

# TRSO: A Tourism Recommender System Based on Ontology

Yan Chu<sup>1</sup>, Hongbin Wang<sup>1(✉)</sup>, Liying Zheng<sup>1(✉)</sup>,  
Zhengkui Wang<sup>2</sup>, and Kian-Lee Tan<sup>3</sup>

<sup>1</sup> Harbin Engineering University,  
Nantong Street No.145, Nangang District, Harbin, China  
{chuyan,wanghongbin,zhengliying}@hrbeu.edu.cn

<sup>2</sup> InfoComm Technology, Singapore Institute of Technology, Singapore, Singapore  
zhengkui.wang@singaporetech.edu.sg

<sup>3</sup> Department of Computer Science, National University of Singapore,  
Singapore, Singapore  
tankl@comp.nus.edu.sg

**Abstract.** In the era of information explosion, the Internet has become one of the most important tools for users to get information. As one of the main applications, most of the tourists, if not all, utilize the search engine to obtain the useful travelling information online which makes tourism recommender systems valuable. However, given a huge amount of online information, it still remains challenging to develop an effective tourism recommender system. To tackle this challenge, in this work, we propose TRSO, an ontology-based tourism recommender system by incorporating different techniques. First, we adopt the association rules to dig out the associated users from a large number of users. By doing so, users in the database are divided into two categories: related users and unrelated users. Second, for the related users, we propose a collaborative filtering algorithm by incorporating the time and evaluation factors. For the unrelated users, we utilize a different collaborative filtering algorithm, which integrates the time factor and the tourism attraction ontology information. Third, we further filter useless information according to the context information. Finally, we expand the tourism attraction with other tourism information such as shopping, eating and traveling based on a tourism ontology. The experimental results on the standard benchmark show that the proposed tourism recommendation algorithm can achieve satisfactory and comprehensive recommendation performance.

**Keywords:** Tourism recommendation · Association rules · Collaborative filtering · Context information · Ontology

## 1 Introduction

Nowadays, with the substantial improvement of people's living standard, tourism has become an increasing popular leisure activity for people. In addition, the

Internet has become an important resource for those who are planning their trips to get tourism information. However, the huge amount information on the Internet always get people overwhelmed. Therefore, it is highly desired to develop an effective tourism recommender system, which is able to provide people the useful tourism information.

Existing recommendation technologies [1] can be roughly classified into four categories: content-based recommendation, collaborative filtering recommendation, knowledge-based recommendation, and mixed recommendation. Among these recommendation technologies, collaborative filtering recommendation has been considered as the most successful recommendation strategy. The basic idea of collaborative filtering recommendation is that if users have the same preferences in the past (such as browsing the same webpages or purchasing the same products), then they are more likely to have similar preferences and hence make the same choices in the future. Considering the same phenomenon arises in tourists behaviors, the user-based collaborative filtering recommendation has been widely used in the tourism recommendation as well [3].

Although existing recommendation algorithms have been adopted in the major e-commerce sites [2], it still remains challenging in developing an effective tourism recommendation algorithms. The first challenge is how to improve the accuracy of the recommendation results to meet users' need. Most of the recommended results remain at the landscape level recommendation, which ignores other tourism factors, such as cloth, food, accommodation and travel. The other challenge is how to develop one dynamic recommendation algorithm, which can take both the context and the personalization into the consideration. To tackle these challenges, this paper aims at proposing a new tourism recommender system which is able to generate accurate result and achieve dynamic recommendation by incorporating the context and personalization.

The main contribution of the paper are two-fold. First, we propose TRSO, a tourism recommender system based on attraction ontology. In particular, we first adopt the association rules to identify the related users which avoids the problem of sparse matrix in collaborative filtering recommendation and reduces the time costs. Meanwhile, we construct the attraction ontology to provide a comprehensive definition and relationships among tourism concepts. Furthermore, we incorporate one hybrid recommendation approach by integrating the results derived from different types of users by applying different algorithms using the time factor, evaluation factor and the ontology. Second, we conduct extensive experiments to evaluate proposed algorithms and the experimental results indicate the superiorities of them over the traditional algorithms.

The rest of the paper is organized as follows. Section 2 introduces the tourism attractions ontology constructions. In Sect. 3, we introduce the TRSO system design as well as all the proposed algorithms. Sections 4 and 5 present our experimental evaluation and conclusion respectively.

## 2 Construction of Tourism Attractions Ontology

The ontology concept originated from the philosophy domain, which refers to the description of the object in existence in the world [4]. Recently, ontology theory has gradually appeared in the field of artificial intelligence, and has been widely applied in various contexts, such as the semantic web [5], intelligent information retrieval [6], and digital libraries [7]. However, there is still no clear definition of ontology today. The most widely quoted definition, proposed by Gruber et al. [8], is that “ontology is a clear specification of the conceptual model.” In a sense, ontology is used to describe concepts and the relationship among these concepts in a particular area or a more general area. By assigning reasonable, clear and unique definitions of concepts and relationships, ontology facilitates the communication between users and computers [9, 15, 16].

Tourism, an integrated industry including shopping, food, accommodation and travel, involves more complex and diverse information. Simply listing the flat tourism information cannot meet users’ need well. Meanwhile, ontology does not only provide clear and rich concepts, but also is capable of characterizing the hierarchical relationship among concepts. Therefore, it is important to take the ontology into consideration to recommend hierarchical tourism information.

Based on existing tourism attraction ontology and the online travelling information, this paper constructs a more comprehensive tourism attraction ontology. It is worth noting that we standardize the names according to the specification. The construction of the tourism attractions ontology includes ten classes: accommodation, services, transportation, food, culture, activity, shopping, environment, natural landscape, and cultural landscape. This ontology is the basis of the whole tourism recommendation and plays a significant role in the subsequent recommendation.

## 3 The Ontology-Based Tourism Recommender System

In the context of tourism recommendation, similar users may share similar preferences. Due to the high cost and exclusive characteristics of tourism, collaborative filtering strategy can be applied for the tourism recommendation. However, in reality, it is inevitable to encounter the problem of sparse matrix by using the collaborative filtering recommendation algorithm. Towards this end, we propose to employ the association rules. Association rules have one important step - minimum support filtering, which guarantees the completeness of users. Considering that the contextual information plays an essential role in tourism, we further incorporate the context filters. Finally, based on the tourism attraction information, we expand the proposed ontology and hence provide comprehensive tourism recommendation.

### 3.1 Mining Association Rules

Association rules were originally designed to mine the products sales correlation in data mining, such as the example of the “beer diaper” [10]. It is worth mentioning that such a relationship is not based on the similarity between products.

This paper aims to take advantage of association rules to find the relationship between tourists instead of the tourism attractions. In particular, based on users' travel history, we aim to find the related tourists for a given tourist. In this way, we can effectively solve the problem of sparse matrix in collaborative filtering, and also improve the efficiency.

Apriori algorithm [17] is one of the most widely used association rules and initially designed to solve database problems. In particular, Apriori algorithm repetitively scans the database to obtain the frequent item sets. Although this algorithm is simple and easy to implement, it is both time and space consuming. Later, FP-Growth algorithm was developed based on the Apriori algorithm, which employs a tree structure to store the data and hence significantly reduces the time costs regarding the repetitive scanning and space costs for the candidate sets.

### 3.2 Collaborative Filtering with Time Factor and Evaluation Factor

Different to existing collaborative filtering algorithm, this paper handles different categories of users differently. Users are first divided into two categories by association rules. The first category consists of the users whose supports are larger than the minimum support, while the second consists of the users whose supports are less than the minimum support. To this end, we propose the TEUCF (user based collaborative filtering with time factor and evaluation factor) algorithm to recommend tourism attractions to users. Detailed algorithm is introduced as follows.

The phenomenon of user's dynamic interests that may change with time, is common in recommendation field. Users' recent behavior can reflect users' current interests better than that occurred long time ago. In other words, users' interests gradually decrease with time. This paper uses Ebbinghaus forgetting curve as the time factor to improve the recommendation performance. Ebbinghaus forgetting curve [11] can simulate the forgetting curve of human brain well, and hence better reflect the normal case of human forgetting. Ebbinghaus forgetting curve is where the horizontal axis represents time (in days) and the vertical axis represents the percentage of the amount of memory (total memory capacity of 100 %).

The Ebbinghaus forgetting curve simulates human memory retention in time. It can be seen from the figure that the amount of memory goes down as time goes on. The curve sharply decreases in the beginning and then gradually levels off. Overall, the curve is gradually approaching the horizontal axis. Due to the fact that the memory curve cannot be expressed by regular functions, several scholars [12] proposed the maintained memory function to fit the memory curve. Maintained memory means the amount of memory maintained in brain. The function of maintained memory curve is:

$$J(t) = \frac{20e^b}{(t + t_0)^c} \quad (1)$$

where  $t$  is the time variable (in days),  $e$  is the natural log base,  $b$ ,  $c$  and  $t_0$  are the constants to be determined. In particular, we empirically set  $b = 0.42$ ,  $c = 0.0225$ ,  $t_0 = 0.00255$ , to make the function most consistent with human memory curve. This paper normalizes the formula, and proposes time factor formula as follows.

$$T(t) = \begin{cases} 1 & t > 1 \\ \frac{e^b}{5^{(t+t_0)^c}} & t \leq 1 \end{cases} \tag{2}$$

Apart from giving scores to tourism attractions, users also provide other evaluation forms, such as the textual evaluation comments. These comments not only illustrate the reasons for the scores but also affect the overall rating of the tourism attraction. If users feel the comments useful, they can give an “thumbs up” to express their agree with the comments. In particular, this paper introduces users’ such “thumbs up” behaviors. The more “thumbs up” an evaluation comment harvests, the more accurate the corresponding score is and the higher weights this score deserves. This paper incorporates users’ “thumb up” behavior as the evaluation factor into the collaborative filtering algorithm by using the below evaluation factor formula:

$$C = \begin{cases} 1 & i = 0 \\ \bigcup_{i=1}^m c_i & i > 1 \end{cases} \tag{3}$$

$$E(u) = 1 + \frac{C_u}{C} \tag{4}$$

where  $C_i$  and  $C$  are the number of reviews for comment  $i$  and the total number of reviews, and  $C_u$  and  $E$  are the number of comments for user  $U$  and the evaluation factor respectively. We incorporate the time factor and evaluation factor into the scoring matrix. In particular, we use Pearson correlation coefficient to measure the similarity. The Pearson correlation coefficient [13] is shown in formula 5.

$$sim_{uv} = \frac{\sum_{i \in I_{uv}} (r_{ui} - \bar{r}_u)(r_{vi} - \bar{r}_v)}{\sqrt{\sum_{i \in I_u} (r_{ui} - \bar{r}_u)^2} \sqrt{\sum_{i \in I_v} (r_{vi} - \bar{r}_v)^2}} = \frac{\sum_{i \in I_{uv}} R_{ui}R_{vi}}{\sqrt{\sum_{i \in I_u} R_{ui}^2} \sqrt{\sum_{i \in I_v} R_{vi}^2}} \tag{5}$$

Therefore, the Pearson correlation coefficient with time factor and evaluation factor is shown in formula 6.

$$R_{ui} = (r_{ui} - \bar{r}) \times T(ui) \times E(u) = \begin{cases} (1 + \frac{C_u}{C})(r_{ui} - \bar{r}) & t < 1 \\ \frac{e^b(1 + \frac{C_u}{C})(r_{ui} - \bar{r})}{5^{(t+t_0)^c}} & t \geq 1 \end{cases} \tag{6}$$

TEUCF algorithm procedure is shown as follows.

---

**Algorithm 1.** TEUCF Algorithm

---

Input: Target users; user-rating matrix M; Time factor; Evaluation factor;

Output: Top-N recommendation;

01: FP-Growth algorithm is used to mine the association rules of matrix M. And get the related user set  $U_b$ , which consist of the related user - attraction rating matrix.

02: Put time factor into rating matrix.

03: Based of step two, put evaluation factor into new rating matrix constructed by step two.

04: Use the new rating matrix which includes time factor and evaluation factor to calculate the similarity of related users to find out Top-N high similarity users, according to these users choice to generate recommendations.

---

### 3.3 Collaborative Filtering Based on Attractions Ontology

Certain tourists with a limited tourism attraction history would be filtered out by the association rules. Many existing algorithms would simply remove these users to achieve better performance which may result in incomplete result for such filtered users. Differently, in this paper, we propose an ontology-based collaborate filtering recommendation algorithm to deal with such users.

It is apparent that directly incorporate users with limited tourism attraction history in the collaborate filtering may devastate the recommendation performance due to the sparse matrix. To solve this problem, we use attraction ontology to classify various attractions into different categories. And we generate a new users-attractions class matrix, where the rating of each class is the average rating of all attractions of this class.

In this part, time is also an important factor which affects users' interests. Recent interests may reflect users current interests better. It is thus reasonable to put time factor into this algorithm. However, it is worth noting that the method to incorporate time factor is different from that in TEUCF. Normally, users may become interested in visiting one type of attractions within one time period and then change their interests after a while. But, it is likely that tourists will visit the similar type of attractions as their previous touring. In other words, the tourists' interests can be regained. On the basis of this observation, this paper uses the last time of visiting one specific type of attractions as the time for the whole continuous visiting for that type to avoid such regain behavior getting overwhelmed by the huge amount of history data. By doing so, each continuous visiting of one specific type of attractions is given one identified time. The rating with time factor formula is shown as follows.

$$R_{uLi} = (r_{uLi} - \bar{r}) \times T(ui) = \begin{cases} (r_{uLi} - \bar{r}) & t < 1 \\ \frac{e^{b(r_{uLi} - \bar{r})}}{5^{(t+t_0)^c}} & t \geq 1 \end{cases} \quad (7)$$

TCUCF algorithm procedure is shown as follow.

---

**Algorithm 2.** TCUCF Algorithm

---

Input: no related users  $U_l$ ; attractions ontology;

Output: Top-N recommendation;

01: Divide the old no related users-attractions rating matrix into the new related users-attractions types rating matrix based on attractions ontology.

02: Find out the last time of visiting this type of attractions. Put this time into the new related users-attractions types rating matrix.

03: Use the new rating matrix which includes time factor to calculate the similarity of related users to find out Top-N high similarity users, according to these users choice to generate recommendations.

---

### 3.4 Hybrid Algorithm

Hybrid algorithm is a common approach in recommender system, which is able to overcome the limitation of a single recommendation algorithm. Existing hybrid algorithms mainly have two kinds of hybrid modes: one for aggregating the results resulted in multiple individual algorithms and the other one for aggregating the algorithm processes. Due to the fact that this work separates users into two categories and handles them respectively, the first hybrid mode is preferred. In addition, the proposed tourism recommender system has two main branches. One branch takes the time factor as well as evaluation factor into consideration, while the other one considers the ontology. In particular, the second branch actually also achieves the hybrid of algorithm processes. The basic idea of the proposed hybrid algorithm is, by adopting the ontology information, to apply different algorithms for different category of tourists individually where the results are further aggregated together. Due to the space limitation and the simplicity of the algorithm, we will omit the detail here.

### 3.5 Multiple Context Information Filtering

Now, we are ready to introduce how the multiple context information filtering is applied in the system. Tourism is a field that is always influenced by context. The context may not only decide whether a tourism activity is feasible, but also affect the performance of tourism recommendation. In this work, the context information is obtained from the tourism attraction ontology. The proposed attraction ontology well fits the characteristics of tourism. The context information are set as 'season', 'location', 'weather' and 'time', which fully takes into account the characteristics of tourism.

To obtain the aforementioned context information, both the explicit and implicit approaches are applied. Specifically, the location and time can be explicitly collected, and the local weather conditions and season information are implicitly obtained online.

Adomavicius et al. [14] proposed two ways for context filtering in 2005, which are context pre-filter and context post-filter. Context pre-filter removes the irrelevant information about users' preferences regarding tourism attractions at first. And then traditional recommendation algorithm can be used to process the

data set. In this way, the recommendation can fit well for users need and context need. Differently, context post-filter first conducts the recommendations by traditional recommendation algorithm. And then filter out the recommendations that do not fit for users' preferences.

Consider that the context pre-filter is much vulnerable to the context granularity. Coarse granularity may lead to the useless attractions in the recommendation, while too fine granularity may lead to the much sparse data set. Therefore, the post-filter method is adopted to ensure the recommendation performance.

### 3.6 Information Expansion Based on the Tourism Attractions Ontology

This work utilizes an ontology-based tourism recommendation algorithm which is able to find out the attractions which users may interested in. Furthermore, we use the context information (e.g., time, season, location and weather) to filter the recommendation results. This part uses a tourism attraction ontology to provide comprehensive tourism recommendation. The attraction ontology in this paper includes ten aspects: transportation, food, culture, activities, attractions, services, shopping, environment and accommodation. This ensures the diversity of tourism information and makes the tourism information recommendation being of more practical value.

## 4 Experimental Results and Analysis

In this section, we first introduce the dataset used for evaluation, and then present the method used for evaluation followed by the experimental results and analysis.

**Datasets:** In this work, the tourism information is crawled from the “Baidu tourism”. In particular, we collect the following information about the attractions: names, attractions ratings, tourism time, weather and evaluation. After preprocessing, we store all these data in the database.

In total, we collected 1975 tourists' records which consists of about 2000000 attraction records. Each attraction has one rating, which can be five levels from low to high. Each record includes attraction, attraction type, location, time and “thumbs up” counts.

In the experiment, we randomly select 80% rating data as the training set and the rest as the test set. This learning process is conducted in multiple rounds. The final result is calculated by averaging the value of all rounds.

**Evaluation Method:** The recommendation algorithm in this paper is developed on the basis of the traditional user-based collaborative filtering algorithm (UserCF), which consists of four steps as shown below:

- The first step is to find the association users whose supports are larger than 10 by the FP-Growth algorithm.



- The second step uses the UserCF algorithm to recommend attractions among related users. This algorithm is referred as FUCF algorithm.
- The third step first uses TEUCF and TCUCF to handle the related users and unrelated users, respectively, and then hybrids the recommendation results. We call this method as the MUCF algorithm.
- The fourth step takes the context information filter into consideration. This algorithm is called CMUCF algorithms.

#### 4.1 Results and Analysis

First, we use the precision and recall to evaluate the quality of the recommendation. Precision represents the ratio of the intersection of the recommended attractions to the user and the actual tourism attractions of the user over the the actual tourism attraction of the user. Recall stands for the ratio of the intersection of the recommended attractions to the user and the user’s actual viewing of the site over the recommended attractions to the user. Note that the larger the precision is, the more accurate the recommendation is. Recall follows the same manner. We change the number of K to do comparative experiments. Precision and recall are shown in Table 1 with different K-value.

**Table 1.** Precision comparison and Recall comparison

k	Pre(UserCF)	Pre(FUCF)	Pre(MUCF)	Rec(UserCF)	Rec(FUCF)	Rec(MUCF)
1	0.3678	0.4059	0.4785	0.1169	0.1236	0.1013
2	0.3029	0.3372	0.4119	0.1862	0.1967	0.1872
3	0.2651	0.2928	0.3779	0.2374	0.2459	0.2462
4	0.2359	0.2627	0.3477	0.2735	0.2836	0.2751
5	0.2157	0.2427	0.3296	0.3049	0.3177	0.3141
6	0.1996	0.2263	0.3149	0.3303	0.3451	0.3600
7	0.1869	0.2111	0.2971	0.3526	0.3633	0.3962
8	0.1767	0.2010	0.2840	0.3728	0.3856	0.4326
9	0.1689	0.1934	0.2709	0.3936	0.4087	0.4638
10	0.1621	0.1884	0.2610	0.4126	0.4356	0.4962

From Table 1, we can see that the precision of FUCF algorithm and MUCF algorithm significantly improved the recommendation performance compared with the traditional collaborative filtering algorithm. In terms of recall, when the number of recommended records is low (i.e., smaller than 6), all the three algorithms perform similarly. As the recommended number increases 6 and above, the MUCF significantly outperforms the other two, and the difference increases continuously. This indicates that the MUCF algorithm proposed in this paper is superior to the traditional algorithms in terms of the precision and recall.

We then employ MAE and RMSE to evaluate the recommendation quality. MAE and RMSE calculate the error between the predicted ratings and the actual

users’ ratings. The smaller the MAE and RMSE are, the more accurate the recommendation is. The performance of different algorithms regarding MAE and RMSE are shown in Table 2.

**Table 2.** MAE and RMSE comparison

k	MAE(UserCF)	MAE(FUCF)	MAE(MUCF)	RMSE(UserCF)	RMSE(FUCF)	RMSE(MUCF)
1	0.4561	0.4652	0.4611	0.6213	0.6165	0.6225
2	0.4648	0.4802	0.4553	0.6468	0.6308	0.6158
3	0.4523	0.4738	0.4544	0.6304	0.6326	0.6132
4	0.4623	0.4701	0.4526	0.6423	0.6275	0.4785
5	0.4690	0.4777	0.4494	0.6551	0.6386	0.6111
6	0.4700	0.4795	0.4516	0.6542	0.6426	0.6135
7	0.4701	0.4755	0.4508	0.6530	0.6407	0.6118
8	0.4673	0.4697	0.4515	0.6501	0.6335	0.6161
9	0.4704	0.4668	0.4502	0.6622	0.6313	0.6150
10	0.4724	0.4654	0.4495	0.6653	0.6296	0.6144

We can see that there is no clear advantages of FUCF over other algorithms regarding MSE when the recommended number is less than 8. The MAE of MUCF algorithm is lower than the other two algorithms only when k equals to 2. Overall, MUCF outperforms the other traditional recommendation algorithms.

Furthermore, we also evaluate the coverage rate as another metric in the recommendation to evaluate whether the personalization recommendation can be achieved. The coverage rate reflects the popularity of recommended attraction. The hot spots often appear on each tourism recommendation, thus the personalized recommendation is poor. The coverage rates are shown in Table 3.

**Table 3.** Coverage rate table

k	Cove(UserCF)	Cove(FUCF)	Cove(MUCF)
1	0.0629	0.0666	0.0988
2	0.1036	0.0962	0.1536
3	0.1480	0.1443	0.1762
4	0.1776	0.1813	0.1973
5	0.2072	0.2035	0.2518
6	0.1996	0.2263	0.2734
7	0.2960	0.2960	0.3506
8	0.3330	0.3404	0.3676
9	0.3478	0.3774	0.3854
10	0.3774	0.4218	0.4313

The corresponding coverage rate curves of different algorithms show that there is no significant difference between FUCF and UserCF in terms of coverage

rate. But MUCF clearly outperforms UserCF which indicates the superiority of MUCF of handling the personalized recommendation.

Next, we evaluate the performance of context filtering algorithms. The context information is obtained from the attraction ontology which includes location, season, weather and time. In the experiment settings, we set the location as “Harbin”, season as the “winter”, time scheduled as the days in December. We obtain the weather conditions from the “[www.tianqi.com](http://www.tianqi.com)”. The sample weather conditions of Harbin in December is shown in Table 4.

**Table 4.** Part of December weather conditions in Harbin

Date	HighTemperature	LowTemperature	DayWeather	NightWeather	WindDirection	WindPower
1st	1	-4	Cloudy	Snow Shower	ES	Gentle Breeze
2nd	-10	0	Snow Shower	Snow Shower	ES	
3rd	-13	0	Snow Shower	Snow Shower	WS	
4th	-7	-13	Sunny	Snow	S	Gentle Breeze
5th	-7	-20	Snow Shower	Sunny	WN	Gentle Breeze
6th	-11	-21	Sunny	Sunny	WS	Breeze
7th	-10	-19	Sunny	Sunny	WS	Breeze
8th	-4	-15	Sunny	Cloudy	WS	Gentle Breeze
...	...	...	...	...	...	.
29th	-11	-24	Sunny	Smog	W	Breeze
30th	-15	-23	Smog	Smog	W	Breeze
31st	-11	-19	Smog	Smog	WS	Breeze

As shown in Table 4, there is no extreme snow or moderate gale in December. Only 30<sup>th</sup> and 31<sup>th</sup> had the smog weather. Therefore, it is reasonable to recommend indoor tourism activities from 30<sup>th</sup> to 31<sup>th</sup> due to the smog weather. And the rest of the month is fine for either indoor or outdoor tourism. We further calculate the precision and recall after the context filtering as shown in Table 5. As what Table 5 indicates, CMUCF achieves the best performance. This also shows that taking the context filtering into consideration can improve the precision and recall simultaneously.

Finally, instead of simply recommending tourism attraction, we expand the recommendation with other tourist attraction information based on the ontology which is able to meet users’ basic necessities (e.g., cloth, food, accommodation and travel). For instance, we have set the location for the “Harbin”, season for “winter”, time for daytime, as the context. We thus based on the context information combined with the ontology to provide the final recommendation. For example, we can get the recommended attraction as “Snow World” and the nearby shopping malls such as “Wanda Plaza”, nearby special restaurant such as “Manhattan restaurant” and nearby hotel such as “seven days inns”. Moreover, considering the season context, we recommend the winter activity “Harbin International ice and snow festival”.

**Table 5.** Precision and recall comparison

k	Pre(MUCF)	Pre(CMUCF)	Rec(MUCF)	Rec(CMUCF)
1	0.4785	0.4877	0.1013	0.1992
2	0.4119	0.4247	0.1872	0.2998
3	0.3779	0.3865	0.2462	0.3592
4	0.3477	0.3645	0.2751	0.4013
5	0.3296	0.3477	0.3141	0.4212
6	0.3149	0.3376	0.360	0.4415
7	0.2971	0.3303	0.3962	0.4598
8	0.2840	0.3243	0.4326	0.4737
9	0.2709	0.3190	0.4638	0.4856
10	0.2610	0.3161	0.4962	0.5019

## 5 Conclusions

Towards a better tourism recommendation service, we proposed TRSO, a tourism recommender system which adopts the ontology. Specifically, we introduced different techniques of identifying related/unrelated tourists using the association rules, constructing the attraction ontology, conducting the hybrid recommendation after applying different algorithms on different types of tourists and filtering the information based on the context information. The experimental results indicate the efficiency and accuracy of the proposed techniques.

**Acknowledgments.** The work is supported by the National Natural Science Foundation of China under Grant No. 61272185, the Natural Science Foundation of Heilongjiang Province of China under Grant No. F201340, the Science Foundation of Heilongjiang Province of China for returned scholars under Grant No. LC2015025, the Fundamental Research Funds for Central University under Grant No. HEUCF160602, and Harbin Special Fund for innovative talents of science and technology research under Grant No. 2013RFQXJ113.

## References

1. Buettner, R.: Predicting user behavior in electronic markets based on personality-mining in large online social networks: a personality-based product recommender framework. *Electron. Markets Int. J. Netw. Bus.*, 1–19 (2016)
2. Xu, W.H., Xiao, L.X., et al.: Comparison study of internet recommendation system. *J. Softw.* **10**, 350–362 (2009)
3. Jannach, D., Zanker, M.: *Recommend System*. Beijing University of Posts and Telecommunications Press, pp. 1–4 (2013)
4. Spiliopoulou, M.: The laborious way from data mining to the web mining. *Int. J. Comput. Syst. Sci. Eng.* **14**(2), 113–126 (1999). Special Issue on Semantics of the Web

5. Du, X., Li, M.: Summary of research on ontology learning. *J. Softw.* **17**, 1837–1848 (2006).
6. Strobbe, M., Van Leare, O.: Interest based selection of user generated content for rich communication services. *J. Netw. Comput. Appl.* **33**, 84–97 (2010)
7. Zghal, H.B., Moreno, A.: A system for information retrieval in a medical digital library based on modular ontologies and query reformulation. *Multimedia Tools Appl.* **72**, 2393–2412 (2010)
8. Gruber, T.R.: A translation approach to portable ontology specifications. Technical Report. 14, 2367–2456 (1993)
9. Ensan, F., Du, W.C.: A semantic metrics suite for evaluating modular ontologies. *Inf. Syst.* **38**, 745–770 (2013)
10. Baby, M., Idicula, S.M.: Apriori-based research community discovery in bibliographic database. *Comput. Netw. Intell. Comput.* **157**, 75–80 (2011)
11. Teyarachakul, S., Chand, S., Ward, J.: Effect of learning and forgetting on batch sizes. *Prod. Oper. Manag.* **20**, 116–128 (2011)
12. Zhiheng, J.: On the function of the past on the psychology of memory. *J. Dyn.* **3**, 3–23 (1988)
13. Orhan, E., Elvan, C., Yusuf, V.: A new correlation coefficient for bivariate time-series data. *Phys. A Stat. Mech. Appl.* **15**, 274–284 (2014)
14. Adomavicius, G., Sankaranarayanan, R., Sen, S., Tuzhilin, A.: Incorporating contextual information in recommender systems using a multidimensional approach. *ACM Trans. Inf. Syst.* **23**, 103–145 (2005)
15. Kang, W., Tung, A.K., Chen, W., Li, X., Song, Q., Zhang, C., Zhao, F., Zhou, X.: Trendsmedia: an internet observatory for analyzing and visualizing the evolving web. In: *ICDE 2014*, Chicago, IL, USA, pp. 1206–1209 (2014)
16. Kang, W., Tung, A.K., Zhao, F., Li, X.: Interactive hierarchical tag clouds for summarizing spatiotemporal social contents. In: *ICDE 2014*, Chicago, IL, USA, pp. 868–879 (2014)
17. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules in large databases. In: *VLDB 1994*, Santiago, Chile, pp. 487–499 (1994)