User Preferences in Tourist Itineraries Recommendation

Pierpaolo Di Bitonto, Francesco Di Tria,
Maria Laterza, Teresa Roselli,
Veronica Rossano, and
Filippo Tangorra

Department of Computer Science - University of Bari "Aldo Moro"
Via Orabona 4, 70125 – Bari, Italy
{dibitonto, francescoditria, marialaterza, roselli, rossano, tangorra}@di.uniba.it

Abstract

One of the most interesting current challenges in the e-tourism field is to offer services that are able to suggest attractions or event itineraries that fit tourist's needs and preferences. In this scenario, we defined a method for generating itineraries of intangible cultural heritage events (processions, special markets, festivals, and so on) based on a theoretical model and the transitive closure computation. The method was, then, implemented, in a knowledge-based recommender system prototype, named T-Path, which is able to suggest event itineraries in Apulia region (in southern Italy). In this paper we propose an approach to identify the itineraries that best fit the tourist/user preferences. At this aim, first of all the type of each itinerary is discovered, and then the itineraries are ordered on the basis of their adequacy to the tourist preferences.

Keywords: Recommender system; Tourist itineraries; Cultural heritage events

1 Introduction

In the latest years the tourism domain has shown a progressive change. More and more people plan their holidays on Internet, surfing on the web in order to search new destinations, flights, accommodations, and so on. Moreover, during their stay the tourists continuously access to the web for searching interesting attraction or events. The current trend in the tourism domain is not only to build systems that are able to suggest and supply information about tourist spots such as temples, churches, museums (Scherp & Boll, 2004), or restaurants (Ricci, & Nguyen, 2004; Yap, Tan, & Pang, 2007), but also to design and implement tourist guides that are able to suggest interesting itineraries to the user (Kramer et al., 2006; Biuk-Aghai et al., 2008; De Choudhury et al., 2010) like a human guide. The main contribution of this research in the tourism field is the proposal of a recommendation process that, indentifying the itineraries that best fit the tourist preferences, supports the educational aspect of the tourism. In other words, the process proposed is able to suggest a set of event that is interesting for the user from a cultural point of view. In this context, we defined a method for automated generation of itineraries of intangible cultural heritage events (such as festivals, special markets, processions, etc.) in a knowledge-based recommender system. The method defined is based on a theoretical model (Di Bitonto et al., 2010a) that consists of a set of functions to characterize each event and a space-time relation. It aims to find chains of events using transitive closure computation. Moreover, the method was implemented in a knowledge-based recommender prototype, named T-Path (Di Bitonto et al., 2010b), that is able to build and suggest chains of intangible cultural heritage events in Apulia region. Up to now our research has been focused only on the itinerary construction, in this paper we propose an approach to identify the type of itineraries and to order them according to the user preferences. The research paper is organized as follows. Section 1 presents an overview of both the proposed method for generating event itineraries and the T-Path prototype. Section 2 presents the approach used to identify the itineraries most suited to the user. Finally, in section 4 some conclusions and future developments are drawn.

2 Automated Generation of Tourist Itineraries

In order to generate tourist itineraries of intangible cultural heritage events (festival, special markets, processions, etc.) occurring in a geographical area, a theoretical model was defined (Di Bitonto et al., 2010a). In the model each event is characterized by means of a set of functions *Start*, *End*, *Location*, and *Distance,* that allow to define when the event occur, where the event occur and how far the event location is from another one. Moreover, in order to build the chains of events a space-time relation was defined. According to this relation, two events are correlated in space-time dimension if they are near both in space and time. Then the transitive closure computation constructs chains of correlated events considering each event as initial. For example, given the set of events $\{e_1, e_2,\ldots, e_n\}$ and an initial event e_i, if there exists another event e_{i+1} that has a space-time relation with e_i, the path $e_i \rightarrow e_{i+1}$ is generated and so on in a recursively way until there is no other event that satisfies the space-time relation. The result is an itinerary composed of a chain of events that have a pairwise space-time relationship ($e_i \rightarrow e_{i+1} \rightarrow e_{i+2} \rightarrow \ldots \rightarrow e_n$). A deeper and more detailed discussion of the proposed method is reported in (Di Bitonto et al., 2010a).
The model for generating itineraries has been implemented in the T-Path prototype. The three-level T-Path architecture is shown in Fig. 1.

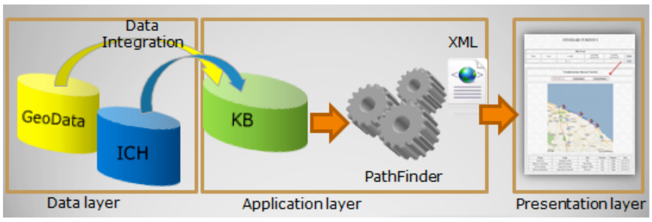

Fig. 1. T-Path architecture

The Data Layer consists of two databases named GeoData and ICH (Intangible Cultural Heritage) and the data integration module. The GeoData database stores geographical information about the territory in which the events occur, ICH database stores information about each cultural heritage event. The integration component uses this information for generating the facts expressed in Predicate Calculus. The Application Layer is composed of the knowledge base, where the facts and the logical program PathFinder are stored. The PathFinder is written in Prolog and generates different itineraries starting from a set of unordered events (using the transitive closure computation as described in the previous section). The result of the computation is a set of itineraries stored in an XML file that are proposed to the tourist by means of the T-Path interface (Presentation Layer).

Fig. 2. Example of event itinerary in Apulia Region

In order to validate both the model defined and the T-Path prototype, the information about the most important religious events occurring between the end of 2009 and the first months of 2010 in Apulia region were used. In particular, we suppose to have three religious events "Bethlehem Nativity" (6 in Fig. 2), "Magi's Night" (7 in Fig. 2) and "Nativity in Marchesale Palace" (8 in Fig. 2), and let us consider 8 be the initial event. We also suppose that 8 and 7 are correlated in space and time, and the same is for 6 and 7. The method allows to construct the itinerary 8→6→7, and T-Path proposes to the user the itinerary depicted in Fig. 2.

3 User Preferences in Itinerary Recommendation Process

In order to improve the suggestion provided by T-Path, the current step of our research is to define an approach to identify the most relevant itineraries for the user. The process consists of two steps: the first step discovers the type of each itinerary according to the type of events that it contains; the latter identifies the itineraries that best fit the tourist preferences. For this purpose, each event is classified into several categories, for example *civil*, *religious*, *cultural*, *artistic*, and so on (for the sake of simplicity, hereinafter only the categories *civil* and *religious* are considered). Moreover, we assume that an event can belong to only one category and the categories are all equiprobable. Let us suppose to have an itinerary P composed of A, B, and C events, where A is *civil*, B is *civil*, and C is *religious*, the probabilistic inference calculated by means of a Bayesian Network (BN) (Pourret, Naim & Marcot, 2008) can be used to discover the type of P. The BN was chosen for two main reasons: first, it is able both to do a causal and diagnostic inference, second, it can discover the itinerary type even if the categories of one or more events are unknown. Fig. 3 shows the BN built after the learning process using the three events listed above. The evidence propagation (A is *civil*, B is *civil*, and C is *religious*) in the BN is depicted in Fig. 4. It discovers that P is civil with the 66.7% of probability and religious with the 33.3%. So, we can represent this probabilistic knowledge about the itinerary P using the vector {0.67, 0.33}. It is important to notice that the BN allows the process to infer the itinerary type even if some event metadata are missing (Fig. 5).

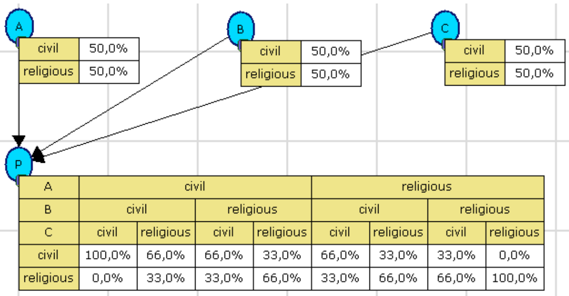

Fig. 3. Bayesian Network for discovering the type of P itinerary

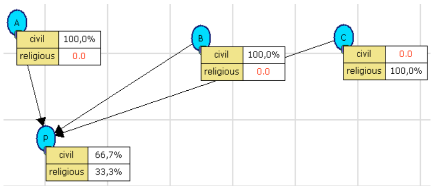

Fig. 4. Evidence propagation in the BN when all event metadata are known

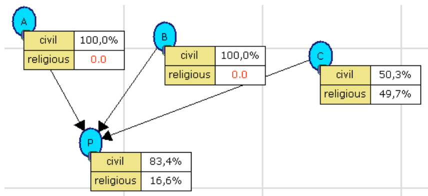

Fig. 5. Evidence propagation in the BN when some event metadata are missing

Moreover, using a BN, rather than a membership function, for classifying the itinerary, it will be possible to assign significance to each event category, assigning different conditional probabilities to use during the inferring process. In this way the itinerary type can be inferred according to the significance associated to each event category. After the identification of the type of the itineraries, it is now necessary to answer the research question: "which itineraries are the most suited for the tourist?" For this aim, also the user preferences are represented by means of a vector. The user vector is built using a user model that collects the information during the user interaction and infers the user preferences. For example, the vector U = {0.6, 0.4} describes the user interests (60% in *civil* events and 40% in *religious*). At this point, in order to discover which the best itineraries are for the tourist, the Euclidean Distance between each itinerary vector and the user vector is measured. Then, the itineraries will be ordered according to the distance measured, the nearest will have priority in the suggestion list. For example, if we consider user vector U = {0.6, 0.4} and two itineraries, P = {0.67, 0.33} and P' = {0.2, 0.8}, the Euclidean Distances *d*(P, U) and *d*(P', U) is measured as follows:

$$d(P,\ U) = \sqrt{|0.67 - 0.6|^2 + |0.33 - 0.4|^2} = 9.89949493661167E\text{-}02 \quad (1)$$

$$d(P',\ U) = \sqrt{|0.2 - 0.6|^2 + |0.8 - 0.4|^2} = 0.565685424949238 \quad (2)$$

Since d(P, U) < d(P', U) the conclusion is that P is most suited for the user than P'.

4 Conclusions and Future Works

In tourism domain the offer of advanced technologies or services to support the tourist before and during the travel is increasing. Most research efforts are focusing on design and development of tourist guides that are able to suggest attractions or event itineraries. In this scenario, we proposed a method for automated generation of itineraries of intangible cultural heritage events. The method has been implemented in a knowledge-based recommender prototype that is able to suggest itineraries in Apulia region. But, until now we did not consider the user preferences in itinerary recommendation. For this aim, the paper proposes an approach to identify the itineraries most suited to the user preferences using a BN and the Euclidean Distance measure between the type of itinerary and the user preferences. The main lack of the proposed approach is the assumption that each event can belong exclusively to one category. Since, this limitation is too restrictive, a future development will be the use of fuzzy Bayesian Network (Fogelberg, Palade, & Assheton, 2008) to infer the type of itinerary. In this way, it will be possible to infer the type of itinerary when an event is, for example, *civil* with 40% of probability and *artistic* with 60%.

References

Di Bitonto, P., Di Tria, F., Laterza, M., Roselli, T., Rossano, V. & Tangorra, F. (2010). A Model for Generating Tourist Itineraries. International Conference on Intelligent *System Design and Applications*. In printing.

Di Bitonto, P., Di Tria, F., Laterza, M., Roselli, T., Rossano, V. & Tangorra, F. (2010). Automated Generation of Itineraries in Recommender Systems for Tourism. In F. Daniel and F.M. Facca (Eds.), *International Workshop on Web Engineering & Tourism.* Vienna, Austria, Springer-Verlag.

Biuk-Aghai, R.P., Fong, S. & Yain-Whar, Si (2008). Design of a Recommender System for Mobile Tourism Multimedia Selection. *Internet Multimedia Services Architecture and Application.* Bangalore, India, IEEE.

De Choudhury, M., Feldman, M., Amer-Yahia, S., Golbandi, N., Lempel, R. & Yu, C. (2010). Constructing travel itineraries from tagged geo-temporal breadcrumbs. *International Conference on World Wide Web*. Raleigh, North Carolina, ACM Press.

Fogelberg C., Palade, V. & Assheton P. (2008). Belief Propagation in Fuzzy Bayesian Networks, *International Workshop on Combinations of Intelligent Methods and Applications.* Patras, Greece.

Kramer, R., Modsching, M., & ten Hagen, K. (2006). Field study on methods for elicitation of preferences using a mobile digital assistant for a dynamic tour guide. *ACM symposium on Applied Computing.* New York, USA, ACM Press.

Pourret, O., Naim, P. & Marcot, B. (2008). *Bayesian Networks: A Practical Guide to Applications*. Chichester: Wiley & Sons, Inc.

Ricci, F. & Nguyen, Q. (2004). User preferences initialization and integration in critique based mobile recommender systems. *Workshop on Artificial Intelligence in Mobile Systems.* Nottingham, UK.

Scherp, A. & Boll, S. (2004). Generic support for personalized mobile multimedia tourist applications. *ACM Multimedia 2004.* New York, USA, ACM Press.

Yap, G. E., Tan, A. H. & Pang, H. (2007). Discovering and exploiting causal dependencies for robust mobile context-aware recommenders. *IEEE Transactions on Knowledge and Data Engineering* 19(7):977-992.