2018 IEEE Fourth International Conference on Big Data Computing Service and Applications

HTARF: A Hybrid Tourist Attraction Recommendation Framework for Trip Scheduling

Yuanfeng Gong, Yushun Fan*
Tsinghua National Laboratory for Information Science and Technology
Department of Automation
Tsinghua University
Beijing 100084, China
gongyf15@mails.tsinghua.edu.cn
fanyus@tsinghua.edu.cn

Bing Bai, Junqi Zhang
Tsinghua National Laboratory for Information Science and Technology
Department of Automation
Tsinghua University
Beijing 100084, China
bb13@mails.tsinghua.edu.cn
jq-zhang15@mails.tsinghua.edu.cn

Zhenfeng Gao, Xing Wu
Tsinghua National Laboratory for Information Science and Technology
Department of Automation
Tsinghua University
Beijing 100084, China
gzf13@mails.tsinghua.edu.cn
wuxing17@mails.tsinghua.edu.cn

***Abstract*—As the users' demand for efficient and personalized services grows continuously, online tourism services are becoming more and more popular. However, many tourism products can only recommend existing tourist packages according to users' profile, which can hardly meet their personalized requirements. In this paper, we propose a novel attraction recommendation method, which takes advantages of quantities of information mined from attraction descriptions and tourist packages. The approach is modified on the basis of content matching and association rule mining and only requests a query from a user to recommend. In addition, the method is applied to our Smart Tourism Services Platform (STSP). Using attractions list offered by our method, a trip plan module of STSP will provide a personalized travel schedule. We examine our approach on a real-world dataset, and an online evaluation result shows that our approach is more effective and performs much better than baselines.**

***Keywords*-tourist attraction recommendation; content matching; association rule mining**

I. Introduction

The amount of information available on the World Wide Web has experienced an enormous increase in the last decade [1]. The Internet provides users with a convenient way to obtain data about tourism. However, how to extract the valid part from the extensive data is very complicated and time-consuming. Popular attractions are more likely to be found online, which will make people miss some little-known attractions that suit them well. Recommendations are a common means of planning in the fields of tourism, traveling and hospitality [2]. A good recommendation can make it easier to design schedule and give users a worthwhile trip, therefore personalized tourism recommendation is essential for a good travel experience.

Most existing recommendation approaches are content matching methods, mainly focusing on keyword search and semantic-based search [3], [4]. However, keyword search is lack of efficiency and semantic-based method is expensive to implement. Context-aware recommendation methods are proposed in [5], [6] to provide more intelligent and useful recommendations with the use of contextual factors like weather conditions. In addition, collaborative filtering [7], [8] has also been widely used in tourism suggestion areas. The hybrid tourism recommendation method also has a considerable proportion. For instance, both [9], [10] propose a hybrid approach based on collaboration and context analysis. Besides, technology based on ontology [11], [12], [13] has been studied in many works.

The majority of existing attraction recommendation methods need user's profile, preferences or ratings on items. In [14], [15] information like user behavior is also used to acquire a better performance in recommendation. Although user profiles are used to represent a common practice in personalization systems, they intrinsically bring several issues [16], [17]. For example, the initial profile, ratings on items and history are often not available, especially for a new user and anonymous user, which will make recommendations unreliable, even unusable. Meanwhile, it is a standard way for travel agents to recommend an existing tourist package according to user's profile and interests, which is a viable but not perfect way, and generating personalized tourist packages for tourists is much better.

In this paper, we address such limitations by proposing a new attraction suggestion approach, based on content matching and association rule mining. The method is applied to STSP[1], a online application of tourism service which is developed on WeChat and is currently under use by over 2000 users in China.

The fundamental idea of our method is that there exists quantities of attractions and tourist packages data on the Internet, from which we can find out the ideal attractions for a tourist without any extra knowledge but a simple query. In our recommendation framework, the travel destination is fixed to Beijing temporarily. As can be seen in Figure 1, the user need to give a sentence as a query, and the system will

* Corresponding Author

[1] https://mp.weixin.qq.com/s/28eaqQVEaUSCZp3aSEzHpw

0-7695-6396-1/18/$31.00 ©2018 IEEE
DOI 10.1109/BigDataService.2018.00022

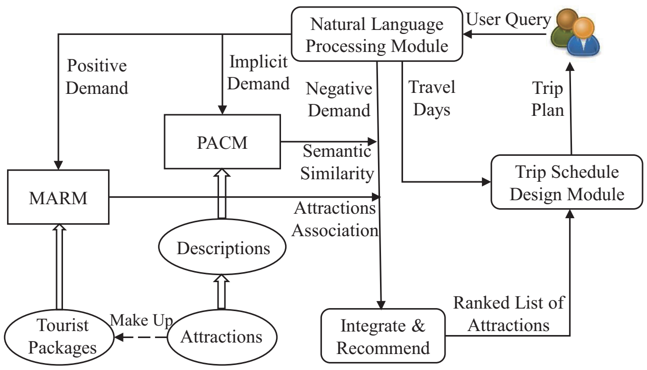

Fig. 1. Tourism recommendation framework of STSP

extract the number of travel days, places the user wants to go (called *Positive Demand*), places the person doesn't want to go (called *Negative Demand*) from the query by natural language processing. And the remaining part of the sentence (called *Implicit Demand*) is a text description of the attractions he or she want to visit. For example, from query $Q1$ "*I plan to play for 3 days, and I want to go to Peking University, and don't wanna go to the Forbidden City. Besides, I'd like to visit places full of cultural atmosphere.*", we can tell the number of travel days is 3, the positive demand is *Peking University*, the negative demand is *the Forbidden City* and the implicit demand is " *I'd like to visit places full of cultural atmosphere.*".

The implicit demand is a refined description of target attractions, by comparing attraction descriptions with it, it is possible to find out suitable attractions. Although the most popular attractions are not necessarily the most appropriate, the popularity of attractions is still one of the most influential travel factors. As a result, the first part of our approach, *Popularity-Aware Content Matching* (PACM), considers the similarity between implicit demand and attraction description as the main influence, and takes attraction popularity as an impact factor.

The positive demand is a set of places that a user wants to visit, using which implicit information in existing tourist packages can be found out to make better recommendations. If many packages include two places, we naturally think the two attractions are strongly related. The user who wants to visit one of the associated two attractions would like to go to the other with significant probability. In the second part of our method, we apply a *Modified Association Rule Mining* (MARM) to work out the most related attractions to the positive demand of travelers.

The negative demand is a set of attractions a user doesn't like, we simply remove it from the final suggested attractions list.

By incorporating the ideas mentioned above, we propose a *Hybrid Tourist Attraction Recommendation Framework* (HTARF). Finally, with the attractions list recommended by HTARF and the number of travel days, trip schedule design module (TSDM) will make a trip plan for the user considering the spatial and temporal characteristics of attractions. Noted that in this paper we focus on recommending personalized attractions, the trip plan part will be only briefly introduced. The main contributions of this paper are summarized as follows.

Firstly, we propose PACM which is capable of selecting the most similar attractions to the implicit demand. By considering attraction popularity, we can recommend attractions more suitable for users.

Secondly, combining PACM with MARM, we propose a hybrid tourist attraction recommendation framework which just need a query of a user to recommend.

Thirdly, we examine our approach on the real-world Beijing attractions and tourist packages dataset. An online evaluation result shows that our method has better precision and performs more efficient than baselines.

The reminder of this paper is structured as follows. Section II introduces the problem definition of our recommendation system. Section III shows the construction of two components PACM, MARM and describes the framework of our approach. Section IV reports the experimental results. Section V lists the related work and Section VI draws the conclusion.

II. Problem definition

In this section, we give some definitions and notations involved in the recommendation process.

Definition 1: Attraction. For attractions, we use A_j to denote attraction j while AD_j to denote the description of attraction j. P_j represents a score of popularity of A_j, and $AD = \{AD_1, AD_2, \ldots, AD_J\}$, $j = 1 : J$, $J = |AD|$. $AD_j = \{w_{j1}, w_{j2}, \ldots, w_{jn_j}\}$ means the set of n_j words used in the description of attraction j. Similarly, implicit demand denoted by ID is associated with a collection of words $ID = \{w_{m1}, w_{m2}, \ldots, w_{mn_m}\}$.

Definition 2: Tourist Package. We use TP_i to denote tourist package i. $TP_i = \{A_{i1}, A_{i2}, \ldots, A_{in_i}\}$ is the collection of n_i attractions appeared in the package, and $TP = \{TP_1, TP_2, \ldots, TP_I\}$, $i = 1 : I$, $I = |TP|$. Besides, $PD = \{A_{p1}, A_{p2}, \ldots, A_{pn_p}\}$ and $ND = \{A_{e1}, A_{e2}, \ldots, A_{en_e}\}$ separately denote the positive demand and negative demand. For the query of a tourist, we can denote it as $Q = \{N_d, ID, PD, ND\}$ where N_d is the number of travel days.

With definitions and notations given above, the tourism recommendation framework of STSP is defined as follows.

Definition 3: Tourism Recommendation Framework. Given AD_j, TP_i, $j = 1 : J$ and $i = 1 : I$. For the query Q, a ranked list of attractions represented by R will be suggested. Top l of list R will be used to generate a tourist package.

III. Model framework

In this section, we introduce the construction of PACM and MARM in our method. Based on the two components, we illustrate how to combine them to support the HTARF.

A. Popularity-Aware Content Matching

The key idea is people like to visit places that match their interests. Moreover, people prefer to visit famous attractions when optional attractions are all in line with their preferences. Based on this consideration, we measure the similarities between the implicit demand and all the attraction descriptions. Then, we introduce a factor which can make popular attractions that have similar match degree rank higher.

Since the implicit demand and attraction descriptions are all textual data, it is natural to use *Vector Space Model* (VSM). VSM is a spatial representation of text documents. In that model, each record is represented by a vector in a n-dimensional space, where each dimension corresponds to a term from the overall vocabulary of a given document collection[18]. We apply tokenization, a standard natural language processing, to the implicit demand and attraction descriptions, and get a dictionary $T = \{t_1, t_2, \ldots, t_n\}$ that is the set of words in the corpus, and have

$$T = (\bigcup_{j=1}^{J} AD_j) \cup ID \tag{1}$$

Then, we can represent implicit demand and attraction descriptions as a n-dimensional vector with T. let $AE_j = \{v_{j1}, v_{j2}, \ldots, v_{jn}\}$ denote the n-dimensional vector of AD_j, and similarly $IE = \{v_{e1}, v_{e2}, \ldots, v_{en}\}$ denote vector of ID.

What has to be aware of is that in tourism domain the user query is far shorter than the attraction descriptions. Therefore, v_{jk} is not the weight of term t_k in AD_j but $\mathbb{1}_{AD_j}(t_k)$ in our approach, where $\mathbb{1}_A(x)$ is an indicator function of set A, and can be formulated as follows.

$$\mathbb{1}_A(x) = \begin{cases} 1, & if\ x \in A \\ 0, & if\ x \notin A \end{cases} \tag{2}$$

That's to say if AD_j contains term t_k, the kth element of AE_j is 1 otherwise 0, and IE follows the same representation. Finally, the similarity between the implicit demand and attraction j can be calculated using the following equation.

$$sim'(j) = \langle IE, AE_j \rangle \cdot [1 + \alpha \log(P_j + 1)] \tag{3}$$

where $\langle IE, AE_j \rangle$ is the inner product of IE and AE_j and $1+\alpha \log(P_j+1)$ is the popularity factor which is greater than or equal to 1 since $P_j \in [0, 4]$ and $\alpha \geq 0$. Parameter α is the power of popularity factor, and the bigger α is, the higher ranking of popular attractions are. The normalized similarity is

$$sim(j) = \frac{sim'(j)}{\max\limits_{j=1:J} sim'(j)} \tag{4}$$

B. Modified Association Rule Mining

An association rule is a statement like "90% of transactions that purchase bread and butter also buy milk". The purpose of association rule mining algorithm is to generate all significant association rules between items in the database [19]. Apparently, association rule mining can be used to make recommendations. For example, [20] puts forward a collaborative recommendation method using association rule mining algorithm and achieves better performance than baselines.

Conditional probability can explain association rule mining. To determine whether A and B are related, we need to calculate confidence which is conditional probability $P(A|B)$ in consecutive cases. According to Bayes formula, we have

$$conf_c(A, B) = P(A|B) = \frac{P(A, B)}{P(B)} \tag{5}$$

where $P(B)$ is the probability that B occurs while $P(A, B)$ is the probability that A and B co-occur.

However, it is not appropriate to apply association rule mining directly in tourism scenarios. Because popular attractions get more attention so that the result of association rule mining will tend to them. It will reduce the precision of the result, and even worse that we can't find the real associated attractions. For instance, no matter what types of attraction people like, they almost always go to *the Forbidden City* unless they've been there before. Suppose A, C are almost the same attractions except for popularity, and popularity of A is greater than C. Due to the influence of popularity, $P(A, B) > P(C, B)$ holds. According to equation (5), we can get $P(A|B) > P(C|B)$. In other words, even A, C are the same except popularity, the confidence of A and B is larger than the confidence of C and B. To solve this problem, we made a slight change to the formula (5), the modified equation is formulated as follows.

$$conf_r(A, B) = \frac{P(A, B)}{P(A)P(B)} \tag{6}$$

Since the more popular A is, the greater $P(A)$ is. Popular attractions are punished when calculating the confidence. For the discrete situation, confidence can be derived by

$$conf_d(A, B) = \frac{Count_S(A, B)}{Count_S(A)Count_S(B)} \tag{7}$$

where $Count_S(A)$, $Count_S(B)$ is the number of occurrences of A, B while $Count_S(A, B)$ is the number of occurrences of both A and B in set S. If confidence is higher than minimum confidence, there is relevance between A and B.

For A_{pk} in PD, $k = 1 : n_p$, the confidence of A_{pk} and A_j, $j = 1 : J$, can be get by following equation.

$$\begin{aligned} conf(k, j) &= \frac{Count_{TP}(A_{pk}, A_j)}{Count_{TP}(A_{pk})Count_{TP}(A_j)} \\ &= \frac{\sum_{i=1}^{I} \mathbb{1}_{TP_i}(A_{pk}) \cdot \mathbb{1}_{TP_i}(A_j)}{\sum_{i=1}^{I} \mathbb{1}_{TP_i}(A_{pk}) \cdot \sum_{i=1}^{I} \mathbb{1}_{TP_i}(A_j)} \end{aligned} \tag{8}$$

The association degree of the positive demand and attraction j can be obtained by

$$asso(j) = \frac{1}{n_p} \sum_{k=1}^{n_p} \frac{conf(k, j)}{\max\limits_{j=1:J} conf(k, j)} \tag{9}$$

Obviously, $asso(j)$ is the mean of $\{conf(k,j)|k = 1 : n_p\}$ after normalization.

C. A Hybrid Tourist Attraction Recommendation Framework

Based on the two components PACM and MARM introduced previously, we illustrate how to combine them to support the HTARF.

When a new user query comes up and is represented as $Q = \{N_d, ID, PD, ND\}$. Firstly, we use PACM to calculate the similarities between the implicit demand and all the attractions through equation (4). Secondly, we use MARM to get the association degrees between the positive demand and attractions through equation (9). Thirdly, combining the two parts, we can get the score of attraction j as follows.

$$score(j) = \beta sim(j) + (1 - \beta) asso(j) \tag{10}$$

where $\beta \in [0, 1]$, finally we recommend a ranked list of attractions R in a descending order with regard to $\{score(j)|j = 1 : J\}$ after removing ND from the set.

Although the complete HTARF has been introduced, there is still one more thing in our tourism recommendation system need to be described. The TSDM is a significant part of our STSP, but we only give it a brief introduction since it's not the focus of this paper. TSDM uses attractions list R, the negative demand ND and the number of travel days N_d as input. We assume a tourist can play 8 hours each day, and the duration of each attraction and the travel time between every two attractions are known. Firstly, we take top l of R as a new list R'. Secondly, we use K-means to divide R' into m categories w.r.t. location. Thirdly, for each day, we select the first attraction with a high score, then add attractions in the same category one by one and make sure the total duration doesn't exceed time limits. Finally, using a greedy algorithm to design the route for each day since it is a small-scale problem.

IV. Experiments

In this section, we introduce the data set for the evaluation, then present a comprehensive performance evaluation of the proposed approach and three baselines used in the experiments.

A. Data Set

We crawled a real dataset of Beijing from *Baidu Travel Website*[2] with 1937 attractions and 3164 tourist packages. Each attraction consists of name, popularity, duration, textual description. In addition, latitude and longitude of each attraction are obtained by *Baidu Map APIs*[3]. The attraction descriptions are processed by words segmentation, stop words removing and meaningless words lifting, and we get a vocabulary size of 15556. Detail information about the data set is shown in Table I.

[2]https://lvyou.baidu.com/

[3]http://lbsyun.baidu.com/

TABLE I
PROPERTIES OF DATA SET

Number of attractions	1937
Number of tourist packages	3164
Average number of attractions in tourist packages	4.78
Vocabulary size	15556
Average number of word tokens in attraction descriptions	74.63

B. Experiment Methodology

To prove the effectiveness of our method, we have designed a questionnaire[4] to acquire evaluation from users. 107 subjects participated in the experiment through the Internet.

We list 10 questions in the questionnaire, for each question there is a query put forward by a user aiming at a particular kind of attractions like $Q2$ "*3 days, I plan to go to the Bird's Nest Stadium, and I'd like to visit ancient architectures and some places of interest*". And using attractions list suggested by four methods, we generate four trip plans for each query in random order. The users are asked to imagine if they were the user in the query, which plan do they think is the best and the worst.

C. Baselines and Parameter Settings

The three baselines used in our evaluation are PACM, MARM used alone and *Popularity-Based Attraction Recommendation* (PBAR) which uses only the most popular attractions.

For all runs of the four approaches, we use $\alpha = 1$, $\beta = 0.55$ and $l = 10N_d$. Then we generate the trip plans associated with four methods by TSDM with the same settings.

D. Experiment Result

The rate of each approach chosen as the best and the worst are shown in Figure 2 and Figure 3. PBAR purely considering the popularity is suitable for users without clear travel targets. Nevertheless, it doesn't work in the situation where the queries have a clear goal. And PBAR is chosen as the worst in 8 out of 10. The implicit demand of $Q1$ is "*I'd like to visit places full of cultural atmosphere.*", and the definition of culture can be academic, artistic, ancient, modern, oriental, western and so on for different people. Also, popular attractions in Beijing tend to be rich in the culture at some level, such as *Prince Gong's Mansion* and *the Imperial College* appeared in the options related to PBAR, which is the reason why PBAR can be selected as the best in Query 1.

PACM makes content matching between ID and attraction descriptions and can find the most related attractions for ID. That PACM is selected as the worst in $Q1$ is also affected by the multiple definitions of culture. MARM can find attractions closest to the positive demand in the aspect of type or geography cause that people have specific preferences and don't like travel in general. Although PACM and MARM are very useful in specific scenarios and perform quite well in

[4]https://www.wjx.cn/jq/18356257.aspx

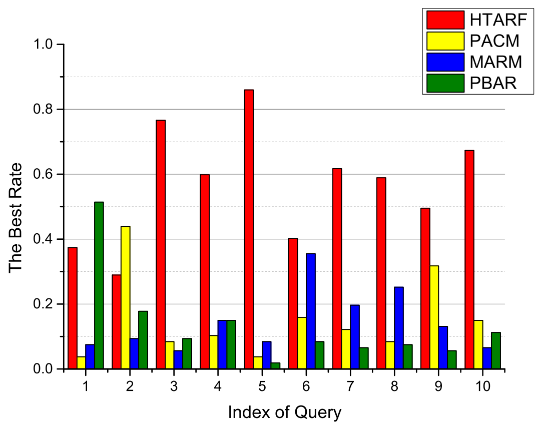

Fig. 2. The best rate of four methods in 10 testing queries

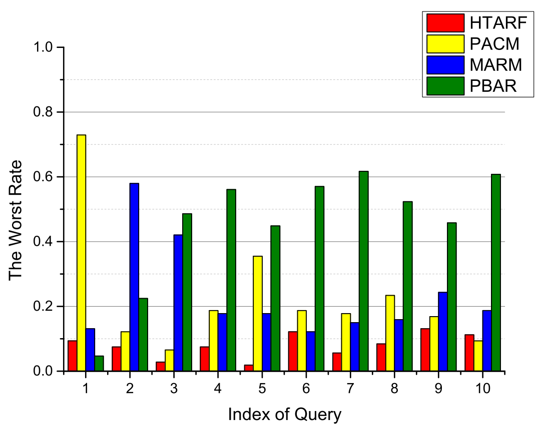

Fig. 3. The worst rate of four methods in 10 testing queries

some queries, they are so unilateral that just consider only one aspect of each query. Due to the combination of advantages of both PACM and MARM, HTARF can make a comprehensive assessment of the user needs, and give the most suitable list of attractions. As we can see in Figure 2, HTARF make the best in 80 percent cases and the second best in the remaining and the worst rate of HTARF is below 0.2 in overall performance.

TABLE II
AVERAGE BEST AND WORST RATE FOR DIFFERENT METHODS

Method	Average best rate	Average worst rate
HTARF	0.566	0.079
PACM	0.153	0.232
MARM	0.146	0.235
PBAR	0.135	0.454

Table II summarizes the average best and the average worst rate of four approaches and shows that HTARF has the best performance among the four methods with the most prominent average best rate and the smallest average worst rate. And the two components of HTARF, PACM and MARM, have similar and moderate performance. PBAR is the most offensive option and receives the worst rating.

In Table III, we list the top five attractions recommended by four methods for $Q4$ except the positive demand, *Fragrant Hills*. And $Q4$ is "*3 days journey, I want to go to Fragrant Hills, and have a look at all kinds of beautiful natural scenery*". PBAR has the worst performance, cause none of the attractions in $S4$ is the similar type to *Fragrant Hills* and only *the Summer Palace* in $S4$ has something to do with natural scenery. Better than PBAR, *Phoenix Mountain* and *Longtan Park* in $S2$ are natural attractions, while *Temple of Heaven* and *Peking University* have some natural scenery. This means PACM can find attractions similar to the implicit demand. All attractions in $S3$ are located around *Fragrant Hills* within 3.5 km, which is the feature of MARM. But the recommendation of MARM is not ideal since none of the attractions in $S3$ is strongly related to the natural scenery. Attractions recommended by HTARF is more satisfying, cause that top three attractions in $S1$ are famous natural attractions and the other two are related to the natural scenery. What's more, *Badachu Park* is a natural attraction and is only 3 km away from *Fragrant Hills* in the meanwhile. That's to say HTARF combines the advantages of both PACM and MARM, and find *Badachu Park* that doesn't exist in both $S2$ and $S3$.

TABLE III
TOP 5 ATTRACTIONS RECOMMENDED BY DIFFERENT METHODS FOR $Q4$

Method	Set of Attractions
HTARF	$S1$={*Longtan Park*, *Badachu Park*, *Phoenix Mountain*, *Temple of Heaven*, *Peking University*}
PACM	$S2$={*Temple of Heaven*, *Peking University*, *Phoenix Mountain*, *XiShenKu Church*, *Longtan Park*}
MARM	$S3$={*Biyun Temple*, *Temple of Reclining Buddha*, *Temple of Great Mercy*, *Jianxin Room*, *Luminous Temple*}
PBAR	$S4$={*the Forbidden City*, *the Summer Palace*, *Tian An Men*, *the Bird's Nest Stadium*, *Great Hall of the People*}

As a conclusion, taking PACM and MARM into consideration can lead to a better performance. HTARF is more efficient and can give tourists better recommendations.

V. RELATED WORK

Attraction recommendation is a hot research spot in tourism domain. Most tourist recommender systems tend to suggest places once the user has decided the destination of the trip [1]. Some works use contextual information in recommendations, work [5] focuses on making attraction suggestion by taking contextual factors like weather conditions into account. Collaborative filtering is one of the state-of-art methods in the recommendation community [4], and is used widely in tourism suggestion. For example, work [7] uses the location data and users' comments at a various location to find interesting locations and possible activities for them. Nevertheless, collaborative-based methods require sufficient users' history, and suffer from cold start problem that a new user doesn't have any history and cannot find other users similar to him.

Recommendation approaches through constructing ontology have also been proposed quite a lot in the tourism field. Work [16] introduces a recommender and information retrieval system based on ontology, to offer personalized suggestions to citizens and tourists, including those with special needs. Besides, using users' location, profile and preferences, work [12] aims to create a platform able to provide personalized services with the use of some standard ontologies. However, what can't be ignored is that standard ontologies may not meet the requirements of particular system while the construction of ontology is time-costing and complicated in practice.

Another group of researchers tries to put forward some hybrid tourism recommendation approach. The paper [9] introduces a hybrid tourism recommender system based on both collaborative method and text analysis, to help travel agents in discovering options for customers. Most of existing tourism recommendation approaches request a lot of user information, such as profile, history, ratings on items, which is not available sometimes and is unfriendly to a new or anonymous user.

In this paper, we propose a method based on content matching and association rule mining to find out the most suitable attractions in massive attractions and tourist packages data for a user with only his query.

VI. CONCLUSION

Customized travel services are in urgent need nowadays, because users' pursuit of personalized services is increasing, and there are many difficulties in finding practical information from the vast amounts of data online.

In this paper, we present a novel attractions recommendation approach used in our STSP to bring people a better travel experience. Users' demand can be extracted from a query and divided into four parts: the positive demand, the negative demand, the implicit demand and the number of travel days. To find the most suitable attractions, we dig out a wealth of information from attraction descriptions and tourist packages by applying PACM and MARM. PACM find attractions that match users' implicit demand while MARM can obtain attractions associated with users' positive demand. Then, we take advantages of the two components and put forward HTARF. Experiments on a real-world dataset show that our approach performs much better than baselines.

Our future work includes incorporating other information, like comments and tags of attractions, to improve the performance. In addition, we plan to implement our method on STSP and test it with a more substantial number of users online.

ACKNOWLEDGMENT

This research has been partially supported by the National Nature Science Foundation of China (No.61673230) and High-Tech Ship Research Project of China (17GC26102.01).

REFERENCES

[1] J. Borràs, A. Moreno, and A. Valls, "Intelligent tourism recommender systems: A survey," *Expert Systems with Applications*, vol. 41, no. 16, pp. 7370–7389, 2014.

[2] T. Berka and M. Plößnig, "Designing recommender systems for tourism." ENTER 2004: 11th International Conference on Information Technology in Travel & Tourism, 2004.

[3] X. Dong, A. Halevy, J. Madhavan, E. Nemes, and J. Zhang, "Similarity search for web services," in *Proceedings of the Thirtieth international conference on Very large data bases-Volume 30*. VLDB Endowment, 2004, pp. 372–383.

[4] Y. Zhong, Y. Fan, K. Huang, W. Tan, and J. Zhang, "Time-aware service recommendation for mashup creation," *IEEE Transactions on Services Computing*, vol. 8, no. 3, pp. 356–368, 2015.

[5] M. Braunhofer, M. Elahi, F. Ricci, and T. Schievenin, "Context-aware points of interest suggestion with dynamic weather data management," in *Information and communication technologies in tourism 2014*. Springer, 2013, pp. 87–100.

[6] G. Adomavicius and A. Tuzhilin, "Context-aware recommender systems," in *Recommender systems handbook*. Springer, 2015, pp. 191–226.

[7] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Collaborative location and activity recommendations with gps history data," in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 1029–1038.

[8] M. Nilashi, O. bin Ibrahim, N. Ithnin, and N. H. Sarmin, "A multi-criteria collaborative filtering recommender system for the tourism domain using expectation maximization (em) and pca–anfis," *Electronic Commerce Research and Applications*, vol. 14, no. 6, pp. 542–562, 2015.

[9] S. Loh, F. Lorenzi, R. Saldaña, and D. Licthnow, "A tourism recommender system based on collaboration and text analysis," *Information Technology & Tourism*, vol. 6, no. 3, pp. 157–165, 2003.

[10] G. Fenza, E. Fischetti, D. Furno, and V. Loia, "A hybrid context aware system for tourist guidance based on collaborative filtering," in *Fuzzy Systems (FUZZ), 2011 IEEE International Conference on*. IEEE, 2011, pp. 131–138.

[11] L. Castillo, E. Armengol, E. Onaindía, L. Sebastiá, J. González-Boticario, A. Rodríguez, S. Fernández, J. D. Arias, and D. Borrajo, "Samap: An user-oriented adaptive system for planning tourist visits," *Expert Systems with Applications*, vol. 34, no. 2, pp. 1318–1332, 2008.

[12] L. Ceccaroni, V. Codina, M. Palau, and M. Pous, "Patac: Urban, ubiquitous, personalized services for citizens and tourists," in *Digital Society, 2009. ICDS'09. Third International Conference on*. IEEE, 2009, pp. 7–12.

[13] V. Codina and L. Ceccaroni, "Taking advantage of semantics in recommendation systems," in *Artificial Intelligence Research and Development: Proceedings of the 13th International Conference of the Catalan Association for Artificial Intelligence*, vol. 220. IOS Press, 2010, p. 163.

[14] Y. Hu and M. Ogihara, "Nextone player: A music recommendation system based on user behavior." in *ISMIR*, vol. 11, 2011, pp. 103–108.

[15] M. Albanese, A. Chianese, A. dAcierno, V. Moscato, and A. Picariello, "A multimedia recommender integrating object features and user behavior," *Multimedia tools and applications*, vol. 50, no. 3, pp. 563–585, 2010.

[16] K. Alonso, M. Zorrilla, H. Iñan, M. Palau, R. Confalonieri, J. Vázquez-Salceda, J. Calle, and E. Castro, "Ontology-based tourism for all recommender and information retrieval system for interactive community displays," in *Information Science and Digital Content Technology (ICIDT), 2012 8th International Conference on*, vol. 3. IEEE, 2012, pp. 650–655.

[17] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE transactions on knowledge and data engineering*, vol. 17, no. 6, pp. 734–749, 2005.

[18] P. Lops, M. De Gemmis, and G. Semeraro, "Content-based recommender systems: State of the art and trends," in *Recommender systems handbook*. Springer, 2011, pp. 73–105.

[19] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," in *Acm sigmod record*, vol. 22, no. 2. ACM, 1993, pp. 207–216.

[20] W. Lin, S. A. Alvarez, and C. Ruiz, "Efficient adaptive-support association rule mining for recommender systems," *Data mining and knowledge discovery*, vol. 6, no. 1, pp. 83–105, 2002.