The University of Hong Kong

STAT3606 Business Logistics

Group Project Report

Hong Kong hotspots closed route simulator using travelling salesman problem model

Ho Yiu Yuen	3035786158	Sy Hei Yeung	3035856877
Lam Chi Shing	3035891305	Lau Yan Chun Chris	3035790941

(Words: 1985)

Introduction

Due to the outbreak of COVID-19 pandemic, tourism in Hong Kong has been severely stroke due to series of quarantine measures. Recently the quarantine arrangements have been shortened, which facilitate tourism to recover gradually by time. To assist the visitors to plan their trip in Hong Kong, we hereby provide a travel route planner with the database contains time and money required for visitors travelling between hotspots. The objective of this project is to provide the defined-optimal route by linear programming with the model of travelling salesman problem to give the best closed route for the visitors while considering travelling in Hong Kong. Further possibilities for expansion of the simulator including sensitivity analysis will also be discussed.

Methodology

To investigate the optimal travelling route, the first step is to determine both the monetary and time cost induced by each travelling path. The process will be described in the following section "Data Collection". Afterwards, we would formulate our problem into a linear programming problem and solve it by using Excel Solver. Finally, we would analyze the results and give a summary.

Data Collection

20 Hong Kong popular hotspots are selected to be recorded in the database. It is flexible to define the expected staying time and expenditure depending on the preference of the visitor. To facilitate the demonstration, here both items are assigned with different values for each hotspot. The table 1.1 shows the expected staying time and expenditure of hotspots.

Ref. no	Hotspot	StayTime (minute)	Exoected expenditure (HKD)
1	Airport	15	0
2	Causeway Bay	30	50
3	Cheung Chau	120	100
4	Chuen Lung Street	30	50
5	Disneyland	240	400
6	HKU	30	30
7	Hong Kong Space Museum	60	30
8	M+	90	100
9	Mai Po Nature Reserve	90	0
10	Mong Kok	45	50
11	Orean Park	150	200
12	Sai Kung	90	0
13	Sham Tseng	60	50
14	Shek O	60	30
15	Stanley	90	30
16	Tai Kwun, Central	45	60
17	Tai On Night market, Sai Wan Ho	60	50
18	Tai Po Insect House	60	0
19	The Mills	90	80
20	Tian Tan Buddha	120	50

Table 1.1 Expected staying time and expenditure for different hotspots

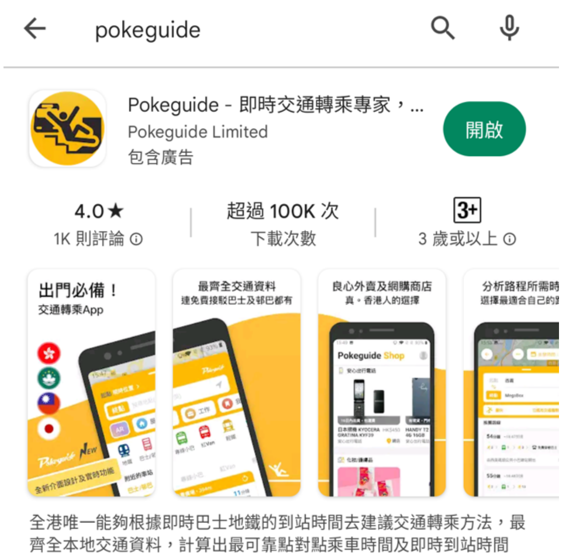
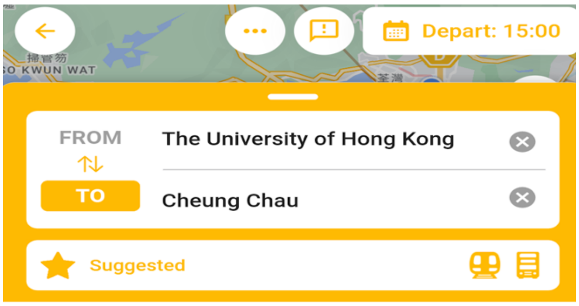


Figure 2.1 app “Pokeguide” in app store

To collect the monetary cost and time required in each path, the app “Pokeguide” shown in figure 2.1 is used. Unlike other popular path searchers, “Pokeguide” not only provides a few possible options in term of time, but also the monetary costs for each possible paths that facilitates data collecting. The figure 2.2 and 2.3 shows the interface of the app that suggest different paths accordingly.



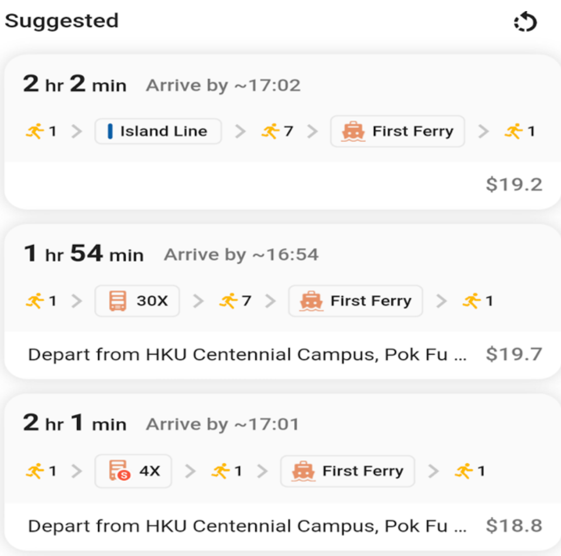


Figure 2.2 interface in app "Pokeguide"

Figure 2.3 path suggested by app "Pokeguide"

Since both time cost and monetary cost may vary in term of time, a local time 15:00 is selected as the departed time for traveling all hotspots. Finally, the best path defined by the app is recorded to our database. For instance, the app can search for the costs induced by the path from HKU to Cheung Chau. All information regarding both directions is recorded. The table 3.1 and table 3.2 explicitly show the time cost and monetary cost for traveling different hotspots respectively. Notice that the first suggestion is clearly not the route with minimal costs. However, as we are investigating the optimal route given the costs of each route, choosing its first suggestion wouldn't matter.

Time cost in minute	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
From \ To	Airport	Causeway Bay	Cheung Chau	Lung Street	Disneyland	HKU	Space Museum	M+	Nature Reserve	Mong Kok	Orean Park	Sai Kung	Sham Tseng	Shek O	Stanley	Kwun, Central	Sai Wan Ho	Insect House	The Mills	Tian Tan Buddha
Airport	10000	60	130	53	44	65	47	56	115	57	101	112	62	98	107	58	82	61	49	70
Causeway Bay	62	10000	118	47	49	27	24	47	115	28	51	69	53	49	45	20	19	71	56	135
Cheung Chau	133	124	10000	141	165	136	115	138	207	139	186	199	175	162	158	116	146	198	153	178
Chuen Lung Street	52	52	141	10000	44	41	42	50	92	26	79	84	30	91	82	54	66	55	16	128
Disneyland	58	58	193	46	10000	64	48	53	135	48	96	186	107	140	148	59	74	97	51	104
HKU	51	29	122	57	58	10000	32	24	121	28	58	86	55	66	64	29	31	70	57	134
Hong Kong Space Museum	57	27	115	43	42	30	10000	25	116	17	51	74	65	66	56	30	39	62	50	129
M+	34	28	138	41	49	30	30	10000	113	31	74	98	40	90	84	39	58	94	46	141
Mai Po Nature Reserve	96	116	159	93	131	118	116	109	10000	101	168	168	132	188	175	137	150	120	167	218
Mong Kok	55	26	138	35	42	33	17	32	104	10000	57	56	41	65	59	42	38	46	40	128
Orean Park	75	20	124	55	58	33	28	54	115	31	10000	75	70	65	98	40	47	79	72	155
Sai Kung	101	76	193	89	91	80	60	91	157	52	116	10000	263	100	104	80	71	91	188	171
Sham Tseng	61	49	137	29	51	57	51	38	104	43	108	174	10000	104	98	52	73	171	29	139
Shek O	98	49	348	91	91	65	67	93	158	64	133	107	111	10000	51	65	40	120	134	187
Stanley	107	54	143	82	99	57	54	78	142	60	80	110	81	52	10000	61	43	36	143	175
Tai Kwun, Central	74	27	115	50	49	25	24	31	117	28	52	89	43	65	62	10000	40	77	45	140
Tai On Night market, Sai Wa	73	22	135	66	66	39	36	61	128	37	63	77	79	40	46	36	10000	77	58	155
Tai Po Insect House	61	61	179	55	91	83	58	74	101	46	100	101	177	102	106	74	89	10000	59	142
The Mills	69	50	154	16	51	45	35	49	139	36	81	170	27	109	145	42	58	72	10000	133
Tian Tan Buddha	137	174	227	128	104	153	137	139	195	137	175	192	148	192	180	148	183	183	133	10000

Table 3.1 table of time cost travelling around hotspots

Transportation Fee	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
From \ To	Airport	Causeway Bay	Cheung Chau	Lung Street	Disneyland	HKU	Space Museum	M+	Nature Reserve	Mong Kok	Orean Park	Sai Kung	Sham Tseng	Shek O	Stanley	Kwun, Central	Sai Wan Ho	Insect House	The Mills	Tian Tan Buddha
Airport	10000	110	123.6	10.8	65	45	65	45	19.5	37	65	33.2	18.9	120	117	65	45	27.7	76.8	173.9
Causeway Bay	65	10000	22.5	14.9	26.5	7	12.5	10.9	33.4	12.5	5.6	25.3	21.8	18.7	10.8	5.6	7	24.7	23	50.1
Cheung Chau	130.8	22.5	10000	28.5	47.9	19.2	16.3	23.7	47	26.1	28	38.2	36.9	29.5	26.4	19.9	22.3	39.2	31.6	40.4
Chuen Lung Street	18.9	14.9	28.5	10000	17.7	14.9	10.5	10.5	19.8	13	14.9	20.2	6.5	24.9	27.7	14.9	14.9	10.7	3.8	42
Disneyland	65	15.2	34.6	17.7	10000	26.5	21	21	42.5	21	26.5	29.9	7	26.5	26.5	26.5	26.5	28.8	17.7	42
HKU	110	7	19.2	14.9	26.5	10000	12.5	11.6	33.4	11.6	7	25.3	26.5	18.7	7.3	5.6	8.7	21.8	18.7	50.1
Hong Kong Space Museum, T	65	12.5	16.3	10.5	21	12.5	10000	6.8	42.2	5.6	12.5	14.4	18.8	24.9	22.9	10.1	14.9	11.7	10.5	44.7
M+, West Kowloon Cultural I	65	10.9	23.7	18.8	21	13	0	10000	26.2	6.2	12.5	18.4	16	23.1	16.2	10.4	19	12.7	10.2	17.7
Mai Po Nature Reserve	39.9	33.4	58.1	19.8	42.5	40.3	42.2	33.1	10000	32.1	51.8	26.2	32.9	39.5	41.8	40.3	41.2	27.3	45.9	69.1
Mong Kok	33	12.5	26.1	8.7	21	13	5.6	5.4	25.3	10000	12.5	18	23	19.2	25.3	14.9	12.5	14	12	44.7
Orean Park	65	5.8	19.4	14.9	26.5	7	12.5	12.5	34.3	12.5	10000	30.5	27	18.7	19.8	5.6	8.7	21.8	14.9	53.5
Sai Kung	33.2	20.7	44.8	17.7	32.8	21.9	15.7	15.7	26	18	25.8	10000	22.7	33.5	33.5	21.9	19.5	17.9	13.6	55
Sham Tseng	18.9	21.8	35.4	6.5	24	26.5	16.9	21.8	26.2	22	28.8	22.7	10000	38.7	35.5	21.8	25.1	22.8	6.5	50.4
Shek O	120	18.7	26.5	24.9	36.5	18.7	24.9	23.1	42.5	19.8	10.5	33.5	35.2	10000	18.8	18.7	14.7	33.8	14.1	61.7
Stanley	117	17.5	2.64	27.7	39.3	16.2	22.9	16.2	46.2	19.5	10.7	33.5	32.1	18.8	10000	9.8	10.6	113	14.1	62.9
Tai Kwun - Centre for Herita	40	5.6	21.8	14.9	26.5	5.6	10.1	21.9	33.4	12.5	5.6	22.1	21.8	18.7	17.5	10000	7.7	18.9	18.8	50.1
Tai On Building Night marke	110	6.7	18.3	14.9	26.5	8.7	14.9	14.9	34.3	13	8.7	20.7	27	10	10.5	7.7	10000	21.8	14.1	53.5
Tai Po Insect House, Tai Po	27.7	24.7	43.3	10.7	28.8	21.8	11.7	11.7	27.3	10.7	21.8	17.1	22.8	37.7	31.7	24.7	21.8	10000	16.9	40.9
The Mills, Tsuen Wan	68.8	23	31.6	3.8	17.7	19	10.5	10.5	45.9	8.7	14.9	13.6	6.5	28.6	14.1	16	17.9	10.7	10000	14.2
Tian Tan Buddha	92	15.2	31.1	42	42	50.1	44.7	17.7	46.5	44.7	26.5	57.7	50.4	61.7	62.9	50.1	53.5	25.8	14.2	10000

Table 3.2 table of monetary cost travelling around hotspots

Notice that the matrices are not symmetrical as the app may suggest an alternative route for the reverse direction. The diagonal entries are supposed to take value 0 but for the convenience of simulation, a relatively large value of 10000 is set.

Model Formulation

We first define the indicator X_{ij} as follows:

$X_{ij} = 1$ if we choose to go from spot i to spot j ;

$X_{ij} = 0$ otherwise

where i and j are positive integers at most 20. The airport, which is the starting spot, is denoted by spot 1.

In addition, we define T_{ij} and C_{ij} to be the time and monetary costs incurred respectively while travelling from spot i to j , which are known from the previous section. The staying time and expected expenditure in spot i are denoted by T_i and C_i respectively.

The model can be visualized by a diagram. An example of 4 spots is shown in figure 4.1:

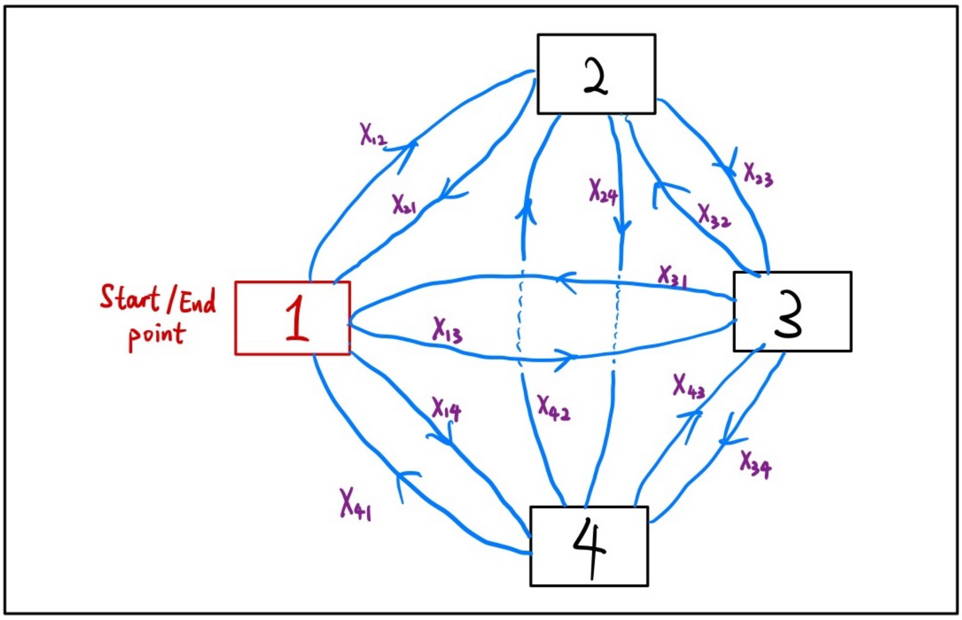


Figure 4.1 example of network graph of model contains 4 hotspots (1,2,3,4)

There are three possible objective functions to be investigated:

1. Minimize the total travelling and staying time incurred, that is:

$$\arg\min \sum_{(i,j)} (T_{ij} + T_i) X_{ij}$$

2. Minimize the total monetary cost and expected expenditure at hotspots incurred:

$$\arg\min \sum_{(i,j)} (C_{ij} + C_i) X_{ij}$$

3. Maximize the number of spots visited:

$$\arg\max \sum_{(i,j)} X_{ij}$$

The constraints are as follows:

1. There is no route going back to the same place, that is, for all i, $X_{ii} = 0$ (this is ensured by setting a value of 10000 in both matrices mentioned in data collection).
2. Every spot has a unique preceding and succeeding spot, that is, for each spot k, $\sum_i X_{ik} \leq 1$ and $\sum_j X_{kj} \leq 1$.
3. We require both the starting point and ending point are the Airport, which is spot 1. Hence, $\sum_j X_{1j} = 1$ and $\sum_i X_{i1} = 1$.
4. If a spot is visited, there is a one and only one path towards AND away from that spot, that is for all spot k, $\sum_i X_{ik} + \sum_j X_{kj} = 0 \text{ or } 2$. As Excel cannot interpret "or" statements, we can reformulate the constraint as $\frac{1}{2}\left(\sum_i X_{ik} + \sum_j X_{kj}\right) \leq 1$ by setting $\frac{1}{2}\left(\sum_i X_{ik} + \sum_j X_{kj}\right)$ to be integer-valued.
5. X_{ij} are integer-valued and $0 \leq X_{ij} \leq 1$.

Notice that the equations in constraint 3 only implies the visitor would "visit" the airport. This does not necessarily mean every path will start and end at the airport. This results in the problem of "subtour" appeared in simulation.

Result of Simulation

Minimum Time Cost

For the first case we simulated, no specific requirement for the hotspots is selected. 5 random hotspots are selected, hotspots (1) Airport are further selected as the starting and ending hotspot of the trip. The objective of the simulation was set as minimum time cost. The information of hotspots and the corresponding transportation time cost are shown in table 5.1 and figure 5.2 respectively.

Selected	Ref. No	Name
1	1	Airport
2	7	Hong Kong Space Museum
3	10	Mong Kok
4	14	Shek O
5	16	Tai Kwun, Central
6	6	HKU

Table 5.1 Table of hotspots selected with reference number.

Transportation Time Matrix							
		1	6	7	10	14	16
		Airport	HKU	Hong Kong Space	Mong Kok	Shek O	Tai Kwun, Central
1	Airport	10000	65	47	57	98	58
6	HKU	51	10000	32	28	66	29
7	Hong Kong Space Museum	57	30	10000	17	66	30
10	Mong Kok	55	33	17	10000	65	42
14	Shek O	98	65	67	64	10000	65
16	Tai Kwun, Central	74	25	24	28	65	10000

Table 5.2 Table of time cost for transportation

Solution Matrix		1	6	9	10	14	16	Objective	679 minutes		
Hotspots Selected	From \ To	Airport	HKU	Mai Po Nature Reserve	Mong Kok	Shek O	Tai Kwun, Central	Constraints			Stay Time
1	Airport	0	0	1	0	0	0	1	=	1	15
6	HKU	0	0	0	0	0	1	1	=	1	30
9	Mai Po Nature Reserve	1	0	0	0	0	0	1	=	1	90
10	Mong Kok	0	0	0	0	1	0	1	=	1	45
14	Shek O	0	0	0	1	0	0	1	=	1	60
16	Tai Kwun, Central	0	1	0	0	0	0	1	=	1	45
	Constraints	1	1	1	1	1	1				
		=	=	=	=	=	=				
		1	1	1	1	1	1				

Figure 5.3 First result of simulation of case (1,6,9,10,14,16) using Excel Solver

The first result is generated by Excel Solver shown in figure 5.3 . Although the optimal solution is reached and all constraints are fulfilled, the path is closed separately. Instead of a complete closed path that contains all hotspots selected, the hotspots, in this case, are paired up and formed as three private closed paths ((1,9), (6,16), (10,14)). Since the starting hotspot is in one of the private closed paths, the traveler can never visit all hotspots selected in this plan.

The private closed path is well-known as *subtour* in travelling salesman problem model (Ahuja, 1993). By observation, the number of paths token in subtour is always equal to the number of hotspots in the subtour. To break the subtours, the available number of paths token should be enforced to be less than the number of hotspots in the subtour. The branch-and-cut method is therefore introduced to cut the subtour manually as follows (Karamanov & Miroslav, 2006):

General Form of Branch-and-cut method:

$$\sum_{(i,j)\in S_k} X_{ij} \leq |S_k| - 1$$

Denoted that S_k be the k-th subtour, $|S_k|$ be the size of k-th subtour.

Constraints for this case:

$$X_{1,9} + X_{9,1} \leq 2 - 1$$

$$X_{10,14} + X_{14,10} \leq 2 - 1$$

$$X_{6,16} + X_{16,6} \leq 2 - 1$$

Solution Matrix		1	6	9	10	14	16	Objective	698 minutes		
Hotspots Selected	From \ To	Airport	HKU	Mai Po Nature Reserve	Mong Kok	Shek O	Tai Kwun, Central	Constraints			Stay Time
1	Airport	0	0	0	0	0	1	1	=	1	15
6	HKU	0	0	0	0	1	0	1	=	1	30
9	Mai Po Nature Reserve	1	0	0	0	0	0	1	=	1	90
10	Mong Kok	0	0	1	0	0	0	1	=	1	45
14	Shek O	0	0	0	1	0	0	1	=	1	60
16	Tai Kwun, Central	0	1	0	0	0	0	1	=	1	45
	Constraints	1	1	1	1	1	1	Subtour Constraint			
		=	=	=	=	=	=	1	<=	1	(1,9)
		1	1	1	1	1	=F6+E9+C8			2	(1,10,9)

Figure 5.4 Final optimal result of case (1,6,9,10,14,16)

After applying the subtour constraint of (1,9), a new solution with subtour of (1,10,9) is reached. With further adding a new subtour constraint, the optimal solution is finally reached with the minimum time cost of 698 minutes shown in figure 5.4 . The optimal closed route is as $1 \rightarrow 16 \rightarrow 6 \rightarrow 14 \rightarrow 10 \rightarrow 9 \rightarrow 1(closed)$ with optimal time required of 689 minutes. Without the Excel Solver, the same optimal solution can be obtained by Hungarian method with branch-and-bound method. After applying the Hungarian method, two optimal solutions are obtained with number of subtours respectively in this case. To break the subtour, the branch-and-cut method is transformed to union relationship to fit in the branch-and-bounds method as follows:

$$\left(\sum_{(i,j)\in S_k} X_{ij} \leq |S_k| - 1\right) \Rightarrow \left(\bigcup_{(i,j)\in S_k} X_{ij} = 0\right)$$

Branch-and-bounds method: $(X_{1,9} + X_{9,1} \leq 2 - 1) \Rightarrow X_{19} = 0 \cup X_{91} = 0$	
Layer 1 (original)	Layer 2
Subtour: (1,9),(6,16),(10,14) Min. time: 679 minutes	**Case 1**: $X_{19} = 0$: Optimal route: 1->16->6->14->10->9->1 Min. time: 698 minutes **Case 2**: $X_{91} = 0$: Optimal route: 1->9->10->14->16->6->1 Min. time: 707 minutes
UB: 679 minutes LB: 776 minutes	UB: 697 minutes LB: 707 minutes

The solution reached by Branch-and-bounds method is same as the one solved by Solver with 697 minimum minutes.

Minimum Monetary Costs

In the sense of mathematics, the model setting with the objective of minimum monetary costs is identical to the one with the objective of minimum time cost. However, one reasonable inference is that the optimal routes with two different objectives may not be the same due to the different value distributions.

$$Optimal\ Route: 1 \rightarrow 9 \rightarrow 10 \rightarrow 14 \rightarrow 16 \rightarrow 6 \rightarrow 1 (closed)$$

$$(Optimal\ minimum\ monetary\ costs: 297.6\ HKD)$$

In our case, the optimal route with objective of minimum monetary costs is reached using Excel Solver. The result validated the hypothesis that different objectives lead to different optimal solution with the same data framework as expected.

Sensitivity Analysis

Maximum number of hotspots

Apart from objectives of minimum cost, it is also interesting to study how many numbers of hotspots the traveler is capable to visit given limited time or budget.

Given the objective of minimum time cost, sensitivity analysis should be conducted to understand how much extra time is required to visit one more hotspot. To achieve such analysis, the model contains one extra targeted hotspot should be solved. Similarly, it also can provide a clear image to the answer of how much time saved for not visiting one of the hotspots targeted by subtracting it from the pool.

Here the sensitivity analysis of adding one more hotspot is defined as *forward analysis*, while the one of subtracting one of the hotspots from the pool is defined as *backward analysis*, come up with the idea of variable selection in regression model.

Sensitiity Analysis	Forward (One more hotspots)			Backward (One less hotspots)		
	Optimal	%Diff.	Select	Optimal	%Diff.	Select
Ori.	698	-	-	698	-	-
2	735	5.30%	Best			
3	1004	43.84%				
4	751	7.59%				
5	973	39.40%				
6	/	/		504	-27.79%	
7	746	6.88%				
8	796	14.04%				
9	/	/		458	-34.38%	Best
10	/	/		645	-7.59%	
11	903	29.37%				
12	883	26.50%				
13	790	13.18%				
14	/	/		536	-23.21%	
15	789	13.04%				
16	/	/		631	-9.60%	
17	736	5.44%				
18	763	9.31%				
19	816	16.91%				
20	949	35.96%				

Figure 6.1 Forward and backward sensitivity analysis for case (1,6,9,10,14,16)

The figure 6.1 shows the result of both forward and backward analysis for the case of (1,6,9,10,14,16), which indicates that hotspot (2) is the best hotspot if the travelers tend to visit one more hotspot, and the hotspot (9) is the best hotspot they should drop if they only tend to visit 4 hotspots.

Meanwhile, it is obvious that hotspot (2) may not always be the best spot for all model with size of 5, complete simulation is required to obtain the optimal solution based on different settings. To obtain the best k hotspots that reached the ultimate minimum time from scratch, the model is required to run $(N_{db}-1)(N_{db}-2)\ldots(N_{db}-k)=\left\langle{N-1 \atop k-1}\right\rangle$ times, which requires great computational cost given a large size of database N_{db}.

Limitation

As for the limitation, the volume of Excel Solver restricts the usage of our model. It can only support 200 decision variables at most that fit at most 10 hotspots for our model. More advanced software, such as GAMS, LINGO, etc., is required to solve model with greater number of hotspots.

Future work

The model can be built in a more personalized way by adding more personalized constraints to it. The hotspots can be further assigned in categories, the best route contains certain number of hotspots in specific categories can be obtained accordingly.

Recalled that all data is collected by setting the departed time as 15:00, the model can become more powerful if the data regarding the cost can be varied automatically with time. If the real time data and additional factors consideration can be included in the simulation, then the model would perform more user-friendly and is turned into an application.

Summary

In conclusion, the closed route travelling simulator using linear programming with the model of travelling salesman problem can assist travel around the world to plan their visit in Hong Kong with optimal result in term of minimum time, minimum monetary cost. With the sensitivity analysis, the traveler also can determine whether specific hotspot is worthy to be included in the plan. This simulator as a fundamental framework of the route planner of hotspots, can be further combined with different features to become a consolidated travelling software application to profoundly change the idea of travel planning.

Reference

1. Ahuja R., T. Magnanti, and J. Orlin, (1993) Network Flows: Theory, Algorithms and Applications, Prentice Hall

2. Karamanov, Miroslav. “Branch and Cut: An Empirical Study.” *Carnegie Mellon University* , Sept. 2006, https://www.cmu.edu/tepper/programs/phd/program/assets/dissertations/2006-operations-research-karamanov-dissertation.pdf