Recommending Ideal Holiday at National Level

Biljana Petrevska\(^1\), Saso Koceski\(^2\)

\(^1\)Faculty of Tourism and Business Logistics, \(^2\)Faculty of Computer Science
Goce Delcev University – Stip, Macedonia
\(^1\) biljana.petrevska@ugd.edu.mk; \(^2\) saso.koceski@ugd.edu.mk

Abstract

The paper underlines the importance of meeting and fulfilling travelers’ and tourists’ preferences by introducing personalized recommendation system. The proposed web-based software model employs the process of collaborative filtering in order to assist tourists in identification of their ideal holiday. The research outcome is creation of generated personalized list of favorable and tailor-made potential items for all visitors of designed tourism portal. The accuracy testing performed highly satisfactory results thus reporting on positive practical experience at national level.

Keywords: Tourism; Collaborative filtering; Tourists’ preferences; Holiday.

Introduction

One may argue the inevitable relationship between tourists, their preferences and information. Moreover, it is a widely-recognized fact that information and decision-making have become the foundation for the world economy [1]. It is more than obvious that whether a potential tourist will be interested in a certain item depends on the preferences. Although may sound fragile, but the vast majority of today’s tourists and travelers know exactly what they are looking for. Yet, they are very demanding and have complex, multi-layered desires and needs. Today’s so called “postmodern tourists” have specific interests and individual motives which results in tailored made tourist products according to their particular preferences. They are often high experienced in travelling and demand perfect tourist products rather than standardized ones. Consequently, they take much more active role in producing diversified tourist products with shorter life cycles enabled by increased usage of the information technology.

However, attracting a bigger number of tourists and travelers is not a trouble-free process, particularly in times of ever-changing travel preferences. Despite the variety of options regarding tourist destination or attraction, visitors frequently are not capable to cope with such a huge volume of choice. Moreover, they need advice about where to go and what to see. In a tourism domain, recommendations may indicate cities to go to, places to visit, attractions to see, events to participate in, travel plans, road maps, options for hotels, air companies, etc. Such scope of work very often is not a trivial task. In this respect, recommendation systems assist tourists by facilitating personal selection and prevent them from being overwhelmed by a stream of superfluous data that are unrelated to their interest, location, and knowledge of a place. So, it is much easier for tourists to access the information they need thus resulting in shorter lead-time for bookings, making last-minute decisions and generally, tailoring their own packages from a suite of options.

Solution is seen in personalization of the information delivery to each traveler, together with the travel history. Yet, the advanced tourist information systems must offer more than just relatively static information about sights and places. Over the past two decades Internet had an enormous impact on the tourism industry, specifically to the way how tourists search for information. A noteworthy transformation was made from just passive searching and surfing to creating content, collaborating and connecting. In this respect, the Web became the leading source of information particularly important in times of increased number of competitors in tourism market. The way out is detected in application of recommendation system as a promising way to differentiate a site from the competitors. So, user-generated content will gain in significance thus enabling developing more accurate recommender.
This study intends to present and elaborate necessity of introducing recommenders in tourism which may assist tourists in finding a way-out in creating their perfect vacation in efficient and transparent way. In order the meet the forth mentioned aim and objective, the paper is structured in several parts. So, Section 2 presents a brief overview on literature review on this issue. The methodology and scope of work are set in Section 3, while the conclusions and future research directions are noted in Section 4.

Related Work

Due to the importance of tourism, recommendation systems applied in tourism have been a field of study since the very beginnings of artificial intelligence. There is a large body of literature regarding their importance and effectiveness of application in tourism, travelling and hospitality. It is a matter of identifying a class of intelligent applications that offer recommendations to travelers, generally as a response to their queries. Hence, without prior knowledge of their preferences, it is groundless to expect efficient tourism development.

Tourists’ Preferences

Many researchers were interested in identifying tourists’ needs, expectations and behavior. In this respect, numerous papers discuss tourist roles in order to define their considerable variations. In mostly, the behavior is related to specific demographic and background characteristics emphasizing the life course as the leading component for investigating tourist role preferences. Yet, attention should be paid to a variety of social structures and processes, including psychological needs and life-course stage.

Cohen [2] was one of the first sociologists who proposed a typology to conceptually clarify the term “tourist” by developing a four-fold typology. Based on that, Pearce [3] identified specific behaviors thus enabling tying the evolutionary nature of tourist role preference and the psychological needs. Moreover he developed 15 different tourist types which allowed creation of several measurement scales. In this respect, the Tourist Roles Preference Scale [4] presents a comprehensive classification of leisure tourists. Additional work resulted in adding two more tourist types to the tourist categorization [5]. Moreover, researchers focused on exploring the experience of tourists as well as the importance of the tourist experience for tourists [6].

Recommendation Systems

Generally, the recommendation systems mostly leverage in-built logical reasoning capability or algorithmic computational schemes to deliver their recommendation functionality. Consequently, the recommenders are an attempt to mathematically model and technically reproduce the process of recommendations in the real world.

Numerous researchers made efforts in their introducing. In this respect the need for developing intelligent recommenders which can provide a list of items that fulfill as many requirements as possible is elaborated [7], [8] and [9]. Also, a recommender system dealing with a case-based reasoning is introduced in order to help the tourist in defining a travel plan [10] and [11]. However, as the most promising recommenders in the tourism domain are the knowledge-based and conversational approaches [12] and [13]. Yet, some other variants of the content-based filtering and collaborative filtering are engaged for recommendation, like knowledge-filtering, constraint-based and case-based approaches [14], [15] and [16]. In the same line, the recommenders based on a text mining techniques between a travel agent and a customer through a private Web chat may easily find an application [17].

Due to rapid expansion of tourism industry, the recommenders for tourism have attracted a lot of interest in academia. Some late research that brought more sophisticated outcomes are referred additionally, like: introducing a personalized tourist information provider as a combination of an event-based system and a location-based service applied to a mobile environment [18]; investigation on sources and formats of online travel reviews and recommendations as a third-party opinions in assisting travelers in their decision making during the trip planning [19]; findings regarding development of a web site in order to enable Internet users to locate their own preferred travel destinations according to their landscape preferences [20] and similar. Furthermore, the usage of the orienteering problem and its extensions to model the tourist trip planning problem was elaborated as efficient solution for number of practical planning problems [21]. It is evidently that the research area is extending resulting in improving the dependability of recommendations by certain semantic representation of social attributes of destinations [22]. Moreover, most recommenders focus on selecting the destination from few exceptions [23] and [24].

Methodology and Scope of Work

The paper aims in developing a web-based tourism portal on national level. To this purpose, the case of Macedonia is empirically investigated. The research outcome is introduction of an efficient and accurate personalized recommendation system which will support tourists and travelers visiting Macedonia by helping them to identify relevant tourist objects to match to their personal interests, preferences and desires.
The research employs dataset from proprietary database collected by the mixed research group composed of researchers from the “Goce Delcev” University. It contains 56320 ratings from 483 users for 818 tourist objects, whereas each user has rated at least 30 objects, and each object has been rated at least once. In order to accomplish the research objective and main aim, a several step methodology was developed. The first step foresees tourists and tourist objects profiling. In the line of modeling tourist personal profile, the system uses tourist types methodology [4]. Namely, the tourist profile indicates the degree to which tourists identify themselves with the given types. Typically, individual tourist cannot be characterized by only one of these archetypes but has unique combination of these personalities, although to varying degrees. Thus, tourist types model the tourists’ generic interests in an abstract form. Vectors are suited to model such tourist profile, whereby each dimension corresponds to a certain tourist type while the value indicates how much the tourist identifies him- or herself with the corresponding type.

Tourist profiling is a two-step process which involves creating the profile and then reviewing the profile to make any necessary adjustments. The initial tourist profile for each system user is created by the user himself during the process of registration, by determining the degree of membership to each of the tourist types. Considering the fact that the human preferences change over time due to various factors, the tourists might change their behavior too. To make the system capable to cope with these changes, the tourist profile adjustment has been enabled. It is based on the ratings the tourist give for each tourist object that he visits after his journey and according to Eq. 1.

\[
U_{ij,t+1} = \frac{1}{2} \left( U_{ij,t} + R_{ik,t+1} \ast w \ast O_{kj} \right)
\]

(1)

Where:
- \( U_i \) denotes i-th user and \( U_i \in U \)
- \( U \) denotes the set of users registered to system
- \( U_{ij,t} \) denotes degree of membership in the moment t of i-th user to tourist type \( T_j \) and \( T_j \in T \)
- \( T \) denotes the set of tourist types according to Gibson and Yiannakis (2002)
- \( O_k \in O \) denotes k-th object in the set of all objects \( O \) registered in system
- \( w \) denotes the weighting factor and

Similarly, the profiles for attractions might be generated and represented in form of a vector. So, every tourist object is modeled through a vector as well. Thereby, this vector describes in a quantitative way how much the object is related to the given types. For example, the famous monastery Saint Panteleimon in the city of Ohrid known as a birthplace for Cyrillic alphabet and used by Saint Clement for teaching the Cyrillic alphabet, might be highly relevant for sightseeing tourists but not for such kind of tourists that would like to do some risky activities.

In the developed system a manual process to link the given tourist types to appropriate tourist objects is proposed. Therefore, for each of the tourist objects, the degree of relationship to each of the tourist types is specified by domain experts. In order to prevent information overload of the tourist and provide only relevant information, the system should recommend a subset of tourist objects according to the personal experiences individual tourist desire and those he/she prefer to avoid. This in turn might lead to an increase of the tourist’s satisfaction of experiencing a relaxed sightseeing trip.

According to this, the next step of the proposed methodology aims to match tourist profiles against the set of tourist objects on the basis of tourist types, thus producing a ranked list of objects for each given tourist and reducing the set of objects. If a tourist profile matches the characteristics of an object, this object should be recommended to the respective tourist. Therefore, the matchmaking algorithm has to examine whether they share similar structures. The more similarities they have in common, the more contributes the tourist object to the tourist’s satisfaction and should be ranked higher.

To estimate the similarity degree between tourist profiles and tourist objects, the system contains a special module based on a vector-based matchmaking function, whereby a given profile and each tourist object constitute vectors and are compared in a vector space model. A common method to obtain the similarity is to measure the cosine angle between two vectors. If the vector space is non-orthogonal, kernel based algorithms can be applied to measure the similarity in such a space. The dimensions of the vector space model correspond to selected tourists types [5], such that each distinct tourist type (e.g., adventure or cultural type) represents one dimension in that space (Eq. 2).

\[
SIM_{cos}(U_i,O_j) = \frac{\sum_{k=1}^{N} U_{ik} \cdot O_{jk}}{\sqrt{\sum_{k=1}^{N} U_{ik}^2} \sqrt{\sum_{k=1}^{N} O_{jk}^2}}
\]

(2)

Where:
- \( U_{ik} \) denotes the degree of membership of the i-th user to the tourist type \( T_k \)
- \( O_{jk} \) denotes the degree of membership of the j-th tourist object to the tourist type \( T_k \)
- \( N \) denotes the number of tourist types.

JOURNAL OF APPLIED ECONOMICS AND BUSINESS, VOL.1, ISSUE 1 – MAY, 2013, PP. 15-22
According to Eq. 2, the degree of similarity between tourist profiles and tourist objects will be calculated. Tourist objects will be ordered by the value of the matchmaking function for a given user, and only those objects that have positive value for this function will be considered for recommendation. Considering the five point Likert scale for rating the objects, to each object in the constructed set, a recommendation mark will be assigned.

Furthermore, another very important fact is considered related with behavior dimension of tourists and travelers planning a vacation or trip. In everyday life, while planning a holiday, people also rely on recommendations from reference letters, news reports, general surveys, travel guides, and so forth. In addition, they desire personal advice from other people with similar preferences or someone they trust. In fact, over 80% of travelers participating in a TripAdvisor.com survey agree that ‘reading other travelers’ online reviews increases confidence in the system is in its early stages. On the other hand, the sparsity problem occurs when available ratings data is rare and insufficient for identifying the similar neighbors. This problem is often very significant when the system is in its early stages. On the other hand, when numbers of existing users and items grow tremendously, traditional CF algorithms will suffer serious scalability problems, with computational resources grown nonlinearly and going beyond practical or acceptable levels.

To reduce the dimensionality of data and avoid the strict matching of attributes in similarity computation the cloud-model CF approach has been adopted. It is constructing the user’s global preference based on his perceptions, opinions and tastes, which are subjective, imprecise, and vague [27]. Hence, it seems to be an appropriate paradigm to handle the uncertainty and fuzziness on user preference.

The main goal of the cloud model CF is to construct the global preference for each user by calculating a triple of three digital characteristics: expected value-Ex, entropy- En and hyper-entropy He [19] (Eq. 3).

\[
Ex = \frac{1}{n} \sum_{i=1}^{n} r_{u,i}
\]

\[
En = \sqrt{\frac{1}{2} \times \frac{1}{n} \sum_{i=1}^{n} |r_{u,i} - Ex|}
\]

\[
He = \sqrt{S^2 - \frac{1}{3} En^2}, \text{where} \ S = \frac{1}{n-1} \sum_{i=1}^{n} (r_{u,i} - Ex)^2
\]

The recommendation function based on the cloud model is defined as in Eq. 4:

\[
R_{u,j} = \bar{r}_u + \sum_{v \in N(u)} \left( r_{v,j} - \bar{r}_v \right) \times \text{sim}(u, v)
\]

Where:

N(u) denotes the k most similar users to active user u ru and rv denote the average rating of user u and v, respectively.

The value of rating rv,j is weighted by the similarity of user v to user u; the more similar the two users are, the more weight rv,j will have in the computation of the recommendation function.

According to the value of total recommendation functions the objects will be ordered and further classified into five categories (Eq. 5).

\[
\text{Cat}_{i,j} = \begin{cases} 
  k = 1, \forall Oj \in O_{i_{rec}} \land 0 \leq \text{Frec}_{i,j} \leq 0.2 \\
  k = 2, \forall Oj \in O_{i_{rec}} \land 0.2 < \text{Frec}_{i,j} \leq 0.4 \\
  k = 3, \forall Oj \in O_{i_{rec}} \land 0.4 < \text{Frec}_{i,j} \leq 0.6 \\
  k = 4, \forall Oj \in O_{i_{rec}} \land 0.6 < \text{Frec}_{i,j} \leq 0.8 \\
  k = 5, \forall Oj \in O_{i_{rec}} \land 0.8 < \text{Frec}_{i,j} \leq 1 
\end{cases}
\]

Design

The proposed web-based tourism portal encompasses national level and is structured in a form of social network. Although official national tourism portal of
Macedonia already exists (www.exploringmacedonia.com) and several other private initiatives act as additional tourism portals, thus supporting country’s tourism profile (www.travel2macedonia.com, www.go2macedonia.com, www.simplymacedonia.com, www.macedialovesyou.com, www.mysticalmacedonia.com, www.macedonia-timeless.com etc), the suggested model is interactive and assisting. Namely, the proposed recommendation system is a significant improvement on existing travel websites and provides tourists and travelers with a customized, unique, and enriching travel experience. It incorporates some standard plugins typical for social networks like Facebook, Tweetter, LinkedIn, MySpace etc. But, it advances the concept by including custom plugins, like the recommended objects plugin which is the core of the portal. It is using the Google Map of Macedonia to visualize static tourist objects (object that are not temporary, like churches, museums, archeology localities, etc) and dynamic objects (object that have limited time duration, like events, expositions, etc.). They are displayed on the map according to their geographical location being grouped in municipalities. Municipalities are recommended to the user in the form of circles as displayed on the map (Fig. 1). The size of the circle indicates the user’s affinity for the municipality; therefore, a large circle indicates a municipality with many tourist objects with high recommendation marks i.e. that match the user profile. By displaying the user’s affinity through the size dimension of the circle, users can easily observe which municipalities would be of most interest to them.

The tourist objects are displayed as icons in the location of the correspondent object (Fig. 2). The image of the icon indicates the type of tourist objects such as a museum, church, or restaurant. The size indicates how closely the object meets the user’s interests. Each attraction also has an information window as displayed in Figure 2. The information window usually includes the name and picture of the attraction, an icon of an umbrella indicating that the attraction is accessible in the rain, and tags. Additionally, it displays a general idea of the time consumption of the attraction, friends who have visited the attraction, and an option to view narratives in either video, audio, or text format. Through this window, the user can also rate the object. This operation is recommended to be done after visiting the object and according to the personal experience and satisfaction. The goal of this operation is two-fold: to help updating the user profile, and to make the process of recommendation more accurate.

Accuracy Testing

In order to measure recommendation accuracy, the information-retrieval classification metrics are used. This step is undertaken for the purpose of evaluating the capacity of the recommender system in suggesting a list of appropriate objects to the user. Such process enables measurement of probability that the recommendation system takes a correct or incorrect decision about the user interest for an item. When using classification metrics, four different kinds of recommendations can be distinguished. If the system suggests an interesting tourist object to the user we have a true positive (TP), otherwise the object is uninteresting and we have a false positive (FP). If the system does not suggest an interesting tourist object we have a false negative (FN). If the system does not suggest an object uninteresting for the user, we have a true negative (TN).

The most popular classification accuracy metrics are the recall and the precision. Recall measures the percentage of interesting objects suggested to the users, with respect to the total number of interesting objects, while precision measures the percentage of interesting objects suggested to the users, with respect to the total number of suggested objects. In order to understand...
the global quality of a recommender system, a combination of recall and precision by means of the F-measure may be done.

Furthermore, to evaluate the system a methodology which uses the k-fold and the leave-one-out together with classification metrics recall and precision was used. The created model is tested on the remaining partition by means of the following algorithm:

Step 1: One user in the testing set is selected (the active user).
Step 2: One rated tourist object (the test object) is removed from the profile of the active user.
Step 3: An order list of recommended tourist objects is generated.
Step 4: If the test item is in the top-3 categories of recommended objects, either the true positive or false positive counter is incremented, depending whether the user liked or disliked the test item.

In this respect, two distinct user groups were considered. Group I contained all users who have rated 30-60 objects (the few raters user group). Group II contained all users who have rated 61-100 objects (the moderate raters user group). Step 1 of the proposed algorithm was repeated for all the users in both groups. Steps 2-4 are repeated for all the objects rated by the active user. In order to understand if a user likes or dislikes a rated tourist object, it is presumed that an object is interesting for the user if it satisfies two conditions. The first constraint reflects the absolute meaning of the rating scale, while the second the user bias. If a rating does not satisfy both conditions, it is assumed that the item is not interesting for the user.

Upon the conducted evaluation, the results for system precision, recall and f-measure were averaged for each of the groups (Table 1).

<table>
<thead>
<tr>
<th>Group</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F-measure (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>75.14</td>
<td>79.18</td>
<td>77.11</td>
</tr>
<tr>
<td>II</td>
<td>81.74</td>
<td>85.32</td>
<td>83.49</td>
</tr>
</tbody>
</table>

According to the obtained results, the suggested national tourism web-based portal with its collaborative recommender system seems to be robust as it achieves good results in both scenarios (group I - users with few and group II - moderate ratings). It also accomplishes a good trade-off between precision and recall, a basic requirement for all recommendation systems. The experimental results show that the proposed approach can provide satisfactory performance even in a sparse dataset.

**Conclusion**

Although the designed national tourism portal is in initial phase of development, resulted in accurate recommendations and guidelines for tourists and travelers in the line of identifying an ideal trip and holiday. In this respect, it must be noted that tourism is defined as one of the most economically-oriented industries in the world due to the fact that enhances and strengthens national economies. Moreover, the development of such software module contributes generally to increasing the awareness of tourist destination that is capable of fulfilling travelers’ preferences, and respectfully in raising net tourism income.

The successful implementation of the proposed recommender (named “MyTravelPal”) based on collaborative filtering notes positive impulses in the line of supporting the national economy through improvement of tourism supply in more qualitative manner. It assists all interested parties in planning their travel on more intelligent way by generating a personalized list of favorable and tailor-made items. Since this portal indicates the motives, preferences and reasons for traveling to Macedonia, it may be of high importance to all key-tourism actors in the process of identifying measures and implementing activities necessary for creating comprehensive tourism policy.

Generally, the contribution of this paper lies in the fact that it proposes methodology for developing a module which relies on efficient and accurate personalized recommendation algorithm that supports tourist consumers to identify relevant tourist objects matching to their personal interests and to plan more efficiently their trips. Additionally, the empirical investigation may alarm the relevant tourism-actors that the time has changed and that the on-line experience has shifted from searching and consuming to creating, connecting and exchanging. Previously passive consumers and web surfers are now generating content, collaborating and commenting thus achieving their rights to enjoy the electronic communication as a fragment of the general economic interest services.

**REFERENCES**

Biljana Petrevska, Saso Koceski:
Recommending Ideal Holiday at National Level

