ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

Tour Plan Using Ontology, Formal Concept Analysis and Bayesian Analysis

Supriya Namdas[1], Sapana Dolas[2], Shubham Sankale[3], Prof.Archana Kadam[4]

Department of Computer Engineering, Pimpri Chinchwad College of Engineering, Nigdipune, Savitibai Phule Pune University, Pune, India

ABSTRACT: in this paper we presents Ontology is playing a crucial role in knowledge management and semantic web. The tourism information ontology is become a core research field in the domain of information retrieval. The intent of this study is to scrutinize the potential role of formal concept analysis (FCA) for integration of heterogeneous tourism information. The ontology for tourists and the integrated ontology for tourism information Providers are mapped in order to evaluate tourist's preferences against the information published by tourism information providers. Two ontologies are developed, one for tourists and the other for tourism information providers. Both of these ontologies are mapped using FCA and Bayesian analysis to evaluate tourist preference against the information published by the tourism information providers. In proposed method user comments also considered for creating tour plan. User comments helps to improve accuracy of system as it contain positive and negative information about places. By using ontologies and user comments user can get queries related recommendation and hence user get more satisfy with results.

KEYWORDS: formal concept analysis, Bayesian analysis, Ontology, recommendation system.

I. INTRODUCTION

In recent years, the number of e-tourists has outstandingly increased in reaction to the expanding availabilities of on-line tourism information. Planning a tour on-line necessitates detailed data (information) regarding many aspects of tourist attractions, such as the activities provided, their open timings, fees for acceptance, and the course between them.

Single tourism web site can hardly supply all required information, while different web sites may provide various and often contradictory information about the identical attraction. For a given section of information, such as admission fees, it may be put forward in various expressions, typically 'cost', 'admission charges', and 'entry charges', as well as the term 'admission price'. For a successful tour plan, the semantically heterogeneous information needs to be integrated and consistently represented, or else tour planning on-line is generally seen as a monotonous, unexciting and frustrating experience [1].

Recent progress in ontology research makes it feasible to integrate diverse information such as that accessible from on-line sources. Ontology consists of a set of concepts and the alliances between them. The definition of on-line information can be interpreted in a uniform way as stated by ontologies, and heterogeneous information can then be integrated.

There are two types of participants involved in tour planning, the tourists and the tourism information providers. Tourists are the people who travel to stay away from their usual habitat for some short period of time for leisure, adventure, for business purposes, to gather information, to study and for other purposes. This paper primarily focuses on those tourists who travel for leisure. The purpose of tourism information providers is to provide all the necessary information regarding the attractions of a place over the internet. The providers can be different government and non-profit organizations, intermediaries (e.g. travel agents), tourism service suppliers and tourists themselves. This paper concentrates mainly on the information provided by government and non-profit organizations as their information tends to be comprehensive, objective, and structured, thus ideal for ontology development. Although the information generated by tourists on social media has become significantly important for tour planning. It tends to be unreliable and unstructured. Tourism web sites, especially those which are provided by the government and non-profit organizations, still remain a better information system. Tourists and the tourism information providers both have their own different approach in tour planning. Each perspective can be represented by ontology, and the two perspectives can be bridged

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

through ontology mapping approaches for producing a tour plan that consists of both tourist's preference and the information published by the tourism information providers [9].

II. RELATED WORK

In this paper, proposed a data mining-based approach, namely Trip-Mine, to efficiently find the optimal trip which satisfies the users travel time constraint based on the user's location. Furthermore, we propose three optimization mechanisms based on Trip-Mine to further enhance mining tech efficiency and memory storage requirement for optimal trip finding. To the best of our knowledge, the advantage of this is the first work that takes efficient trip planning and travel time constraints into account simultaneously [9].In this paper they proposed the Collaborative Location Recommendation (CLR) framework for location recommendation. Considering activities (i.e., temporal preferences) and different user classes (i.e. Pattern Users, Normal Users, and Travellers) in the recommendation process, CLR capable of generated more precise and refined recommendations to the users compare to the existing methods. CLR employs a dynamic clustering algorithm CADC to cluster the data into groups of similar users, similar activities and similar locations efficiently by supporting incremental update of the groups when new GPS system trajectory data arrives. Authors evaluated CLR with a real-world GPS dataset, and confirm that the CLR framework provides accurate result location and more proper recommendations compared to the existing methods [10].This paper provides a conceptual foundation for analysis of spatio temporal data of places visited by people worldwide using community contributed geo tagged photo collections. First, they define several types of spatio temporal cluster of people's visits. Second, discussed methods that can be used for analysis of these clusters. Third, they offered an analysis of tourist activities [3].Formal Concept Analysis (FCA) is a method for data analysis, knowledge representation and information management that is widely unknown among all information scientists in the USA even though this technology has a multiple significant potential for applications [4] [6].In this paper, they selected tourist attractions to visit at a destination is a main stage in planning a trip. Although various online travel recommendation systems has been developed to support users in the task of travel planning during last decade, few systems focus on the recommending specific tourist attractions for users. In this paper, an intelligent system is used to provide personalized recommendations of tourist attractions in an unfamiliar city is presented. Through a tourism ontology, the system allows integration of heterogeneous online travel information. Based on Bayesian network technique and the analytic hierarchy process method, the system recommends tourist attractions to a user by taking into account the travel behaviour both of the user and of other users. Spatial web services technology is embedded in the system to provide GIS functions. In addition, the system provides an interactive geographic interface for displaying the recommendation results as well as obtaining users feedback. The experiments show that the system can provide personalized recommendations on tourist attractions that satisfy the user [2].Internet had significantly influenced the tourism sector providing a great variety of services and products online. However, the number of choices had increased so dramatically that is very difficult for the consumers to find what they are looking for. For this purpose, recommendation systems for tourism have attracted a lot of research energy and interest. The main characteristic of these systems is that they can personalize their recommendations to each user interacting with the system. Personalization is even more essential for tourism recommendation systems used in handheld devices where the screen is even smaller and the presentation capabilities are Limited. This paper addresses these problems and provides some development steps for a tourism recommendation system by making a state of the art in personalized e-tourism services both in computers and hand held devices as well as a review of the user modelling and personalization techniques used in these systems. Furthermore, the theories used for the improvement of the personalization procedure in tourism recommendation systems; their applications and evaluation are discussed [10].The Attraction System: an empirical analysis of tourist behaviour Greg Richards Attractions have been viewed by many as central to the tourism process. Tourist attractions are often the reason for visiting a particular destination; they provide activities and experiences at the destination and a means of collecting the signs of tourism consumption. It is not surprising, therefore, that a great deal of attention has been lavished on the production and consumption of tourist attractions [12].This paper presented an ontology-based approach and a formal concept analysis (FCA) approach to tourism information for online tour planning. Two ontologies, first ontology for tourists and second ontology for tourism information providers, are developed with respect to own perspective. The ontology for tourists is developed from the tourism literature survey. The ontology for tourism information providers is developed by integrating heterogeneous online tourism information using a FCA approach. The two ontologies are then mapped using the FCA and Bayesian analysis to evaluate tourist's preferences against information publish by tourism information provider. An example of planning a tour to New York City is used

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

to illustrate the proposed ontology approach. An analytic hierarchy process is used to rank the tourism attractions suggested by the ontology and FCA-based approaches [1].In this paper, we propose an ontological recommendation multi-agent for Tainan City travel. The core technologies of the agent contain the ontology model, fuzzy inference mechanisms, and ant colony optimization. The proposed agent can be recommend the tourist a personalized travel route to enjoy the Tainan City according to the tourist's requirements and there preferences. It includes a context decision agent and a travel route recommendation agent. The context decision agent needs a suitable location distance, counts the context relation and infers the context information based on the tourists requirements and Tainan City travel ontology. The travel route recommendation agent is responsible for ending a personalized tour and plotting this travel route on the Google Map. Finally, the tourist can follow the personalized travel route to enjoy the cultural heritage and the local gourmet food during his stay at Tainan City. The experimental results show the proposed approach can effectively recommend a travel route matched with the tourist's requirements [8].

III. PROPOSED ALGORITHM

1. System Architecture

a) Ontology Development

The ontology for tourism information providers represents the perspective of various tourism information providers, and consists of descriptions of tourist attractions. Corresponding to the diverse and informal nature of on-line tourism information, a two-step approach is proposed for its development. The first step develops a large number of 'local ontologies,' one for each web site. In the second step, an integrated ontology is derived from all of these local ontologies to present a unified representation of on-line tourism information.

b) Mapping between the two ontologies

The ontology for tourists and the integrated ontology for tourism information providers are mapped in order to evaluate tourists' 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.

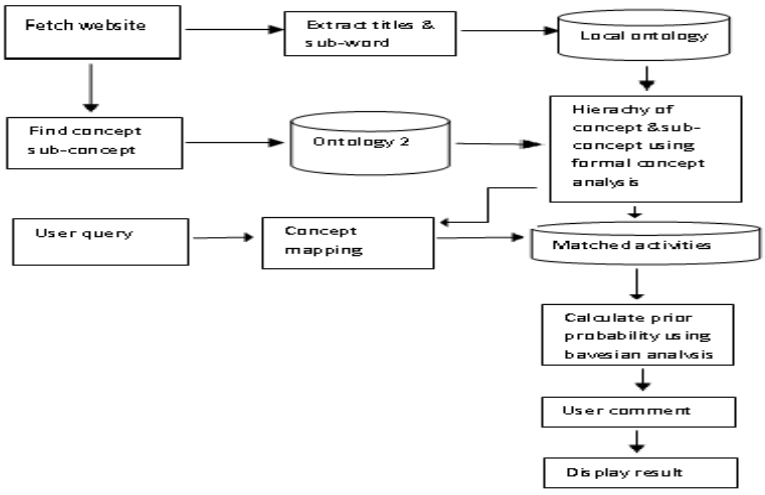

Fig 1: System Architecture

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

c) Property Comparison

For each set of mapped concepts, the second level of ontology mapping compares and matches their property values. The result of this process helps select appropriate attractions that satisfy the preferences of a tourist. Since out of the four categories of preferences (tour time, budget, transportation mode, and activities) preferred activities plays a primary role in the selection of attractions, property comparison between the pair of mapped concepts, 'preferred activities' and 'provided activities', is discussed first. Because the two concepts are both expressed in nominal values, a Bayesian analysis is used to compare their values. In a Bayesian analysis, a hypothesis of a concept, presented as the prior probability (P (H)), represents an initial belief. Additional evidence (P (ejH)) represents the likelihood of the hypothesis and is used to update the prior probability to the posterior probability (P (Hje)). The property comparison for other mapped concepts is through simple quantitative comparison [1].

d) Formal concept analysis

Formal Concept Analysis(FCA) is mainly use for the analyzing of data the main motivation of formal concept analysis was the concrete representation of complete lattices and their properties by means of formal context data tables that represent binary relations between objects and attributes. The aim and meaning of Formal Concept Analysis as mathematical theory of concepts and concept hierarchies is to support the rational communication of humans by mathematically developing appropriate conceptual structures which can be logically activated. Formal context in FCA is a triple $K = (G, M, I)$ where G is a set of *objects*, M is a set of *attributes* and the binary relation $I \subseteq G \times M$, called *incidence* shows, which objects possess which attributes [4] [6].

e) Shortest path algorithm

In Dijkstras algorithm, maintain two sets, one set contains vertices included in shortest path tree, other set includes Vertices not yet included in shortest path tree. At every step of the algorithm, we find a vertex which is in the other set (set of not yet included) and has minimum distance from source. Below are the detailed steps used in Dijkstras algorithm to find the shortest path from a single source vertex to all other vertex.

1. Create a set sptSet (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. Assign a distance value to all vertices in the input graph. Initialize all distance values as INFINITE. Assign distance value as 0 for the source vertex so that it is picked first.
3. While sptSetdoesnt include all vertices

(a) Pick a vertex u which is not there in sptSetand has minimum distance value.
(b) Include u to sptSet.
(c) Update distance value of all adjacent vertices of u. To update the distance values, iterate through all adjacent vertices. For every adjacent vertex v, if sum of distance value of u (from source) and weight of edge u-v, is less than the distance value of v, then update the distance value of v [1].

f) Bayesian analysis

Bayesian analysis is a statistical paradigm that answers research questions about unknown parameters using probability statements. Bayesian inference uses the posterior distribution to form various summaries for the model parameters, including point estimates such as posterior means, medians, percentiles, and interval estimates known as credible intervals. Moreover, all statistical tests about model parameters can be expressed as probability statements based on the estimated posterior distribution. Unique features of Bayesian analysis include an ability to incorporate prior information in the analysis, an intuitive interpretation of credible intervals as fixed ranges to which a parameter is known to belong with a pre-specified probability, and an ability to assign an actual probability to any hypothesis of interest.

Bayesian inference is a one of the method of statistical inference in which Bayes theorem is used to update the probability for a hypothesis as more evidence and information becomes available. Bayesian inference is an important technique in statistics, and especially mathematical statistics. Bayesian inference it is closely related to subjective probability, often known as Bayesian probability. Bayesian inference uses the posterior distribution to form various summaries for the model parameters, including point for estimate such as posterior means, medians, percentiles, and interval estimates known as credible intervals. all statistical tests are about model parameters can be expressed as probability statements which is based on the estimated posterior distribution. Unique features of Bayesian analysis include an ability to perform incorporate prior information in the data analysis, an interpretation of credible intervals as

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

fixed ranges to which a parameter is known to belong with a pre-specified probability, and an ability to assign an actual probability to any hypothesis of interest. Bayesian inference derives the posterior probability as a consequence of two antecedents, a prior probability for the observed data. Bayesian inference calculates the posterior probability according to Bayes theorem

$$P(H \mid E) = \frac{P(E \mid H) \cdot P(H)}{P(E)}$$

g) User Comments

In proposed method, we are using user comments for creating tour plan. User comments contain very important information related to respective places. User comments give us actually tourist positive and negative feedback. It is one way to check out exact detail information about respective place. Using tourist actual reviews we easily predict real thing about location. But many websites may manipulate user comments so it get difficult to use reviews, which may be misguide tourist. In our system we first break user sentence in to words and then classify user comments into two categories manipulated and non-manipulated using naïve based classification algorithm. This values store with respective location. If positive feedbacks are present then that place get recommended.

IV. MATHEMATICAL MODEL

1) 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…… }

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

3) 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.

4) OF={r, n , c}

OF is output from FCA which use for creating hierarchy concepts

5) r is reference concepts

6) n is new concepts

7) c is local concepts

8) I is input, I ={U,Q}

9) U is users of Systems = {U1,U2,U3….)

10) Q is queries of users = {q1,q2,q3…..}

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

11) T is preferred tour time

12) B is preferred budget

13) A is preferred activities

14) M is preferred mode

MC= {A,OF,OM}

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

15) A is preferred activities

16) OF is sub concepts hierarchy

17) OM is user attraction places which map by preferred activity.

18) 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

19) Where $P(e)$ is a normalization constant

20) P={p1,p2,p3……}

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

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

O={pa, pb , pc…..}
22) Pa= {d1,d2,d3….}
d=day
23) Pb= {a1,a2,a3…..}
a= city
24) D={G,SP,So,En}
G={V,E,ds}
25) V={v1, v2, v3,…..}
Where v1,v2,v3 are vertices
26) E={e1,e2,e3….}
Where e1,e2.. are edges
27) Ds distance between vertices
28) So is source point of tourism
29) En is end point of destination
30) SP is shortest path

V. RESULTS

This system focuses on tourism attractions it may be extended to different 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 in our system we are improving accuracy level which is shown in fig.

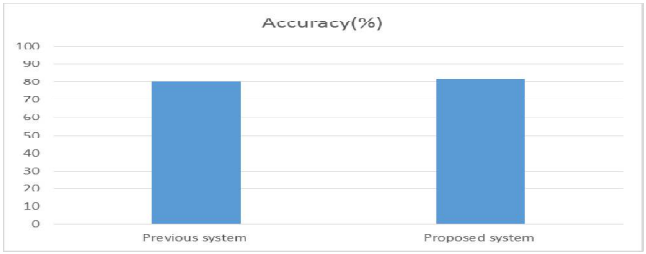

Fig 2: accuracy of both system

Here we are getting tour route from system. With minimum path. Using users preferred activities and users budget. Which is shown in following snap.

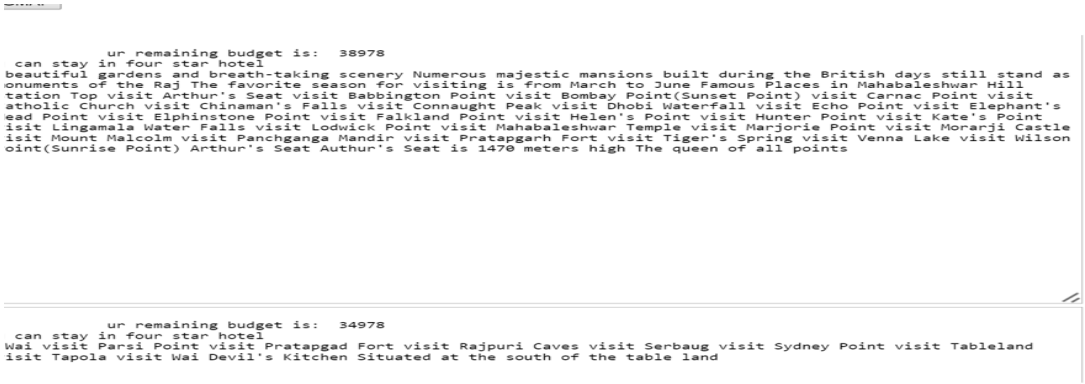

Fig 3: result of system

The above fig shows the can we visit in our tour by using users selected preferences as well as shows in which hotel we can stay it may be two, three or four star hotel.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

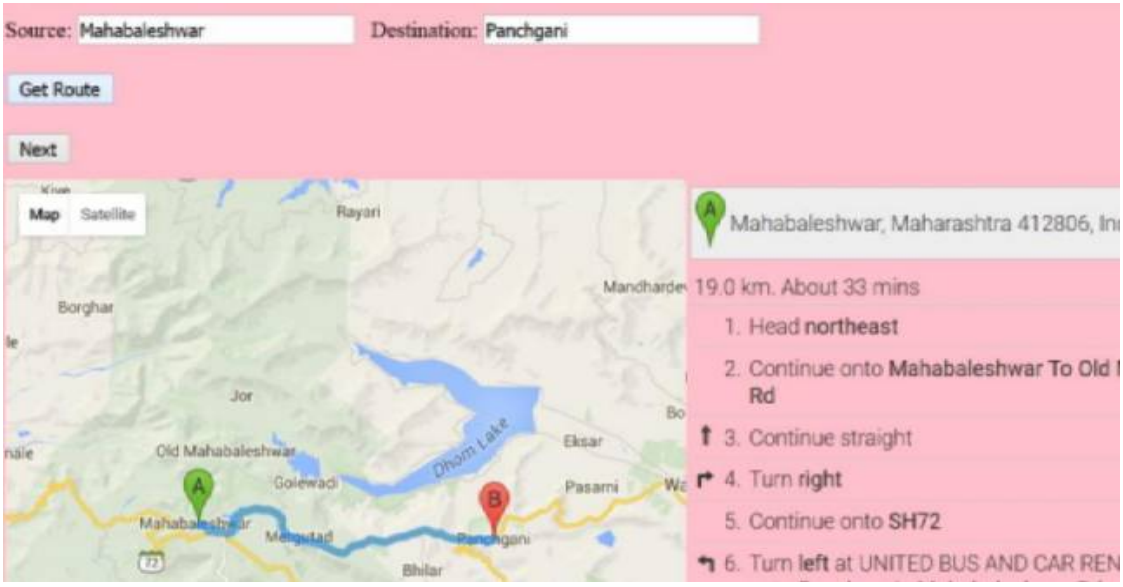

Fig 4: Google map showing result

From above diagram we are getting a shortest path from source to destination and it will shows us with Google map API.

VI. CONCLUSION AND FUTURE WORK

Ontology for tourists and in particular an integrated ontology for tourism information 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 direction for on-line tourism information systems, including tour planning. The integrated ontology for tourism information providers can help for development of semantic web-based tourism information systems The approaches discussed in this paper is a way to more advanced research. For e g, concepts of personal profile can be added to the ontology for tourists. Personal profile describes the tourists' characteristics related to their choices of attractions. E.g. age, occupation, income, personality, and interests. In a case where a traveller does not identify preferred activities, the personal profile will help to get a tourist's preferred activities. It is recognized as an emerging source of information for on-line tour planning. The utilization of unstructured information in the development of ontologies presents both challenges and opportunities in future research

ACKNOWLEDGEMENT

We would like to thank our project guide Prof. ArchanaKadam for her enormous co-operation and guidance. Wehave no words to express our gratitude for a person whowhole heartedly supported the project and gave freely ofher valuable time while making this project. The technicalguidance provided by her was more than useful and made theproject successful. We would also like to thank our Department of Computer Engineering, PimpriChinchwadCollege of Engineering.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

REFERENCES

1. Yuxia Huang, Ling Bian "Using Ontologies and Formal Concept Analysis to Integrate Heterogeneous Tourism Information" School of Engineering and Computing Sciences, Texas AM UniversityCorpus Christi, Corpus Christi, TX 78412 USA,2015.
2. Y. Huang and L. Bian, A Bayesian network and analytic hierarchy process based personalized recommendations for tourist attractions over the Internet, Expert Syst. Appl., vol. 36, no. 1, pp. 933943, 2009
3. 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. 5166, 2009.
4. J. Poelmans, D. I. Ignatov, S. O. Kuznetsov, and G. Dedene, Formal concept analysis in knowledge processing: A survey on applications, Expert Syst. Appl., vol. 40, no. 16, pp. 65386560, 2013.
5. L. Bian and S. Hu, Identifying components for interoperable process models using concept lattice and semantic reference system, Int. J. Geograph. Inf. Syst., vol. 21, no. 9, pp. 10091032, 2007.
6. U. Priss, Formal concept analysis in information science, Annu. Rev. Inf. Sci. Technol., vol. 40, no. 1, pp. 521543, 2006.
7. Z. Xiang and U. Gretzel, Role of social media in online travel information search, Tourism Manage., vol. 31, no. 2, pp. 179188, 2010.
8. C.-S. Lee, Y.-C. Chang and M.-H. Wang, Ontological Recommendation Multi- Agent for Tainan City Travel, Expert Systems with Applications, vol. 36, no. 3, pp. 6740-6753, Apr. 2009.
9. Eric Hsuch-ChanLuandCluh-Yuan Lin 'Trip-Mine:An Efficient trip planning approach with travel time constraint'2010.
10. C.-C. Yu and H.-P. Chang, Personalized Location-Based Recommendation Services for Tour Planning in Mobile Tourism Applications, E-Commerce and Web Technologies, vol. 5692, pp. 3849, and 2009.
11. D. L. Applegate, R. E. Bixby, V. Chvatal and W. J. Cook, The Traveling Salesman Problem: A Computational Study, Princeton University Press, 2006.
12. S. Kisilevich, D. Keim and L. Rokach, A Novel Approach to Mining Travel Sequences Using Collections of Geotagged Photos, in Proceedings of the 13th AGILE International Conference on Geographic Information Science, pp. 163-182, May2010.