

Sairmais.com, a new web 2.0 portal for **tourism information** with a **Recommender System**

ARMANDO VIEIRA * [armando.vieira@sairmais.com]

RITA LAVOS ** [rita.lavos@sairmais.com]

Abstract | One of the largest problems in using the Internet is the excess of information and the lack of personalization. These are particularly acute problems in tourism planning and booking. The Web 2.0, or social web, has the capability to solve these difficulties by filtering the information through a network of trustful connections managed by the user. In this work we will present a Web 2.0 portal for tourism, developed upon this personalization concept where the user is the centrepiece. Information is filtered through a social network of recommendations. The recommendation system implemented at sairmais.com is explained.

Keywords | Recommender System, Web 2.0, Personalization, Artificial Intelligence, Social Networks, Folksonomies, Sairmais.com.

Resumo | Um dos maiores problemas da Internet é o excesso de informação e a falta de personalização. Estes problemas são particularmente graves no sector do turismo e das reservas online. A Web 2.0, ou Web social, tem a capacidade de resolver estas dificuldades através da filtragem da informação numa rede de conexões fiáveis centradas no utilizador. Neste trabalho apresentamos um portal Web 2.0 para o turismo desenvolvido com base neste conceito de personalização onde o utilizador desempenha um papel decisivo. Toda a informação é filtrada através de uma rede social de recomendações. Finalmente é explicado o sistema de recomendação implementado no sairmais.com.

Palavras-chave | Sistemas de Recomendação, Web 2.0, Redes Sociais, Personalização, Inteligência Artificial, Folksonomies, Sairmais.com.

* **Doutorado em Física** pela Universidade de Coimbra e Director da Sairmais Lda.

** **Licenciada em Comunicação Social e Educação Multimédia** pela Escola Superior de Educação de Leiria e **Directora de Comunicação** da Sairmais Lda.

1. Introduction

The Internet has reached an unimaginable splendour with everything being virtually available to everyone. However, this enthusiasm easily turns into frustration when one is faced with thousands of web pages to dig in. From the era of information scarceness we have entered into the era of information flooding.

In the tourism framework many questions are still waiting to be solved. How to search successfully for a trip? How to find a low budget trip? How to integrate the hotel with the activities? How to find places that fits my special tastes? What to see, where to eat, schedules, route suggestions? How to share comments, suggestions, opinions? If the user doesn't have a clear target in mind, looking for a dream trip can easily become a cumbersome nightmare, especially in big cities.

Web portals like Booking and Trip Advisor have come out with solutions that are good starting points. However, information is still fragmented, insufficient for an assertive answer and does not guarantee a reliability source. TripWolf or Virtual Tourist goes beyond a simple Web 1.0 approach in creating a Web 2.0 platform for sharing of experiences. However browsing through countless comments and opinions can be as bad as having none.

The field of Artificial Intelligence has developed several important tools that can be useful in this context. Just to cite a few: Advisor Systems, Artificial Neural Networks, Genetic Algorithms or Automated Routing Planning. Unfortunately, most of these data mining tools have not being fully applied in e-tourism.

In this work we will present a web portal, sairmais.com, where, for the first time, a Recommender System (RS) algorithm, first used in Artificial Intelligence, is applied in e-tourism. Recommender Systems can propose the most appropriate results for each user from an essentially unlimited choice of products.

We will show that, although still in its infancy, RS are likely to replace the traditional travel agents.

The RS can provide the visits a reasoned set of options, from destinations, the hotels, museums and places to visit, which fit to your profile and budget.

The RS are very important tools to deal with the modern tourist, with their demands and needs ever more complex each day. The new tourist is not satisfied with any destination, but the best. They are experienced and demand diversity and perfection. They also change rapidly, thus requiring dynamic products with short life cycles. This requires the adaptation and optimization of tourism enterprises.

In this paper, we will present an innovative recommendation system using tags and a evaluations from a social network. The matrix containing elements (tags, users, items, ratings) will be used to obtain detailed recommendations for users. While relatively easy to implement, our approach is very efficient while not computational demanding.

The remaining of the paper is organized as follows. Section 2 the sairmais.com portal is presented with emphasis on its major highlights: the geo-referenced information its personalization and integration. In Section 3 the recommender system is explained in detail and finally in Section 4 we present some conclusions.

2. The Sairmais.com portal

The Sair+ project emerged from an essential problem of the online travel market still to be solved. How to choose a destination, a restaurant, and outdoor activity or event that is totally adjusted to each user's needs?

Internet offers a multitude of possibilities. Googling "hoteis lisboa" wil come out with thousands of possibilities. How to choose the right one? The user will have to browse through suggestions, establishing evaluation parameters, rank the hotels and eventually decide. Quite a cumbersome and time consuming task!

To make things worst, the user has to dig several websites in order to collect information about the surroundings, such as nearby restaurants and activities as, most often than not, the information is scattered. The *sairmais.com* portal has emerged to ease the nightmare that can be the preparing of a family trip. *Sairmais.com* is designed to be the first tourism portal using a recommendation system capable to offer personalized proposals of hotels, restaurants, pubs, experiences and outdoor activities to each user.

At *Sairmais.com* the user is at the center of the action. Users rates, identifies and selects hotels and restaurants preferences. These rates are used to filter information during the search process. In fact, *Sairmais.com* pretends to be nothing more than the automation of the “word of mouth” comments about our trip are exchanged through a network of friends. We want to stand for tourism recommendation as Amazon’s is for book recommendation.

The portal (still on beta stage) is a one stop platform with relevant touristic information

(geo-references) about every interesting spot.

2.1. Web 2.0 – the new era of Internet

Sairmais.com appears in a transition stage on the World Wide Web: from the passive web, where users are mere consumers (web 1.0), to the social web where users create content and are active players (web 2.0) – see Figure 1.

Social Networks are a more natural and meaningful way to gather personal information which can be explored in a recommendation system. They can gather and spread very detailed information fast.

The *Sairmais.com* is a Web 2.0 portal designed to share personalized geo-referenced and integrated information between users. The portal has an inbuilt social network that allows interactivity between the website and users, and within users. Although it doesn’t rely on a wiki format, information can be directly inserted and edited by users.

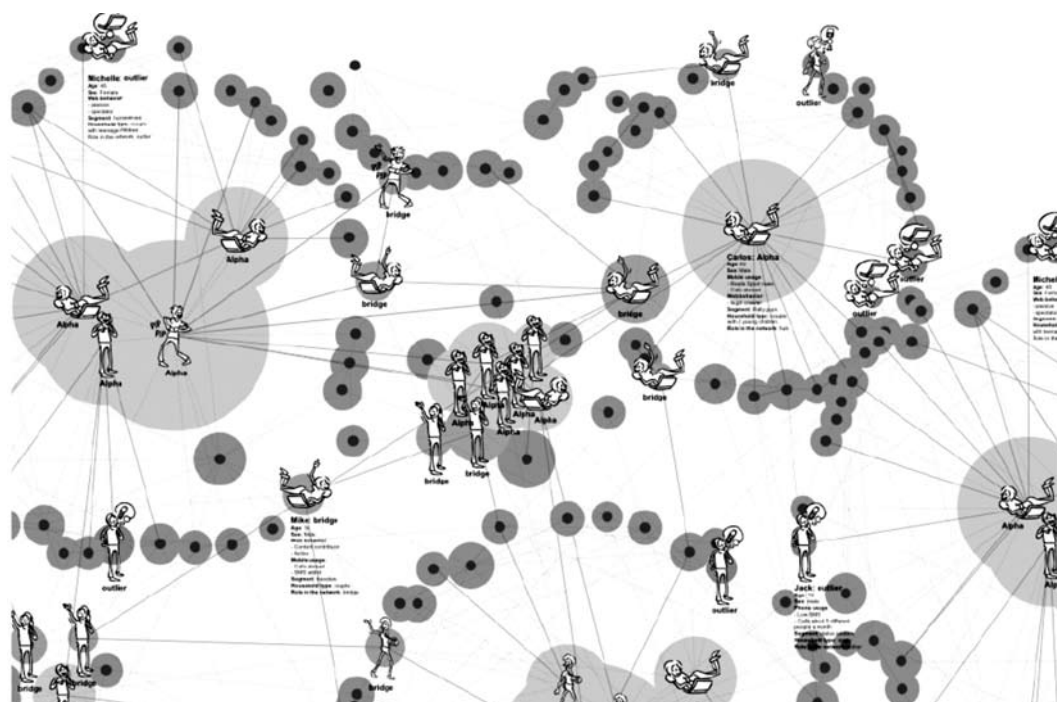


Figure 1 | The new concept of interactive web: the Web 2.0.

2.2. Where does Sairmais.com stands?

This emergent internet paradigm has introduced new concepts in the digital world which previously existed only in the real world. In order to better understand the Sairmais.com project, we will analyze other landmark portals.

Amazon

Founded in 1994 by Jeff Benzons as an American online book store, Amazon has fast become the largest internet warehouse selling almost everything worldwide. One of the appealing innovations of Amazon.com is its powerful recommendation system highlighting the comments, ratings and the purchase users’ history, thus creating a customized recommendation list by crossing each user’s data. See Figure 2.

Booking.com

Booking is the leading hotel booking portal in Europe, with more than 30 million unique visitors per

month and about one million monthly reservations. Booking.com provides geo-referenced information and the ability to make online reservations at thousands of hotels around the world. It does not have a recommendation system, but clients can rate and review hotels.

Facebook

Mark Zuckerberg created the Facebook.com when he was a student at Harvard University. Facebook is the most used social network. Here the user is the focus, like any social network. It also offers users the ability to create applications freely.

TripAdvisor

Founded in February 2000 by Stephen Kaufer, Tripadvisor.com is a travel guide and a search engine of tourist information, where user’s can post their opinion and integrate interactive travel forums. The websites of TripAdvisor has more than 25 million visitors, 10 million members and 20 million comments. However, this website doesn’t

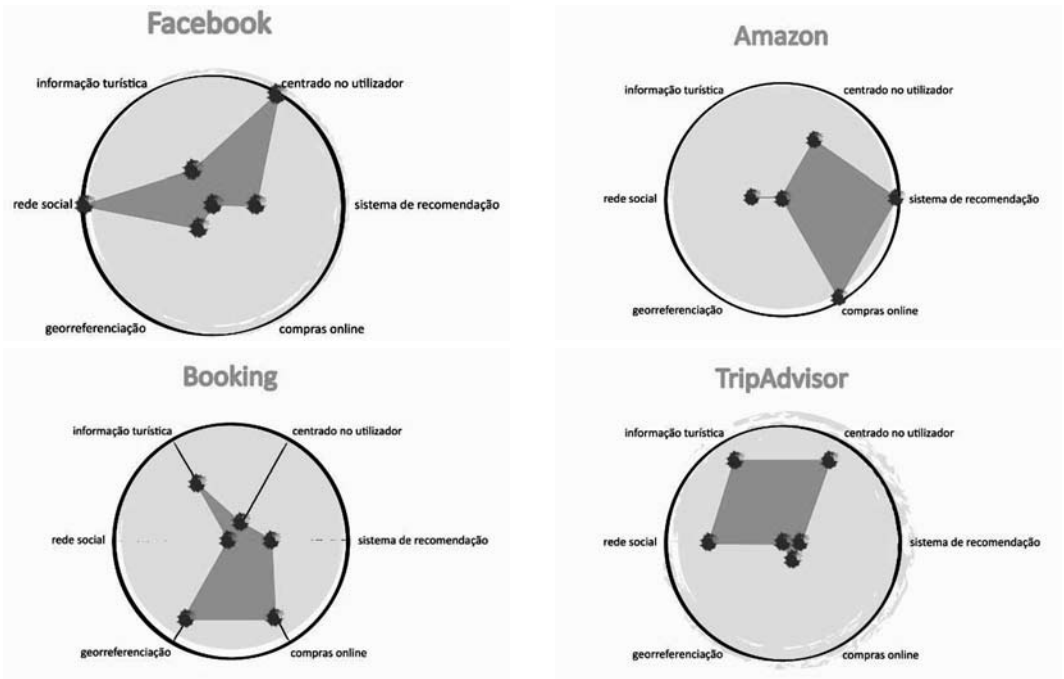


Figure 2 | Some landmark web portals with their strengths and weakness.

highlight the recommendation systems and the geo-referenciation, a key subject in a travel site.

Sairmais.com brings together all key features of the most visited sites: The social network and Facebook's user centralizing value, the recommendation system and online shopping Amazon integrating and linking these features with the tourist information and georeferenciation – see Figure 3. Our is to bring value to national tourism, by integrating and linking technology, previously scattered, in a single platform.



Figure 3 | Sairmais.com targets a reference in the national online tourism market.

2.3. Tools

The project's aim is to fill the gap between hotels and their clients, creating a user social network. It will be a network within the network, where users can share information about places they liked or disliked, and in a self-organized way, they can have access to a personalized offer.

Personalization

Each internet user is unique, with different interests and he looks for different things on the Internet. An online platform offering the 'information' users precisely want to find was something considered a science fiction fact, only possible to reach by Artificial Intelligence.

The concept Info Personalizer Filter (IPF) was developed to select all the necessary information using the parameters above:

- Features and idiosyncrasies each user sets filling out his profile on the website.
- The social network user builds and the friends shared in the platform.
- The historical information user builds while his network grows.

Geo-referenciation

During many years the maps were a travel sign. How could a tourist start a journey without a map? With a map, the tourist could situate himself; he could know where was his hotel, the tourism office, the cinemas, the monuments, the beautiful sights and he could choose the shortest way to reach these places.

Nowadays, maps are overpast. The development of the GPS (Global Positioning System) technology provided the positioning, navigation and timing services to worldwide users, anytime, anywhere on the Earth.

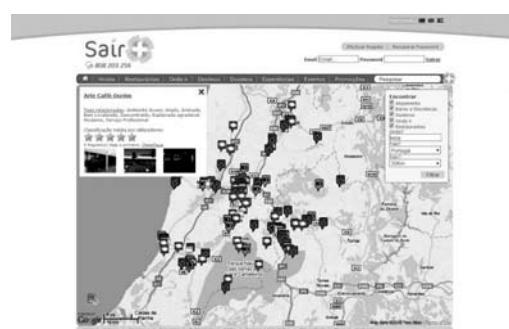


Figure 4 | A snapshot of a local map research in Sairmais.com. The user can search by the name of a city or location, define the radius longitude and choose the type of information he wants to find (hotels, pubs and discos, destinations).

Having the technology to support Google Maps, Sairmais.com will be located geographically on the map tourist destinations of interest to individual users (hotels, restaurants, cinemas, bars, nightclubs, tourist attractions, etc.).

Interactivity

Sairmais.com also offers users information. Each user can register his opinions and feelings on the website, which will be available to other users. That's the network paradigm based on the free information share.

Sairmais.com is developing a framework to transfer the social relationships, the feelings and the information sharing to digital web, applying them to the tourism context.



Figure 5 | Sairmais.com social network allows the sharing of friends, opinions and reviews of users and it has a critical role in the recommendation system.

Integration

The information in Sairmais.com is all integrated and related in a links web. For instance, clicking on "Dossiers Restaurantes", the user is redirected to this page content and he can access other related "Dossiers", like "Rest. Românticos". Clicking in this last one, he can access the content and another related content ("Restaurantes de Autor em Lisboa" or "Restaurantes Japoneses"). The social network will provide comments on these contents, reviews and classifications.

On the next stage of this project, Sairmais.com is predicting to create personalized roadmaps, using hotels, restaurants and other business associated collaboration.

2.4. The product

Tourism is one of the largest industries worldwide. Online reservations have had a huge growth and

now represent almost one billion Euros, only in the Iberian Peninsula. However this large online reservation centers still work in a model based on "books": series of products displayed by category, price and region.

Though, web technology has evolved tremendously in recent years, as the demands of consumers, especially younger ones. Today's consumers want to express their opinions. They want interactivity, speed, personalization.

Sairmais.com is aware about the power of networks, and will match strategic partnerships with associations promoting tourism, business organizations, and other entities related to tourism.

Sairmais.com to hotels

We're offering hotels facilities and competitive rates using our portal:

- Access to backoffice where each unit can edit and update their information content continuously and make it available to all users;
- Personalized feedback from all users on the information available on Sairmais.com;
- Hotels promotion areas;
- Sales team keeping up with Hotels during the entire contract process;
- More competitive prices.

Sairmais.com to users

The Internet is built to users and by users. That's why they are on focus. At first, Sairmais.com is a tourist information website. Putting to these two features, we propose the customization of information. Entering the website, user will have the most varied tourist information.

Sairmais.com to restaurants

Sairmais.com wants to preserve Portuguese gastronomy which is a rich national heritage. Therefore we didn't forget restaurants:

- Access to backoffice where each unit can edit and update their information content continuously and make it available to all users;

- Personalized feedback from all users on the information available on Sairmais.com;
- Promotion areas.

2.5. The future

If this is a website made to users and by users, how can Sairmais.com reach them? Sairmais.com knows that the internet community is quite extensive, composed of many smaller communities, who share common interests. To grow as a tourist information sharing community, Sairmais.com will be known by the most ancient communication process in history – “the word of mouth” – moving on the existent social networks, disseminating directly to users by launching challenges and inviting users to share a little of themselves in our website, taking benefit from our advantages.

At the moment, the phenomenon is Facebook, one of several existing social networks on the Internet. According to Nielson Online, Facebook was the most visited site in the United States on April, 2009. This analysis confirmed 700% growth comparing to April 2008. This year, Facebook dethroned MySpace to second place ranking. During the period under review the site has absorbed 13.9 billion minutes of Internet users. The platform has become so popular that several world leaders, as well as the head of the Catholic Church, Pope Ratzinger, have a profile on Facebook.

The potential of this social network for tourism is invaluable. We highlight two applications: Trazzler and Backpacker.

The first one (Trazzler) is a tool that helps the traveler to answer the question “Where do I go?”. Entering your travel preferences, the Trazzler identifies the traveler’s personality and gives user personalized advice. By this application user can know his friends and other users of Facebook destination preferences. User can send travel tips to friends and family, discover new sites and fantasize about his dream holidays.

Backpacker is an application that allows facebookers access to a full guide to cheap accommodation (hostels). Once user clicks a particular country or destination, he finds a list of cities and in each city a detailed list of lodgings and places to visit. This application also allows a reservation service, not yet operational, but with potential to become a great tool for travelers using Facebook.

Many available applications and network flexibility to implement marketing strategies and promotion make from Facebook a rich and varied market, impossible to ignore.

3. The Recommender System

Recommender Systems (RS) are algorithms designed to predict the rating of items or to suggest a list of items that are unknown to the user.

First generations of Recommender Systems (RS) rely on a behavioural pattern and tastes of each user to build a profile specific. In this so called profile-matching RS, where the profile is matched to each product. However, leading the user through a series of questions in a sort of self-assessment module is very cumbersome and time consuming.

The lack of purchase information, infrequent use, and the pronounced variety-seeking tendencies of users are serious problems for tourism adviser systems. Modern RS use social information: the relationships we establish with other people and how we relate with them: the social network.

One of the most used RS are collaborative tagging systems, also known as folksonomies, which are web-based systems that allow users to upload their resources and to label them with arbitrary words, the so-called tags. These systems are becoming very common among web users in popular web services such as Flickr, del.icio.us, Last.fm or Gmail. These sites give the possibility for users to tag or label an item of interest.

In general, tagging is associated to the Web 2.0 and is becoming the new trend enabling people to easily add metadata to content. Hence, these additional metadata can be used to improve search mechanisms, better structure the data for browsing or provide personalized recommendations fitting the user interests.

3.1. User and item based Recommender Systems

Formally RS work in a two dimensional array (User, Item) that are connected through some criteria matching. Two typical recommendation tasks are normally considered by a RS:

- predict the rate of an item unknown to the user and
- predict the set of items that the user may be interested in visiting or buying.

Most recommender systems derive recommendations based on opinions from people who have alike tastes. This is done by creating a neighbourhood for each user containing classifications from other users with similar tastes. The prevalent method to perform such selection is Collaborative Filtering (CF).

The idea is basically to reach the nearest neighbour: given some users profiles, it predicts whether the user may be interested in a certain item based on other users preferences. There are two types of RS based on CF: *user-based* and *item-based*.

In *user-based*, recommendations are generated by considering solely the rating of users on the items computing pair-wise similarities between users. The similarity between two users u and v , $S_{u,v}$, is defined as:

$$S_{u,v} = \frac{\langle R_u, R_v \rangle}{\|R_u\| \|R_v\|}$$

where R is the array containing the recommendations of each user. The recommendation for user u of the

item i , R_{ui} , is evaluated computing a weighted sum over of the ratings over the neighbour users:

$$R_{u,i} = \frac{\sum_{v=1,N} w(u,v)(R(v,i) - \bar{R}_v)}{\sum_{v=1,N} w(u,v)}$$

where the matrix $w(u,v)$ contains the similarities between user u and user v . Note that the importance of each recommendation is weighted by the users similarity.

The *item-based* RS is similar to the user-based RS the only difference being that the similarity is measured between pairs of items instead of pairs of users. In this case the prediction is based on computing the similarities over similar items. This approach is more time consuming, specially for big databases containing hundreds of thousands of items.

3.2. Collaborative tag recommendations

Content information used in attribute-aware RS algorithms is typically attached to the items and is usually provided by domain experts. Therefore, an item always has the same attributes among all users. On the other hand, tags are provided by various users. Thus, tags are not only associated to the items but also to the users. Although attributes and tags are both metadata and could act as additional background knowledge to improve RS algorithms, they should be handled differently.

Despite the considerable amount of research done in attribute-aware RS algorithms, the specific problem of integrating tags to RS algorithms is still in its infancy, particularly in tourism business. Most existing works on RS with tags are limited to recommending tags, i.e. assisting users for annotation purposes, while using tags as supplementary source for recommending items has never been investigated.

As collaborative tagging is getting more widely used, this information could also be employed as background knowledge in RS. At Sairmais.com we are implementing a collaborative tagging system. This system has the advantage of enabling a semantic reach structure on the data. Furthermore, tags are more reliable and give extra diversity that a simple rating system cannot afford. To our knowledge, this is the first tourism information portal that will have implemented this advanced recommendation system.

Our generic mechanism allows tags to be integrated to standard CF algorithms such as user- and item-based algorithms. However, unlike attributes, which only have a two-dimensional relation $\langle \text{item}, \text{attribute} \rangle$, tags hold a three-dimensional relation $\langle \text{user}, \text{item}, \text{tag} \rangle$.

We cope with this three dimensionalities by projecting it as three two-dimensional problem, $\langle \text{user}, \text{tag} \rangle$ and $\langle \text{item}, \text{tag} \rangle$ and $\langle \text{user}, \text{item} \rangle$. This can be done by augmenting the standard user-item matrix horizontally and vertically with user and item tags correspondingly. User tags, are tags that user u , uses to tag items and are viewed as items in the user-item matrix. Item tags, are tags that describe an item, i , by users and play the role of users in the user-item matrix (See Figure 1). Furthermore, instead of viewing each single tag as user or item, clustering methods can be applied to the tags such that similar tags are grouped together.

Tag information has a different nature than attribute information. Attributes are attached to items that appear the same to all users. Tags of an item are descriptions of the item. They are not only attached to the item but depends on the user's preference. They are "local" descriptions of an item that might change from users to users. This suggests that a RS algorithm that is able to capture both user's and item's aspect of tags would be superior to other recommender systems.

Most systems that use collaborative tagging do not contain rating information. For example, Last.

fm, a popular internet radio and music community website, allows user to tag the music but it does not support users to provide explicit ratings of the music. We will consider both rating and tagging.

Our recommender system is designed not only to predict ratings of new items, restaurants or hotels, and also to suggest new items to the user. We use a fusion algorithm that tackles the predicting item problem that takes tags into account. Fusion for the predicting item problem is done by combining the predictions of user and item-based. Predictions are computed based on recommendations of a list of items that is ranked by decreasing frequency of occurrence in the ratings of his/her neighbours.

3.3. Social networks

Social Networks (SN) are a more natural and meaningful way to gather personal information which can be explored in a recommendation system. Facebook, Myspace and Linkedin have become the dominant platforms where users share about everything about themselves. These spaces have important information that can be used to filter and rank items on our database and recommend properly places to visit or restaurant to have a romantic dinner. Our recommendation system will take advantage of these social relations.

Sairmais.com has an inbuilt SN where users can connect or just follow other users. This SN will be used to circumvent on the major problems in recommender system: the sparseness of the rating matrix. This difficulty is very well known: most users rate only a very limited set of items and the vast majority of items do not receive enough recommendations. This is the so called, long tail effect. With the Social Network, users are naturally clustered and the number of useful recommendations is effectively increased. This approach is designated Social Collaborative Filtering algorithm (SCF).

3.4. Recommendations in action

Sairmais.com has two types of users: the normal user and the tourist ambassador. The last type is constituted by users with a deep knowledge of a specific city or region. Other users have the option either to become friends (a bidirectional relationship) or just follow (a unidirectional relationship) this ambassador.

The user has three possibilities to evaluate an item:

- rate it on a scale from 0 to 5,
- tag it, using a pre-defined set of tags, and
- Comment it.

Only information about rates and tags will be used to create recommendations. Items can be ranked based on different criteria:

- Georeferenced search based on distance to a specific point in the map;
- A search based on tag specification of the items;
- A profile matching between user preferences and tags associated with the items;
- A ranking based on our Collaborative Filtering (CF) algorithm and
- Ranking using the Social Network Social Collaborative Filtering (SCF).

These filtering criteria can be combined to obtain an optimal selection of more fine tuned solution. Please note that some of these recommendations options are not still implemented in our portal.

4. Conclusions

In this paper we have described the Sairmais.com portal and some of its most innovative aspects. Sairmais.com is intended to be a breakthrough on

tourism information: it has a Web 2.0 implementation, it is completely centred on the user, all information is integrated and easy to navigate and, above all, it will have a powerful recommendation system.

Sairmais.com is intended to be a major step towards a fully functional web semantic portal for touristic information and automatic Dynamic Packages Management: see Xu and, for an implementation see www.yourtoure.com.

Finally we have described the tag aware recommendation system that will be used at sairmais.com and show its flexibility for rating items and users.

References

- Adomavicius, G., Sankaranarayanan, R., Sen, S., Tuzhilin, A., 2005, Incorporating contextual information in recommender systems using a multidimensional approach in *ACM Trans. Inf. Syst.*, Vol.23(1), pp. 103-145.
- Basilico J., Hofmann, T., 2004, Unifying collaborative and content-based filtering, in *Proceedings of the twenty-first international conference on Machine learning*, pp. 65-72, New York, NY, USA, ACM Press.
- Basu C., Hirsh H., Cohen W., 1998, Recommendation as classification: using social and content-based information in recommendation, in *AAAI '98: Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence*, pp. 714-720, Menlo Park, CA, USA, 1998, American Association for Artificial Intelligence.
- Deshpande M., Karypis, G., 2004, Item-based top-n recommendation algorithms, in *ACM Transactions on Information Systems*, Springer-Verlag, 22/1, 2004.
- Jaeschke R., Marinho L., Hotho A., Schmidt-Thieme L., Stumme, G., 2007, Tag recommendations in folksonomies, in *Proceedings of the 11th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD)*, Warsaw, Poland (to appear), 2007.
- Kabbaj, D., Mohamed, Y., 2003, June, Strategic and Policy Prospects for Semantic Web Services Adoption in US Online Travel Industry, Retrieved September 12, 2006 from the MIT Center for Digital Business.
- Mishne, G., 2006, Autotag: a collaborative approach to automated tag assignment for weblog posts, in *WWW '06: Proceedings of the 15th international conference on World Wide Web*, pp. 953-954, New York, NY, USA, ACM Press.
- Xu, Z., Fu, Y., Mao, J., Su, D., 2006, Towards the semantic web: Collaborative tag suggestions, in *Proceedings of the Collaborative Web Tagging Workshop at the WWW*, Edinburgh, Scotland.