A Label Correcting Algorithm for Dynamic Tourist Trip Planning

Jin Li
School of Computer and Information Engineering,
Zhejiang Gongshang University, Hangzhou, P.R.China
Email: jinli@mail.zjgsu.edu.cn

Peihua Fu
School of Computer and Information Engineering,
Zhejiang Gongshang University, Hangzhou, P.R.China
Email: fph@mail.zjgsu.edu.cn

Abstract—One of the most important considerations for tourist in tourism is how to design an optimal trip planning. Selecting the most interesting points of interest and designing a personalized tourist trip is called a tourist trip design problem (TTDP) and, it can be modeled as an orienteering problem. However, most previous researches solve this problem in static network, which leads to unreasonable results. In this paper, to formulate this orienteering problem in a time-dependent network, a mathematical programming model is built with introducing a time-aggregated graph. Then a novel label correcting algorithm (LCA) is presented to solve this problem based on the idea of network planning and dynamic programming. Finally, a numerical example is given to show the effectiveness and feasibility of the algorithm.

Index Terms—tourist trip planning, orienteering problem, time-dependent network, label correcting algorithm, tourism

I. INTRODUCTION

Recently, China has held various sightseeing and exhibition activities, such as Beijing Olympic Games, Shanghai World Expo and many types of commercial exhibitions, etc. However, many tourists visit a region or a city during one or more days. Obviously, it is impossible to visit everywhere and tourists want to select what they believe to be the most attractive points of interest (POI). Nowadays, their selection can be made according to information of web sites, articles in magazines, or on guidebooks available in bookstores. Once the decision is made, the tourists will decide on the time to visit each point and choose a route between the points. However, it is difficult for tourists to design schedule, as many factors, such as many places are crowded, the traffic accidents may happen, or the road is under construction or closed, will lead to uncertainties for trip plans. The goal for tourists is to make a decision on trip plan where the POIs are visited as many as possible within a time budget. In fact, the traditional tourist trip design in static network has not been able to meet the demand. With the help of mobile and sensing technology, and taking the real-time traffic information into account, only tourist trip design in dynamic network is able to provide personalized and high-quality services for tourists.

Nowadays, the tourist trip design problem (TTDP) is commonly seen as an extension of orienteering problem (OP)[1]. The OP is also known as the Selective Traveling Salesperson Problem (STSP)[2-4], the knapsack problem[5], the Maximum Collection Problem (MCP)[6] and the bank robber problem[7]. Furthermore, the OP can be formulated as a special case of the Resource- Constrained TSP (RCTSP)[8]. Since these kinds of problems are NP-hard, most researches have focused on heuristics and metaheuristics [9-14], e.g. guided local search, the ant colony algorithms, etc. Dynamic trip design relies mainly on the mobile communication technology to provide tourist location-based real-time and fast navigation services. In the meanwhile, it can also help the tourists choose target points and find the shortest route to reach there. This kind of mobile tourist guides include Cyberguide[15], Gulliver’s Genie[16], GUIDE[17], CRUMPET[18], etc. However, tourists preferences are diverse[19] with the rising quality of life and aesthetic taste. Static trip design does not take into account the specific situation of the tourist, e.g. tourist’s start and end visiting time, current time, total time budget and weather conditions, etc. So the personalized mobile tourist guides based on context and location become an inevitable trend. Ten Hagen et al.[20] designed a mobile tourist guide, i.e. Dynamic Tour Guide (DTG) to decide on the tour of visiting a city during a period of time. He saw the tour as a set of Tour Building Blocks (TBB), which denotes attractions, hotels and something else chosen by tourist. Then, TBB receive interest matching points by a semantic matching algorithm, which calculates the degree of similarity between the TBB and the profile of the user. Finally, a heuristic approximation algorithm is used to calculate the tour. Souffria et al.[21] proposed a personalized tourist trip design algorithm for mobile tourist guides. By combining an artificial intelligence and a guided local search metaheuristic approach, this algorithm provided fast decision support for tourist using the mobile devices, and was compared with other
algorithms making use of a real data set from the city of Ghent.

The above researches just study the TTDP in a static network based on the start point, destination point and time budget without involving tourist trip design problem in a dynamic time-dependent network. However, the tourist trip network is actually time-dependent in real exhibitions or tourism due to the crowded places, emergencies, temporary shows, and promotions, etc. For example, the tourist travel time is significantly different between peak time and valley time in exhibitions. With the development of advanced sensor network, information system and databases, dynamic traffic information is available. Therefore, to formulate the tourist trip design problem in time-dependent network, based on the previous research on time-aggregated graph[20][21], this paper proposes a mathematical formulation for tourist trip design problem in time-dependent network, and puts forward a label correcting algorithm (LCA) to solve this problem.

This paper is organized as follows. The problem formulation for tourist trip design problem in time-dependent network is described in Section 2. In Section 3, the label correcting algorithm is given. Moreover, Section 4 validates the proposed label correcting algorithm by numerical example. Finally, the concluding remarks and further research are included in Section 5.

II. PROBLEM DESCRIPTION

A. Problem Assumptions

Tourist trip design problem in time-dependent network can be expressed as follows. Given a set of visiting points and the tourist preference value, each visiting point has an average stay time. And each visiting point can only be visited once. Then the objective of this problem is how to decide on the time to visit each point and select a route between the points in order to maximize the total utility of tourist trip (that is, the sum of preference value of all selected visiting points) within a given time budget. This paper will solve the tourist trip design problem in the time-dependent network based on the following assumptions:

- In time-dependent network, the travel time on an edge depends on the time entering it, i.e., once an edge is entered, the travel time is given by the start time of entering it. It is assumed that time is discretized into small units (such as 1 hour or less than 10 minutes).
- A tourist has a preference value for every visiting point, which is set as an random integer from interval [1, 10]. This value represents the tourist’s interest in the point. To be practical application, this preference value can be obtained using information retrieval or semantic identification technology via mobile tourist guide devices [20][21].
- Tourist often decides trip plan in terms of the schedule, set the time budget to be fixed and constant.

B. Time-aggregated Network

Given a transportation network \( G = (V,E) \), where \( V = \{v_i | i=1,2,\cdots,n\} \) is the node set, \( E = \{e_{ij} | v_i, v_j \in V \} \) is the edge set, if \( v_i \) is adjacent with \( v_j \), then there is one edge \( e_{ij} \) between them, \( |E| = m \). Since the edge and the travel time on a edge are time-dependent, it is not easy to denote the tourist moves on each edge. We use a time-aggregated graph[22] to represent the network change at each time instant.

In time-aggregated graph, each arc \( e_{ij} \) has two properties: time series \( ET_\mu \) indicating the time instances at which they are present and travel time series \( TT_\mu \) representing the travel times at various instances. \( ET_\mu = [1,2,\cdots,T] \) represents the existing state of arcs at each instant, and is a set of the corresponding time when the arcs are connected, which indicates a network topology varies with time. Travel time series \( TT_\mu = (T_\mu(1), T_\mu(2),\ldots,T_\mu(T)) \) denote the travel time when the edge is present, \( T_\mu(t) \) represents the travel time on edge \( e_{ij} \) when the entry time is \( t \). If the edge \( e_{ij} \) isn’t connected at time \( t \), then \( T_\mu(t) = \infty \), which means this edge cannot be passed and must wait. For example, a time-dependent network at time instant \( t = 1,2,3 \) is shown in Figure 1. The topological structure and travel time of this network vary with time. For edge \( e_{21} \), it is present at time \( t = 1,2 \), but disappears at time \( t = 3 \), i.e., no passing. Moreover, the travel time of edge \( e_{21} \) equals 1 at time \( t = 1 \), and becomes 5 at time \( t = 2 \). This time-aggregated graph is illustrated in Figure 1(d), where edge \( e_{21} \) has two series properties: \( [1,2] \) represents the time series when the edge is present, and \( (1,5,\infty) \) is the travel time series representing the travel time at time instant 1 and 2.

C. Trip Denotation

In the time-dependent network, temporal and spatial dimensions are used to represent the tourist trip, where time is described as discrete unit and space is expressed as \( V = \{V_i | i \in [1,|V|]\} \). So, the tourist trip is a list consisting in elements \( V_i(t_i) \), where \( V_i(t_i) \) denotes that the node \( V_i \) is reached at time \( t_i \). The tourist trip is presented as \( P(V_1,V_2)=\{V_1(t_1),V_2(t_2),\ldots,V_n(t_n)\} \), where
$V_1$ is the source node and $V_n$ is the destination node. According to the time-aggregated graph, if the stay time at node $iV$ is $ivt$, then the earliest arrival time at edge $e_{ij}$ is $t_e = \min\{t | t \geq t_i + vt_i, t \in ET_y\}$. For the two nodes $V(t_i)$ and $V(t_j)$ on edge $e_{ij}$, the time constraints is expressed as: $t_j = t_i + vt_i + T_y(t_e)$.

![Figure 1. Denotation of time-dependent network using time-aggregated graph.](image)

### D. Mathematical Model

Notations and Meanings:
- $G = (V, E)$: Time-dependent network;
- $P(i)$: Set of predecessor nodes of node $i$;
- $S(i)$: Set of successor nodes of node $i$;
- $T$: The total time budget of the trip;
- $t_0$: The traveling start time;
- $p_i$: Tourist preference value for node $i$;
- $vt_i$: Stay time on node $i$;
- $t_i$: Arrival time at node $i$;
- $t_{ij}(t)$: Travel time on edge $e_{ij}$ when the entry time is $t$;
- $x_{ij}(t) = \begin{cases} 1, & \text{if edge } e_{ij} \text{ is entered at time } t \\ 0, & \text{otherwise} \end{cases}$

Choose node $V_1$ as the source point and node $V_n$ as the destination point. And we establish the following mixed integer programming model as follows.

$$\max \sum_{i=0}^{T} \sum_{j=1}^{n-1} p_i x_{ij}(t)$$

subject to

$$\sum_{i=0}^{T} \sum_{j=1}^{n} x_{ij}(t) = \sum_{i=0}^{T} \sum_{j=1}^{n} x_{ji}(t), \forall k = 2, \ldots, n-1$$

$$\sum_{i=0}^{T} \sum_{j=1}^{n} x_{ij}(t) \leq 1, \forall i = 2, \ldots, n-1$$

$$\sum_{i=0}^{T} \sum_{j=1}^{n} x_{ij}(t) = t_j, \forall j = 2, \ldots, n$$

$$\sum_{i=0}^{T} \sum_{j=1}^{n} t_{ij}(t) = t_i + vt_i, \forall i = 1, \ldots, n$$

$$t_i = t_0, t_j \leq T$$

$$x_{ij}(t) = 0, 1, \forall e_{ij} \in E, \forall t = 1, \ldots, T$$

The objective of the problem is to maximize the total utility, as shown in (1). In this formulation, constraint (2) and (3) are flow-conservation constraints. Constraint (4) ensures that every point is visited at most once.
Constraint (5) and (6) guarantees that if one edge is visited in a given tour, the arrival time of the edge following node is the sum of the preceding arrival time, visiting time and the edge travel time. Constraint (7) is the start time and end time constraint. Constraint (8) and (9) are the variables constraint.

III. LABEL CORRECTING ALGORITHM

A. The Ideas of Algorithm

The objective of the time-dependent tourist trip design problem is to decide an optimal trip to maximize the tourist’s total utility within time budget taking into account tourist preference and time-dependent network, and provide the tourist with real-time navigation service via the mobile communication devices (e.g. PDA, cell phone). In this dynamic network, the trip design and travel time depend on the start time of the source node. Consequently, starting from an optimal start time, a best trip is produced to help the tourist visit his interesting points within a time budget. We present a label correcting algorithm to solve this problem. This algorithm can produce an optimal trip plan and achieve maximum utility within a given time budget.

Some notations used in the algorithm are defined in the following.

Definition 1: $Q$ is the priority queue of the node to be processed and satisfies first-in-first-out (FIFO) principles.

Definition 2: $C_i(t)$ is the cost of the source node $V_i$ at time $t$, which represents the travel time from node $V_i$ to the destination node at time $t$.

Definition 3: $U_i(V_j,t)$ represents the total utility arriving at node $V_j$ at time $t$ ($t \in ET_y$), where $V_i$ is the predecessor of $V_j$.

The start time is variable. In order to seek the optimal start time and travel route, we use a label correcting algorithm from the destination node back to the source node. As the edge travel time series represent the edge travel time in every time unit, the label records the travel time and total utility every time from the current node to the destination node. This algorithm calculates the node's trip utility and updates the pre-node's label. Repeat these iterations until we get the optimal route with the best start time.

B. The Detail Steps of Algorithm

According to the idea of network planning and dynamic programming[25], a novel label correcting algorithm is presented in the following.

Step 1: Initialization

Given every edge $e_y$ two properties: the edge time series $ET_y$ and travel time series $TT_y$. $V_s$ is the source node and $V_n$ is the destination node. $p_i$ represents the tourist’s preference value for each visiting point $V_i$, the stay time is $vt_i$. Let $p_i = p_n = 0$, $vt_i = vt_n = 0$.

Step 2: For the destination node $V_n$, $C_n(t) = 0$, $U_n(V_n,t) = p_n \cdot t = 1, 2, \ldots, T$, where $V_n$ is the virtual successor node of source node $V_s$, $Q = \{V_n\}$.

Step 3: Processing the node in priority queue $Q$

3.1. For $V_j$ in $Q$, delete $V_j$ from $Q$;

3.2. For $V_j \in P(j)$, calculate $C_j(t) = C_j(t + vt_j + T_y(t)) + T_y(t) + vt_j$, $\forall t \in ET_y$. If $C_j(t) > T$, then $T_j(t) = \infty$, else if $C_j(t) \leq T$, then go to step 3.3;

3.3. Calculate the tourist’s total utility and label

Calculate the total utility from the source node $V_j$ to the destination node at time $t$.

$U_j(V_j,t) = U_j(V_s,t) + T_y(t) + p_j$, $V_j \in S(i)$, $V_j \in S(j)$, $\forall t \in ET_y$;

Label $(V_j,t,C_j(t),U_j(V_j,t))$;

3.4. Node $V_j$ is inserted into $Q$, and update it;

Step 4: Judgment of whether all the nodes have been processed

If $Q \neq \phi$, go to step 3, else go to step 5;

Step 5: Calculation of the route’s total utility

Calculate $U(V_i) = \max_{t \in (t,\infty)} U_i(V_i,t), \forall t \in ET_i, e_{tr} \in E$;

Step 6: Backward the tourist’s route with maximum total utility

According to the label with maximum utility, we can get the optimal start time $t_{best}$. And go back based on this start time, we will obtain the optimal route from the source node $V_1$ to the destination node $V_n$ in the following:

$P(V_1,V_n) = \{V_1(t_{best}), V_2(t_1), \ldots, V_n(t_n)\}$. If there are several routes with maximum total utility, choose the route with minimum $C_i(t)$ as the final optimal route.

C. Time Complexity Analysis of Algorithm

According to the steps of the algorithm, we can see that when the travel time and label of the node is a scalar, the time complexity in worst case is $O(|V||E||P|)$. As each label is generated when the length of time series is $T$, the total computational time is $O(|V||E||P|)$.

The above analysis indicates that our algorithm is a polynomial algorithm, and can meet the real-time applications.

IV. NUMERICAL EXAMPLE

The efficiency and feasibility of the algorithm would be demonstrated by the following numerical example in this section.

Example: Given an exhibition graph as shown in Figure 2, where $V_1$ is the entrance (source node) and $V_n$ is the exit (destination node), other nodes are visiting point; The first figure at each node denotes the stay time and the second figure denotes the preference value of a
tourist. The time budget of the trip is $T = 10$ and the figures on the edges are travel time series. To simplify the problem, all the time series on the edges are set as $T_{ij} = \tau_{ij}$, $\forall e_{ij} \in E$. If $t > T$, then $T_g(t) = T_g(T)$.

We shall use the proposed label correcting algorithm to solve the above example, and the optimal trip is calculated in detail as follows.

**Figure 2.** Time-dependent exhibition graph.

**TABLE I.**

| Trace of the label using label correcting algorithm |
|---|---|---|---|---|---|---|
| steps | $V_1$ | $V_2$ | $V_3$ | $V_4$ | $V_5$ | $V_6$ |
| 1 | | 2 | 3 | 4 | 5 | 6 |
| 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 6 | 12 | 13 | 14 | 15 | 16 | 17 |

© 2012 ACADEMY PUBLISHER
First, initialization:

Given the time series \( ET \) and travel time series \( TT \) on each edge \( e \); \( V_i \) is the source node and \( V_n \) is the destination node. For every node \( V_i \), the preference value is \( p_i \) and the stay time is \( v_t \), \( p_i = p_n = 0 \), \( v_t = v_{n_t} = 0 \).

For the destination node \( V_n \), \( C_{n}(t) = 0 \), \( \bar{U}_n(v_{n+1},i) = p_n \), \( i = 1,2,\cdots ,T \), where it is assumed that the successor node of \( V_n \) is \( V_{n+1} \) (virtual node), \( Q = \{ V_n \} \).

Each node in priority queue \( Q \) is processed in order, and the trace of the label is shown in table 1. It is noted that the label for each node is denoted as a triple \( \{ Sub, TS ,TU \} \), where \( Sub \) represents the sub index of the successor of the labeled node, \( TS \) is the travel time series from the labeled node to the destination node, and \( TU \) is the total utility from the labeled node to the destination node. For each travel time series, if the travel time exceeds the time budget at any time, it is denoted as "- - ."

According to the label at the source node \( V_1 \), the maximal total utility is \( U(V_1)=17 \) and the best start time is \( t_{new}=3 \).

Backward the route based on the node label, the optimal route is obtained in the following:

\[ P(V_1,V_6) = \{ V_1(3), V_2(4), V_3(6), V_4(8), V_5(10), V_6(13) \} \]

From the above implementation of the algorithm, it is clear that the labeled correcting algorithm is effective to decide on optimal start time, and provide a best trip to improve the tourist satisfaction.

V. CONCLUSION

A variety of sightseeing and exhibition activities develop rapidly in China, such as urban tourism, business exhibitions, theme parks etc. And they have played an important role in promoting the economic growth. The tourist trip design problem is not only the vital content of their trip plans, but also the key to provide high-quality service for tourists. With the development of the mobile communication technology and the popularization of cell phone, PDA, etc., the mobile tourist guides turn into possible. They are able to provide the tourists with the real-time and personalized services based on context and location. The previous researches ignored the time-dependent characteristics in exhibition networks. The traffic network in tourism is one of the complex systems, which has great uncertainties. The tourist travel time in the exhibition network may change because of many things such as crowded places, temporary shows, promotions, and so on. Therefore, by introducing a time-aggregated graph, this paper establishes a time-dependent tourist trip design model, and proposes a label correcting algorithm. The time complexity is also analyzed. Then the efficiency and feasibility of the algorithm are demonstrated by a numerical example.

Further studies will focus on multi-trips design problem during several days, and provide the tourists more satisfied and high-quality services, simultaneously considering more realistic factors such as the location open time and capacity constraints.

ACKNOWLEDGMENT

The authors wish to thank the reviewers for their valuable comments. This work was supported in part by the National Natural Science Foundation of China (Grant No.71171178), Humanities and Social Sciences Foundation of Ministry of Education of China (Grant No. 12YJC630091), Zhejiang Provincial Natural Science Foundation of China (Grant No. LQ12G02007), and Zhejiang Provincial Commonweal Technology Applied Research Projects of China (Grant No.2011C23076).

REFERENCES


Jin Li was born in Jiangsu province, China. He received the B.S. and M.S. degrees from Nanchang University, Nanchang, Jiangxi Province, China in 2004 and 2007 respectively. He received the Ph.D. degree in management science and engineering department, School of management, at Fudan University, Shanghai, China in 2010.

From July 2010 until now he works as an assistant professor in the Department of Management Science and Engineering, School of Computer Science and Information Engineering, Zhejiang Gongshang University, Hangzhou, Zhejiang Province, China. He has authored/coauthored more than 20 scientific papers. His research interests include logistics and supply chain management, system modeling and simulation, network optimization, and emergency management.

Dr. Li is a fellow of China Society of Logistics (CSL), Chinese Computer Federation (CCF), and Chinese Association for System Simulation(CASS).