Tour Plan Using Ontology, Formal Concept Analysis and Bayesian Analysis

[1]Prof. Vishal R. Shinde, [2]Vinay R. Marathe, [3]Arun R. Kamani, [4]Prem A. Kalekar

[1]Asst. Professor, [2,3,4]UG Student, [1,2,3,4]Computer Engg. Dept. Shivajirao S.Jondhle College of Engineering & Technology, Asangaon, Maharshatra, India.

[2]imvinaymarathe007@gmail.com, [3]arunkamani@gmail.com, [4]premkalekar8@gmail.com

Abstract - This paper presents that Ontology is playing a crucial role in knowledge management and semantic web. The tourism data ontology has become a core research field in the domain of data retrieval. The intent of this learning is to scrutinize the capable role of Formal Concept Analysis (FCA) for integration of different tourism information. The ontology for tour commuters and the integrated ontology for tourism data Providers are mapped in order to obtain tourist's predilections against the information published by tourism data providers. Two ontologies are developed, one for tour commuters and the other for tourism data providers. Both of these ontologies are mapped using Bayesian analysis and Formal Concept Analysis to evaluate tourist predilection against the information presented by the tourism data providers.

Keywords— Formal concept analysis, Bayesian analysis, Ontology, recommendation system.

I. INTRODUCTION

Data Mining consists of a following sequence of steps: data preprocessing, data mining, knowledge presentation and pattern evaluation. The mining step may a knowledge base or interacts with the user. The really interesting patterns are presented to the user and may be created and saved as new information in the knowledge database.

Due to internet, growth of data happen from terabytes to petabytes, but today's real world databases are highly susceptible to inconsistent data, noisy, and missing, due to their typically big size (often several gigabytes or more) and their likely origin from multiple, different sources. Low quality data will lead to low quality data mining results. In order to help improve the quality of the information and, consequently, of the data mining results, and to improve the efficiency and ease of the mining process, need data preprocessing. Therefore, as one of the most important background knowledge, data preprocessing plays a fundamentally important role in Data Mining. There are a Number of data preprocessing techniques. Concept hierarchy is one of them.

In a proposed method to extract user interest information ontology concept is used. Ontology has been defined as "the branch of metaphysics that studies the nature of existence or being as such". Ontology is a system of conviction that reflects an analysis of an individual about what constitutes a fact. In simple terms, ontology is associated with a central question of whether social entities need to be perceived as objective or subjective.

Recent developments in ontology make it promising to integrate different information such as that available from on-line sources. Ontology includes a set of ideas and the associations between information. The meaning of the online information can be easily arranged in a uniform manner according to ontologies, and dissimilar information can then be integrated. Applying concept hierarchy on tourist information user can efficiently plan there tour with proper information. Planning a tour on-line requires detailed statistics about several characteristics of tourist fascinations, such as the activities provided by them, admission fees, open hours, and the routes between them.

The ontology for tour commuters and the integrated ontology for tourism data providers are mapped in order to evaluate tourists' predilections against the information published by tourism information providers. The mapping in between two ontologies is performed at two levels. The First level maps the conceptions between them, and the second level matches the properties for a given set of mapped conceptions.

The First objective is to build an ontology to characterize the perspective of tourists from the tourism literature. The second is to develop an ontology for tourism statistics providers that integrates heterogeneous on-line tourism statistics using a Formal Concept Analysis (FCA) approach. The third thing is to map between the ontology for tour commuters and the integrated ontology for the tourism statistics providers, using an FCA and a Bayesian approach in order to devise a tour plan. The first step matches provided activities with user preferred activities using the Bayesian analysis. In the second

step, the attractions in city that correspond to users preferred provided activities are retrieved. The third step ranks these retrieved attractions based on user's predilections.

In a proposed method, implementing ontology of tourist information. Perceptual constraints, functional constraints are applied on ontology development. For implementing hierarchy of tourist information using ontology, use concepts and sub concepts which are decided and organize according to user interest and tourist place ranked. The proposed method works for only tourist information.

II. LITERATURE SURVEY

1. DEVELOPING A WEB-GIS BASED JUDGEMENT SUPPORT SYSTEMS FOR TOURISM PLANNING. AUTHORS: A.Mansourian, M.Taleai, Z.Bahari sojahrood. Preparing a tour and selecting tourist's attractions at a destination is an important spatial problem in tourism planning activities. In this approach, a web-based spatial judgment support system (SDSS) is developed to provide modified tour plan for each tourist. Based on TOPSIS method, as one of multi criteria judgment making methods, the system can rank tourist fascinations regarding her/his predilections. Multi-objective judgement making method is utilized for grouping selective tourist's attractions and changing those to different days that tourist will stay at the destination. Developed systems can make some recommendations related to tour plan and best route to visit places based on GIS tools and several criteria such as: opening and closing time of each place, less time to stay at each location, distance between each two places and etc. The system is inserted over a Web GIS. Developed system is tested in a case study at Tehran, Iran.

Once a destination has been selected, we can find many sites which giving us information about places to visit in a city, activities to do during the travel, restaurants, etc. But this information is static in most cases and is presented to all users in the same way. Also, the tourist is faced with much information and must select manually useful and interested information from them. Current tourist websites do not automatically provide plans and schedules, according to user needs and sites schedules. So, from the user point of view, it is useful to have a recommender system that tells him/her which places may be interesting to visit in a certain city taking to his/her profile and favourites, computing a tourist daily plan indicating which places to visit in the given timeframe, and also how to go from one place to another.

TOPSIS model is used to rank the attractions. TOPSIS (technique for order predilection by similarity to a perfect solution) methodology is conferred in. The fundamental principle is that the chosen different ought to have the shortest distance from the best resolution and therefore the farthest distance from the negative-ideal resolution.

2. BAYESIAN NETWORK AND ANALYTIC HIERARCHY PROCESS BASED PERSONALIZED RECOMMENDATIONS FOR TOURIST ATTRACTIONS OVER THE INTERNET.

AUTHORS: Ling Bian, Yuxia Huang. Selecting tourist attractions to visit at a destination is an important stage in an arrangement of a tour. Although various online travel recommendation systems have been developed to support users in the task of travel planning during the last decade, few systems focus on mentioning specific tourist attractions. In this approach, a smart system to provide recommendations of tourist attractions in an unknown city is presented.

Through a touristry ontologies, the system permits combination of heterogeneous on-line travel data. supported Bayesian network technique and also the analytic hierarchy method (AHP) methodology, the system recommends traveler attractions to a user by taking into consideration the travel behavior each of the user and of different users. GIS functions square measure provided by spatial internet Service Technology. Additionally, it offers a cooperative geographic interface for displaying the approval results also as getting users' opinion.

The purpose is to estimate a user's preferred activities of visiting attractions in a Bayesian network based on that person and other users' travel behaviours. Three stages are involved to complete this task: the qualitative stage, the quantitative stage, and the updating stage. Tourist attraction selection is a multi- criteria decision problem, which deals with the evaluation of a set of alternatives in terms of a set of decision criteria. The AHP method deals with decision-making problems by determining the relative importance or weight of criteria through pairwise comparison of the criteria. The pairwise comparison matrix expresses the relative importance of each criterion over others. Recommendations of tourist attractions involve spatial data and spatial handling functions.

3. CASE BASE QUERYING FOR TRAVEL PLANNING RECOMMENDATION.AUTHORS: Hannes Werthner, Francesco Ricci. This approach describes the final design and performance of associate degree intelligent recommendation system aimed toward supporting a leisure mortal within the task of choosing a holidaymaker destination, bundling a group of product, and composing an idea for the travel. The system allows the user to spot his/her own destination and to individualize the pass by aggregating elementary things (additional locations to go to, services, and activities).

Case-based reasoning techniques alter the user to browse a repository of past travels and alter the ranking of the elementary things enclosed during a recommendation once these are selected from a list.

The system integrates information and knowledge originating from external, already existent, holidaymaker portals exploiting associate degree XML-based intermediary design, information mapping techniques, similarity-based retrieval, and on-line analytical process.

Coming up with a travel towards a commercial enterprise destination could be a complicated downside determination activity. The term "destination" itself, i.e. the ultimate goal of the itinerary, refers to a thought that's unclear and lacks a unremarkably in agreement definition. 1st of all, the spatial extension of the destination is understood to be a perform of the someone distance from the destination. for example, Italy could be a destination for a Japanese, however not for a au mortal WHO could concentrate on a selected region, like Italian region, instead of on a historical town or one thing else. Moreover, a destination isn't a mere geographical entity; it's going to be a would like or a group of activities or experiences. There in sense a destination can't be conceptually separated from a itinerary to the destination, i.e., a destination is reshaped within the holidaymaker perspective in some ways per the activities that he can perform in (towards) that destination. Modeling this obscure conception and therefore the call method that leads totally different users to their most well- liked destination continues to be associate degree open and difficult analysis downside.

III. AIMS & OBJECTIVES

Aims

1) By using ontology help tourisms for planning
2) Search user satisfied results.
3) Design a tour plan according to user requirements

Objective

1) The First objective is to develop an ontology to represent the perspective of tourists from the tourism literature.
2) The second is to build an ontology for tourism information providers that integrates heterogeneous on-line tourism information using a Formal Concept
Analysis (FCA) approach.
3) The third objective is to map between the ontology for tourists and the integrated ontology for the tourism information providers, using an FCA and a Bayesian approach in order to devise a tour plan.

IV. PROPOSED SYSTEM

A single tourism website scarcely offers all needed information, while different web sites may provide diverse and often conflicting information about the same attraction. For a given piece of information, such as admission fees, it may be presented in various terms, typically 'cost', 'admission price', and 'entry fee', as well as the term

Table: Comparative Study

Sr.No	Paper Title	Author's Name	Approach used	Merits	Demerits
1	A Bayesian Network and Analytic Hierarchy Process Based Personalized Recommendations for Tourist Attraction.	Y. Huang and L. Bian.	Analytic hierarchy process.	Proposed for recommendation of tourist attractions at a given destination and recommendation based on Location.	Not work on problem of how to plan these attractions as a tour.
2	Ontological Recommendation Multi-Agent for Tainan City Travel.	S. Lee, Y. C. Chang and M.-H. Wang.	Ontological recommender.	Proposed to recommend a tour plan for the traveller.	Not find optimized solution.
3	Personalized Location-Based Recommendation Services for Tour Planning in Mobile Tourism Applications.	C. Yu and H.-P. Chang.	Location-Based Recommendation.	Proposed for tour recommending, including travel agent.	The problem of automatically finding semantically annotated Sequences.
4	Tour-Mine: An Efficient Tour Planning Approach with Travel Time Constraints.	Eric Hsueh-Chan Lu and Chih-Yuan Lin Vincent S. Tseng.	Tour Mine Approach.	Propose a novel data mining based approach, namely *Tour-Mine*, to efficiently find the optimal tour which satisfies the user's travel time constraint based on the user's location. Achieves significantly high efficiency in optimal tour finding.	Not develop data mining-based approaches to automatically find the interesting attractions and estimate the stay time of attractions.
5	A Novel Approach to Mining Travel Sequences Using Collections of Geotagged Photos	S.Kisilevich, D.Keim and L.Rokach.	Geotagging approach.	Addresses the problem of automatically finding semantically annotated sequences based on geotagged photos.	It did not consider the travel constraints and user predilections.

'admission fees' itself. For an effective tour plan, the semantically heterogeneous information needs to be integrated and evenly, consistently represented, otherwise tour planning on-line is often seen a annoying, boring and frustrating experience.

Recent developments in ontology research make it possible to integrate diverse information such as that available from on-line sources. A set of concepts and the relationships between them is what that is included in ontology. The meaning of the on-line information can be understood in a uniform manner according to ontologies, and heterogeneous information can then be integrated. Tourism information providers provide information about attractions at a destination over the Internet.

The suppliers are often numerous government and non-profit organizations, intermediaries (e.g. travel agents), business enterprise service suppliers and tourists themselves. Focuses on data provided by government and non-profit organizations as their information tends to be inclusive, objective, and structured, so ideal for ontologies development.

Though tourists-generated social media info has become progressively necessary for tour coming up with, it tends to be fly-by-night and amorphous. Business enterprise internet sites, particularly those provided by the govt. and non-profit organizations, stay as higher info systems. Tour commuters and tourism information providers, as the represented by an ontology, and the two perspectives can be bridged through ontology mapping approaches for deriving a tour plan that deals with both tourist's predilections and the information published by the tourism information providers.

ONTOLOGY FOR TOUR PLANNING

Ontologies are utilized in tourism research. One kind of research is to develop a globally adoptable standard for tourism info suppliers to uniformly represent their info. Associate example is that the wordbook on tourism and Leisure Activities outlined by the world tourism Organization (WTO). This standard will facilitate unify tourism terminology to assist Ontology consists of a collection of ideas and also the relationships between them. It's a proper illustration of the common vocabulary of an information area. These ideas and their relationships are underneath an agreement inside a section of information or across totally different areas of information effectively seek for tourism info. However, there's a scarcity of central authority to enforce such a standard. Instead, varied localized tourism ontologies are developed and these ontologies exist over the web. These efforts primarily concentrate on the event of one ontology for the tourism info suppliers. Presently, the challenge is to integrate these native ontologies. One planned approach is to link a pair of native ontologies at a time, however this is often a tedious and cumbersome task. The event of a central ontology will effectively integrate multiple native ontologies. This integrated ontology is a reference to perceive the heterogeneous tourism info from varied tourism info suppliers. Moreover, an efficient tour set up needs 2 ontologies to represent 2 totally different views. Matching between the 2 ontologies is important to support the tour set up. Sadly, there has been less attention paid to mapping the ontologies of various views, specifically during this case, between the ontology for tourists and therefore the ontology for tourism info suppliers.

MAPPING OF ONTOLOGIES

Ontologies for tour commuters and tour information providers are mapped in order to evaluate tourist preferences against the information published by tourism information providers. The mapping between the two ontologies is performed at two levels. The first level maps the concepts between them, and the second level matches the properties for a given set of mapped concepts.

FORMAL CONCEPT ANALYSIS (FCA)

Formal concept analysis (FCA) may be a technique of information investigation with growing quality across varied domains. FCA analyzes data that describe relationship between a specific set of attributes and a specific set of objects. Such information typically seem in several areas of human activities. Formal concept Analysis produces 2 sorts of output from the input file. The primary may be a concept framework. A set of formal ideas within the information that are hierarchically ordered by a sub-concept- super-concept relation is named conception framework.

BAYESIAN ANALYSIS

Bayesian analysis is an arithmetical criterion that gives answers for research questions about unknown parameters using probability statements. Bayesian inference uses the posterior distribution to form various analysis for the model parameters, including point estimates such as posterior means, medians, interval estimations known as credible intervals, and percentiles. Moreover, all statistical tests about model parameters can be conveyed as probability declarations based on the predicted posterior distribution. Distinctive features of Bayesian analysis include an ability to incorporate prior information in the analysis, a perceptive interpretation of credible intervals as fixed ranges to which a parameter is known to reside with a pre-specified probability, and an ability to appoint an actual probability to any hypothesis of interest.

V. ALGORITHMS ALGORITHM FOR FINDING SHORTEST PATH

1. Create a set spsSet (shortest path tree set) that keeps track of vertices included in shortest path tree, i.e., whose minimum distance from source is calculated and finalized. Initially, this set is empty.
2. Allocate a distance value to all vertices in the input graph.
3. Initialize all distance values as INFINITE.

4.Assign distance value as 0 for the source vertex so that it is picked first.

5. While spsSet doesn't include all vertices

(a) Pick a vertex x which is not there in spsSet and has minimum distance value.

(b) Include u to spsSet.

(c) Update distance value of all adjacent vertices of x. To update the distance values, iterate through all adjacent vertices.

6. For every adjacent vertex y, if sum of distance value of x (from source) and weight of edge x-y, is less than the distance value of y, then update the distance value of y.

VI. SYSTEM ARCHITECTURE

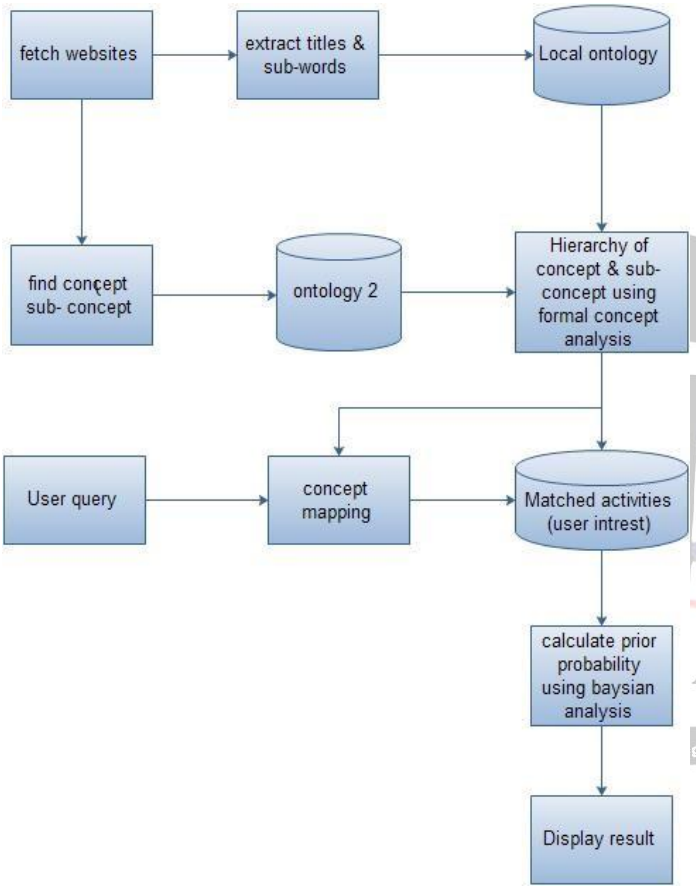

Fig. 1 System Architecture

VI. MATHEMATICAL MODEL

Let, system S={I,OC,MC,B,BA,O,D}

• OC is ontology creation function
• OC= {LO,FCA,IO}
• LO is local ontology consist of user's attraction

LO={ L1,L2,L3…… }

FCA = {T,OF} is Formal Concept Analysis method applied on local ontology for creating integrated ontology

T={G,M,I}

T is triplet where

G and M are two sets of elements called objects and attributes respectively, and I is a binary relationship between them.

OF={r, n , c}

OF is output from FCA which use for creating hierarchy concepts

r is reference concepts

n is new concepts

c is local concepts

I is input, I ={U,Q}

U is users of Systems = {U1,U2,U3….)

Q is queries of users = {q1,q2,q3…..}

q1 contains four values ={T,B,A,M}

T is preferred tour time

B is preferred budget

A is preferred activities

M is preferred mode

MC= {A,OF,OM}

MC is concept mapping in which users preferred activities is map with provided activities.

A is preferred activities

OF is sub concepts hierarchy

OM is user attraction places which map by preferred activity.

BA is Bayesian analysis is used to compare preferred activities and provided activities values. In a Bayesian analysis, a hypothesis of a concept, presented as the prior probability (P(H)), represents an initial belief. Additional evidence (P(e | H)) represents the likelihood of the hypothesis and is used to update the prior probability to the posterior probability (P(H)) , as in

Where P(e) is a normalization constant

P={p1,p2,p3……}

O is Output extracted from provided activities which have highest rank with tour planning.

O={pa, pb , pc…..}

Pa= {d1,d2,d3…..}

d=day

Pb= {a1,a2,a3…..}

a= city

D={G,SP,So,En} G={V,E,ds}

V={v1, v2, v3,…..}

Where v1,v2,v3 are vertices

E={e1,e2,e3….}

Where e1,e2.. are edges

Ds distance between vertices

So is source point of tourism

En is end point of destination

SP is shortest path.

VII. EXPECTED OUTPUT

The tourist registers his personal information in the website then he login into the website. He is then requested for input regarding a particular tourist destination. After giving input, Ontologies and Formal Concept Analysis (FCA) generates a tour plan which is suggested to the tourist. Also the shortest path from the Source to destination is provided. The output generated is shown is the given figures.

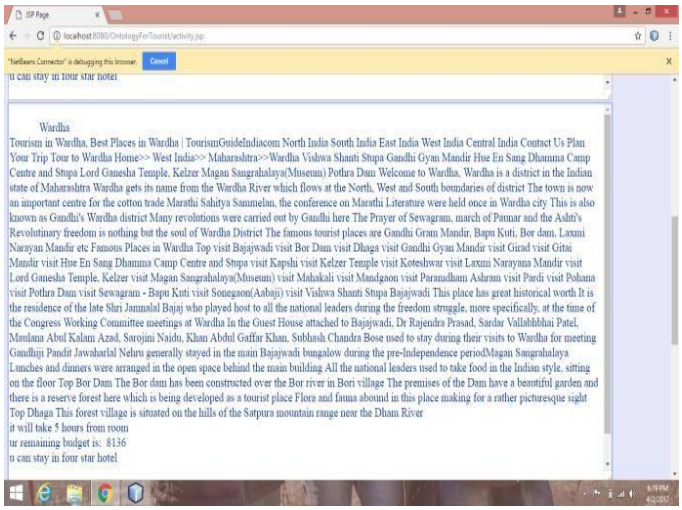

Fig.2 Generated result.

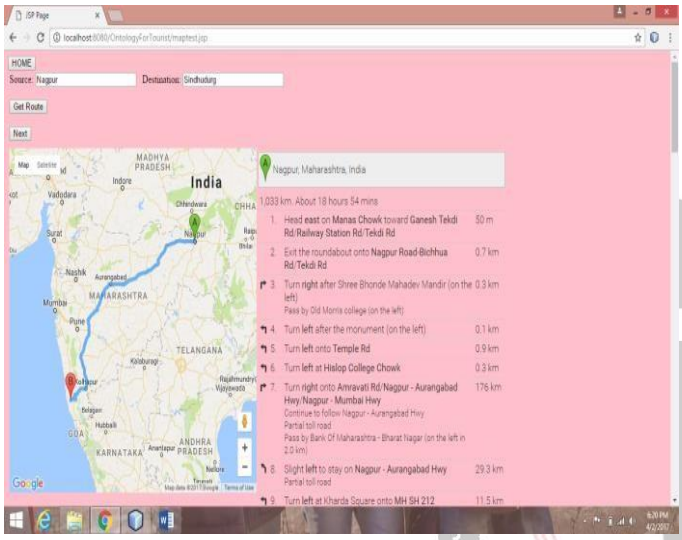

Fig.3 G-Map

VIII. CONCLUSION

We have tried to implement the paper "Using Ontologies and Formal Concept Analysis to Integrate Heterogeneous Tourism Information "by Yuxia Huang and Ling Bian , IEEE, June 2015. And according to the implementation conclusion is that Ontology for tour commuters and in particular an integrated ontology for tourism data providers are developed, and the two ontologies are mapped to match the perspectives between tourism service users and providers. In the context of ontology research, the approaches developed. The semantic web, the next generation of web, has been considered the potential direction for on-line tourism information systems, including tour planning. The integrated ontology for tourism information providers can help develop semantic web-based tourism information systems. This system focuses on tourism attractions it can be stretched to other services that are also commonly involved in a tour planning, such as selecting accommodation and dining services. Similar to suggesting tourism attractions, a pair of ontologies, one for the service users and the other for the service providers are needed and then mapped between them to select appropriate hotels or restaurants.

REFERENCES

[1] "Using Ontologies and Formal Concept Analysis to Integrate Heterogeneous Tourism Information" yuxia huang and ling bian1School of Engineering and Computing Sciences, Texas AM University Corpus Christi, Corpus Christi, TX 7841 USA,2015.

[2] J. Barzilai, ''Deriving weights from pairwise comparison matrices,'' J. Oper. Res. Soc., vol. 48, no. 12, pp. 1226–1232, 1997.

[3] L. Bian and S. Hu, ''Identifying components for interoperable pro- cess models using concept lattice and semantic reference system,'' Int. J. Geograph. Inf. Syst., vol. 21, no. 9, pp. 1009–1032, 2007.

[4] P. du Boucher-Ryan and D. Bridge, ''Collaborative recommending using formal concept analysis,'' Knowl.-Based Syst., vol. 19, no. 5, pp. 309–315, 2006.

[5] J. L. Crompton, ''Motivations for pleasure vacation,''

Ann. Tourism Res., vol. 6, no. 4, pp. 408–424, 1979.

[6] M. Dell'Erba, O. Fodor, W. Hopken, and H. Werthner,

''Exploiting seman- tic Web technologies for harmonizing e-markets,'' Inf. Technol. Tourism, vol. 7, nos. 3–4, pp. 201–219, 2005.

[7] G. Ergun and P. R. Stopher, ''The effects of personality on demand for recreation activities: Some preliminary findings,'' Transp. Res. A, vol. 16, no. 1, pp. 55–63, 1982.

[8] D. R. Fesenmaier, B. Pan, U. Gretzel, and Y. Huang, ''Tourist judgement model report,'' DIETORECS, Trento-Wien- Urbana, Tech. Rep. IST- 200-29474, 2002.

[9] R. J. Gitelson and D. L. Kerstetter, ''The relationship between sociode- mographic variables, benefits sought and subsequent vacation behavior: A case study,'' J. Travel Res., vol. 28, no. 3, pp. 24–29, 1990.

[10] U. Gretzel, N. Mitsche, Y. H. Hwang, and D. R.

Fesenmaier, ''Tell me who you are and I will tell you where to go: Use of travel personalities in destination recommendation systems,'' Inf. Technol. Tourism, vol. 7, no. 1, pp. 3–12, 2005.

[11] T. R. Gruber, ''A translation approach to portable ontology specifications,'' Knowl. Acquisition, vol. 5, no. 2, pp. 199–220, 1993.

[12] T.-C. Huan, ''Monitoring and exploring U.S. international travelers' behaviors in Canada: Development and application of index scores,'' Ph.D. Dissertation, Dept. Forestry Natural Resour., Purdue Univ., West Lafayette, IN, USA, 1997.

[13] Y. Huang, ''Toward semantic interoperability in travel planning on the semantic Web: Learning a reference ontology from online heterogeneous tourist attractions classification systems,'' Inf. Technol. Tourism, vol. 11, no. 1, pp. 51–66, 2009.

[14] Y. Huang and L. Bian, ''A Bayesian network and analytic hierarchy pro- cess based personalized recommendations for tourist attractions over the Internet,'' Expert Syst. Appl., vol. 36, no. 1, pp. 933–943.

