

***e-Tourism* : a tourist recommendation and planning application**

Laura Sebastia, Inma Garcia, Eva Onaindia, Cesar Guzman
 Universidad Politecnica de Valencia
 Dpt. Computer Science
 Valencia, Spain
 {lstarin, ingarcia, onaindia, cguzman}@dsic.upv.es

Abstract

e-Tourism is a tourist recommendation and planning application to assist users on the organization of a leisure and tourist agenda. First, a recommender system offers the user a list of the city places that are likely of interest to the user. This list takes into account the user demographic classification, the user likes in former trips and the preferences for the current visit. Second, a planning module schedules the list of recommended places according to their temporal characteristics as well as the user restrictions; that is the planning system determines how and when to perform the recommended activities. This is a very relevant feature that most recommender systems lack as it allows the user to have the list of recommended activities organized as an agenda, i.e. to have a totally executable plan.

1 Introduction

Nowadays, most people who plan a trip or a day-out will first initiate a search through the internet. More and more people realize the advantages of the new technologies for planning leisure activities [17] as an increasing number of companies and institutions offer tourist information which is easily accesible through web services. However, travelers usually have a limited knowledge of the city to visit and they are unaware of the local artistic, social or entertainment places. A user may find a large amount of information about the city, but he may invest a long time selecting the activities he prefers and organizing them to profitably spend a day-out.

e-Tourism is a web application that generates recommendations about personalized tourist tours in the city of Valencia (Spain). It is intended to be a service for foreigners and locals to become deeply familiar with the city and plan leisure activities. *e-Tourism* makes recommendations based on the user's tastes, his demographic classification, the places visited by the user in former trips and, fi-

nally, his current visit preferences. One of the main components of *e-Tourism* is the planning module which is aimed at scheduling the recommended activities according to their duration, the opening hours of the places to visit and the geographical distances between places (time to move from one place to another). Thus the *e-Tourism* output is a real agenda of activities which not only reflects the user's tastes but also provides details on how and when to perform the recommended activities.

e-Tourism uses comprises the *Generalist Recommender System Kernel (GRSK)* module, a domain-independent, taxonomy-driven recommender system that uses a mixed hybrid recommendation technique, fed by the recommendations obtained from different algorithms.

This paper describes the main characteristics of *e-Tourism*. Section 2 summarizes the state-of-the-art of similar recommenders. Section 3 gives an overview of *e-Tourism* and introduces an scenario that will be used as an example of the *e-Tourism* working model. Section 4 describes the knowledge representation. Sections 5 and 6 detail the GRSK and the planning subsystems, respectively. We finish with some conclusions and future work.

2 Background

A **recommender system (RS)** [14] is a personalization tool that attempts to provide people with a list of information items that best fit their individual tastes. A RS infers the user's preferences by analyzing the available user data, information about other users and information about the environment. In summary, a RS offers the possibility of personalizing the information filtering so that only information tailored to the user's needs and preferences is shown. The adequacy of recommendations depends on the amount of available information [14]. However, the task of introducing information should not be too tedious for the user, so the RS must be able to infer new data items and enrich the user profile as the person interacts with the system [12].

Four **basic recommendation techniques** are distinguished in RS [2]: demographic, collaborative, content-based and knowledge-based techniques. They exhibit some advantages and disadvantages [1]; a common solution adopted by many RS is to combine these techniques into an hybrid RS [11] thus improving recommendations by alleviating the limitations of one technique with the advantages of others. Some hybrid recommendation techniques are described in [2]. The difference between them lies in the way the different RS are combined.

Tourism is an activity strongly connected to the personal preferences and interests of people. For this reason, travel, leisure and tourism web sites tend to incorporate RS for simulating the interaction with a human travel agent [4]. Some examples of tourist web services that use a RS are: DieToRecs [5], ITR [15] or Trip@dvice [16].

However, the use of a RS for traveling and leisure presents several limitations. For example, visiting a same city is not a frequently done activity so that it is difficult to gather different impressions about a particular place and build accurate users's profiles. Moreover, once decided the places to visit, it is interesting to have them organized as an agenda in order to visit as many places as possible and to minimize the movements. The definition of a tourist plan is a time consuming task that involves managing different kinds of information as opening hours, distance between each place to visit or the time spent on the visit. So, the task of the tourism web service is not only to help selecting the places to visit but also to help organizing a plan.

However, it is not so common to find services to generate a tourist route in a city. Some projects (Guide [3], Crumpet [13] or DeepMap [6]) support the user mainly during the tour itself, to move around the city and often provide context-aware information about tourist attractions. But, the organization of a coherent agenda is totally left to the user. Star [9] is a web-based system that supports the tourist to build a personalized agenda for a tour in the city, but the user must specify the tourist attractions to visit in the tour.

The above points reveal that the use of RS in tourism and leisure together with the integration of planning techniques raises a long-term challenge in this particular field.

3 *e-Tourism* Overview

We are developing a tool, called *e-Tourism*, whose goal is to compute a leisure and tourist plan for a user, taking into account his preferences and the information of the context where the visit will take place. Our system does not solve the problem of traveling to an specific place but it focuses on recommending a list of the activities that a tourist can perform in a city. It also considers activities timetables and distances between the activities in order to compute a leisure and tourist agenda. *e-Tourism* is composed of three

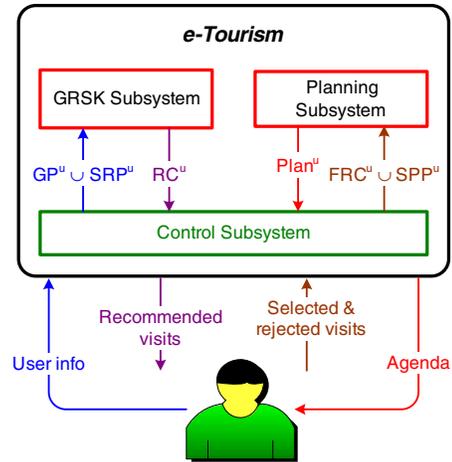


Figure 1. *e-Tourism* system.

subsystems (figure 1): the control subsystem, the *Generalist Recommender System Kernel (GRSK)* subsystem and planning subsystem. The GRSK is a general-purpose module whilst the remaining subsystems depend on the specific application.

To show the working model of *e-Tourism*, this section introduces a scenario that will be used as an example throughout the paper. John, 40 year-old, lives in a city near Valencia with his wife and two children (5 and 8 year-old). He usually goes out with his family. John was in Valencia some months ago and visited two churches: *San Miguel de los Reyes* and *San Nicolas*. He is planning a new visit to Valencia.

The *e-Tourism* first step is to **build the user profile**. The first time John uses *e-Tourism*, he must register and enter his personal details and general preferences. As general preferences, John likes visiting "*Churches*". With this information the system builds an initial user profile which will be updated accordingly with the relevance feedback obtained from the user; i.e. the activities which have been finally performed by John.

Besides this general information, each time John enters the system for a new visit he will be requested to introduce his specific preferences for the current visit (recommendation query): like dates of the visit, his time schedule, whether he is on his own or with children, etc. John usually spends a day out with his wife and children, but his mother joins them in this new visit. John's mother likes "*Gothic Architecture*", so he introduces it as a particular preference for this visit. Moreover, their available time slot is from 12 to 18h.

The module in charge of building the user profile is the **control subsystem**, the core of *e-Tourism*. It works as an user interface, initiates the execution of the other subsys-

tems and centralizes the exchange of information. This includes to convert the user data into a recommendation query, show the user the list of recommended activities, record the activities that are finally selected by the user, and, finally, show the user the calculated plan, i.e. the tourist agenda.

The second step is to **generate a list of activities that are likely of interest to the user**. This list is computed by the **GRSK subsystem**, whose input are the user’s general and specific preferences. The GRSK computes the recommendations according to the current user’s profile and other users’ profiles (depending on the recommendation technique). The calculated recommendations constitute the list of proposed items to visit. Section 5 explains the functional behaviour of the GRSK.

Figure 4 shows the list of recommended activities for John’s visit. Each item is associated a priority and the GRSK selects the ones that better suit John, i.e. the highest priority items. The final list of recommendations is composed of the items which appear in shadow. From this list of recommendations, John picks up the activities he is really interested in (marked with \checkmark in figure 2), and discards those ones he does not want to be included in the final agenda (X). The remaining items are marked as indifferent (\sim).

The third step is to **compute the tourist agenda** with the selected activities of the previous step. At this stage the system schedules the activities according to the time restrictions of the user and the environment. The module in charge of computing the plan is the **planning subsystem**. The input of this module is the set of activities positively selected by the user (\checkmark), the activities left as indifferent (\sim), and other preferences necessary for planning like the user available time. The result is a tourist agenda containing the activities to perform by the user together with the time when the activities should start and the estimated duration of each activity. Section 6 explains how this plan is computed.

The final step is to **process the user feedback**. When the user logs again in the system, he is asked to rate the activities in the last recommended plan. The information obtained from these ratings is used to improve the user profile and gain more suitable recommendations. The management of the user feedback is an ambitious task that is not still finished. We intend the system to be able to dynamically adjust itself to offer higher quality recommendations by obtaining more information from the current user as well as learning from the decisions of all users.

4 e-Tourism Knowledge Representation

This section illustrates the different knowledge *e-Tourism* needs to provide an accurate tourist agenda.

4.1 Taxonomy

The *e-Tourism* behaviour relies on the use of a taxonomy to represent the user’s preferences and the items to recommend. The entities in a taxonomy are arranged in a hierarchical structure connected through a *is-a* relationship where the classification levels become more specific towards the bottom. In the tourism taxonomy, entities represent **concepts** that are commonly managed in a tourism domain like architectonic styles or types of buildings. The leaf nodes of the taxonomy are the **items** *e-Tourism* will recommend to the user and they are categorized by the lowest-level concept in the hierarchy, i.e. the most specific concept. Edges that connect an item to a concept are associated a value to indicate the **degree of interest** of the item (an activity in a tourism taxonomy) under the concept, i.e. as a member of the category denoted by the concept. More formally:

Definition 1 *The taxonomy T is a directed graph (C, E) , where C is the set of nodes of the graph which represent the taxonomy concepts and E is the set of edges that connect a concept with their successors. We distinguish two types of edges: $e_{c-c} = (c_j, c_k)$, which links a concept c_j with a successor concept c_k ; and $e_{c-i} = (c_j, i, r_j^i)$, which links a concept c_j with an item i with a degree of interest $r_j^i \in \mathbb{R}$.*

e-Tourism utilizes a hand-crafted taxonomy that encompasses a set of concepts to describe tourist and leisure activities. These concepts are also used to classify the user’s preferences and interests. The taxonomy is based on the “*Art & Architecture Thesaurus*”¹ which provides terminology about objects, concepts and places important to art, architecture and culture disciplines.

Figure 2 shows part of the tourism taxonomy. The leaf nodes of the graph are the activities that can be performed in the city of Valencia, e.g. “*Visit the Botanical Garden*”. As indicated above, the values of the edges connecting an activity with its most-specific concept show the interest of visiting that place. For example, both the *Botanical Garden* and the *Turia Garden* are parks, but the *Botanical Garden* is more worth visiting as a park than the *Turia Garden*.

Each leaf node of the graph represents an item to recommend (figure 3), which in the particular case of the tourism taxonomy is described by a name, a short description and its location (address). It is also necessary to record the opening hours of places and buildings -which may be different depending on the season or the day in the week- and a standard duration of each activity -which can be personalized for an specific user. This information will be mainly used by the planning subsystem.

¹www.getty.edu/research/conducting-research/vocabularies/aat.

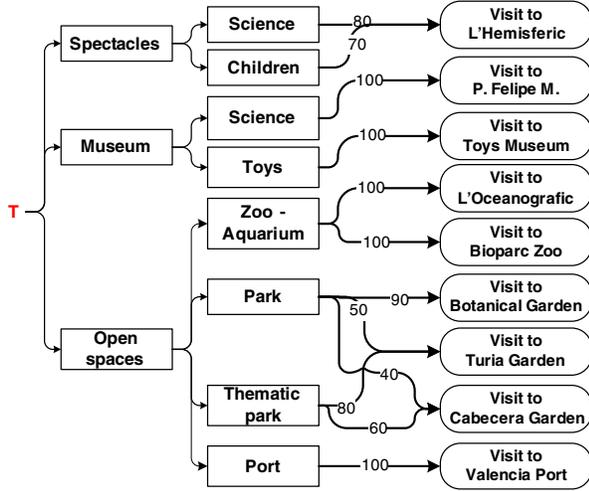


Figure 2. Part of the *e-Tourism* taxonomy.

Visit to L'Oceanografic		
Opening hours:	July, august: 10:00 – 24:00 Others: 10:00 – 20:00	
Address	Camino de las Moreras s/n. 46013 Valencia	
Standard duration:	User Classification	Duration
	1	1h 30'
	2	2h 15m
	3	4h 15m
	4	6h 45m

Figure 3. Item.

Additionally, an activity or item $i \in I$ is defined by a list of features F^i , which represent the incoming edges of each leaf node.

Definition 2 A feature is a pair on the form (c_n^i, r_n^i) , where $c_n^i \in C$ is a concept defined in the taxonomy; $r_n^i \in [\kappa_1, \kappa_2]$ is the degree of interest of the item i under the concept c_n and $\kappa_1, \kappa_2 \in \mathbb{R}^2$.

For example, according to the taxonomy in figure 2, if $i = \text{"Visit the Turia Garden"}$ then F^i could be set to $\{(Park, 50), (Thematic Park, 80)\}$. On the other hand, if $i = \text{"Visit the Botanical Garden"}$ and this garden is considered to be a more interesting park than the *Turia Garden* then $F^{i'}$ will contain that feature with a higher degree of interest, i.e. $\{(Park, 90)\}$. The degree of interest can be dynamically updated through the user feedback.

²In the examples along this paper we assume that $\kappa_1 = 0$ and $\kappa_2 = 100$.

Additionally, items are associated a numeric value AC^i (acceptance counter) to represent how popular the item i is among users; this value indicates how many times this item has been accepted when recommended.

4.2 User Information

This section is devoted to the user information which is necessary in *e-Tourism*. Some part of the user information managed in a tourist domain belongs to the user's profile (e.g. tastes and preferences, historical interaction, etc.). On the other hand, there is information important for the planning process; for instance, information related to the current visit such as the date, the user available time or the user current location.

User Profile

The profile [12] of a given user u defined in *e-Tourism* records, in first place, personal and demographic details about the user like the age, the gender, the family or the country. In our scenario, the demographic details are: John, 40 year-old, lives in a city near Valencia with his wife and two children (5 and 8 year-old). Second, the user general preferences model, denoted by GP^u , contains the description of the types of items the user u is interested in. More formally: $GP^u = \{(c_n, r_n) : 1 \leq n \leq |C|\}$. John has defined as general preference: (*"Churches"*, 30).

A user profile in *e-Tourism* also contains information about the historical interaction of the user with the RS, that is, the set of items the user has been recommended and his degree of satisfaction with the recommendation. In our scenario the historical information contains two visits, $\{(\text{"San Miguel de los Reyes"} , 20), (\text{"San Nicolas"} , 50)\}$, both classified in the taxonomy as *"Churches"*.

Recommendation Query

Each time the user enters the system for a new visit he will be requested to introduce his specific preferences for the current visit (grouped into a recommendation query), which may differ from his general preferences. A recommendation query contains the maximum number N of recommendations the user desires and the set of specific preferences, formally denoted as SP^u . SP^u is divided into specific recommendation preferences $SRP^u = \{(c_n, r_n) : 1 \leq n \leq |C|\}$ and specific planning preferences $SPP^u = \langle date, (Ts, Te), dur_{lunch}, dur_{dinner}, user\ location \rangle$, where $date$ denotes the visit date, (Ts, Te) represent the user available time, dur_{lunch} and dur_{dinner} represent the time reserved for lunch and dinner respectively if the user wants the plan to include the meals and $user\ location$ is the current geographical location of the user. In our scenario: $SRP^{John} = \{(\text{"Gothic Architecture"} , 100)\}$ and $SPP^{John} = \langle 10/8/2008, (12, 18), 1h30', 0, Astoria\ Hotel \rangle$.

5 The Generalist RS Kernel (GRSK)

The task of the *Generalist Recommender System Kernel (GRSK)*[7] is to generate the list of activities to recommend to the user. The GRSK uses the information in the taxonomy to classify the data of the user’s profile and generate the list of recommended activities. It has been designed to be *generalist*, that is independent of the current catalog of items to recommend. The GRSK can work with any application domain as long as the data of the new domain can be defined through a taxonomy representation.

The GRSK uses a **mixed hybrid recommendation technique** that combines the following basic RS techniques: demographic and content-based recommendations and current preference-based filtering. We have defined an independent module for each basic technique and a module that controls the hybrid recommendation techniques. The GRSK has been designed to easily adding new basic or hybrid techniques by simply developing new recommendation modules. The recommendation generated by each basic module is independent from the others.

The **demographic RS technique** classifies the user into a demographic category according to the details of his profile and his general preferences (GP^u). This technique associates a list of the taxonomy concepts to a user type. In our scenario, John is classified as a “*Person with Children*”, because this is the main characteristic of his profile. Therefore the system considers, among other things, the following features to recommend activities: $\{(Zoo-Aquarium, 100), (Thematic Park, 80), \dots\}$. We opted for a demographic RS because it is able to provide recommendation to the problem of having a new user. In addition, it can recommend items which contain different characteristics from other previously recommended items.

The **content-based RS technique** recommends a set of items by taking into account the features of the items previously accepted by the user. Our aim when using this recommendation technique is to increase the user satisfaction by recommending similar items to those already accepted. John has previously visited churches, so this RS will recommend other churches; more precisely, it recommends the *Valencia Cathedral*, as figure 4 shows.

The **current preferences-based filtering** is an information filtering technique [10] that works with the specific user preferences for the current interaction. Basically, it analyzes and stores the specific preferences (SRP^u) that differ from the general preferences (GP^u). This technique allows the user to define some preferences different to his general preferences without modifying his profile. John has defined for the current visit “*Gothic Architecture*” as a specific preference and, for this reason, this RS recommends visiting the *Miguelete Tower* and the *Lonja* (see figure 4).

Each recommendation technique calculates a separated

Recommended Item	Taxonomy		RS ratios			Acceptance Counter	Pr
	Concept	r	DRS	CBRS	CPBF	%AC	
✓ <i>L'Oceanografic</i>	Zoo – Aquarium	100	100			70	270
x <i>P. Felipe Museum</i>	Science museum	100	100			70	270
✓ <i>Valencia Port</i>	Open Spaces	100	90			50	240
✓ <i>L'Hemisferic</i>	Science spectac.	80	80			65	225
x <i>Miguelete Tower</i>	Gothic	88			100	33	221
~ <i>Lonja</i>	Gothic	90			100	24	214
~ <i>BioParc</i>	Zoo – Aquarium	100	100			10	210
x <i>Turia Garden</i>	Thematic Park	80	100			20	200
x <i>Valencia Cathedral</i>	Church	100		30		68	198
<i>Reales Atarazanas</i>	Gothic	80			100	9	189
<i>Serranos Gate</i>	Gothic	70			100	12	182
<i>Quart Gate</i>	Gothic	60			100	11	171
<i>Cabecera Garden</i>	Thematic Park	60	100			10	170
<i>Jardín botánico</i>	Park	90	50			20	160
<i>The Virgin Basilica</i>	Church	80		30		40	150
<i>Santa Catalina Tower</i>	Church	78		30		26	134
<i>Toys Museum</i>	Toys museum	100	30			3	133
<i>San Juan de la Cruz</i>	Church	50		30		4	84
<i>San Juan del Hospital</i>	Church	44		30		5	79
<i>Santos Juanes</i>	Church	35		30		3	68

Figure 4. Items priority.

list of items. These lists of items are then processed by the **mixed hybrid RS technique**. First, it computes a priority for each item in those lists:

$$Pr^a = \frac{AC^a}{\sum_{\forall i \in I} AC^i} * (\kappa_2 - \kappa_1) + r_{RS}^a + r_{taxonomy}^a$$

where AC^a is the acceptance counter of the activity a , r_{RS}^a is the degree of interest of the activity a obtained from the RS technique and $r_{taxonomy}^a$ is the degree of interest of the activity a under the concept of the taxonomy.

The hybrid RS combines the items in each RS list to obtain a single list of recommended items, which is ordered according to the computed priority. In case an item appears in more than one list (that is, it has been selected by several RS techniques), we only consider the appearance with the highest priority. The hybrid RS selects the N best recommendations, which are the set of recommended items for the user u (RC^u). Each recommended activity in RC^u is denoted by a pair of the form $\langle a, Pr^a \rangle$.

Figure 4 shows the initial set of items to recommend to John and the priority calculus. Columns *DRS* (demographic RS), *CBRS* (content-based RS) and *CPBF* (current preferences-based filter) show the ratios applied to the items depending on the recommendation technique. The shadowed items are those recommended to John, that is, they belong to RC^{John} .

6 Planning Subsystem

Following with our scenario, once the GRSK has computed the set of recommended activities RC^{John} , John is shown these activities and asked to mark as selected (✓) those activities he wants to perform and as rejected (X)

those activities he does not want to perform in this occasion. The remaining activities are considered as indifferent (\sim). As figure 4 shows, John selects visiting *L'Oceanografic*, the *Valencia Port* and *L'Hemisferic*, and rejects visiting the *Miguelete Tower*, the *Turia Garden* and the *Valencia Cathedral*. This list of recommended activities filtered by John is sent to the planning subsystem joint with his specific planning preferences. The planning subsystem analyzes this information and builds the user-adapted planning problem whose solution will be the tourist agenda.

6.1 Building the planning problem

The planning subsystem manages three groups of different data:

1. The user's specific planning preferences SPP^u (see section 4.2).
2. The filtered recommended activities FRC^u which is a list of tuples of the form $\langle a, Pr^a, si^a \rangle$, where Pr^a is the priority computed by the GRSK for activity a and si^a is a value in the set $\{selected, indifferent\}$ which indicates whether the user has selected activity a or has no preference for it (the rejected activities are not considered at this stage).
3. The information about each activity a in FRC^u , which is a tuple of the form $\langle a, open_a, close_a, location_a \rangle$, where $open_a$ and $close_a$ indicate the opening hours of activity a (taking into account the date of the visit) and $location_a$ is the address of the place where the activity takes place. These values are extracted from the item information (see figure 3).

All these data are properly analyzed and combined to build the user-adapted planning problem. It is important to note that not all the activities in FRC^u will be likely included in the plan since the scheduling will depend on the user available time, his temporal constraints and the time restrictions of the environment (i.e. opening hours of places). Therefore, the planning subsystem must select which activities to include in the plan. Also, it should consider the user initial location and the time to move from one place to another, that is the distance between two consecutive activities in order to estimate the start time of the latter.

The resulting planning problem [8] can be specified as $\langle actions, init, goal \rangle$. $actions$ is the set of possible actions in the domain and it contains two different sets: the set A of selected or indifferent activities and the set M of movement actions to move from one activity to another. Each action $a \in actions$ is defined by $a = (dur^a, u^a)$. For each activity $A_j \in A$, the duration dur^{A_j} is a value obtained according to the type of user. That is, each user is classified into a different category which determines the duration

of each activity. For example, figure 3 shows four different durations, one for each type of user. On the other hand, the utility u^{A_j} of an activity A_j is a value computed in the following form:

$$u^{A_j} = \begin{cases} Pr^{A_j} & \text{if } si^{A_j} = \text{selected} \\ Pr^{A_j} * \alpha & \text{if } si^{A_j} = \text{indifferent} \end{cases}$$

where $\alpha \in [0, 1]$ is a parameter to weigh the relative importance of the indifferent activities.

Additionally, we add three dummy actions to the set A :

1. $A_0 = (0, \infty)$ to represent the initial action of "being at the initial location".
2. $A_{|A|+1} = (dur_{lunch}, u_{lunch})$ and $A_{|A|+2} = (dur_{dinner}, u_{dinner})$, to represent the actions "having lunch" and "having dinner". If the user wants the plan to include the meals, u_{lunch} and u_{dinner} are set to ∞ and the duration of both actions are specified in SPP^u . Otherwise, the utility of these actions is set to 0. Moreover, we build a tuple of the form $\langle lunch, open_{lunch}, close_{lunch}, location_{lunch} \rangle$ (resp. for *dinner*), where $open_{lunch}, close_{lunch}$ (resp. *dinner*) are set to the typical start/end hours of meals in the city to visit and, for the sake of simplicity, we consider that the locations of these activities coincide with the location of the last performed activity before lunch/dinner.

M is the set of movement actions, where each $M_{j,k} \in M$ represents the action of moving from the place where activity A_j is performed to the place where activity A_k is performed. The duration $dur^{M_{j,k}}$ of an action $M_{j,k}$ is computed by taking into account the distance between activities A_j and A_k . For calculating this distance, we store the city map which comprises all the streets in the city and the intersections of the different sections of each street, which will give us the path to go from A_j to A_k . The utility $u^{M_{j,k}}$ is treated as a penalty, because spending a lot of time in movements is not desirable; it is computed as $u^{M_{j,k}} = -dur^{M_{j,k}} * \beta$, where $\beta \in \mathfrak{R}$ is a value calculated according to the utility of the selected activities.

$init$ is the initial state of the planning problem, which contains the user's available time (Ts, Te) and the opening hours of each activity ($open_{A_j}, close_{A_j}$), whereas $goal$ is an optimization function.

6.2 Resolution of the planning problem

Given the user specific planning preferences, the set A of all the activities that can be performed (including the three dummy actions) and the set M of movement actions, we formulate the resulting planning problem as a *Constraint Satisfaction Problem* (CSP) [8] as follows. We define a set

of variables $a_{i,j} \in \{0, 1\}, \forall i, j \in [1, |A|]$ to denote whether activity A_i is performed in j^{th} place in the plan or not. We post two constraints over these variables: $\sum_{\forall j} a_{i,j} \leq 1$, to prevent activity A_i from being performed twice in the plan and $\sum_{\forall i} a_{i,j} \leq 1$ to avoid performing two activities in the same plan position. In addition, we set $a_{1,1} = 1$ to denote the dummy action *initial location* is the first action to be executed.

From the set of variables $a_{i,j}$ we can infer the necessary movements in the plan. In order to do so, a set of variables $m_{i,j,k}, \forall i, j, k \in [1, |A|]$ are introduced in the CSP. $m_{i,j,k}$ will be set to 1 if activity A_i is performed in the k^{th} position and A_j is performed in the $(k + 1)^{th}$ position, that is, $m_{i,j,k} = a_{i,k} * a_{j,k+1}$.

We also define $ts_i, te_i \in [\max(Ts, open_i), \min(Te, close_i)], \forall i \in [1, |A|]$ to denote the start and end time of activity A_i . The domain of these variables is determined by the opening hours of the different places and the user's available time. Table 1 summarizes the information we manage of each selected/indifferent activity in the planning problem as well as the dummy actions.

Following we specify some additional CSP constraints.

The values of the start and end time of each activity must be consistent with the activity duration:

$$te_i = ts_i + dur_i \quad \forall i \in [1, |A|]$$

For example, the constraint $te_2 = ts_2 + 4$ will set the values of ts_2 and te_2 , which represent the start and end time of activity 2, respectively.

The activities and movements in the plan must not overlap.

$$te_i + dur_{i,j} \leq ts_j + (1 - \sum_{\forall k \in [1, |A|]} m_{i,j,k}) * \Lambda \quad \forall i, j \in [1, |A|]$$

That is, if activity A_i is performed immediately before A_j (that is, $\sum_{\forall k} m_{i,j,k} = 1$), then the start time of A_j must be greater or equal than the end time of A_i plus the time to go from A_i to A_j . Otherwise, we use a constant Λ which takes on a high enough value so as to satisfy the constraint. For example, the following constraint indicates that activities 2 and 3 cannot overlap, provided that it takes 15 minutes to move from *L'Oceanografic* to *Valencia Port*:

$$te_2 + 15 \leq ts_3 + (1 - \sum_{\forall k \in [1, 8]} m_{2,3,k}) * \Lambda$$

The total duration of activities and movements cannot exceed the available time of the user:

$$Te - Ts \geq \sum_{\forall i, j \in [1, |A|]} a_{i,j} * dur_i + \sum_{\forall i, j, k \in [1, |A|]} m_{i,j,k} * dur_{i,j}$$

The system offers the user the choice of selecting the most desirable plan. Thus, we consider two optimization

	Activity	Priority	Opening hours	Dur	Utility ($\alpha = 0.5$)
1	Initial location	∞	0-24	0'	∞
2	L'Oceanografic	270	10-24	4h	270
3	Valencia Port	240	8-24	1h30'	240
4	L'Hemisferic	225	16-21	1h15'	225
5	Lonja	214	10-15	1h	107
6	BioParc	210	10-21	2h15'	105
7	Lunch	∞	13-15	1h30'	∞
8	Dinner	0	0-24	0'	0

Table 1. Activities in our planning problem.

functions: the first one maximizes the utility of the whole plan, whereas the second is aimed at maximizing the time spent in preferred activities. These optimization functions are defined as:

$$\begin{aligned} \max U &= \sum_{\forall i, j \in [1, |A|]} a_{i,j} * u_i + \sum_{\forall i, j, k \in [1, |A|]} m_{i,j,k} * u_{i,j} \\ \max T &= \sum_{\forall i, j \in [1, |A|]} a_{i,j} * dur_i * u_i - \sum_{\forall i, j, k \in [1, |A|]} m_{i,j,k} * dur_{i,j} \end{aligned}$$

In our scenario, when the CSP resolution uses the U optimization function, the obtained plan contains more activities than a plan obtained with the T function. This seems obvious as, in general, the more activities the more utility. In the current example, the plan consists of $a_1 \rightarrow a_5 \rightarrow a_3 \rightarrow a_4$. The corresponding agenda is shown in figure 5. On the other hand, the plan obtained when using the T optimization function contains fewer actions because in this case longer actions are preferable. For the current example, the obtained plan is $a_1 \rightarrow a_2$.

6.3 Tourist agenda

When the system solves these two CSPs, it obtains (at least) two plans which contain a subset of the activities in FRC^u joint with the time when the activities should start and end. This plan is shown as an agenda of activities. A **plan** is defined as a tuple of the form: $\langle date, (a, ts_a, te_a)^* \rangle$, where *date* is the visit date, *a* refers to the scheduled activity, and ts_a and te_a are the start and end time of that activity, respectively.

7 Conclusions and further work

Nowadays there exists an increasing interest on tourism recommender systems as more and more people use travel web services to obtain information for their trips. However, most of the existing services simply provide specific travel items to the user; the generation of personalized tourism tours require, among other things, the incorporation of planning capabilities to properly combine and relate the different travel items.

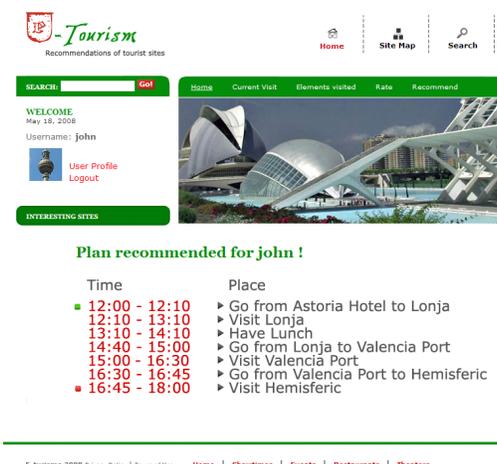


Figure 5. Tourist agenda for John when using maximum utility

e-Tourism is a web service that generates recommendations about personalized tourist tours in the city of Valencia (Spain). It is intended to be a service for foreigners and locals to become deeply familiar with the city and plan leisure activities. *e-Tourism* makes recommendations based on the user's tastes, his demographic classification, the places visited by the user in former trips and, finally, his current visit preferences. The tool shows the user an agenda of recommended activities which reflect the user's tastes and takes into account the geographical distance between places or the opening hours of those places.

e-Tourism is currently under development and there will be an accessible web version very soon.

We also plan to incorporate new hybrid techniques and good metrics in the GRSK to measure the effectiveness of recommendations. Moreover, we are interested in group recommendation as people usually travel on group trips. This introduces a new problematic as now recommendations must adapt the preferences of the majority of users or be in accordance with the common likes of all users.

On the other hand, we are also working on some improvement in the planning process. Our objective is that the planner considers the preferences of the user about how he would like his agenda to be organized. For example, some people prefer visiting museums in the morning and other people prefer to have a relaxed visit with no many activities in the same day. The user preferences can be extracted from the user feedback by analyzing which activities he has finally performed in his former visits, their duration, at which time activities were done, the density of visits or the elapsed time between visits.

Acknowledgments

This work has been partially funded by Consolider Ingenio 2010 CSD2007-00022 project and by the Spanish Government TIN2005-08945-C06-06 project.

References

- [1] Adomavicius G., Tuzhilin A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), 2005.
- [2] Burke R. *The Adaptive Web*, chapter Hybrid web recommender systems. Springer Berlin/Heidelberg, 2007.
- [3] Cheverst K., Mitchell K., Davies N. The role of adaptive hypermedia within a context-aware tourist guide. *Communications of the ACM*, 2002.
- [4] Delgado J., Davidson R. Knowledge bases and user profiling in travel and hospitality recommender systems. In *Proc. of the ENTER*, 2002.
- [5] Fesenmaier D.R., Ricci F., Schaumlechner E., Wober K., Zanella C. Dietorecs: Travel advisory for multiple decision styles. In *Proc. of ENTER*, 2003.
- [6] Fink J., Kobsa A. User modeling for personalized city tours. *Artificial Intelligence Review*, 18(1), 2002.
- [7] Garcia, I., Sebastia, L., Onaindia, E., Guzman, C. Using a generalist recommender system for a tourist planning application. Technical Report TR-GRPS-AI-1, Universidad Politecnica Valencia, 2008.
- [8] Ghallab M., Nau, D., Traverso P. *Automated Planning. Theory and Practice*. Morgan Kaufmann, 2004.
- [9] Goy A., Magro D. Dynamic configuration of a personalized tourist agenda. In *Proc of the ICWI*, 2004.
- [10] Hanani U., Shapira B., Shoval P. Information filtering: Overview of issues, research and systems. *User Modeling and User-Adapted Interaction*, 11, 2001.
- [11] Pazzani M.J. A framework for collaborative, content-based and demographic filtering. *Artificial Intelligence Review*, 13, 1999.
- [12] Pazzani M.J., Billsus D. *The Adaptive Web*, chapter Content-based recommendation systems. Springer Berlin/Heidelberg, 2007.
- [13] Poslad S., Laamanen H., Malaka R., Nick A., Buckle P., Zipf A. CRUMPET: Creation of userfriendly mobile services personalised for tourism. In *Second International Conference on 3G Mobile Communication Technologies*, 2001.
- [14] Resnick P., Varian H. Recommender systems. *Communications of the ACM*, 40(3), 1997.
- [15] Ricci F., Arslan B., Mirzadeh N., Venturini A. ITR: A case-based travel advisory system. In *Proc. of the ECCBR, LNCS*, volume 2416, 2002.
- [16] Ricci F., Cavada D., Nguyen Q.N. Integrating travel planning and on-tour support in a case-based recommender system. In *Proc. of the Workshop on Mobile Tourism Systems (with Mobile HCI)*, 2002.
- [17] Werthner H. Intelligent systems in travel and tourism. In *Proc. of the IJCAI*, 2003.