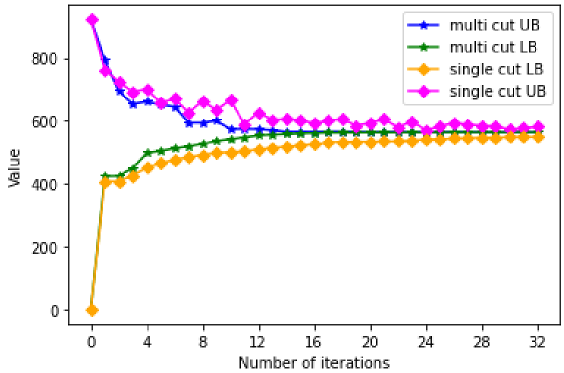
	$Station_i$											
	A	B	C	D	E	F	G	H	I	J	K	
$\mathcal{U}$	21	10	10	17	19	12	20	8	8	10	10	
$\mathcal{E}$	10	10	9	10	11	10	11	8	8	10	10	
$\mathcal{N}$	10	9	9	10	10	10	10	8	8	9	10	
$\mathcal{L}$	22	10	10	17	19	12	20	8	8	20	10	
	$Station_i$											
	L	M	N	O	P	Q	R	S	T	U	V	total
$\mathcal{U}$	19	8	10	10	10	18	10	8	12	12	10	272
$\mathcal{E}$	10	8	9	10	10	10	10	8	9	10	10	211
$\mathcal{N}$	10	8	9	10	10	10	10	8	10	10	9	207
$\mathcal{L}$	19	8	10	10	10	18	10	8	12	12	10	283

5.3 Two-Stage in Benders Cut

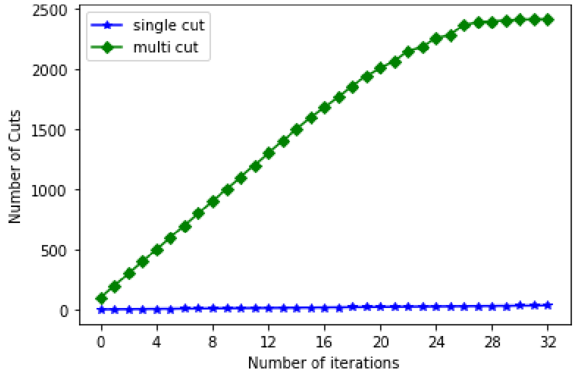
In this part, 100 scenarios data generated from normal distribution is used. In particular, we compare the convergence rate of multi-cut and single-cut version. With cut violation tolerance set to 0.0001, multi-cut algorithm finishes in 33 iterations and generates 2414 cuts in total. The single-cut algorithm doesn't finish in 400 iterations, but the lower bound and upper bound are close. We visualize the convergence rate of two algorithms in Figure 2. From the graph, it can be observed that the multi-cut algorithm converges faster than single-cut version, although single cut algorithm also restricts the optimal objective value to a reasonably small region. As a trade-off, multi-cut algorithm adds far more constraints, causing computational inefficiencies.

Table 5: Optimal bike station assignment (100 scenarios, $k_i = 50$)

	$Station_i$											
	A	B	C	D	E	F	G	H	I	J	K	
$\mathcal{U}$	21	21	21	21	21	21	21	21	21	21	21	
$\mathcal{E}$	12	12	11	11	11	12	12	12	11	12	11	
$\mathcal{N}$	11	10	10	10	10	10	10	10	10	10	10	
$\mathcal{L}$	44	44	45	45	44	44	44	45	45	45	45	
	$Station_i$											
	L	M	N	O	P	Q	R	S	T	U	V	total
$\mathcal{U}$	21	21	21	21	21	21	21	21	21	21	21	462
$\mathcal{E}$	11	12	10	11	12	11	11	12	10	11	11	249
$\mathcal{N}$	10	11	10	10	10	10	11	8	10	11	10	224
$\mathcal{L}$	45	44	45	44	45	44	44	44	45	44	44	978



(a) Convergence rate of Benders Cut



(b) Number of cuts

Figure 2: Convergence rate for Benders Cut

5.4 Multi-stage Node-Based Extensive Form

The multi-stage formulation considers three stages: morning, afternoon and evening. Constraints are almost identical to those of the two-stage extensive formulation. One exception is that the bikes are only transshipped in the last stage. In our computational experiment, branching factor is 8. Hence, there are 585 nodes in total and 512 leaf nodes. The data again is generated from the normal distribution. The optimal bike assignment is displayed in the Table 6. The result is not very different from the two-stage solution since the demand is sampled from the same distribution.

Table 6: Multi-stage optimal bike station assignment

	$Station_i$											
	A	B	C	D	E	F	G	H	I	J	K	
x_i	11	10	10	11	12	10	11	8	8	11	9	
	$Station_i$											
	L	M	N	O	P	Q	R	S	T	U	V	Total
x_i	11	8	10	10	10	11	10	8	10	10	10	219

5.5 VSS & EVPI

In this section, we analyze the Value of the Stochastic Solution (VSS) and Expected Value of Perfect Information (EVPI). VSS indicates benefit gained by considering the stochasticity of the demand rather than its deterministic counterpart (where the random demand is replaced by its expectation). Mathematically, it's defined as

$$VSS = v^{MV} - v^{S}$$

where v^{MV} denotes the long run cost using the optimal solution to expected demand and v^S denotes the objective of our stochastic program.

Table 7: VSS with normal distribution

	Scenario numbers					
	100	300	500	700	900	1000
$\mathcal{N}$	358.785	361.098	359.326	361.500	359.085	361.362

EVPI measures the maximum amount a decision maker would be willing to pay in return for complete and accurate information about the future. It's defined as

$$EVPI = v^S - v^{PI}$$

where v^S denotes the value of the stochastic program and v^{PI} denotes the expected value of the optimal 'wait-and-see' solutions.

Table 8: EVPI with normal distribution

	Scenario numbers					
	100	300	500	700	900	1000
$\mathcal{N}$	40.412	38.253	39.663	39.038	39.424	39.268

From the analysis from two stage extensive formulation, the results from one hundred to one thousand scenarios are close. It can also be observed from EVPI & VSS that one hundred scenario is a good approximation of reality. However, VSS is lightly larger when there are more scenarios. As number of scenarios increases, the deterministic equivalent form captures more information of random demand. Hence, it makes sense that the value of the stochastic solution is larger. For EVPI, the same analysis applies. As more randomness is incorporated in the model, the value of perfect information decreases.

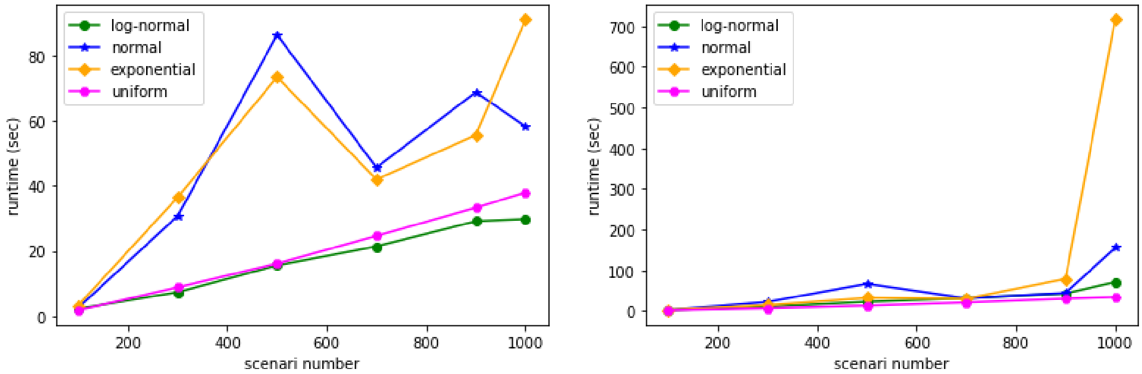
5.6 Run Time

The runtime under different use cases by solving the extensive form has also been investigated in Figure 3. The plots show a general trend of linearly increasing computation time with the number of scenarios, which is expected. The normal and the exponential solution experienced an immediate raise of the computation time at 500 scenarios. Based on the optimal solution computed from the previous section, with 500 scenarios, the model produced the most optimal results among the use cases of different scenario numbers when using the normal and exponential distribution. Therefore the unexpected rise of computation time at 500 scenario numbers could be explained. It is also notable that using the exponentially distributed stochastic demand generally resulted in the greatest computation time among the four distributions. Using a uniformly distribution stochastic demand results in the most consistent computation time history, which could be used as a baseline.

5.7 Discussion

Based on the computational results in the previous sections, while the scenario number did not affect the results significantly, using a normally distributed stochastic demand gave the best optimal solutions among the four potential probabilistic distributions. Given the setting for this project, this is considered to be an optimistic result since the realized overflow and slack amount of bikes at each station were considered to be a decision variable. In other word, the realized overflow and slack amount were "selected" by the system such that the objective value is minimized. Bike sharing systems' rental actions run in a continuous time frame in reality, which is impossible to be modelled in a two-staged stochastic problem. The set up of the two-stage problem in this project was modelling the realized overflow and lack of bikes using a single time frame. Such setting will loose the continuity of rental actions and the intermediate results from the realistic point of view of BSS problems.

The team encountered the issue of distribution parameter selection during the data generation process. This is caused by the capacity limits given by the original paper. A relatively large parameter value would push the optimal bike assignment to the limit of the stations capacity. This also suggests that the BSS problems should be modelled in a way that is closer to the reality of performing rental actions under a continuous time frame (e.g. multi-staged stochastic programming). Alternatively, the team suggested that the realized overflow and slackness can also be generated using stochastic simulation instead of being modelled as decision variables. In this way, the optimal solution would grant a more realistic view of the problem.



(a) Runtime for provided station capacity

(b) Runtime for fixed station capacity

Figure 3: Runtime for BSP

During the re-implementation of the multi-stage formulation for the BSS problem, the team noticed that the decision variables at each stage were modelled independently from the previous stage (i.e. number of realized rented bikes are not limited by the previous stage results). Further adjustment for the multi-stage model should be applied.

Conclusion and Future Work

In this paper, we presented a two-stage formulation of the bike-sharing problem to optimally decide the number of bikes for each station and to minimize the cost caused by overflow, shortage, and transshipment. Computationally, we solved the problem in the extensive form and Benders Cut (both multi-cut and single cut). Although the extensive formulation is solved faster than Benders Cut in our experiment, we expect the extensive form fails when more scenarios are involved. Utilizing two-stage extensive formulation, we analyze the effect of demand distribution on our decision variables and objective values. We conclude that the normal distribution with 100 scenarios is a good representation of reality. In addition to two-stage model, an extension to multi-stage is made by considering morning, afternoon and evening. However, as discussed in the previous section, the multi-stage model fails to build connections between stages, which can be extended in the future. The model in the paper considers only static transshipment and doesn't consider the profit generated by the travelling distance between stations. Usually, the service provider gains more profit when the user takes a longer ride. In the future, we could incorporate the distance information between stations and dynamic transshipment to avoid shortage and overflow problem.

References

Francesca Maggioni, Luca Bertazzi Stein W. Wallace, Matteo Cagnolari. 2018. Stochastic optimization models for a bike-sharing problem with transshipment .

Qiong Tang, Dezhi Zhang Hao Guo, Zhuo Fu, Minyi Li. 2020. Addressing the bike repositioning problem in bike sharing system: A two-stage stochastic programming model .

Ramon Alvarez-Valdesa, Jose M. Belenguera Enrique Benaventa Jose D. Bermudeza Facundo Mu~noza Enriqueta Verchera Francisco Verdejo. 2014. Optimizing the level of service quality of a bike-sharing system .