EVALUATION OF ONLINE ITINERARY PLANNER AND INVESITAGATION OF POSSIBLE ENHANCEMENT FEATURES

H.M. Tam[1] & Pun-Cheng, L.S.C.[2]

Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University
([1]tamhoming@gmail.com & [2]lspun@polyu.edu.hk)

KEY WORDS: Online Itinerary Planner, Location-based Service, Web-based application, Geo-visualization, Tourist guide

ABSTRACT:

Current information available to the tourists visiting Hong Kong seems to be abundant and fragmented. Usually, individual tourists tend to plan their itinerary well ahead they arrive, either through tourist guide or online information sources (e.g. existing itinerary planning websites). With the enhancement of information technology, it is expected that online itinerary planning would be a complete supplementary to the hard copy travel guides and magazines in the future. However, current online itinerary planner overlooks the transportation link and optimum travel plan. For example, several tourist spots can be visited in a sequential manner so as to optimise the time spent. It is the place where this project plays its role. This project aims to develop a scheduling algorithm based on the Greedy Algorithm that helps prepare an itinerary for tourists visiting Hong Kong on an individual basis. Tourists' limited knowledge in the spatial extent of and transportation facilities of Hong Kong have always been obstacles of arranging an efficient (optimal) travelling plan. With the view to minimize the travelling time and maximize the time of sightseeing, shortest travelling time between tourist spots have been adopted as the principle in deriving the solution. The benefit brought by the presence of this system and the current availability of tourist information (both transport and tourist spot) will be assessed and evaluated.

1. INTRODUCTION

This chapter aims to let readers to understand the background, objectives and scope of the system. Such information will be important for readers' further understanding in the system architecture.

1.1 BACKGROUND

Current information available to tourists visiting Hong Kong is abundant. However, it is rather fragmented: you may find transport or tourist spot information available separately on different web sites. For instance, the Hong Kong Tourism Board has set up a website titled "Discover Hong Kong" for some time. It provides information from tourist spots, accommodations to dining and shopping. The information available on sites like "Discover Hong Kong" is indeed quite diverse in content. You may need some time to integrate the data you have received on your own. A tourist may have no idea how to travel between two tourist spots - they may finally come up with an itinerary that covers all the destinations, but such itinerary is not necessarily cost efficient.

Sequencing a time and cost efficient itinerary can be a hard question to tourists, especially at times when they don't have much knowledge about the complex transportation and the spatial proximity of their desired visiting places. They will not be able to make good use of the well developed transport facilities throughout their trips. It is the place where this project plays its role to help. It will help tourists schedule the optimal path (with sequence) of visiting their desired spots so as to achieve the highest efficiency and save their time and money. It is their limited time of stay in Hong Kong making every second become important. It is beyond question that different tourists will visit the same list of tourist spots at different sequence. Optimizing their travelling sequence and preparing transport suggestion will bring much convenience to them.

1.2 OBJECTIVES

This project aims to develop a scheduling algorithm that helps prepare an itinerary for tourists visiting Hong Kong on an individual basis. Tourists may select their own tourist spots. Besides deriving a solution which includes all the tourist spots selected by the user, the itinerary solution also includes suggested transport means (and also the time taken for the means), time spent on each tourist spot and so on. In short, the objectives of the system are:

1. To provide a tool for tourists to help them schedule their visit
2. To provide an itinerary solution
3. To provide transport information between tourist spots and the hotel
4. To provide maps showing the vicinity of the points of interest (both tourist spots and the hotel)
5. To suggest the time to be spent at a particular tourist spot

It is the ultimate objective of the system to serve as a complete substitute to all other collaborations of tourist information sources so that users can have their own itinerary. They may also print their itinerary far before they set off.

1.3 SCOPE OF STUDY

Hong Kong is a city with well developed transport network. This transport network includes a great variety of transport modes, e.g. railway, bus, mini-bus, taxi, tram etc. Given two places, you can usually name more than 1 means of transport connection available. As the aim of the study is to investigate the tourist scheduling algorithm, we would restrict the route chosen to be the shortest possible travelling time between two spots.

Another issue that the system should cater is the number of places (tourist spots and hotels). It is certainly impossible to include all the tourist spots in Hong Kong - there are too many. Besides, classifying a place as a "tourist spot" is a subjective decision. The system is a pilot study and therefore limits the number of tourist spots selected, yet, spreading throughout the territory. It is not only the number of tourist spots that matters, but also their spatial extent (in terms of travelling time among spots), especially when it involves cross-district travelling. With the reasons stated above, the system developed would be restricted to two issues:

1. Number of tourist spots, and;
2. Routes with the shortest time in between tourist spots

Other criteria (other than travelling time and cost) should also be put into consideration. The rationale of the scheduling algorithm will be discussed in later sections.

2. LITERATURE REVIEW

The context in this section is actually giving answers to a very simple yet difficult question: "What constitutes an itinerary planner?"

2.1 ITINERARY

An itinerary is usually referred as tourist guide, or guide book. It is a book that aims to provide information of high practical value to tourists. These may include geographic location, transport, shopping, dinning or any other information of tourist spots limited to a particular area. An itinerary may also solely refer to the way of getting from one place to another. The term "itinerary" in this paper is confined as the one that tourists can rely on during their stay in Hong Kong.

2.1.1 Information provided: "Lonely Planet" is a renowned brand of travel guidebook in the world. In the contents of the book "Hong Kong & Macau – Pick & Mix Chapter", the key contents include eating, shopping, entertainment, sports & activities and transport. (Lonely Planet, 2008) This classification is indeed sensible and reasonable – tourist spots and activities have always been tourists' utmost concern. The system developed should serve this principal function.

It is obvious that not every single tourist will purchase a guide book like "Lonely Planet". Instead, they look for free resources available on the internet or other channels. The information availability is therefore assessed.

2.1.2 Tourist spot information: The Hong Kong Government is always keen in promoting Hong Kong to the tourists and branding Hong Kong as the "Asian's World City". The site titled "Hong Kong Fun in 18 Districts" set up by the Home Affairs Department includes a list of renowned tourist spots. Hong Kong Tourism Board (HKTB) is an organization which is responsible for promoting Hong Kong to the world. Since Hong Kong is always referred as the "Shopping Paradise", the Board has set a web site titled "Where to Shop". The site includes textual information, supplemented by photos that help tourists learn about the details of individual shopping area.

2.1.3 Transport information is crucial to a tourist. Some cities, e.g. Sydney, have detailed arrival and departure timetables for every transport. It is not the case in Hong Kong – only the frequency of transport (usually given as interval, say, 8-12 minutes) will be shown publicly (though detailed departures exist for managerial purpose). Most of the transport companies in Hong Kong have made their transport information online so that the public can get the information anytime. In early 2009, the Transport Department launched the Public Transport Enquiry Service (PTES). This service provides transport route suggestions once user specifies their origin and destination (Figure 1).

Figure 1 Transport route suggestions by PTES

2.1.4 Itinerary planning service provider: There are websites that provide itinerary planning service, TripIt is one of them. It processes travel confirmation emails, weather, driving directions etc into an itinerary. The objective of the site is to save the time that travellers spent on arranging their own itinerary. TripIT may sound great to cities like San Fransico where driving between spots are usual practices. Yet, the place where the pilot study take place – Hong Kong – is a small place with 7 million people. Car rental is certainly not a good option for travellers in Hong Kong - public transport is too well developed. Narrow and congested road conditions of Hong Kong also discourage tourists from renting a car.

HKTB launched a website which provides an interface for the user to plan their itinerary. User first selects the duration of stay, time of the day for arrival and departure. This itinerary is impracticable since its time interval is only up to morning, afternoon and night sessions only. It does not cater any dining needs nor shopping needs of the tourists.

2.2 SCHEDULING ALGORIHM

The scheduling of an itinerary that takes the shortest time is actually a practical case of the Travelling Salesman Problem (TSP). Given a list of places, a salesman must find a path to visit all the places for exactly one time. TSP is a well-known NP-complete problem, meaning that there is no algorithm that can solve TSP efficiently. The exact algorithm to solve TSP is to compute all the permutation and look for the one with lowest cost. It is therefore the time taken for scheduling an itinerary with n cities is actually $O(n!)$. It works fine for a small number of cities, but becomes impracticable for cases with more than 20 places.

Heuristics and approximation algorithms are therefore devised. They give good solutions quickly. With these algorithms, extremely large problem can also be solved within a reasonable amount of time. These solutions are probably 2-3% away from the optimal solution. The common approach of these algorithms is to first construct a possible path. Then, improve the solution using different improvement algorithms, like iterative improvement or randomized improvement.

3. METHODOLOGY

An online itinerary planner should provide solutions quickly and promptly since the user is too impatient to wait. Therefore, the system devised in this paper employs the approximation algorithm with improvement tactics to compute the solution in a timely manner. Assumptions are made for better modelling of the system.

3.1 ASSUMPTIONS

Below is a list of assumptions that have been considered in the initial stage of the system design:

1. User values their time in Hong Kong of the utmost importance. They want to keep their travelling time as short as possible.
2. Users will only need to specify the hotel they stayed and the tourist spots they would like to visit. System will then schedule the itinerary accordingly.
3. Users will always set off their trip from their hotel. They also end their trip there. The hotel specified by the user will always be set as the origin and the destination of the trip.
4. Users will specify the start time and end time for their visit.
5. In case of the time slot (duration between start time and end time) is too short to finish all the visiting. The system will put the rest of the spots on another day, starting from the same time stated by the user.
6. User may need to know the time they need to get from one spot to another. The suggested staying time at a particular tourist spot should also be made available.
7. User may want to know the environment of the vicinity of the tourist spots.

3.2 IMPLEMENTATION PLATFORM

To implement the system, a free and open source cross-platform web server package - XAMPP is employed. It is a solution stack of software with open source software to run web sites on servers. It is easy to install and therefore suitable to be used as the platform of the development. Although, the time taken to run the page may differ from loading a page on a WWW server, it is the features of the system to be assessed. The efficiency (or the networking issue) is left for further enhancement of the project.

3.3 DATABASE

3.3.1 Design: With respect to the assumptions listed, the database schema of the system is to be designed accordingly. Since the basic function of the system is to schedule the itinerary and display the public transport data between tourist spots, the database developed should have fields storing corresponding information. Three tables are constructed in the database, namely SPOT, JOURNEY and SUBJOURNEY. The three tables are designed with different purposes: (Table 2 & Figure 3)

Table 2 Purposes of the three tables in the database

Table	***Purposes***
SPOT	1. Keeps details of individual tourist spots and hotels (e.g. name, district) 2. Facilitates information storage and retrieval 3. Latitude and Longitude are stored to enable the use of Google Map for accurate location display
JOURNEY	1. Stores the transport route suggestions of all combinations of origin-destination 2. Time taken to travel using the respective transport route suggestion
SUBJOURNEY	1. Stores the information of individual transport route – pickup and dropoff point will also be given. 2. Details of different legs within a journey – interchange is unavoidable in some cases.

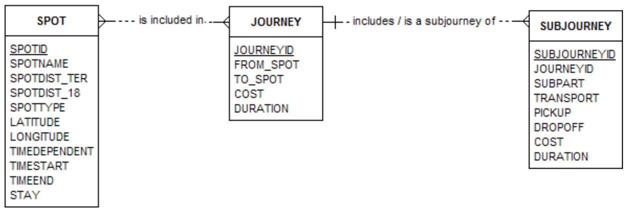
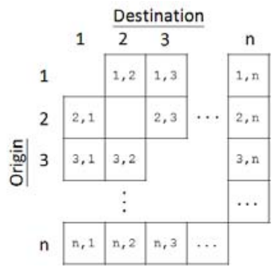


Figure 3 Entity-Relationship diagram of the system

3.4 DATA PREPARATION

3.4.1 Data Preparation: The scheduling algorithm adopted in the system uses the travelling time as the means of scheduling. Real transport data was collected through the PTES (see Section 2.1.3). Throughout the collection of data from the PTES, the fastest route is chosen since time of utmost concern to tourists.

Hotels and tourist spots can be treated as the same during the collection of transport data. They can be considered as Points Of Interest (POIs). Schematically, the cost (time taken) between the POIs data can be represented like a square matrix (Figure 4). The first cell on the 2^{nd} row, named with (2,1), should hold the cost of getting from POI2 to POI1. If (2,1) is considered as the inbound journey, then (1,2) would be the outbound of the same POI pair. If the POIs are drawn as vertex in a graph, the two edges between each POI pair would be the cost of the inbound and outbound journey between them (Figure 5). The graph is an asymmetric graph.



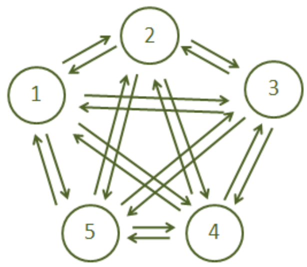


Figure 4 Schematic representation of the transport routes to be prepared (Left)

Figure 5 Asymmetric graph formed upon completion of the transport route (Right)

3.5 SCHEDULING

There are two ways to schedule the list of tourist spots selected by the user.

1. Schedule all the tourist spots with the use of greedy algorithm together with the k-opt heuristic search (General Solution)
2. Place the time constrained spots at particular position that satisfy the constraints, then schedule the rest of them (Time-constrained solution)

(1) has been devised in the prototype platform while (2) is still on theoretical side. Real data would be used for (1) and the cost matrix is given here (Table 6).

Origin \ Destination	1	2	3	4	5	6
1	/	15	10	0	10	15
2	15	/	5	15	20	5
3	10	25	/	10	25	20
4	0	15	10	/	15	15
5	15	20	30	15	/	20
6	15	5	10	15	20	/

Table 6 Cost (Time) matrix of possible tourist spot pairs

3.5.1 General Solution: Consider a case where the user has selected the hotel with the SPOTID of 1, and some tourist spots with SPOTID from 2-6. With the use of the greedy algorithm (GA) alone, the solution will be like Figure 7. User shall start from POI A and end at POI A again. sequence (numbers inside circles represents tourist spots) while the figures between ten with :

Figure 7 The scheduling solution with Greedy Algorithm

Since there is a possibility that the path generated may be the longest path (Jergen et. al, 2004), the solution needs to be passed through an improvement step. The tour improvement heuristic will be adopted as the improvement tool. The result generated with improvement tool applied will be:

Figure 8 The scheduling solution with Greedy Algorithm with tour improvement heuristic

Attention is to be drawn on the difference in the position of the 3 spots located in the middle before and after the application of tour improvement heuristic. The difference in position of the 3 spots is marked by the rectangles in Figure 9.

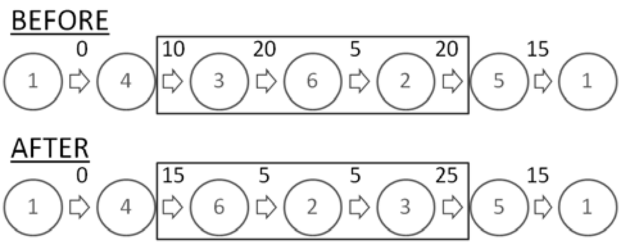


Figure 9 Difference in position before and after tour improvement heuristic. (3-Opt adopted)

The difference resulted in Figure 9 is actually a result of the K-Option improvement heuristic. At this case, the value of K is set as 3. The repositioning took place since a shorter path is available (note only the figures above the arrows in both rectangles). There should be at least 5 spots in the itinerary (4 if start and end point are the same). The first node and the last node will be fixed (because of their nature). In an arbitrary list with spots A, B, C, D, E. It would yield *3!* (a total of 6) permutations (only three spots in between can be swapped).

For a list with more than 5 nodes, the 3-Opt will be first performed on the 2nd, 3rd and 4th spot. Then the floating window will be shifted to the 3rd, 4th and 5th. It proceeds and processes until the last node of the floating window touches the $(n\text{-}1)^{th}$ node.

By illustrating how the floating window works, the difference in the sequence generated solely by greedy algorithm and greedy algorithm with 3-Opt employed is now justified.

The reason of setting the K value as 3 in the improvement heuristic is due to the spread of the tourist spots at different districts – there are about 3 tourist spots at the same district in the database. Since it is rather unlikely that the tour will be improved (in terms of reduction in total travelling time) by having tourist spots, the value 3 is justified.

3.5.2 Time-constrained Solution: The time slot fitting approach is the algorithm which first deals with time constrained tourist spot first. To make it easier for understanding, the illustration only includes ONE time constrained spots only. Lan Kwai Fong is selected as the time constrained spot so as to demonstrate this approach. It is assumed that the bars in Lan Kwai Fong only operates from 18:00 – 24:00 daily. (Figure 10)

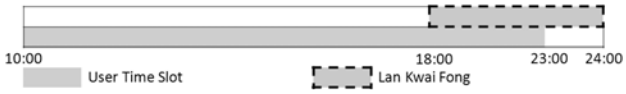


Figure 10 User defined time slot and operating hours of time constrained spot

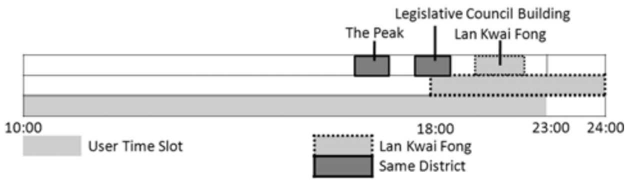


Figure 11 Time constrained spot with some same district spots inserted

Then, Lan Kwai Fong would be inserted as late as possible in the user defined time slot (Figure 11). Visit to the Lan Kwai Fong should end some time before the user defined time slot since it takes sometime for the tourist to back to the hotel. Tourist spots of the same district will then be inserted before the time that Lan Kwai Fong resides. (Figure 11)

Once the tourist spots (of the same district) have been inserted, the algorithm shall look for the possibility of making the whole trunk to be switched to an earlier time. (Figure 12) The rationale of this move is due to the consideration that there is a chance that the schedule is subject to delay (by traffic congestion). If there is a delay of 30 minutes for the day, the maximum time that the user can spend on Lan Kwai Fong would reduce by 30 minutes.

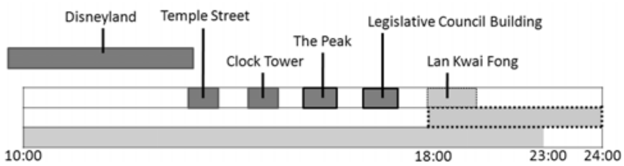


Figure 12 Time constrained spot and two other spots shifted forward

Once the shifting is completed, tourist spots of other districts will be inserted before the scheduled tourist spots. If the amount of time that takes to finish the trip is loner than the current time available. The system shall be looking for the possibility for a shift to a later time (Notice that the time visiting the time constrained spot has been shift earlier, so, there are some rooms for another shift).

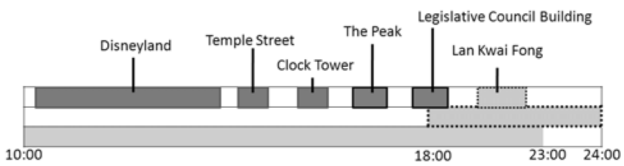


Figure 13 Forward shift of the whole route in order to fit in an extra tourist spot

Extra tourist spots will then be fitted in (Figure 13). Optimization will then take place by implementing the k-opt heuristic iterative search. The sequence before the time constrained spot is said to be fixed after the optimization. (Figure 14) Please be noted that the swapping in position between Temple Street, Disneyland and Clock Tower is fictional. It only serves the illustration of optimization (3-Opt) only.

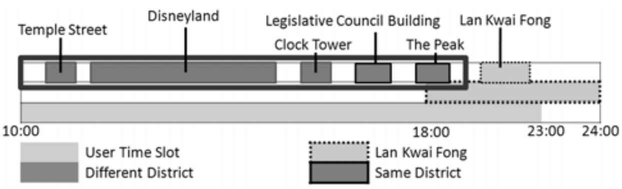


Figure 14 Optimization completed for the spots before the time constrained spot

The scheduling of the spots later than the Lan Kwai Fong will then be continued. The scheduling mechanism is much simpler – time constrained spot (Lan Kwai Fong) as the start point and the hotel as the end point.

4. SYSTEM IMPLEMENTATION

This chapter aims to provide details on the issue considered during the preparation of the user interface and other relevant issues.

4.1 SYSTEM WORKFLOW

It takes 4 steps to build the customized itinerary (Figure 15). Hotel should be chosen first since it is the first and last node of the itinerary. Secondly, the arrival and departure time would let the system know how much time the user would be make himself available each day. User should then select the tourist spots. The system would then click the "Schedule" button. The system would then process the information and generate the itinerary using the algorithm specified in the section 3.5.

Figure 15 The 4-step system workflow of the database

4.2 USER INTERFACE & REPORT PRINTING

It is the intention of the system to be developed with a user-friendly user interface. Short text description will be given in the first page to show the procedure of the scheduling function.

The whole program will be divided into 3 steps:

1. User selects a hotel
2. User specify the duration of visits to be done each day.
3. System will print the information specified by user for checking (Figure 16)
4. User selects a number of tourist spots.

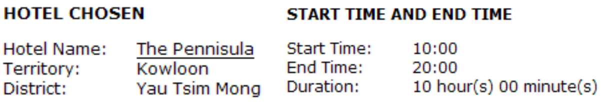


Figure 16 Information shown upon the completion of step 1 and 2

Figure 17 Itinerary generated (extract)

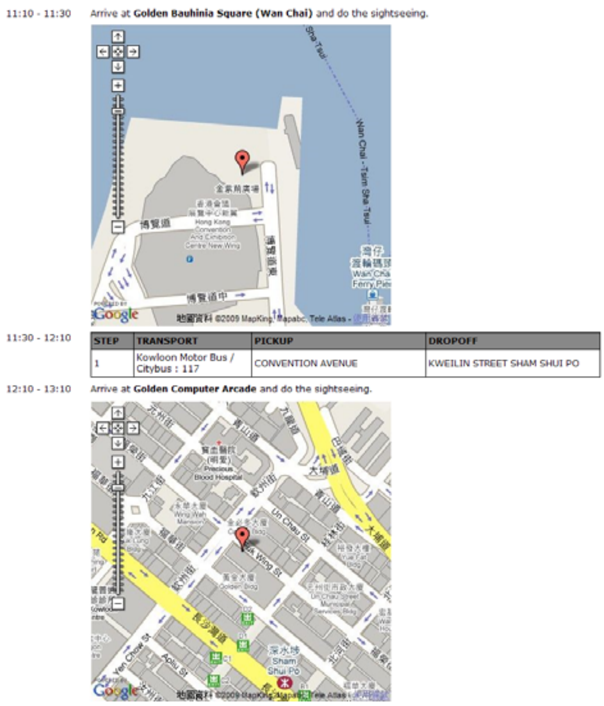


Figure 18 Detailed transportation means tourist spots

The user may then select the spots and let the system to do the scheduling. Figure 17 and 18 show a typical layout of the itinerary generated. Besides, the system is capable of handling exceptional case. For example, an ambitious user may select numerous tourist spots that cannot be visited within the time slot specified. It is therefore the system prompts the user of this fact and continue the journey on another day. (Figure 19)

Figure 19 Prompt of overtime visiting

5. SYSTEM EVALUATION & ENHANCEMENT

5.1 SYSTEM PERFORMANCE

The system receives users input and generates the path without any error. Yet, the execution time was not recorded. The limited size (in terms of number of tourist spots) does not put much stress in the time of computation.

Since there is only one user in the development platform, it is rather hard to perform either stress test or any similar testing. Normally, the system is able to return a result within 5 seconds at a case where all the tourist spots have been selected. Around 500 database transactions were taken place.

5.2 SYSTEM LIMITATION & POSSIBLE ENHANCEMENT

5.2.1 Programming language: PHP is easy to program and implement. It is possible for later system to consider the compilation time of different languages and choose a better and more efficient one. Compatibility between the scripting language and the server support should also be considered.

5.2.2 Development Platform and Implementation Platform: In this project, the combination of PHP and MySQL is used to implement the scheduling. It should be noted that this combination is adopted because of its convenient and readily available technical support.

5.2.3 Scheduling algorithm: The rationale of setting the K value of K-opt heuristic as 3 has been discussed earlier. This "magic number" is to be revised once necessary. It should be noted that the selection of the k value should be careful and be limited to small numbers or the number of combination will be increasing geometrically. The system currently uses the greedy algorithm together with the heuristic search to seek for possible improvement of the result. Though the problem size in this study is small enough to perform exhaustive search, it is not the intention of the system to compute the shortest solution, but an optimal solution instead. The real time computation of the total travelling time of an exhaustive search possesses too many queries to the database.

To avoid doing exhaustive search in deriving an optimal solution, an alternative is to screen out some of the "impractical cases" from the exhaustive search. For example, given 5 tourist spots, the tourist spots can be further divided district-wise. It is to filter out certain impracticable combinations so as to save the computation time.

5.2.4 Solution with time constrained spots: It is the limitation in time and difficult in terms of coding to implement the time dependent solution. The current scheduling ignores the time effect of the spots – the time effect refers to the opening hours (e.g. the Hong Kong Disneyland), operating hours (e.g. the Museums), or favouring hours (e.g. visiting the Peak at night is more preferable than afternoon time). The algorithm to solve the time constrained spots have been discussed earlier.

5.2.5 Which weighs higher: time or distance? The pilot study only consider the predetermined travelling time when performing scheduling. It is however that traffic condition differs a lot in peak and non-peak hours. The distance factor may be applied in order to find out the chances of having traffic congestion during the peak hours.

5.2.6 From the nearest district to the farthest district? The system currently prepares the initial solution using the Greedy Algorithm, which makes the itinerary resulted visiting the nearest spot first. That means the solution will end up starting from less transport time to longer transport time. It is questionable if the total time of the trip will be lower if the trip can be started another way round – the system first pick the furthest spot from the origin and do the greedy algorithm, then the improvement heuristics.

5.2.7 Mapping of results: Currently, there is only a map showing each spot. Yet, the map does not provide detailed information to enable user to get to the transport facilities that the system suggests.

5.3 SUGGESTIONS

The computation of the initial solution (which needs to be improved at later part) is derived from the greedy algorithm. Some studies have pointed out the greedy algorithm may not come up with a good solution, or even a decent solution for later improvement to take place. The choice of the algorithm was mainly due to its simplicity. Further enhancement should explore the possibility of using geographical relationship as aid to schedule. Better utilization of Goggle Map may help user learn more about the environment surrounding the area.

6. CONCLUSION

As stated in the very beginning, the initiative of the project is to devise a system that capable to schedule an itinerary that satisfies the conditions specified by the users. The scheduling algorithm and the database architecture is rather simple. It is expected that further enhancement is to be made on both the algorithm and the database architecture. The use of hardware to accommodate / suit the needs of the system is to be assessed afterwards.

REFERENCE

1. CORMEN T.H., LEISERSON C.E. & RIVEST L.R. (1989) Introduction to Algorithms. The Massachusetts Institute of Technology
2. EVANS R. J. and MINIEKA E. (1992). Optimization Algorithms for Networks and Graphs. Marcel Dekker, Inc.
3. MICHAEL S. (2002). Internet Travel Planner : How to Plan Trips and Save Money Online. Globe Pequot Press
4. PUN-CHENG Lilian, MOK E.C.M. et.al. (2007), EASYGO-A public transport query and guiding LBS. Lecture Notes in Geoinformation and Cartography, Location Based Services and TeleCartography, Springer, pp. 545- 554
5. WONG, C.S. (2004). A web-based GIS for Scheduling Travel Journey in Hong Kong. The Hong Kong Polytechnic University
6. HOME AFFAIRS DEPARTMENT, HKSAR (2008). Hong Kong Fun in 18 Districts. Hong Kong Special Administrative Region. (http://www.gohk.gov.hk/, Accessed 31st March, 2009)
7. Public Transport Enquiry Service (Pilot Version) http://ptes.td.gov.hk/ (accessed 27th Nov, 2009)THE HONG KONG TOURISM BOARD (2009) Discover Hong Kong (http://www.discoverhongkong.com, Accessed 1st February, 2009)
8. TripIt - Online travel itinerary and trip planner http://www.tripit.com/ (accessed 27th Nov, 2009)

APPENDIX

Implementation of the pilot study is available at:
http://myweb.polyu.edu.hk/~06159619d/v5.php