A Semantic Social Recommender System Using Ontologies Based Approach For Tunisian Tourism

Mohamed FRIKHA; Mohamed MHIRI; Faiez GARGOURI
med.frikha@gmail.com, med.mhiri@gmail.com, faiez.gargouri@isimsf.rnu.tn
University of Sfax - MIRACL Laboratory - Sfax, Tunisia

KEYWORD
Recommender System;
Social Network;
collaborative filtering;
Social Semantic Web,
Ontology; Tourism

ABSTRACT
Tunisia is well placed in terms of medical tourism and has highly qualified and specialized medical and surgical teams. Integrating social networks in Tunisian medical tourism recommender systems can result in much more accurate recommendations. That is to say, information, interests, and recommendations retrieved from social networks can improve the prediction accuracy. This paper aims to improve traditional recommender systems by incorporating information in social network; including user preferences and influences from social friends. Accordingly, a user interest ontology is developed to make personalized recommendations out of such information. In this paper, we present a semantic social recommender system employing a user interest ontology and a Tunisian Medical Tourism ontology. Our system can improve the quality of recommendation for Tunisian tourism domain. Finally, our social recommendation algorithm is implemented in order to be used in a Tunisia tourism Website to assist users interested in visiting Tunisia for medical purposes.

1. Introduction

The spread of online social networks around the world has given rise to the creation of a new place of interaction and communication amongst people where users can share their knowledge, experiences, and opinions thanks to the online social networks and may have an impact on people’s behavior in terms of traveling and purchasing products. In fact, a social network is a set of people or groups of people whose social relationships (e.g. friendship, collaborative work, information exchange….) make them interact together. Social network analysis, therefore, could be defined as the study of social and economic entities such as people in organizations (actors) and of their interactions and relationships. Those very interactions and relationships could help us know the user preferences an extract his/her interests to improve recommendations. Generally speaking, those networks are represented by the graphs’ theory and the matrix technology. Since social networks are developed by several different kinds of relationships, it remains impossible for the graphs’ edges and the numerical values to explain all the semantic relationships. Having mentioned this problem, we propose, as a solution, a method that makes it possible to represent a social network based on ontology. Using an ontology-based method could allow us to describe all the semantic relationships and the interactions in a social network. The importance of the use of ontologies to represent the types of actors and the relationships in a social network for the purpose of semantically visualizing the databases has been demonstrated by (Correa et al., 2011). Based on the fact that ontologies can be used to semantically describe social relationships and interactions and that social networks can generate and
suggest several varied recommendations with reference to the user needs (Middleton et al., 2004; Middleton et al., 2009), this paper aims to exploit the impact of online social network (Facebook) on the consumers’ purchasing decision process in Tunisian Medical Traveling Website. To achieve this purpose, we have developed a user interest ontology (Frikha et al., 2014) that represents all the preferences and interests of a user.

Based on mainstream literature on recommender systems in social networks (Zhou et al., 2012), social networks already propose to their users recommendations based on their profiles. In a similar fashion, they help users find people for sharing common social activities and preferences. Being aware of the importance of the user and of his/her preferences in recommendations, we study, in this paper, the impact of incorporating semantic user profile (derived from past users’ behaviors and preferences) on the accuracy of a tourism recommender system (Sieg et al., 2010; Su et al., 2011). To obtain a satisfactory recommendation, the characteristics of the recommended activities should match tourists’ preferences. The information about the user is usually captured with three ways: first, it can be explicitly captured by directly asking the user for it such asking the user to fill a questionnaire. Second, the system tries to associate the user with a pre-defined social network that has well-known preferences. Third, new information about the user can be obtained in an implicit way by observing his/her interaction through the social network. Our tourism recommender system use these three types of information about the user. All these information are represented in the user interest ontology that tries to build a user profile using semantic knowledge representations. Accordingly, Artificial Intelligence representation languages and inference tools can lead to an improved quality of the Tunisian medical tourism recommendations. Our primary goal is to strengthen the quality of medical tourism services and contribute to serve a positive image about Tunisia.

The rest of the paper is organized as follows. Section 2 explains a work preview on ontology-based recommendations. Section 3 describes the need of a social network-based recommender system, while Section 4 shows and accounts for the use of a personalized social recommender system for Tunisian Medical Tourism information; a new paradigm of tourism recommender systems developed by using information in social networks. Section 5 presents the need for a semantic user profile in the social network and explains our method for modelling a user interest ontology to personalize social recommender system. In the next Section, we describe the Tunisian Medical Tourism ontology (TMT ontology) that will be used by our tourism recommender system. Section 7 explains the functioning of our TMT e-tourism Website. Finally, section 8 concludes the paper and describes the orientations of our future work.

2. Related Work

2.1. Representation and Management of User Interest

Traditional methods of recommendation rely on keywords to model user’s interests and information requirements (Tan et al., 1998). Due to the vagueness of keywords, the user interest model cannot represent user’s interests and the relationships between different items of the target domain accurately. Besides there are cold-start and data sparsity problems (Sarwar et al., 2001), resulting in poor performance. Many studies have used ontology to design user interest model. Their work has, however, been restricted to using it in traditional recommendation methods.

Pretschner and Gauch (1999) are among the first to make use of ontology to build user model and provide personalized document access. Domain ontologies have been employed to organize documents. Based on a user's surfing history, the user model is built and the personalized document access is performed by referring to the interest degree of each document’s concept. The semantics of concept relations and ontology structure, however, is not taken into consideration during the calculation of the user’s concept interest degree.

The creation of a user model based on ontology to make improvement for information recommendation has also been suggested by (Middleton et al., 2004). They advocate that when a low-level concept in the concept hierarchy is selected by the user, this implies that the user may also be interested in a high-level concept. The major limitation of this method is that the inference algorithm is rather simple, which makes it unable to handle complex ontologies.

Jiang and Tan (2006) suggest the concept of user ontology, and develop a set of statistical methods to relate individual user ontology to domain ontology. When building user interest model, a value is assigned to every concept and
relation. It contains more abundant semantics and more precise interest description. Nevertheless, dealing with complex ontologies requires a longer learning cycle to gain the user ontology.

Su and al. (2011) propose a user interest model based on user interest ontology. In this work, recommendation process is presented by using the ontological user interest model. The biggest shortcoming of this method is that it does not use the advantage of social network in the recommendation process. (Su et al., 2006) use only collaborative recommendation in the process of recommendation, but collaborative recommender systems fail to help in cold-start situations, as they cannot discover similar user behavior because there is not enough previously logged behavior data upon which to base any correlations.

Ferraro and al. (2014) describes an adaptive recommender system which adopts a semantic approach to assist the user both in the travel planning phase and, on-site, during the trip. (Ferraro et al., 2014) propose a general architecture of an adaptive recommender system that uses ontology to assist users in the travelling. In this work, extraction of the user interest is not automatically from the social network, but users are required to answer a set of questions about personal interests and to provide detailed information regarding their requirements.

2.2. Ontology in Tourism Systems

Through ontologies, we can solve a number of problems related to tourism. As ontologies allow the sharing of domain knowledge based on a common vocabulary across heterogeneous platforms, they can also be used to solve interoperability problems (Dell’Erba et al., 2002). In addition, ontology allows for the sharing of common understanding of the structure of information among people and software agents (Noy and McGuinness, 2003); this also can help to standardize business processes and knowledge architectures in tourism. Moreover, ontology serves as a model of knowledge representation out of which knowledge bases, that describes specific situations, can be built (Kathrin et al., 2007). These reasons motivated our decision to develop a tourism-related ontologies recommender system for the medical tourism in Tunisia. This is based on the assumption that an ontology-based framework, enabling the leveraging of factual knowledge about a specific tourism context for recommendations, is likely to enhance the quality and credibility of tourism recommendation services.

Recently, there is a widespread development of tourism ontologies where some of them have reached a considerable level of consolidation, allowing the representation of generic aspects as well as specific sub-domains that describe detailed scenarios. Harmonize (Höpken et al., 2006) was one of the first ontologies that was built to face the interoperability problems of Tourism, focusing on data exchange between organizations. Four main topics of the Tourism domain are outlined: attractions, events, food and drink, and accommodation. Mondeca (Prantner et al., 2007), developed an ontology with around 1000 concepts, most of them were mentioned in the Thesaurus on Tourism and Leisure Activities developed by the World Tourism Organization (WTO).

Another ontology, QALL-ME (Ou et al., 2008), was constructed to establish a shared structure for multimodal and multilingual Tourism question answering. The DERI e-tourism ontology (Hepp et al., 2006) was developed in the OnTour project and focused on three main topics: accommodation, activities and infrastructures. Some classes of this ontology were used as a test-bed for an automatic system of ontology population (Ruiz-Martinez et al., 2011). The Core Domain Ontology for Travel and Tourism (Barta et al., 2009), was another ontology developed in 2009, and it was based on the Harmonize ontology. Its main objective was to define a common ontology for the tourism sector in order to support the interoperability of the agents in low-level operations.

Finally, (Moreno et al., 2012) developed SigTur/E-Destination Web-based system which provides personalized recommendations of touristic activities in the region of Tarragona. The activities are classified and labeled according to a specific ontology to guide the reasoning process. The development of SigTur/E-Destination requires a knowledge management side, represented in a specific domain ontology to provide a classification of the main types of activities and to guide the knowledge-level inference process needed to assess the preferences of the user on each of them.
2.3. Tourism Recommendation Types

Garcia et al. (2011) argued that current Tourism recommenders could suggest either a trip destination or a set of activities to perform once the user arrived at a certain place. These recommender systems face the problems of integrating heterogeneous sources of information (such as Web resources associated to flight and hotel companies) to find a trip that matches the user’s constraints and preferences. In fact, information extraction techniques and semantic technologies such as tailored ontologies are used to parse, interpret and integrate information stored in heterogeneous sources (Ambite et al., 2002; Camacho et al., 2005). Those systems store a set of preferences of the user and this allows them to rate and rank a collection of destinations. They also suggest places to visit in a certain geographical area taking into consideration the user profile by referring to a daily plan (Huang and Bian, 2009).

With reference to the recommendation strategy, recommender systems can be classified in three major types: content-based, collaborative and demographic approaches (Montaner et al., 2003), (Borràs et al., 2014). In collaborative filtering approaches, recommendations are made by relating users with similar preferences and suggesting items they like (Castillo et al., 2008; Lee et al., 2009). Several ratings are, thus, required from the users before the system gives useful recommendations. This is known as the cold-start problem (Basiri et al., 2010). Content-based approaches (Pazzani and Billsus, 2007) recommend items by taking into account the properties of the activities that users have enjoyed previously and only items closely related to those the user liked in the past are recommended. To make this kind of recommendation, one has to build a user profile that shows the degree of interest in each of the different criteria describing an activity. Demographic recommenders classify users according to stereotypes typical classes and base the recommendations on general features related to those classes. In order to avoid the limitations of individual recommendation schemas, hybrid approaches were proposed (Schiaffino and Amandi, 2009; Burke, 2000; Salter and Antonopoulos, 2006; Christakou et al., 2007). These approaches combine some of the above-mentioned types.

Social networks have recently become of particular importance in the tourism industry. When planning their journeys, travelers, in addition to the personal experience and opinion of relatives and friends, refer to social networks during and after their trips, thus generating interest in viewers that can become travelers themselves (White et al., 2007). In our recommender system, we consider an important role to the social network to make recommendation based on the user’s preferences as well as his/her friends’ preferences.

3. Social Network-Based Recommender Systems

Traditional recommender systems suffer from many issues. For example, in order to measure item similarity, content-based methods rely on explicit item descriptions. However, such descriptions may be difficult to obtain for items like ideas or opinions. Collaborative filtering has the data sparsity problem and the cold-start problem (Adomavicius et al., 2005). In contrast to the huge number of items in recommender systems, each user normally only rates a few. Therefore, the user/item rating matrix is typically very sparse. It is difficult for recommender systems to accurately measure user similarities from those limited number of reviews. A related problem is the cold-start problem. Even for a system that is not particularly sparse, when a user initially joins, the system has none or perhaps only a few reviews from this user. Therefore, the system cannot accurately interpret this user’s preference.

As we can see, traditional collaborative filtering and its variations do not utilize the semantic friend relations among users in recommender systems. However, this is essential to the buying decisions of users. The impact that social networks has on product marketing has been recognized by many researchers (Subramani et al., 2003; Yang et al., 2003). When we think of buying a particular and unfamiliar product, we most often tend to seek immediate advice from some of our friends who have come across this product or experienced it. We, similarly, tend to accept and welcome a friend’s recommendation because we trust them.

Integrating social networks in recommender systems can result in more accurate recommendations. That is to say, the information obtained about the users and their friends makes it unnecessary to look for similar users and to measure their rating similarity as the fact that two people are already friends can imply that they have things in common (He et al., 2010). The data sparsity problem can be solved in this case. The cold start issue can also be overcome because even if a user has no interests’ history, recommendations can still be made based on a friend’s preferences.
If we recall the decisions that we make in our daily life, such as finding restaurants, buying a house, and looking for a travel, many of them are actually influenced by some factors. Intuitively, a customer’s buying decision or rating is decided by both his/her own preference for similar items and his/her knowledge about the characteristics of the target item (He et al., 2010).

All of these intuitions and observations motivate us to adopt this approach of social network-based recommender system that can take advantage of information in social networks. Our goal is to represent all these information extracted from social network in a user interest ontology (Frikha et al., 2015) that can help us in the recommendation process. In fact, in order to be able to recommend an object to the user, one must know the interests that represent a part of the user profile.

4. Personalized Recommender System for Tunisian Medical Tourism

The growth of medical tourism in Tunisia emanates from receiving an estimated 12,000 foreigners in 1998 (Chanda, 2001) to an estimated 42,000 in 2003 (Lautier, 2008) while 155,000 foreign patients went to Tunisia in 2013 (Tunisian Ministry of Health). We can say, therefore, that medical tourism is an important economic driver for developed countries as well as developing countries as its contribution to the growth of the economy, services and job creation is widely recognized. Additionally, in a competitive market, providers’ services should be able to offer an affordable cost for medical tourists. This means that cost saving is important to encourage foreign patients to choose Tunisia for their treatment and holidays.

Personalized recommender systems are now playing an important role in providing better tourism experience to many tourists. Our primary objective in this study is to apply a personalized recommender system to Tunisian Medical Tourism. The reason behind this application is to provide tourists with better information searching experiences and to improve the individuation of tourism information.

Recommendations are selected from TMT ontology as the most similar to the user interest. User interest are extracted from the user profile and represented in the user interest ontology. Our semantic social recommender system will calculate similarity between the user interest ontology and the TMT ontology. Then, the system will suggest the most similar items to the user. Figure 1 illustrate an overview of the system.

![Fig. 1. Overview of our semantic social recommender system](image-url)
in the next section). Additionally, the Tunisian Medical Tourism ontology has been created using the Travel\(^1\) ontology as a reference guide to represent medical tourism activities in Tunisia. The decision about which concepts and relationships should be represented in the TMT ontology has been taken by a committee of experts in the Tunisian tourism domain at the TMT Forum\(^2\). The populating of the TMT ontology has been made from Tunisian medical tourism providers’ services. TMT ontology is used to explicitly classify the medical activities in order to recommend among a predefined set of distinctive main concepts which, are going to be used by the intelligent recommender system to perform its reasoning processes. Our semantic social recommender system will search the most adequate medical activities for the user taking in consideration all information about the user represented in the user interest ontology.

4.1. Framework of the Personalized Social Recommender System

The major objective behind applying personalized recommender system to medical tourism information services is to allow tourists to have efficient information searching experiences and to boost the individuation of tourism information. In our personalized social recommender system, the user accesses to our Website and give answers to some questions guide then connects to his/her social network, and then the system extracts the preferences and relations to determine the user’s interest as well as his/her friends’. After that, the system generates specific recommendations by employing a recommendation algorithm which leads to the presentation of the results.

The system collects all the data from the user’s profile and extracts his/her interests. Those user interests are presented in the form of an ontology. Our social recommender system, however, needs the user interest’s ontology to calculate similar items to the user’s preferences in the TMT ontology. These items will be recommended to the user. Figure 2 further illustrates the functioning of our framework.

![Figure 2. Framework of the social tourism recommender system based on user interest ontology](image)

- **Presentation layer**

The purpose behind a presentation layer is to allow the interaction between the user and the system and to present the recommended items to the user. This layer picks the user’s profile when the user answers to some health interest questions and connects to his/her account, gathers data and displays information through the Website. It acts as the interface for the user to access the system. By using Tunisian Medical Tourism Website as interface, it enables the user to visit

---

\(^1\) http://protege.stanford.edu/plugins/owl/owl-library/travel.owl

\(^2\) http://tmt-forum.com
items which are recommended to him/her by the system. The presentation layer mainly consists of Graphical User Interface and the results of the recommendation system.

- **Intelligent analysis layer**

  Intelligent analysis layer is made up of intelligent analysis module and the social recommendation system. Intelligent analysis module is composed of two modules of user’s preferences capture and trust friends’ preferences capture. User preference is collected from user’s profile, and friends’ preference is collected from the user social graph presented by Facebook after the determination of the trust friends. Before capturing the user preferences, we must first analyze the entire user profile, the personalized friends’ list; the News Feed; the relations and the “likes” of the user and their trusted friends. All these preferences will be represented in a user interest ontology (detailed in the next section). Then the social recommendation system generates recommendations with the help of the user interest ontology. By using social recommendation algorithm, our system looks for similar items to this user’s interest ontology in the Tunisian Medical Tourism Ontology. Our intelligent recommendation system picks out information in agreement with the user’s needs from the TMT Ontology and produces recommendation set. In this layer, the system can offer a dynamic guidance for the user’s choice when he/she desires to travel to Tunisia for medical tourism.

- **Ontologies layer**

  Ontologies layer consists of the user interest ontology and the Tunisian Medical Tourism Ontology. TMT ontology contains information about medical activities and centers, thalassotherapy cures and centers, hotels SPA, travels and entertainments of Tunisia. The aim behind these types of information is to ensure satisfaction on the users’ part when searching for information about medical tourism in Tunisia. It is the storehouse for operation data. In other words, ontologies layer stores and manages data that will be used in our social recommender system.

  For each user session, a user interest ontology classes are loaded into memory to enable the recommender system to associate a preference degree to each of the classes with reference to the explicit and implicit information provided by the user. The system will search for similar items to the user preferences in the TMT ontology. These preferences are important to decide which activities to recommend to the user.

  Ontology is defined as an explicit specification of conceptualization (Gruber, 1993). In ontology, the concepts are grouped together in a semantic graph and the ties between these concepts are referred to as semantic relations or specification relations, such as ‘is-a’ relation in taxonomy. In ontology, concepts or terms are grouped into different classes. Different semantic relations can relate those classes and group them together. Besides, each class has several kinds of related properties and values. Furthermore, ontology is characterized by reasoning that is used to analyze the relationships between ontology classes and to reason about their properties and attributes. Our ontologies have been developed with the Protégé editor and represented in the OWL language. Jena is an open source Semantic Web framework for Java used in our system. It provides an API to extract data from and write to RDF graphs. It is used to provide tools to manage these ontologies and to apply inference mechanisms based on specific rules.

  Presentation layer, intelligent analysis layer and ontologies layer are interconnected to accomplish the function of providing personalized recommendation to the user.

4.2. Algorithm of the Social Recommender System

In this section, we show how the semantic information can be used in the semantic social recommender system. Based on our approach, the semantic information is broken down into two major parts: the ontology representation of user interest and the semantic similarity between the user interest ontology and the item description in the TMT ontology.

The performance of the semantic social recommender system are based on knowledge base usually defined as the user interest ontology and on the other hand on social network analysis measures (like trust friends, degree centrality, etc). The limitation of the classical recommender systems approach lies in the tight coupling with the collaborative filtering recommendation. Although there are several graph based recommender systems, these recommender systems have
never employed the social network analysis measures in recommendation algorithms. That is why we suggest involving social network analysis measures in recommendation algorithms. We also propose to involve the user’s semantic preferences in this recommendation algorithm in order to obtain a semantic social recommendation algorithm.

We, then, present our algorithm of semantic social recommendation to reduce the list of Medical Tunisian Tourism items that will be recommended to the user; the most similar items to the user interest ontology will be recommended and rated by the user for the purpose of updating our user interest ontology. The recommendation algorithm is shown in Table 1. We can define the algorithm input as all the items in the TMT ontology, and the algorithm output as a list of most items similar to the user interest ontology.

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>User accesses to the TMT Website and answers their health interest questions</td>
</tr>
<tr>
<td>2.</td>
<td>User connects with his/her Facebook account in our TMT Website</td>
</tr>
<tr>
<td>3.</td>
<td>Picks up user information and friends’ information</td>
</tr>
<tr>
<td>4.</td>
<td>Creates the User Interest Ontology</td>
</tr>
<tr>
<td>5.</td>
<td>Starts the search in our Tunisian Medical Tourism Ontology</td>
</tr>
<tr>
<td>6.</td>
<td>For each item in the TMT ontology (a) Compare the user interest ontology with the description of this item (Similarity) (b) Compare the similarity degree between user interest and item with a threshold. (c) If the similarity degree &gt; threshold i. add item to the proposed list (d) Else Go to (5)</td>
</tr>
<tr>
<td>7.</td>
<td>Sort out the proposed list</td>
</tr>
<tr>
<td>8.</td>
<td>Recommend the most similar item in the list to the user</td>
</tr>
<tr>
<td>9.</td>
<td>Ask the User to rate the recommended item</td>
</tr>
<tr>
<td>10.</td>
<td>If User doesn’t like the recommended item (User Rating ≤ 2) (a) Recommend the next item in the proposed list (b) Go to (8)</td>
</tr>
<tr>
<td>11.</td>
<td>Else Update user interest ontology with the proposed item</td>
</tr>
<tr>
<td>12.</td>
<td>If User wants another recommendation (a) Go to (5)</td>
</tr>
</tbody>
</table>

**Table 1.** Steps of the recommendation algorithm

Our proposed algorithm uses semantic information described in a user interest ontology. This ontology contains social information represented in the social network as heuristics integrated in our recommendation algorithm to find out the item most interesting to this user in the TMT ontology. When a user has access to the Tunisian Medical Tourism Website and connects with his/her account, the system will automatically gather the user information and the friends’ comments and likes to create the user interest ontology. After that, the system starts the search for similar items to this ontology that represent the preference of the user in our TMT ontology. The system will prepare a proposed item list and will sort out items in the list. An item is similar to the user interest ontology if it is superior to a prefixed threshold. The top ranked items will be, then, recommended to the user. User may rate recommended items with an integer value between 1 and 5, where 5 corresponds to the best. If the user does not like the recommended item (User Rating ≤ 2), the system will recommend the next item in the proposed list and will ask him/her to rate it. Furthermore, if the user likes the recommended item, our system will update the user interest ontology and will search again in the TMT ontology taking into consideration this update.
5. User Interest Ontology to Personalize Social Recommender System

The user profile contains information about the user, such as name, age, friends; and about his/her preferences, such as items that the user liked in the past. Indeed, the user profile contains information about the users, their preferences, behavior in the network and interactions with friends. In social networks based recommender systems, the user profile is an excellent source of information to retrieve interests of the user and to determine the trusted friends, so it is very important that this profile be treated appropriately. According to the literature (Susan et al., 2007), the user profile has several types of representations: a set of words, a matrix of weighted keywords, the use of taxonomy (Sulieman et al., 2014) or by an ontological representation of the profile through the use of ontologies. In our work, we will build an ontology to represent user preferences and his/her interests.

Our method for extracting user interests consists of several steps. After giving the user’s answers and obtaining the initial profile, we collect user data from his/her social profile and behaviors in the network. Then, we analyze the user data to extract the interests of the user. Finally, we present all these data in the form of ontology. The ontology’s main concepts are the user’s interests and relationships. The conceptual and semantic relationships between the user’s interests are determined in the data analysis step. Figure 3 further illustrates the functioning of our method to represent the user interest ontology.

Fig. 3. Ontology user interest modelling from a social network
5.1. User Questions Answers

When the user accesses to our TMT Website, the system will propose questions to know about his/her health interests in order to create the initial profile and to take explicit information about his/her health state. The information provided are about the country of origin of the user, the location of the accommodation, the type of accommodation, an initial estimation of the budget, and the travel period. With these answers, our system will know the initial user profile of this user to represent it in the user interest ontology. Then the user will connect to his/her social network (Facebook) and the system collects data from the user profile and completes the construction of user interest ontology from user’s preferences and his/her trust friends’ preferences.

5.2. User Data collection

The user is identified in a social network by his/her traits and behaviors (personal specifications), and his/her interactions with the various system services. In our approach, data generated by the user are used primarily to explore their interests. These data can be of different formats: photos, links pages, texts, etc. In our work, we treat the textual form, which presents the major form of data, to identify user interests. In this context, the content generated by the user can be regarded as explicit and implicit data. First, the explicit data are the data about users which are explicitly declared and which represent user information and preferences. The user information may be either static or dynamic. Static user information are mainly general information such as name, first name, sex, date of birth, etc. while dynamic user information relate to marital status, level of education, place of residence, occupation, email address, phone number, nickname, photo, etc. Then the implicit data are usually discovered through actions provided by the user or by its behavior such as comments, quotes “like”, publications, events, etc. and also deduced from explicit data contained in the ontology.

5.3. Data Analyzing and User Interest Extraction

In this step, we analyze the preferences and behaviors of the user (shares, comments, events, etc...). Thus, in this phase, we try to filter the raw data and display them in a representation in order to further process them. Indeed, these textual data are not structured and written in a formal language, which makes their use very difficult in their raw state.

- User preferences analyzing

To analyze the preferences of the user, which are presented by a set of XML files extracted from the social network. These XML files are obtained with the Facebook API Open Graph that gives us the authorization to accede to the user profile. First, we begin by removing all empty concepts. Then, we construct a matrix of occurrence that can indicate the number of occurrences of each concept. We afterwards, determine the semantic relations between the concepts using “WordNet” which covers most common English words, and a base of terms containing the common concepts that do not belong to “WordNet”. We use a calculating similarity between concepts when the concept does not belong to “WordNet” or to the base of terms.

After extracting the various concepts used by the user, we will try to determine the semantic relationships that exist between these concepts. To do this, we use the language tool “WordNet” to determine if there are relationships between the concepts of the matrix. In many cases, we can find concepts that do not exist in “WordNet”. To remedy this, we use a base of terms that contains commonly used concepts and acronyms in the medical tourism domain. When the concept does not exist in the base of terms, we use a function that allows the calculation of structural similarity to check if the concept is similar to another concept in the matrix or in the base of terms.

Techniques based on strings are not sufficient when the concepts are semantically close and when their names are different. The query language resource such as “WordNet” may indicate that the concepts are similar. In our work, we use the proposed function (Fellah et al., 2008) to calculate the linguistic similarity between two concepts. For the
calculation of this similarity, the $\text{Syn}(c)$ function calculates the Synsets of the concept $C$ in “WordNet”. When $S = \text{Syn}(c1) \cap \text{Syn}(c2)$ is the set of common sense between the to-be-compared $c1$ and $c2$. The cardinality of $S$ is then:

$$\lambda(S) = |\text{Syn}(c1) \cap \text{Syn}(c2)|$$  \hspace{1cm} (1)

When $\min(|\text{Syn}(c1)|, |\text{Syn}(c2)|)$ is the minimum between the cardinalities of the two sets $\text{Syn}(c1)$ and $\text{Syn}(c2)$. The similarity between the two concepts is then $c1$ and $c2$ and is defined as follows:

$$\text{Sim}_{\text{ling}}(c1,c2) = \frac{\lambda(S)}{\min(|\text{Syn}(c1)|, |\text{Syn}(c2)|)}$$  \hspace{1cm} (2)

This measure return 1 if at least $c1$ is the only synonym to $c2$ and $c2$ is the only synonym to $c1$.

We have built a base for the frequent terms taking into consideration the terms that do not exist in “WordNet” along with the abbreviations and special characters of the tourism domain. Sometimes, we find terms that do not exist neither in “WordNet” nor in the base. We, accordingly, use a structural similarity measurement between each of these terms and the other matrix terms.

- User behavior analyzing

The user provides a set of actions that enrich our knowledge about these interests. These actions can be comments (on photos, videos, articles, etc.), publications (such as photo or video sharing), and events of various interests or “I like” signs on the pages or groups. However, we begin by extracting sentences of each document (comments, status, etc.). Then, we apply the same method in the user preference analysis to extract concepts and relations between them. In this step, we use the TF-IDF method (Salton et al., 1975) to measure the weight (density) of each term in a document. The TF-IDF algorithm is based on term weighting and has the advantage of being easily used with statistics and generic techniques. In our approach, to obtain the interests of the user, we remove from the matrix constructed in the previous step the concepts with a lower number of occurrences of a predetermined threshold. For this, we only get the most interesting concepts for the user.

5.4. Determination of Trusted Friends List from the User Social Graph

To determine the list of friends near a user, we begin, first of all, by all the friends who have made interactions or shared information with the user in a very specific period of time. Thus, interactions are of different types. They can be comments or mentions “Like” on objects in the profile (from the user or from friends). Then we calculate the number of occurrences of each friend in these interactions. Finally, we choose friends who have a number of occurrences beyond a predetermined threshold. However, the extraction of the trusted friends’ list of a user is performed to determine the preferences and interests of each friend. Indeed, these interests can be useful to know the interests of our user.

5.5. Concept Weight in the User Interest Ontology

A key factor in this ontology is the representation of semantic relations obtained in the step of determining the types of relationships between concepts. In other words, the ontology of user interests consists of a set of concepts and a set of connections using a logic axioms language. Each concept in our ontology has a degree of interest (weight) to express the importance of this concept for the user. This weight is calculated using the number of occurrences and the last occurrence time of the concept. This weight represents the degree of interest of the concept to the user. As user interests may change over time, we have chosen to weigh the concept according to both the freshness and the frequency in order to unify the recent and persistent interests, accordingly, and to make difference between long-term interests and temporary interests. To achieve this purpose, the time sensitive user profile is taken into consideration (Kacem et al., 2014). Kacem et al. propose the formula $W(C_i)^{t}$ to measure the freshness of a term. They review the notion of term frequency
by adjusting it with a temporal-biased function. The weight $W(C_i|D_c)$ of a concept $C_i$ in the current Date $(D_c)$ is the sum of its time-biased relative frequency defined as follows:

$$W(C_i|D_c) = \sum nTF(C_i|D_c) \cdot K(D_c, D_j)$$  

(3)

Where $nTF$ is the normalized term frequency (Kacem et al., 2014) of the concept $C_i$ at a current date $D_c$. $K(D_c, D_j)$ is the Kernel Gaussian function as a temporal-biased function (Lv et al., 2009; Gerani et al., 2010) between the current date $D_c$ and another date $D_j$.

6. Tunisian Medical Tourism Ontology

Hundreds of SPA (Santa Per Acqua) and thalassotherapy centers are located across Tunisia and this makes Tunisia the second wellness and SPA destination in the world. At present, the SPAs, the Turkish Hammans and more recently, the thalassotherapy centers, are known for healing by water. A successful health cure in Tunisia, therefore, focuses on: the diversity of resorts across the country, a wide selection of thalassotherapy centers, a generous dose of professional care, a squeeze of human kindness, peaceful and cool surroundings, etc.

As travelling to a foreign country to receive treatment should be safe, it is important to reduce the risk to attract medical tourists. Thus, it is suggested to provide the necessary information using informational and social functions of the internet as a means to reduce this risk and make medical tourists feel comfortable when choosing Tunisia. Furthermore, it is important to present successful experience for medical tourists in order to help them evaluate the opportunity of visiting Tunisia (Ben Cheikh et al., 2015).

Medical Tourism is a domain rich of data that are stored in many hundreds of data sources. Many of these sources should be used in concert during the development of tourism information systems. That is why, it is very important to create a referential model that represents the medical tourism in Tunisia (SPA centers and cures, clinics, hotels, thalassotherapy activities, surgical acts for medical tourism, etc.). In our research team, we have created a Tunisian Medical Tourism Ontology that represents all concepts and relations of the medical tourism in Tunisia. The e-tourism ontology provides a way to achieve integration and interoperability through the use of a shared vocabulary and meanings for terms with respect to other terms (Cardoso et al., 2006). Ontology plays an important role in facilitating a semantic integration of heterogeneous data in the Tunisian tourism domain. Indeed, this ontology is extended from the travel ontology. We have made some pruning and filtering operations for the travel ontology, then a refinement activity is added to extend the obtained ontology with particular attributes of Tunisian medical tourism domain. The decision about which concepts and relationships should be represented in the TMT ontology has been taken by a committee of experts in the Tunisian tourism domain at the TMT Forum.

We constructed our TMT ontology using the travel ontology that is represented in the OWL. The TMT ontology was implemented using Protégé Version 3.5 Ontology Editor and represented in the Ontology Web Language (OWL). The OWL is one of the most recent and popular ontology languages. It is the semantic web standard for formally specifying knowledge on the web. OWL is a markup language for sharing and publishing data using ontologies on the Internet. OWL is a vocabulary extension of the RDF (Resource Description Framework) and it is derived from the DAML+OIL Web Ontology Language. It facilitates machine interpretation of Web contents in a better way than XML, RDF, and RDF Schema by using additional vocabulary along with formal semantics (Knublauch et al., 2004).

Figure 4 presents a partial view of the TMT ontology represented in OWL Language description for medial tourism in Tunisia. It gives an idea about the TMT ontology in the form of a hierarchical tree relationship. The figure demonstrates that there are first degree classes defined in this ontology which are as follows: treatment_center, Activity, Accommodation, surgical_Act, Destination and Cure. All these Classes are the subclasses of Thing, Hotel and Clinic are subclasses of Accommodation, which has object’s property hasDestination and hasActivity. Classes Balneotherapy_Center, Spa_Center and Thalassotherapy_Center, on the other hand, are subclasses of class Treatment_center.

---

While classes *Fitness, Swimming, Sauna, Hammam* and *Massage* are subclasses of class *Activity*. Figure 4 further clarifies our *TMT ontology* in OWL description.

**Fig. 4.** Partial view of the TMT ontology
The populating of the TMT ontology has been instantiated from Tunisian medical tourism providers’ services. It is used to explicitly classify the medical activities in order to be able to recommend via a predefined set of concepts, which are going to be used by the intelligent recommender system to perform its reasoning processes.

7. TMT e-Tourism Website

As a demonstration of our semantic social recommender system, an e-tourism Website prototype for the promotion of Tunisian medical tourism was built, which encapsulated two knowledge-based recommender systems (the user interest ontology and the TMT ontology) to help user find the best advice for his/her health and promote medical tourism in Tunisia. At first, the user accesses to the Website. The system will ask him/her to complete a forum and answer some questions to know his/her health situation. The main objective is to obtain as much information about his/her health interest as possible with the smallest number of questions. The second task of the user is to connect to his/her Facebook profile (from our Website) to pick up the user’s preferences and his/her friends’ preferences. Then the system creates the user interest ontology. Afterwards, our semantic social recommender system will search the most adequate activity as well as the tourism provider service by applying the semantic social recommendation algorithm (Table 1) taking in consideration all information about the user represented in the user interest ontology.

The prototype user interface developed for the user recommendation is based on three steps. The first screen in the user interface will ask the user to choose his/her degree of interest for every type of medical tourism (Plastic Esthetic, dental Cure, Heart surgery, Eye surgery, Thalasso Cure, Balneo Cure). Then, we move to the second screen where the user will answer some questions (Country of origin, location of the accommodation, Kind of Accommodation, travel budget and Travel period). Finally, the system asks the user to access to his/her Facebook account to finish the construction of the user interest ontology for this user. When the steps are performed, the user clicks the “Recommend Me” button and the system will search the most adequate medical activities and providers’ services in the TMT ontology for the user taking into consideration all information about this user. Our system shows the most interesting offer for the user and asks him/her to rate it. If the user does not like this offer, the system will propose to the user the second offer in the recommended list given by the recommender system algorithm.

The TMT e-tourism prototype implementation is built based on Java Servlet technology, using the NetBeans integrated development environment (IDE). The Web Graphical User Interface and functionalities were implemented using Java Server Pages and Dream Weaver web design software. The social recommender system was implemented as Enterprise Java Beans components implanted in the web interface, which were mapped using the Protégé ontology Java API to trigger ontology querying and reasoning capabilities.

8. Conclusion And Future Work

In this work, we have presented an overview of recommendation approaches and their utility to define how to extract information from social networks. We present the role of social networks as sources for the development of recommendation systems. In addition, we proposed a method that personalized social network-based recommender system with user interest ontology. Finally, we integrate the user interest ontology in a semantic social recommender system to deal with the lack of semantic information in personalized recommendation system in tourism domain. The updating of our user interest ontology could be developed in future research work to encompass user interest changes. Currently, we are going to test users’ satisfaction about our prototype. In our future work, we are working towards using metrics to evaluate our recommendation algorithm and to compare it against other algorithms. Also, we search to finalize our TMT e-Tourism Website and make it available online to improve the medical tourism sector in our country and to assist users interested in visiting Tunisia.
9. References


Kathrin Prantner, Ying Ding, Michael Luger, Zhixian Yan, Christoph Herzog, 2007. Tourism ontology and semantic management system : state-of-the-arts analysis.


