The On-Demand Tourist Paths Problem

Teodor Gabriel Crainic
Dept. management et technologie
Université du Québec à Montréal
and
Center for Research on Transportation
Université de Montréal

Nicoletta Ricciardi

Giovanni Storchi
Dept. di Statistica, Probabilità e Statistiche Applicate
Università degli Studi di Roma “La Sapienza”

Abstract

The aim of this paper is to study how to organize the paths or activities on demand of tourist in artistic towns, with the objective of satisfying the demand of visit. Museums, churches, streets and places of interest or attraction are considered, in order to define a "desirability" for the tourist. The index of desirability of a place is defined by parameters such as the importance, popularity, typology, the average visiting time and some descriptive attributes. This index is associated to each nodes or link of the graph as a label. The current city representation has to be processed such each visiting type of monument or activity is defined.

Two main set of paths may be required: *Sequential-activity paths*, where each of the activity specified by the user have to be performed in the order given by the user; and *Mixed-activity paths*, where there is no particular request on the sequence on activities.

The goal is to find the best path according to the user-specified preferences. The label of each activity is updated using the usual path labelling techniques, adjusted to account for the type of path computed.

Key words : Tourist paths, Routing, Shortest paths, A*.

1 Introduction

The aim of the On-Demand Tourist Paths Problem (DTPP) is to identify visit itineraries on demand according to particular requests made by a tourist that he wants to know an historic-artistic town. Museums, churches, restaurants and other places of interest or curiosity must be considered, as well as particular streets and places. All the considered sites are important in their own right due to their cultural, historical or shopping interest. In order to define a degree of "desirability", to each touristic place, it is associated a score that points out the importance and/or interest, popularity, type of place and so on. The score is not an absolute value but it depends on the type of visit that the tourist wanted to, it depends from the intensity and from the type, i.e. visit rapidly, normal, or in depth or only to pass in front look and photograph or if he prefer the ancient one rather than the modern one.

The city representation has to be processed to yield a graph that helps finding such paths rapidly. In our approach, on each of the nodes and links of the resulting network, only one instance of each type of monument or activity is explicitly represented. Thus, for example, if three museums are located on the same street, this street will be represented as a sequence of three nodes and links added to separate the museums. A set of attributes are associated to each nodes and to each links. Attributes are of two types: descriptive, such as length, walking travel time, global interest, main type of activity and so on, and touristic, such as type of attraction, the epoch of construction or architectural type, visit time for various types of visitors, etc. The user specifies one or several requests in terms of total time available, starting and finish location, and the desired type of itinerary. The goal of the DTPP is to propose for each request a set of paths that answer the request by different path composition. In others words is to find the best path for each of the specified categories within the set of user-specified temporal constraints and preferences.

In a previous work [5], in 2000, has been studied a problem to determine some particular paths called the Jubilee itinerary. In fact during the Catholic Jubilee in Rome the problem to be studied has been that how to organize the flows of visits of the pilgrims in the city with the objective was, not only, determine a religious path but also to avoid phenomena of congestion. In this formulation we had only one type of activity: visit church. Each church received an interest score both religious, either historical, artistic, architectural. Moreover for each site was known, the mean time of visit, the duration of Mass, Holy lectures, the timetable and capacity and the travelling (walking time) between each pair of churches. The goal is find itineraries to maximize total religious interest with time limits respecting capacity and timetable [3].

The On-Demand Tourist Paths Problem may be considered as a variant of the Travelling Salesman Problem, where one is not obliged to visit all the nodes of the graph, because the time available for the visit is limited. Rather, one collects rewards at the nodes and on the links visited. The objective is to select the nodes visited and to build a tour to maximize the total reward, while respecting the time available and a number of tour feasibility constraints [2].

The paper is organized as follows. In Section 2 we introduce the problem and formalize the objective function. In Section 3 we present the input structure. The mathematical formulation is presented in Section 4. In Section 4.1 we present the methodological model. Finally, in Section 5 we propose the heuristic and we discuss the results.

2 The problem

The DTPP is addressed to all that tourists that having available time and they wants to get organize himself a visit of the city. Tourist requests are related to the knowledge the tourist has of the city that is visiting. The tourist can formulate requests that range from the generic walk with only a time constraint to a specifics request to visit certain particular monuments. In fact many and various can be the request on the type of paths and on the sequence of the visits. First visit a museum and then make shopping or only to take a walk in the famous places or only visit a particular architecture and to go to the restaurant.

All possible places of interest must be considered and described in terms of parameters. Museum, churches and other places of interest have been selected, as well as restaurants, shops, theaters and other various sites. Parameters such as the average visiting time, the index of desirability and a "celebrity" index have been defined for each of them.

The first problem is to analyze which can be the demands for the tourists. Really happen, that a tourist have time in hand in a non well known city and they ask himself how to use the available time in the best one ways conform to own preferences. Someone can be interested to particular artistic aspects, or a religious places or a selective shopping. They can make requests with various degrees of precision related to the knowledge of the city and of the own interest. In fact, there are many types of tourists, from those who only watches or photograph monuments and to carry on or those to immerse themselves into every particular.

The general problem can be characterized by the following characteristics:

- On-line computation
- Points of interest - monuments and streets - have list of attributes, each attribute has a relative order by number of "stars" like in classical tourist guide.
- Request for an Origin to Destination path with a maximum available time and a list of preferences.
- Desired output: set of paths by *category* to maximize:
  - The sum of score of request
  - The sum of score of monuments visited
  - The number of monuments seen and/or streets passed on).

- Paths are homogenous with regards to customer preference criteria.

Two different *types* of paths may be considered:

- *Sequential-activity* paths, where each of the main activities specified by the user have to be performed separately and in the *order* specified by the user; e.g., museum visiting, lunch, shopping.
- *Mixed-activity* paths, where there is no particular order on activities. A more complex problem considers constraints on the total time for some (all) activities or, alternatively, on their relative proportions.

Moreover, for each type it is possible to do a complete visit of all site of interest or only a cursory glance by passing. In this second hypothesis the related score will be considered proportionally to the user-specified preferences. Post-optimization procedures might be designed to offer additional information or counsel to the tourist.

3 Input description

The scenario is described by two groups of data. The first group describes the physical structure of the city, that is the network of the streets through which it will be possible to characterize different avenues, distances, public squares with their physical characteristics. The second group includes for all the touristic sites considered a synthetic description of the characteristics. The street network is represented by a graph where every arc is defined by its extreme nodes and is described by its physical characteristics. The arcs of the graph are split up, if necessary, to insert the nodes representative of the places that may be visited. In fact every nodes and arc of the graph has to represent only a site of visit. Thus, for instance, if on a street there are a church and a museum, then the arc is decomposed by inserting two new nodes. In an analogous way, if there are more than one point of interest at a given location (e.g., a plaza), a node can be split to allow independent visits to each sites. Museums, churches, archaeological sites, panoramic points, commercial streets or typical restaurants, are identified as places of tourist interest. To every site is associated a set of qualitative and quantitative parameters, such as the relevance, time of visit for different types of tourists, the type of the site. Moreover, the main cultural or architectural characteristic must be identified, for example, if a church is baroque or gothic, or if a museum is modern or classic art. More refined classifications may be introduced, of course. The site capacity, the opening hours, the cost, and the normal congestion at various time periods of the day and year are other important characteristics that can be considered. A score is used for each parameter to determine the relative relevance of the site.

Every site therefore is characterized through a set of attributes associated to each node to describe the characteristics both physical and touristic.

Attributes are of two types. *Descriptive*, such as length, travel time by foot, global interest, main type of activity and so on, and *touristic*. The latter are made up of a table where for each type of activity (e.g., museums, monuments, archeological places, etc.) are indicated the *time to visit*, the *interest to visit*, and the *interest to see* (that is, to pass by and admire from the outside).

a) Type: church, museum, site archaeological, scenic view, ...

b) Class: classic, modern and ancient art, gothic, renaissance,

c) Interest degree: average of the values (stars) present in the most common tourist guides.

d) Estimation of average visit time (by type of visit).

e) Opening hours.

f) Cost.

Every link (arc of the graph) is characterized by:

a) Length.

b) Walking times.

c) Type: pedestrian, stairs, facilitated access.

d) Characteristics: shopping, visit, connecting street,

e) Relevance: degree of interest relatively to the characteristics.

A combination of the parameters of visited site, weighed on the preferences of the tourist, determines the score of *attractiveness* or *desirability* of the path to propose to the tourist.

For each type of path and site of interest, it is possible to do a complete visit or only to give a casual glance while passing by. The appropriate interest measure corresponding to each activity must therefore be used in the computation of the score of the site and path. The particular intensity of the type of visit request by the tourist determines a parameter said *modality* of visit that will proportionally influence the scores of the single visited sites. We will call *modality* k the type as the interest of visit of the user (no to be confused with the mode on travelling on arc). The score of a path is determined as the sum of the scores obtained from the visit of the tourist places along the path. The scores of the sites and streets, are then be used to compute a set of paths to maximize the objective function.

4 The model

The resulting network $G = (\mathcal{N}, \mathcal{A})$ is then composed of a set of nodes $\mathcal{N}$, $|\mathcal{N}| = n$ that include street junctions and separating nodes, and a set of directed links $a \in \mathcal{A}$, with the set of attributes associated to each nodes S_i and to each links S_{ij}, $(i, j) = a \in \mathcal{A}$.

To compute path labels for mixed-activity paths, a simple sum on the appropriate activities should be sufficient.

The mathematical scheme is:

$$\max Z = \sum_j S_j^k y_j + \sum_i \sum_j S_{ij}^k y_{ij} \tag{1}$$

subject to:

$$t_{\max} \leq \sum_j t_j^k y_j + \sum_i \sum_j t_{ij}^k y_{ij} \tag{2}$$

$$y_j, y_{ij} \text{ boolean } [0, 1] \tag{3}$$

Where:

S_j^k score of the node$j$ with modality k,

S_{ij}^k score of the arc (ij) with modality k,

t_j^k visit time with modality k,

t_{ij}^k arc time with modality k,

$t_{\max}$ maximum time available.

The score of a path is then determined as the sum of the scores obtained from the visit of the tourist places and streets along the path. The scores of the nodes and streets have been used to compute a set of paths weighed with the modality k.

4.1 The methodology

As tell precedence this problem is difficult as computational complexity therefore we resort to an efficient heuristic algorithm.

In a tourist city with a lot of things to see and to do, the graph assumes very great dimensions, in fact all the places of interest must singly be represented. Our attention is focused then to reduce the graph trying to consider only those parts of the city that that are really compatible with the request of the tourist mainly with the time constraints. The first and fundamental request concern the starting place of the path (what it can coincide with the place where the tourist is), then the final destination and the available time. To reduce the graph we consider only the part of graph, that we will call "active graph", that it is inside the ellipse with the focus in the starting end ending points and as sum of the "distances" the maximum available time. For the reduction of the graph in rapid way we use in first approximation the geographical distance as the crow flies. Operating in this way we are sure that all the nodes outside of the ellipse belong to non admissible paths and therefore eliminable from the analysis.

First problem is to define how the graph should be structured. Every point of possible tourist interest will be placed on a node or an arc of the graph depending to the specifics characteristic. All information must be stored in a comprehensive and flexible data base for an efficient treatment. The only common parameters that must be always present in every request of tourist are where and when the visit start, the time available and if he have a fixed destination.

In our problem it is necessary to generate a path (or a set of paths) under tourist requests and evaluate them. To determine the path we use a labelling techniques A* shortest paths algorithm structure [4] . The choice of A* algorithm is determined since is one of the best and fastest algorithm to determine the shortest path between two different nodes.

The label of each activity is updated using the usual path labelling techniques, possibly adjusted to account for the type of path computed.

The objective function is to maximize the total score of the paths. The score is determined from the sum of the scores obtained from the visit of the tourist places along the path. From an algorithmic point of view, we envision *labelling*-type procedures.

One may envision a multi-activity, multi-category label as a table where the *score* of each path - its current *label* - is computed using an *aggregation* of the individual labels of the appropriate activities at the node. The label of each activity is updated using the usual path labelling techniques [1], possibly adjusted to account for the type of path computed.

The score of a path is determined as the sum of the scores obtained from the visit of the tourist places along the path respect de modality k choose. The scores of the nodes and streets have been used to compute a set of paths.

5 The heuristic approach

Two main issues have to be addressed. The first concerns the organization of the data for easy access, sharing by many request-answering searches, and algorithmic performance. The second issue is related to the efficient computation of several "best" paths according to user-specified preferences.

The essential data in this model are, at the first the starting point s as the place where the tourist is or from that where he decides to start. If the tourist have a knowledge of final point f, otherwise the ending point can be coincide with the start. Let be t_{max} the max time available, $t_{i,j}$ the travel time along the path used from a node i to j, and t the residual current time: (i.e. t= t_{max}- $t_{i,j}$). Let $E_{i,j}$ set of nodes inside the ellipse with foci in i and j and let $w^k_{i,j} \in E_{i,j}$ be the node with the best score relatively the k modality. The best nodes found in each ellipse are stored in a ordered list Q containing, this list permit to manage.

$P^*(i,j)$ is shortest path determined by A^* form i to j and $P(s,f)$ the tourist path obtained from the sum of all shortest paths determined. Let t^k_i the visit time of i site with modality k.

All nodes are georeferenced, to allow an immediate localization of the nodes inside the ellipse using distance as the crow flies.

5.1 Procedure

Procedure $Tourist - Path(s, f, t_{max}, k)$;

begin

$Initialization : i = s, j = f, S = 0, t = t_{max}, Q = \emptyset$;

build the shortest path $P^*(i, j)$ with time $t^k_{i,j}$;

$t = t - t^k_{i,j}$;

insert in Q all nodes w inside the ellipse $E_{i,j}$ with the related score S^k and the foci of related ellipse i, j;

$P(s, f) = P^*(i, j)$;

while $Q \neq \emptyset$ ***do***

begin

repeat

Let $\overline{w}$ the node of maximum score in Q;

retrieve and *extract* $\overline{w}$: $Q = Q \backslash \{\overline{w}\}$;

build the shortest path $P^*(i, \overline{w})$;

build the shortest path $P^*(\overline{w}, j)$;

$t^* = t - t^k_{i,j} + t^k_{i,\overline{w}} + t^k_{\overline{w},j} + t^k_{\overline{w}}$;

until $t^* < t_{max}$ or $Q = \emptyset$

if $Q \neq \emptyset$ ***then***

begin

$t = t - t^k_{i,j} + t^k_{i,\overline{w}} + t^k_{\overline{w},j} + t^k_{\overline{w}}$;

remove from Q all nodes $w \in E_{i,j}$ inside the ellipse $E_{i,j}$;

insert in Q all nodes w inside the ellipse $E_{i,\overline{w}}$;

insert in Q all nodes w inside the ellipse $E_{\overline{w},j}$;

$P(s, f) = (P(s, f) \backslash P(i, j)) \cup (P^*(i, \overline{w}) \cup P^*(\overline{w}, j))$;

end

end

For all the nodes and arcs of the path $P(s, f)$:

$S = S^k_j + S^k_{i,j} \quad \forall\, i, j \in]s, f]$

end.

6 Conclusion

The work is addressed to help the tourist to determine a set of paths of visit using a friendly technologies as PC, cellular phone, or PC palmar and so on, for this reason the necessity to handle a very fast and light data base and algorithms.

Even if the tourist don't know the city, but with a simple technology and a very simple set of information every tourist can able to obtain personalized itineraries.

We want to offer a service for a single tourist that ask a preferred itinerary. Our objective is that to maximize the interest of the itinerary respect the request of the tourist. Methodology to compute the paths to maximize all "score" functions related to a request (visit time, preferences, etc.) within the available response time.

It is possible to offer to the tourist more than an only path. To do this we have calculated several objective function. The first one respecting the choices and the applications made. The other ones respond to general criterions. In this way we propose to the tourist some alternatives.

The city representation has to be processed to yield a graph that helps finding such paths rapidly. In the current approach, on each of the links of the resulting network, only one instance of each type of monument or activity is explicitly represented. Thus, for instance, if on a street there is a church, (or a museum), then the arc is decomposed by inserting a new node. In an analogous way, if there are more than one point of interest at a given location (e.g., a square), a node can be decomposed to allow independent visits.

References

[1] R. K. Ahuja, T. L. Magnanti, J. B. Orlin; *Network Flows*, Prentice Hall, 1993.

[2] I-M Chao, B. Golden, E.A. Wasil; A fast and effective heuristic for the orienteering problem, *European Journal of Operational Research*, 88, 475-489, 1996.

[3] M. Gentili; Visiting a network of services with time constraints, *Computers and Operations Research*, 30, 8, 1187-1217, 2003.

[4] N.J. Nilsson; *Problem-Solving Methods in Artificial Intelligence*, McGraw-Hill, New York, 1971.

[5] A. Scozzari, G. Storchi; An O.R. Application for the Catholic Jubilee in Rome, *Ricerca Operativa*, 87, 49-55, 1999.