Context-based semantic matchmaking to enhance tourists’ experiences
Carlos Lamsfus², Christoph Grün¹, Aurkene Alzua-Sorzabal² and Hannes Werthner¹
¹Institute of Software Technology and Interactive Systems
Electronic Commerce Group – Vienna University of Technology
{gruen, werthner}@ec.tuwien.ac.at
²mugiLab - Laboratory for Studies on Human Mobility and Technology
Competence Research Centre in Tourism, CICtourGUNE, Spain
{carloslamsfus, aurkenealzua}@tourgune.org

Abstract
Tourists, preparing for a journey, suffer from information overload when they use the Internet to look for information about their next potential destination. Although approaches to support tourists in decision making exist (e.g., in form of recommendation systems), providing the right information for each type of tourist is still a challenging task. This is especially a major issue when tourists are already at a particular destination and desire to use their mobile devices to consume up-to-date travel-related information tailored to their current situation, i.e. context. This paper presents a context-based matchmaking approach that addresses the needs of tourists during their trip and aims to provide a more satisfying visit experience. In order to identify a set of tourism objects (e.g., attractions) that are most attractive for tourists, the CONCERT framework is presented exploits contextual information such as location and time as a filter to select relevant tourism objects. Within a second step, the matchmaking framework VMTO is introduced that acts on top of CONCERT and ranks the selected tourism objects according to personal tourist preferences.

Keywords: tourism, semantic Web, ontologies, tourist classification, matchmaking, context.

1. Motivation
One of the industries that has benefited enormously from the use of the Internet is the tourism sector. Internet technology has created an online travel market where travel organizations are able to sell their products and communicate with their customers through electronic media. On the other hand, the richness of information that is available online has empowered tourists to exploit the Internet for researching travel-related information and even partially book objects for their trip online. This way, stakeholders (suppliers and consumers) benefit from the use of the Internet for both information research and additional selling channel. Therefore, the provision and consumption of online travel services have become for both nearly a “daily” business.

The penetration of high-end mobile devices equipped with GPS together with decreasing mobile data prices have resulted in an increased usage of mobile services. Tourists therefore like to access travel-related services not only in the pre-trip phase, but especially in the on-trip phase of the tourist life cycle (cf. Figure 1).

![Figure 1. Tourists’ Life Cycle (Werthner & Klein, 1999)](image)

In the on-trip phase, tourists are mobile and act in unknown environments where they especially need personalized, up-to-date on-trip assistance in the form of information about tourism objects (e.g., attractions, museums, restaurants). Mobile tourism services, accessible through mobile handsets, provide the opportunity to cope with the temporal and special constraints. However, using mobile devices to obtain the right piece of information at a given moment of time represents a real challenge. One reason is the limited interaction possibilities due to the small mobile phone screen size and lack of a keyboard, which demand more cognitive work from the tourist.
In order to prevent information overload of the tourist and provide only relevant information, these services should sense and react to the current situation of the tourist. This in turn might lead to an increase of the tourist's satisfaction of experiencing a relaxed sightseeing trip. Thus, given that human mobility is the essence of tourism, and that tourists intensively use mobile devices, how can the right piece of information be delivered to tourists on the move? This goal can be achieved by adapting an application towards the current context of tourists.

The main objective of the context-based matchmaking framework presented in this paper is to support tourists’ mobility at a particular destination by helping them to identify relevant tourism objects matching their personal interests. This objective will be accomplished in a two-step process: firstly, contextual information is exploited to eliminate those tourism objects that do not fit the current situation of the tourist (Lamsfus et al., 2009a). For example, it might be the case that they are too far away and not reachable from the current position of the tourist or that they are temporarily closed. Secondly, a matchmaking process is presented that ranks the selected tourism objects by matching the tourist preferences against tourism objects on the basis of tourist types that are found in scientific tourism literature (Gibson & Yiannakis, 2002). The more similarities they have in common, the more contributes the tourism object to the tourist’s satisfaction and therefore should be ranked higher.

The structure of the paper is as follows. Section 2 outlines relevant related work. Then, Section 3 presents the CONCERT framework that exploits context information to select relevant tourism objects. Section 4 introduces the VMTO approach that ranks the selected set of tourism objects based on tourist preferences. Finally, Section 5 concludes this paper.

2. Related Work

In the following, we report on state of the art in the area of context-awareness, tourism ontologies as well as user profiling.

Context-awareness

Context-aware mobile systems have a long tradition in tourism, which is a very well suited application domain for these kinds of systems (Grün et al., 2008a; Beer et al., 2007). In fact, some studies show that tourists will soon require mobile context-based services (Gretzel & Wöber, 2004; Bernardos et al., 2007; Höpken et al., 2008). However, most of the existing mobile tourism guides and research prototypes do not fully exploit context information in order to adapt the information to the individual situation and requirements of tourists (Grün et al., 2008b). Only few examples exist that provide more personalized information by taking into account various context information together (Höpken et al., 2006; Beer et al., 2007; Höpken et al., 2008). Even though the CAIPS system (Beer et al., 2007) provides rule-based push information, it does not present a general framework to support mobility. It is rather focused on the modelling and definition of rules that can be used by destination management organisations in order to push personalized content to visitors.

The study of context in tourism plays a crucial role, since it is an information facilitator in the negotiation process between all available tourism information (offered by travel organizations) and the information that visitors require at a given moment of time based on their situation. Context is the link between the need for information and the information itself.

Since research in context-awareness began almost 20 years ago, two phases can be clearly distinguished. The first one covered the 1990s and was primarily focused on studying the notion of context under different disciplines. The main objective of these applications was to enhance human computer interaction by providing the application with context information. Although context was studied as a secondary variable in these applications, intensive work was conducted in setting its theoretical foundations, and thus the most important definitions of context were originated in that decade (Want et al., 1992; Schilit and Theimer, 1994; Dey et al., 2001).

During the 1990s ontologies were still not very well known in the computer science community. Some ontology-based applications appeared during these years in different domain areas, such as knowledge management, information integration etc., but still the benefits of ontologies had not been fully discovered and their applicability in context-aware systems had not been experienced yet.

However, as of the year 2000, the second phase in context-aware research began, aiming at establishing a standard context management model that leverages semantic technologies. In fact, one of the greatest differences between the two phases is the use of semantic technologies to model context and to manage context information. However, in the second phase researchers have been more concerned in
developing an ontology-based standard context management framework (Chen et al., 2004; Gu et al., 2004; Chen et al., 2005; Strimpakou et al., 2006; Preuveneers et al., 2006) rather than working on its theoretical foundations.

Existing approaches with respect to context-awareness in the field of human mobility have more or less re-used the conception of context and ignored the requirements for context modelling within this specific field. Not surprisingly, the consequent models do not completely suit its requirements.

In most recent years ontologies have been widely used in pervasive computing and have been pointed out as adequate tools for context modelling and management (Strang and Linnhoff-Popien, 2004; Bettini et al., 2009). They can be used to integrate, share and re-use context knowledge stemming from distributed and heterogeneous sources of information and in addition, facilitate the usage of logical reasoning capabilities that can be used both to check the context model’s consistency and to make implicit information explicit.

Tourism ontologies

Recently, there is a proliferation of ontologies that have been developed in the area of e-tourism. QALL-ME (Ou et al., 2008) is an EU-funded project that aims at establishing a shared infrastructure for multilingual and multimodal question answering in the tourism domain. Thereby, it allows users to pose natural questions and returns a list of answers in the most appropriate modality.

The HARMONISE (Fodor & Werthner, 2005) ontology is now the central element within the Harmonisation Network for the Exchange of Travel and Tourism Information that aims to create an International network for harmonization and data exchange in the tourism industry. The ontology focuses on two sub-domains of the tourism domain, namely events and accommodation.

CRUZAR’s ontology (Mínguez et al., 2009) is based on the upper-ontology DOLCE in order to model visitor’s profiles, travel routes and POIs. To describe POIs, it further reuses properties from the Dublin Core, FOAF and SKOS-Core. SPETA exploits concepts from the e-tourism ontology (Cardoso, 2006) in order to describe tourist services. In addition, it links to concepts from DBPEDIA and YAGO to describe concepts such as attractions or activities and FOAF to describe social links of tourists. DTG's ontology (Hagen et al., 2005) is built leveraging some existing taxonomies from DAML and GETTY.

User profiling

As tourists have individual preferences, tourist profiling plays an essential role in the provision of personalized travel information. With this purpose in mind, the classical user model employed for personalization (Kobsa, 2001) needs to be adapted to model the needs and interests of tourists.

Based on Maslow’s hierarchy of individual need (Maslow, 1954) a huge number of studies of tourist motivation have been conducted that investigate different factors why certain tourists undertake specific types of travel (Page, 2009). Related to this field is the work of (Vogt & Fesenmaier, 1998) that proposes a model of information needs of tourists forming a categorization of different types of needs, including functional, recreational, as well as aesthetic needs. Such needs, wants, behaviour and expectations of tourists can be further classified into tourist types. Cohen (1974), distinguished four different tourist types, comprising the organized mass tourist, the individual mass tourist, the explorer and the drifter. A set of 15 different tourist types such as the action seeker, active sport tourist or thrill seeker have been proposed (Gibson & Yiannakis, 2002). For example, the thrill seeker is described as type of person “interested in risky, exhilarating activities which provide emotional highs for the participant”. As pointed out in Page (2009), such an absolute classification to classify tourists may not take into account the diversity of holidays tourists undertake and the inconsistencies in tourist behaviour. Over time, tourists might change their behaviour or have a mixed profile. It is examined whether such predefined travel types can be used to obtain personalized recommendations (Gretzel et al., 2004). Thereby, each of the predefined travel types is linked to certain activities. The study showed that travel types are indeed a useful means to capture the interests of tourists with respect to certain activities.

3. CONCERT: Contextual Computing in Tourism

CONCERT’s objective is to study the context of visitors of a particular destination with respect to the field of human mobility. The goal is to determine in a more precise way the information that formally describes the mobile visitor context and to define requirements of such applications. In this sense,
context is regarded as the main entity, not an auxiliary variable for studying something else as in earlier approaches (cf. Section 2). Hence, it is of high importance to study the factors that define the context of people on the move and to find out, which kind of information is at least needed to describe a mobile visitor’s context (Lamsfus et al., 2009a).

On a practical level, CONCERT proposes not to use sensor infrastructure to gather contextual information. Populating cities, regions and/or open areas with networks of sensors is not affordable.

\[\text{Figure 2. The double interoperability level provided by CONCERT}\]

In order to tackle this barrier, the CONCERT framework gathers both contextual and tourism data from the Internet as well as from mobile embedded sensors (e.g., GPS), and does not require further complex infrastructural deployments (cf. Figure 2b). In this way, the need to populate a particular area of interest with sensors is avoided, and in addition, the usage of an application based on the CONCERT framework is not limited to those areas (cf. Figure 2a). The idea behind this approach is to increase the level of abstraction of context (thus, increasing the level of portability -interoperability- of the model), by providing the model with the means to access context information in an anywhere-anytime manner.

Finally, context information has to be translated into a consistent computing model so that it can effectively assist tourists in their mobility, and additionally enhance and improve their visiting experience. Different context models have been identified and analyzed according to ubiquitous computing environments requirements (Strang et al., 2003) and context-aware applications (Chen and Kotz, 2001; Bettini et al., 2009). Both studies indicate that ontologies clearly address all context modelling requirements and are an adequate tool for that purpose. Besides, ontologies have proven to be useful in information integration, re-use as well as sharing, which are essential to provide interoperability at the model level.

In an attempt to increase the level of scalability, modularity and interoperability at the model level, context is modelled in CONCERT by means of a network of ontologies (Barta et al., 2009; Haase et al., 2006). This network of ontologies, called ContOlogy, incorporates the requirements identified in the field of human mobility and implements them in terms of motivation, preferences-demographics and role, according to standard parameters established by the tourism scientific community (UNWTO, 2008; Lamsfus et al., 2009b). At the moment the ContOlogy integrates 11 ontologies. All together there are 86 classes, 41 object properties, 22 datatype properties and 43 restrictions. The language used to specify each of the ontologies has been OWL (W3C, 2004b) in its DL sublanguage. The level of expressivity shown by the network of ontology is SHOIN(D). The following table depicts ContOlogy’s main components.

<table>
<thead>
<tr>
<th>Ontology</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visitor (UNWTO, 2008)</td>
<td>Characteristics of the human being in mobility.</td>
</tr>
<tr>
<td>Preferences (UNWTO, 2008)</td>
<td>Information that describes the visitor’s personal characteristics, demographics, etc.</td>
</tr>
<tr>
<td>Role (UNWTO, 2008)</td>
<td>The role a visitor plays at a given moment.</td>
</tr>
</tbody>
</table>
Activity (Strimpakou et al., 2005) Represents what the visitor is doing at a given moment. This information can be taken from the mobile device’s agenda. (COMANTO)

Environment (Preuveneers et al., 2005) Represents the surroundings of the visitor (CoDaMoS) as well as the weather conditions at the location of the visitor.

Device (W3C, 2004a) Physical object the visitor carries with him/her.

Network (Cadenas et al., 2009) Infrastructure to connect devices and convey information.

Motivation (UNWTO, 2008) Represents the reason why the visitor is travelling.

Location (Preuveneers et al., 2005) Coordinates that define where a visitor is at a given moment of time.

Time (W3C, 2006) Physical dimension that measures span between facts.

Tourism Objects Represents the services provided in a certain environment.

Table 1: The ContOlogy Network of ontologies

A rule-based information engine built on top of the ContOlogy network filters the incoming tourism objects with respect to certain context factors at a given moment. These factors represent the situation of a visitor, i.e. his/her preferences, current weather condition and location, etc and used to filter personalized information as shown in the example in Figure 3.

For instance, the second rule depicted in Figure 3 takes into consideration the food preferences of a particular visitor. Given the visitor’s location, environment, preferences, time of the day, etc., a number of restaurants are offered to the user, that are in his/her surroundings and match his/her context.

4. VMTO: Vector-based Matching of Tourists and Tourism Objects

The task of the VMTO framework is to define a function that matches tourist profiles against the filtered set of tourism objects (obtained from CONCERT), in order to produce a ranked list of objects for each given tourist. If a tourist profile matches the characteristics of an object, this object should be recommended to the respective tourist. Therefore, the matchmaking algorithm has to examine whether they share similar structures.

To estimate the similarity degree between tourist profiles and tourism objects, VMTO follows a vector-based matchmaking approach, whereby a given profile and each tourism object constitute vectors and are compared in a vector space model. A common method to obtain the similarity is to measure the cosine angle between two vectors. If the vector space is non-orthogonal, kernel based algorithms can be applied to measure the similarity in such a space. The dimensions of the vector space model correspond to selected tourists types found in scientific tourism literature (Gibson & Yiannakis, 2002), such that each distinct tourist type (e.g., adventure or cultural type) represents one dimension in that space.
Thereby, a tourist profile vector indicates the degree to which tourists identify themselves with the given types. Typically, individual tourists cannot be characterised by only one of these archetypes but have mixed profiles as they show attributes of several types, although to varying degrees. Thus, tourist types model the tourists’ generic interests in an abstract form. Vectors are suited to model such tourist types, whereby each dimension corresponds to a certain tourist type while the value indicates how much the tourist identifies him- or herself with the corresponding type.

![Figure 4. Modeling the tourist profile in a vector](image)

Figure 4 depicts an exemplary tourist, who likes to enact in the role of an adventurer, followed by sport and sightseeing, and rather dislikes cultural activities.

In the same way as the tourist profile is represented in form of a vector, every tourism object is modelled through a vector as well. Thereby, this vector describes in a quantitative way how much the object is related to the given types. For example, the famous cathedral “Stephansdom” (Staint Stephen’s cathedral) in Vienna might be highly relevant for sightseeing tourists but not for such kind of tourists that would like to do some risky activities.

As destinations usually offer a huge number of different tourism objects to visitors, we propose a semi-automatic process to link the given tourist types to appropriate tourism objects. Therefore, in a first step, domain experts mark for each of the prototypical tourist types (e.g., adventure or cultural types) a small sample of typical tourism objects that are closely related to these types. The degree of relationship is specified with different weightings (cf. Figure 5).

![Figure 5. Linking prototypical tourist types to typical tourism objects](image)

That is done individually for each of the tourist types. After this step, certain tourism objects (e.g., the Spanish Riding School) are linked to their most corresponding tourist type (e.g., the cultural type). Some tourism objects (e.g., the clock museum) might not have been linked to any of the tourist types by the domain expert. In addition, some information might be missing. For example, we do not know whether the Spanish Riding School is also relevant for other tourist types such as the explorer. To tackle these issues, ontological knowledge is exploited to define a similarity metric between the different tourism objects. This similarity metric can then be used to propagate the weightings throughout the semantic network (of tourism objects). For example, based on the fact that the Albertina Museum is highly relevant for cultural tourists and given the statements #Albertina_Museum rdf:type Museum and #Museum_Modern_Art rdf:type Museum it can be derived that the Museum of Modern Art might be relevant for cultural tourists as well. Not only rdf:type and rdfs:subClassOf relations can be exploited to derive the similarity between objects, but also user-defined properties. If two objects have the same architectural style or are linked to the same people (e.g., house of Habsburg) they also have a higher similarity degree. The weightings are thus dependent on the relationships within the semantic network.

For establishing a similarity metric, all tourism objects are semantically annotated based on a tourism ontology. For defining a similarity metric, ontology-based similarity approaches can be used, including taxonomy-based and feature-based similarity measures. After this step, every tourism object has now a vector that expresses the correlation to each of the given tourist type in a quantitative way.
5. Conclusion

eTourism represents an active field of research in several disciplines, comprising computer science, human-computer interactions, recommendation systems, as well as mobile pervasive computing. This paper presents a context-based semantic matchmaking framework in order to propose personalized tourism objects to tourists on their trip, thus enhancing their experiences while they are at their destination. The framework combines two existing approaches, i.e. the contextual computing framework in tourism, CONCERT, and the VTMO ranking framework. Although the emphasis of this paper is on promoting context-based services in the realm of tourism, it also aims to provide new aspects in the realm of ubiquitous computing, semantic technologies and recommendation systems.

The CONCERT framework is based upon a thorough review of relevant literature and its main contributions can be summarized as follows. First, it provides a new theoretic approach to study the notion of context, which is framed within the field of human mobility. In addition, CONCERT provides a double interoperability level: at the infrastructural level by not using sensors to gather contextual information and at a model level by using networks of ontologies in order to be able to share, re-use and integrate contextual knowledge.

However, tourists have individual preferences, which CONCERT does not address in a holistic way, since its profiling approach is based on the concept of roles, which are defined by the UNWTO and describe different types of purposes. User profiling and classification are an important research issue in tourism and especially tackled by recommender systems. Understanding users’ interests is an important prerequisite for delivering personalized information. Thus, the VMTO framework defines a function that matches tourists’ profiles against filtered sets of tourism objects in order to produce a ranked list of objects for each given tourist. This way, the ranking produced by the VTMO enhances the context-based functionality provided by the CONCERT framework.

References


