AGENT-BASED PERSONALIZED TOURIST ROUTE ADVICE SYSTEM

Yuxian Sun[a], Lyndon Lee[b]

[a]Department of Geo-Information Processing, International Institute for Geo-Information Science and Earth Observation (ITC), P.O.Box 6, 7500 AA Enschede, The Netherlands
(email: sun@itc.nl)
[b]Department BT Exact, Adastral Park, Martlesham,   IPSWICH, IP5 3RE, United Kingdom
(lyndon.lee@bt.com)

ABSTRACT

This paper proposes a tourist route model that is based on the integration of GIS spatial analysis functions and a kind of heuristic search algorithm based on local optimisation. A vector-based model is used to represent tourists' personal interests and the available tourist resources. Based on these two representations, the system can build a user model that indicates the attractions values of the tourist features in a particular area to individual tourists. The route agent then uses this user model to generate personalized tourist routes. This research adopts a *Tabu search* method, called *extension/collapse* algorithm, that is applied in the operations research for maximizing some utilities under certain constraints. It is a heuristic search method that progressively selects each tourist site based on both its attraction value and the cost value. An empirical evaluation has been carried out on the personalized route advice system. The result suggests that the tourist route model generates tourist routes that are empirically consistent with the tourists' preferences.

INTRODUCTION

As mobile devices decrease in price and size, and increase in power, storage, connectivity and positioning capabilities, tourists will increasingly use them as electronic personal tour guides. However, customers are not satisfied by being dumped with a lot of non-relevant information on their mobile devices. "Context-aware" and "personalisation" have become the buzzwords in location-based services.

The context of a tourist may include external factors, such as physical geographical environment and social-cultural events, and internal factors, such as the physical conditions of a tourist. Personalisation allows tourists to express what they like and the computer program will use this information to plan tour routes and activities for tourists. Therefore, knowing what each tourist likes and using this information to provide a personalised services become the main challenge for the service providers.

One of these services is the personalised tourist route recommendation. Most of the GIS network analysis tools, such as ArcGIS, and navigation systems are able to recommend routes to travellers. However, they mostly apply the shortest-path based on Dijkstra's shortest distance algorithm in time or distance. These functions cannot be directly applied to tourist route calculation problem because of the following two reasons:

- A shortest or a minimum-cost path is not, in most of the cases, what a tourist needs. Instead, he would like to follow a route that can give him the most satisfaction by including as much as possible those features that he likes.
- A tourist usually has certain kinds of resources to be consumed during a tour, which can be the time or distance that the tourist can afford during a tour. This resource is usually expressed as an upper bound that cannot be exceeded, but should be consumed as much as possible during a tour.

To overcome these problems, this research adopts a different algorithm, e.g. a kind of *Tabu search* method, that is applied in the operations research for maximizing some utilities under certain constraints. It is a heuristic search method that progressively selects each tourist site based on both its attraction value and the cost value.

SYSTEM ARCHITECTURE DESIGN

The design of such a personalised tourist route advice system is based on agent-based technology. Agent technology has been mostly adopted by the software that simulate individuals' behaviour in a social or physical environment (Itami and Gimblett, 2001; Lee *et al.*, 2002; Wasson *et al.*, 2001; Zipf, 2002). The individuals can be human, animals, plants or other objects that possess certain behaviour

and interact with other individuals and/or their environment.

The main characteristic of agent-based technology is that the structure of the software is represented by a group of agents who collaborate in achieving the goal of the task in hand. Among different types of agents, the personal assistant agents are particularly interesting to this research. This type of agents operates at the user interface level and actively assists users by offering information and advice to the users (Wasson *et al.*, 2001). These agents usually apply a kind of intelligent learning algorithm so that they can intercept the user input, examine it and take actions that are more specific to that particular users' needs at that moment. These agents are also called learning or adaptive agents. However, the assistant agents are different from the conventional autonomous agents in that they may not act truly autonomously and their goals must be adjusted constantly based on their interception of the users' goals. According to Wasson *et al.* (2001), the key questions for the design of any assistant agent are how to model the users' intent, how to determine the appropriate actions given that intent and how to develop an appropriate interface between the users and the assistant agents.

The functions required for the tourist route advice system proposed in this research are far too complicated for a single agent to accomplish. It must rely on the co-operation/collaboration among multi-agents, each of which is specialised in solving a particular aspect of the complex problem. An agent platform, such as the one specified by FIPA-OS(FIPA, 2002), provides an infrastructure in which agents can be deployed, communicate and collaborate to achieve the goal of the tasks. Figure 1 shows the agent-based architecture of the proposed tourist route advice system.

The user interaction agent takes the responsibility for dialoguing with users, *i.e.,* presenting questions and information to users and record the users' responses to the system for further analysis. The presentation agent determines what information and map to be presented to the user based on his current position and other information.

The monitoring agent silently observes the users' spatio-temporal behaviour during a trip. It uses the dynamic positioning information of the users' movement to detect the changes of the users' spatio-temporal behaviour. When the change is detected, it informs the inference agent for further analysis and actions.

The user model agent is responsible for establishing, maintaining and updating the user profiles for each individual user. It updates a user profile when it is informed by the inference agent that changes in users' interests have occurred. Since the user profiles are maintained by the user model agent, the only way for the other agents to access the information recorded in user profiles is through the intermediary of this agent.

The Inference agent is the kernel of the whole system. Based on the information received from the user interaction and monitoring agents, the inference agent determines the appropriate actions to be taken, such as generating a new route, updating the user model, presenting updated information to users and/or starting a new dialog sessions for further information when necessary.

The responsibility of the route agent is to generate routes based on the request from the inference agent. To do this, it requires the cooperation of the spatial agent to perform necessary spatial operations and of the user model agent to retrieve the information on user profiles.

The database agent performs all the database management tasks, such as retrieving and updating the data stored in a GIS and/or Database Management System (DBMS).

The user interaction and monitoring agents operate in an active way: they keep acting from the moment a tourist starts using the system and reside on the user's light-weight mobile devices. The other agents act in a more passive way because they are activated only by those two agents when necessary and they reside on the server side for heavy tasks.

The personalized tourist route advice system developed in this research takes advantages of agent-based technology. The personal assistant agents resides on the light-weight mobile devices and they can proactively communicate with the agents on the server that perform heavy tasks such as route generation.
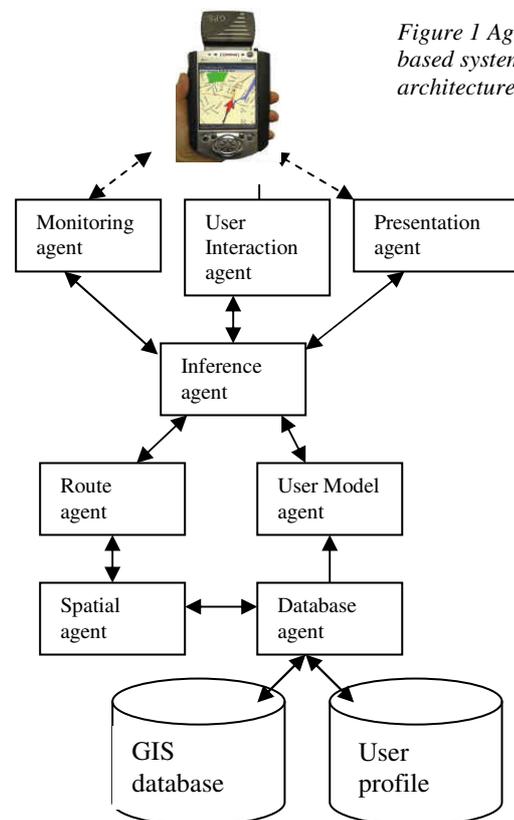

Figure 1 Agent-based system architecture

USER MODEL BUILDING

To build a personalised tourist support system, it is important for the software agent to know what an individual tourist likes. User preferences has been widely studied and applied in e-commerce, where vector representation is mostly used to represent both the products to be recommended and the user's preferences of these products (Lee *et al.*, 2002; Rogers *et al.*, 1999; Yuan and Tsao, 2003).

In the proposed tourist route advice system, the products to be recommended are the tourist routes. Each route is composed of a series of connected road segments. The construction of the route is an iterative process by selecting the most attractive road segment at each step based on the tourist's personal interests. This personal interests is stored as the following vector for each tourist:

$$(W_1, W_2, W_3, \ldots W_i \ldots W_m)$$

where $1 <= i <= m$, m is the total number of attributes characterising the tourist resources. Table 1 gives an example of this vector for tourist X, in which value 0 indicates that he is not at all interested in the corresponding attribute and value 100 means the highest priority.

Table 1 Attributes of tourist resources in a study area

Attributes category	Attributes	Tourist weights on attributes
Landscape	forest	90
	heath	60
	meadow	30
	water area	100
	agriculture field	0
Road type	main road	0
	local road	30
	cycle path	70
	foot path	100
Points of interest (POI)	church	70
	wind-mill	100
	Farm house	70
	restaurant	0
	children-playground	0

The next step in building the user model is to calculate the so-called "attraction value" for each road segment. The attraction value is calculated based on both the tourist resources available along the road segments and tourist's personal interests. The classification and the assessment of the tourist resources is an important data and is usually carried out by tourist experts. For example, the attributes identified in Table 1 is typically related to recreational walk or cycling. In addition, the experts assess the attraction values of each individual tourist resources based on the general tourist evaluation and the (special) characteristics of the tourist features. The assessment is also expressed as a numeric value between 0 and 100, with 100 indicating the highest tourist attraction value. The final attraction value of a particular tourist feature is the multiplication of the tourist's interests and the expert's assessed value. For example, if an area of forest is assessed by experts as having attraction value of 70 and a tourist gives a value of 90 to the forest features in general, the final attraction value of this forest to this tourist would be 70 x 90 = 63.

DATA PREPARATION

Before constructing a route for a tourist, the database has to be prepared. In a GIS, a road network is modelled as geo-referenced network and, for the efficiency of computation, the geo-referenced network can be abstracted as a planar graph that preserves the structure of the geo-referenced network. As explained later, the route construction algorithm is entirely based on the road network. Therefore, we have to link the tourist attraction values calculated before with the road segments.

Calculation of the attraction values for the attributes under the landscape category in Table 1 is complicated. They are related to the scenery attractions along the tourist routes and are represented as area features in the GIS database. However, the algorithm applied in our tourist route advice system considers only road network. Therefore, a link must be built between the landscape area features and the road network elements, *i.e.,* edges and nodes. This link is al

ready implicitly recorded in the GIS database because all features in a GIS database are geo-referenced and the topological relationships of these features is also implicitly stored. The buffer operation provided in a GIS tool is used here to reveal this relationship:
– A buffer polygon is generated along each road segment with a distance, for example 15 meters, along both sides of a road segment(see figure 2). The value of this buffer distance can be adjusted depending on the situations, such as the scenery visibility along tourist routes.
– The percentage of the area of each landscape feature within the total buffered area is then calculated. These percentage values represent the availability of each landscape feature on the corresponding road segment.

To derive the attraction values of each road segment for each tourist, we multiply the percentage values with the attraction values calculated in the previous step. The sum of these values represents the total attraction value of the landscape on each road segment.

Figure 2 The buffer polygon for landscape attraction calculation

ROAD CLUSTERING

Besides some attraction points (Points of Interests – POIs), a tourist route should also pass through as much as possible those road segments with high attraction values. Therefore, we first select those road segments with high values, for example, > 75 and then group them based on their geographic closeness.

This process is called *clustering* and those features that are close together form one cluster. Thus, a cluster includes road segments that have high attraction values and that are also close in geographic space. The dark lines in Figure 3 shows an example of the result of this clustering process, in which 22 clusters are created.
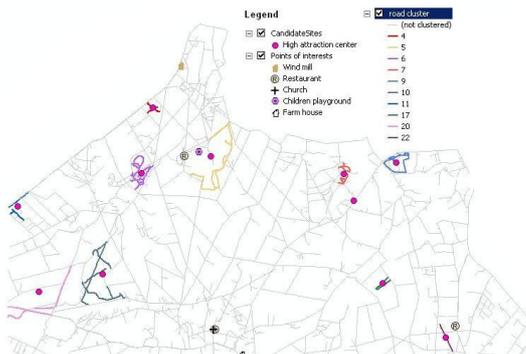

Figure 3 Road segment clustering

After clustering, a point is created for each cluster that represents the location, the weight and the cost values of the cluster. This is achieved by, using the GIS spatial analysis operations, generating the centroid of the bounding polygons of each clustered feature. These centroids are shown as round dots in Figure 3.

These cluster centres together with the existing POI features that have weight or attraction values greater

or equal to a predefined value (50, for example) become the candidate sites, call it "candidate list". The elements in the candidate list will participate in the selection process to construct tourist routes in the next sections.

PATH FINDINGS

One final step in the data preparation is to generate paths between any pair of nodes in the candidate list. The output of this process is a new complete graph, in which nodes represent the candidate tourist sites and edges represent the paths between pairs of tourist sites.

To reduce the complexity, the shortest path finding function in the GIS software is used to generate and record a path between any pair of nodes in the candidate list. However, the objective of the tour route algorithm is to maximize the attraction value while maintaining the distance, time or other cost constraints. This objective is taken into consideration during the shortest path generation by using the weight attribute that represents both the attraction value and the cost value. It is calculated by

$$roadSegmentWeight = \frac{Cost}{totalAttractionValue}$$

where *totalAttractionValue* is calculated in the previous steps, and *Cost* can be distance, time or other cost factors. With this weight attribute, the shortest path function is more likely to select the road segments that have higher attraction and lower cost values. Figure 4 shows an example of such a complete graph, where the nodes represent the attractions points and the edges represent the path between a pair of them.

While generating the paths, the program also calculates the total attraction and the total cost values for each path and these data are stored in the database to be used during the route construction phase described next.
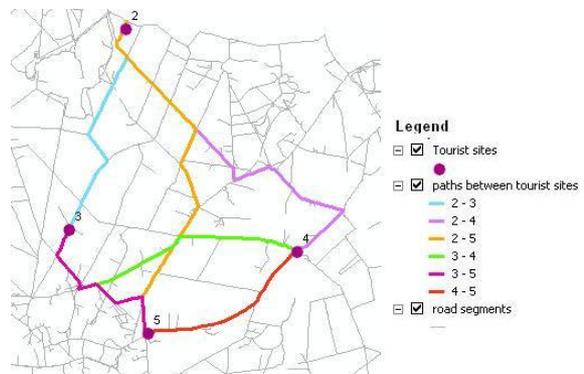

Figure 4 Path generation between pairs of candidate sites

ROUTE CONSTRUCTION

Many relevant studies have already been carried out in the field of operations research, such as *Travel Salesman Problem (TSP), Prize Collecting Travel Salesman Problem (PCTSP),* and *Orienteering Tour Problem (OTP).* These problems are generally called the *Travelling Salesman Sub-tour Problem (TSSP)* (Mittenthal and Noon, 1992).

As Silver (2002) pointed out, all algorithms currently available for finding optimal solutions to TSSP problems require a number of computational steps that grows exponentially with the size of the problem. "It is rather difficult, if not impossible, to obtain the optimal solution of the mathematical model representation of the problem under consideration". Therefore, most of the studies on optimisation look for heuristic solutions rather than for a mathematical model. Another difficulty in the optimisation is that there are often a large number of decision variables, which are combinatorial in nature. Even when the objective function is as simple as linear, there can be many local optima. Applying heuristic methods often results in a reasonable solution to a problem, but one that cannot be guaranteed to produce the mathematically optimal solution.

The *Prize Collecting Travelling Salesman Problem (PCTSP)* is considered a variant of the well-known *Travelling Salesman Problem (TSP)* and has been studied in many researches, especially in the field of operations research (Balas, 1989; Dell'Amico *et al.*, 1998; Mittenthal and Noon, 1992). The TSP is the problem of a salesman who wants to find, starting from his hometown, the shortest possible trip through a set of customer cities and to return to his hometown; visiting exactly once each city. The PCTSP is a relaxed variant of the TSP in which a salesman does not have to visit every city. This may be because he has limited resource available, such as time, distance or budget of travel, and his goal is to reduce cost as much as possible while collecting the minimum required prizes. Each city is associated with a value that represents the prize that the salesman may collect if he visits the city and/or the penalty that he has to pay when he does not visit the city.

A similar optimisation problem is also studied in another application domain - *Orienteering Tour Problem (OTP).* Orienteering is a sport in which each participant is given a compass and a map, on which a number of places, called control points, are shown. One form of orienteering is known as a score orienteering event, in which each control point has a score allocated to it and all participants must reach the finish point within a given time. The control points close to the start and finish points usually have lower scores and those further away have higher scores. The participants do not have to visit all the points. The aim of each participant is to try to collect as high a score as possible within the given time. This requires each participant to solve an optimisation problem to plan his route, to realize his limitations as to which and how many control points he should reach and to use up almost all the available time.

Although the problem definition of PCTSP is different from that of OTP, *i.e.,* minimizing the cost while collecting a sufficient total points versus maximizing the total points while maintaining the cost constraint, respectively, they both consist of two types of decisions, namely, selecting and sequencing. Because not all the nodes need to be visited, decisions have to be made on which node to be visited and which one not. On the other hand, the sequence of visiting the nodes plays also an important role to achieve the objective functions of the problem. A generally applied approach to solve both problem is so-called *Tabu search* – a heuristic method (Silver, 2002). The Tabu search starts with a complete feasible solution, and it continues developing additional complete solutions from a sequence of neighbourhoods, *i.e.,* local improvement. There are variations of Tabu search and the main difference lies in the approach during the improvement phase.

Tourist routing problem is essentially similar to PCTSP or OTP, in which each point in the candidate list becomes a city in PCTSP or a control point in OTP. Dell'Amico *et al.* (1998) proposed an *extension/collapse* algorithm for solving PCTSP and this research adopts this algorithm in constructing tourist routs. For details of this algorithm see Dell'Amico *et al.* (1998).

EVALUATION

Ideally, an evaluation of such a system should be carried out in a real tourist environment, in which the testers would be asked to follow several routes recommended by the software. However, due to resource constraints, mainly the time required to follow the routes, a simulated test environment has been designed for use on computers.

In this simulated test environment, testers enter a weight for each attribute of the tourist attractions. The software then generates four tourist routes – one is generated based on the weights given by the tester, *i.e.,* the "original" preference, and the other three are generated based on the weights adjusted at different levels, *i.e.,* "noisy" preferences, from the weights given by the tester. The objective is to find out if the testers prefer the routes generated based on their "original" preferences to the others that were derived using "noisy" preferences.

In total, forty-five test cases were generated. Besed on the analysis of these forty-five cases, it can be concluded that the preference of the routes by the testers is quite consistent with the weight adjustment given to each route by the software, *i.e.* that the bigger the difference of the weight adjustment is, the less likely the route is favoured by the testers.

CONCLUSIONS

This research designed and developed a personalised tourist route advice system. The agent-based technology offers many advantages for developing such a system. The vector-based user model is simple and efficient in representing tourist personal interests. However, the suitability of vector representation for complex tourist area, such as historical cities, still remains to be evaluated.

GIS is a powerful tool in storing and processing spatial data that are needed for tourist route construction. As it is computationally impossible to derive a mathematical optimum solution for problems such as the one of tour route construction, the heuristic local optimisation approach is applied in this research and the result is shown to be feasible.

REFERENCES

Balas, E., (1989). 'The prize collecting travelling salesman problem'. *Networks*, 19.

Dell'Amico, M., Maffioli, F. and Sciomachen, A., (1998). 'A lagrangian heuristic for the prize collecting travelling salesman problem'. *Annals of Operations Research*, 81.

FIPA (2002) *Agent management specification* http://www.fipa.org/specs/fipa00023/SC00023J.html 2003)

Itami, R.M. and Gimblett, H.R., (2001). 'Intelligent recreation agents in a virtual GIS world'. *Complexity International*, 08.

Lee, W.P., Liu, C.H. and Lu, C.C., (2002). 'Intelligent agent-based systems for personalized recommendations in Internet commerce'. *Expert Systems with Applications*, 22.

Mittenthal, J. and Noon, C.E., (1992). 'An insert/delete heuristic for the travelling salesman subset-tour problem with one additional constraint'. *Journal of the Operational Research Society*, 43(3): 277-283.

Rogers, S., Fiechter, C.N. and Langley, P., (1999) 'An adaptive interactive agent for route adivce', *the third annual conference on Autonomous Agents*, Seattle, Washington, United States

Silver, E.A., (2002) 'An overview of heuristics solution methods', *The 7th Annual International Conference on Industrial Engineering*, Busan, Korea

Wasson, G., Gunderson, J., Graves, S. and Felder, R., (2001) 'An assistive robotic agent for pedestrian mobility', *Fifth international conference on Autonomous agents*, Montreal, Quebec, Canada

Yuan, S.T. and Tsao, Y.W., (2003). 'A recommendation mechanism for contextualized mobile advertising'. *Expert Systems with Applications*, 24.

Zipf, A., (2002). 'Adaptive context-aware mobility support for tourist'. *IEEE Intelligent Systems, special Issue on "Intelligent Systems for Tourism"*(November/December,).