IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Selection and ranking of activities in the social web

This is the author's manuscript

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/152899 since 2016-06-30T17:55:49Z

Publisher:

attuale: Springer Netherlands Dordrecht ; Tokyo ; Lancaster ; Boston : Reidel Kluwer Academic

Published version:

DOI:10.1007/978-3-319-03524-6_32

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

12 October 2024

IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

This is the author's final version of the contribution published as:

Ilaria Lombardi; Silvia Likavec; Claudia Picardi; Elisa Chiabrando. Selection and ranking of activities in the social web, in: Proceedings of AI*IA 2013: Advances in Artificial Intelligence - XIIIth International Conference of the Italian Association for Artificial Intelligence, attuale: Springer Netherlands Dordrecht ; Tokyo ; Lancaster ; Boston : Reidel Kluwer Academic Publishers:Journals Department, PO Box 322, 3300 AH Dordrecht Netherlands:011 31 78 6576050, EMAIL: frontoffice@wkap.nl, kluweronline@wkap.nl, INTERNET: http://www.kluwerlaw.com, Fax: 011 31 78 6576254, 2013, 9783319035239, pp: 372-284.

The publisher's version is available at:
http://link.springer.com/content/pdf/10.1007/978-3-319-03524-6_32

When citing, please refer to the published version.

Link to this full text:
http://hdl.handle.net/2318/152899

This full text was downloaded from iris - AperTO: https://iris.unito.it/

iris - AperTO

University of Turin's Institutional Research Information System and Open Access Institutional Repository

Selection and Ranking of Activities in the Social Web

Ilaria Lombardi, Silvia Likavec, Claudia Picardi, and Elisa Chiabrando

Università di Torino, Dipartimento di Informatica, Torino, Italy

Abstract. The paper presents a framework for the *selection* and *ranking* of *activities* proposed to users in a social networking service, as for example Facebook's "social events". Our proposal takes into account the peculiarities of this application area, considering a variety of factors including the spatiotemporal relations between activities and between activities and users, and the social opinions weighted by the user's interest in the specific themes and type of an activity.
To this aim, we describe a semantic model for the involved entities, and especially for the available activities. We describe a process that can be easily configured according to user preferences, extended to take into account additional ranking factors, and adapted to include existing recommendation/ranking strategies. Finally, we specifically introduce ranking factors that can be used within the specific model and process we describe. The results of a preliminary evaluation show the effectiveness of our approach.

1 Introduction

In this paper we tackle the problem of selecting and ranking real-life activities proposed to people within a social networking service. By "activity" we intend anything that: (a) has a participatory aspect, i.e. it is offered and people can decide whether to join or not, (b) requires the physical presence of people in the location where the activity takes place, (c) is offered within a given time interval and requires a certain amount of time to be completed. For example, Facebook[1] "events" can be regarded as activities according to this definition. Activities, as "one-and-only items" [4] have some peculiarities:

- Activities have a specific start and end in time; once an activity is finished there is no need to recommend it. Therefore, the opinions of users on past activities cannot be used to recommend them. On the other hand, the recommendations are needed before the users had a chance of participating in the activities.
- Joining an activity is time consuming: we cannot expect a user to join two activities that occur simultaneously.

Many standard ranking strategies used in social networks and mobile applications can be applied to activities, e.g. content-based recommendation [10], where the main topics or themes of activities can be compared to user's interests or to her friends' interests, and context-aware recommendation [1], where the activities can be ranked by proximity, assuming that the user can be geo-localized, or by time of day depending on what is available at what time. Still, many recommender systems fail to combine the basic information about the activity with the additional factors which could influence

[1] http://www.facebook.com

the recommendation process. This is where the motivation for our work comes from. We consider the *themes* of the activity (i.e. the topics the activity is related to, for example "wine" or "english literature") and the activity type (for example "movie" or "concert"), but also the spatial-temporal aspect of the activity, as well as other additional factors which can influence users' decisions to participate. Our goal is to reduce the number of activities shown to the user, and to rank them in a way that is meaningful with respect to the user's context and objective.

To this aim we propose a framework for activity recommendation that takes into account these peculiarities and is composed of (i) a data model giving a semantic description of the actors and their data and (ii) a process for selection and ranking that considers several ranking factors and can consequently be implemented on top of existing ranking strategies. The process we describe can be adapted to the operating context and is flexible to user configuration. We discuss the general framework but we sometimes mention, as examples, the specific configuration choices we selected in our system.

We present the results of a user study, in which our system was tested, showing that:

R1 additional knowledge on activities, besides their themes and type, affects users' choices, so that

R2 recommender systems need to incorporate such knowledge in order to improve the accuracy of their recommendations.

The paper is organized as follows: we describe our data model in Section 2 and our process for selection and ranking in Section 3. Then, we present the results of our experiments in Section 4 and position our work within pertinent research in Section 5. Section 6 concludes the paper.

2 Data Model

In our framework we distinguish three types of entities:

- *spatial-temporal objects* or STOB's for short, the social events which are to be recommended to users;
- *users* which are the entities that use the system and receive the spatial-temporal recommendations or STOB's;
- *activities* that are STOB's selected by the user and inserted in her calendar.

Notice that selection and ranking of activities is strongly influenced by the user's goal when requesting a recommendation. We can distinguish two contexts of usage:

- **daily life context**: the selection and ranking process is aimed at "here and now". In this case the selection and ranking is **proactive**: whenever the user connects to the system, she is presented with a properly ranked selection of activities.
- **journey context**: the user makes an explicit query specifying the spatiotemporal constraints she is interested in (for example, the user is planning a trip somewhere). In this case the selection and ranking process happens **on demand**.

Next we describe STOBs, users and activities and their formal representation.

2.1 STOB

The **STOB's** are stored in the STOB *Ontology*, that describes the structure of a social event and all its instances. STOB *Ontology* is connected to other ontologies (Domain Products Ontologies, *Time Ontology* and STOB *Type Ontology*). To represent the STOB class we used the Web Ontology Language (OWL 2[2]).

The structure of a STOB s can be represented as follows:

$$\mathsf{s} = \langle Id(\mathsf{s}), \mathsf{T}(\mathsf{s}), \mathsf{RQT}(\mathsf{s}), \mathsf{S}_t(\mathsf{s}), \mathsf{TYP}(\mathsf{s}), \mathsf{THM}(\mathsf{s}), \mathsf{RAT}(\mathsf{s}) \rangle$$

where:

- $Id(\mathsf{s})$ is an **identifier** which univocally identifies s in the system;
- $\mathsf{T}(\mathsf{s})$, called **temporal description**, defines the temporal existence of the STOB: $\mathsf{T}(\mathsf{s}) = \mathsf{I}_\mathsf{s}$, where I_s is a time interval in which the given STOB is happening; using the *Time Ontology* it is possible to define the *start time* and the *end time* with a specific moment or an interval, and the *duration* of a STOB with a temporal length;
- $\mathsf{RQT}(\mathsf{s}) = t_{rq}$ is the **required time** to participate at STOB s - notice that in general $\mathsf{RQT}(\mathsf{s}) \leq \mathsf{T}(\mathsf{s})$, for example an art exhibit may be available from 10am to 5pm, but require only an hour to visit;
- $\mathsf{S}_t(\mathsf{s})$, called **spatial description**, associates to a STOB s a spatial position in a certain time instance. It is a circular area characterized by its *center* $c(\mathsf{s})$, given with its latitude and longitude coordinates, and its *radius* $r(\mathsf{s})$;
- $\mathsf{TYP}(\mathsf{s}) = p$ is the **type** of STOB s (concert, seminar, course, birthday, etc.). The available values are described in the STOB *Type Ontology*;
- $\mathsf{THM}(\mathsf{s}) = \{\mathsf{th}_1, \ldots, \mathsf{th}_q\}$ are the **themes** that describe more closely the STOB s; they will be matched up with the user model to calculate the similarities between users and STOB's in a classic recommendation system. We rely upon an*ontology-based user model* [17] in which user interests are recorded for each class in the domain ontology. Each user profile is an instance of the domain ontology where every domain entity e has an interest value $i(e)$ associated to it. The themes of the STOB's are a subset of the user model categories, hence the user's interest for a STOB can be computed based on her interest for the themes;
- $\mathsf{RAT}(\mathsf{s}) = \{\mathsf{r}_1, \ldots, \mathsf{r}_m\}$, $\mathsf{r}i \in [0, 10]$, is the set of all **ratings** of STOB s. They allow users to express how much they consider a STOB interesting, before it happens.

2.2 User

The **user** is defined in the knowledge base by generic user information (name, surname, gender, age, etc.), some particular spatial-temporal information (spatial-position and area of interest), her social network and a set of activities that make up her *calendar*.

We can denote a user with u and characterize it by:

$$\mathsf{u} = \langle Id(\mathsf{u}), \mathsf{T}(\mathsf{u}), \mathsf{S}_t(\mathsf{u}), \mathsf{SN}(\mathsf{u}), \mathsf{UM}(\mathsf{u}), \mathsf{CAL}(\mathsf{u}) \rangle$$

where:

- $Id(\mathsf{u})$ is an **identifier** which univocally identifies u in the system;
- $\mathsf{T}(\mathsf{u})$, called **temporal description**, that defines t_u as the point in time in which u finds herself and I_u as the time interval for which user u wants to receive the recommendations, obtaining $\mathsf{T}(\mathsf{u}) = \langle t_\mathsf{u}, \mathsf{I}_\mathsf{u} \rangle$;

[2] http://www.w3.org/TR/owl-overview/

- $\mathsf{S}(\mathsf{u})$, called **spatial description**, which associates to user u a spatial position in a certain time instance. As for the STOBs, user's spatial description is a circular area characterized by its *center* $c(\mathsf{u})$, given with its latitude and longitude coordinates, and *radius* $r(\mathsf{u})$. But, differently from STOB's, for users it is not always possible to define a fixed center that describes where the "user position" indeed is. Therefore, the area in which the user wants to obtain recommendations can vary, depending on user's preferences. Also, if the user is geo-positioned, $r(\mathsf{u})$ may be very small. If, however, the user is not geo-positioned, for example we know she is in "San Francisco" and nothing more, than $c(\mathsf{u})$ and $r(\mathsf{u})$ are estimated to be the center and radius of San Francisco. Moreover, a user has a propensity to move, expressed as a distance $r_P(\mathsf{u})$ she is willing to cover to reach an event. This is a value that can be frequently changed by the user herself. In the computation of reachability we actually consider $r(\mathsf{u}) + r_P(\mathsf{u})$.
- $\mathsf{SN}(\mathsf{u})$ is the social network of the user u;
- $\mathsf{UM}(\mathsf{u}) = \{\langle \mathsf{e}_1, i(\mathsf{e}_1)\rangle, \ldots, \langle \mathsf{e}_n, i(\mathsf{e}_i)\rangle\}$ which represents the user model defining a set of entities in which the user u is interested in, together with the interest values $i(\mathsf{e}_i) \in [0,1], i = 1, \ldots, n$ for each e_i. The entities in $\mathsf{UM}(\mathsf{u})$ can be classes or instances of the domain ontology[3];
- $\mathsf{CAL}(\mathsf{u}) = \{(\mathsf{s}_1, p_1, t_1), \ldots, (\mathsf{s}_k, p_k, t_k)\}$, is the *calendar*, which defines a set of activities in which the user u wants to participate, together with the specification if the participation is public or not and the time of participation.

2.3 Activities

We assume that the user can express the intention to participate in a STOB by adding it to her calendar. When a STOB is added to the calendar, it is instantiated with respect to the specific user (so that she may for example change the start/end time) and the STOB becomes an activity. If, for example, an activity is offered regularly (e.g. every Friday, 8-10pm), the user may want to specify she wishes to attend on Friday the 26th.

We distinguish the following activities in the calendar:

- *joined activities*: the user has decided to join these activities; a significant overlapping of a new activity with a joined activity considerably reduces its feasibility;
- *prospective activities*: the user is considering to join these activities but has not decided yet whether to participate. An overlap of a new activity with a prospective activity is shown but has a minor impact, or no impact at all, on its feasibility.

If we denote an activity with a, the basic structure of a can be represented as:

$$\mathsf{a} = \langle Id(\mathsf{a}), \mathsf{T}(\mathsf{a}), \mathsf{STOB}_{id}(\mathsf{a}), \mathsf{USER}_{id}(\mathsf{a}), \mathsf{PAR}(\mathsf{a})\rangle.$$

where:

- $Id(\mathsf{a})$ is an **identifier** which univocally identifies a in the system;
- $\mathsf{T}(\mathsf{a})$, called **temporal description**, defines the temporal interval when the user wishes to attend the event;
- $\mathsf{STOB}_{id}(\mathsf{a})$ is the **STOB id** which had originated the activity;
- $\mathsf{USER}_{id}(\mathsf{a})$ is the **user id** that inserted this activity in her calendar;
- $\mathsf{PAR}(\mathsf{a}) = \{0, 1\}$ defines the **type of participation**, 1 for joined, 0 for prospective.

[3] How interest values are calculated is out of the scope of this paper, see for example [17].

3 Reasoning

The reasoning process we propose has two separate phases: selection and ranking. In the selection phase the system discards the STOB's that are not at all interesting for the user, or that do not satisfy her spatiotemporal constraints, creating a *pool* of potentially interesting STOB's. Selection is described in Subsection 3.1. Ranking sorts activities in the pool, by taking into account several factors, detailed in Subsection 3.2.

Selection and ranking are executed asynchronously, communicating by means of data repositories. In particular, when the process is executed in **proactive** mode, as in the *daily life* context, selection phase is performed on a cyclical basis on all the valid (i.e. non expired) STOB's within the system. The ranking phase takes place when the user connects to the system to view its proposals, and it is refreshed whenever the user tunes her settings. When the process is executed **on demand**, as in the *journey* context, selection runs when the user queries the system to plan for her journey, and is immediately followed by the ranking phase. The selection phase however is not run again unless the user changes relevant spatiotemporal constraints, while the ranking phase is again refreshed whenever the user edits her settings.

Figure 1 describes the whole recommendation process, showing how it interleaves with user interaction. In fact, although we neither explicitly address the problem of how the recommendation results are displayed to the user, nor we investigate interaction modalities that allow her to fine tune the recommendation, we believe that a successful recommendation scheme must provide enough information to the UI Module so that it can handle these tasks in a flexible and affordable way.

As we already said, selection creates a *working pool of STOB's* that will be the input for the rest of the process. However, we assume that the user may want either to manually add STOB's she likes to this pool, to be later reminded of them, or to remove once and for all those she discards, in order not to see them anymore. Since we do not want a re-run of the selection process to overwrite these operations, the process saves the add/remove actions performed by the user in a *user operations repository*.

Ranking works on the pool created by selection and it has two additional inputs: *(i)* the *User calendar*, as our spatial-temporal recommendation takes into account the availability of the user in order to suggest other STOB's, and *(ii)* the *Filter configuration*, that allows the user to change the weight of each ranking factor in the final ranking.

The output of ranking goes directly to the UI Module, which provides the means for the user to perform the operations she prefers. Those relevant to our process are: *(i)* manual addition or removal of STOB's into/from the working pool; *(ii)* changing the filters configurations; *(iii)* creating a new activity in the calendar from a STOB.

3.1 Selection

The goal of selection is to build, for each user, a *pool* of potentially interesting activities to recommend, ruling out everything that is definitely out of the user's interest. Of course, an activity that had been previously ruled out, may be added back to the user's pool (depending on the privacy settings for the activity). However, the rationale of the Selection phase is to prevent the user from being flooded with activities proposals and to make recommendation more effective by focusing it on a well-defined group of items.

The pool determined through the selection phase goes through an additional filtering process as more transient user constraints are applied to it. The output of this phase is the *working pool*. To create an efficient *pool* some criteria have been selected:

- all STOB's that have already occurred are removed from the list of STOB's in the case of proactive context, while in on demand context only STOB's in the requested timeframe are considered;
- all STOB's geo-localized outside the maximum radius $r(\mathsf{u}) + r_P(\mathsf{u})$ of the user u are not considered;
- all STOB's with a user's interest, that does not exceed a minimum threshold $\mathsf{RAT}(\mathsf{s}) < \psi$, are eliminated from the list.

In some cases there are STOB's that have a user's interest with a high value. If this interest exceeds a maximum threshold $\mathsf{RAT}(\mathsf{s}) > \phi$ the criterion of spatial distance could not be considered.

The *working pool* can change dynamically as a consequence of the user's interaction with the system (e.g. she decides to change the recommendation time frame) or it can be edited manually (e.g. the user decides to discard an activity she is not interested in).

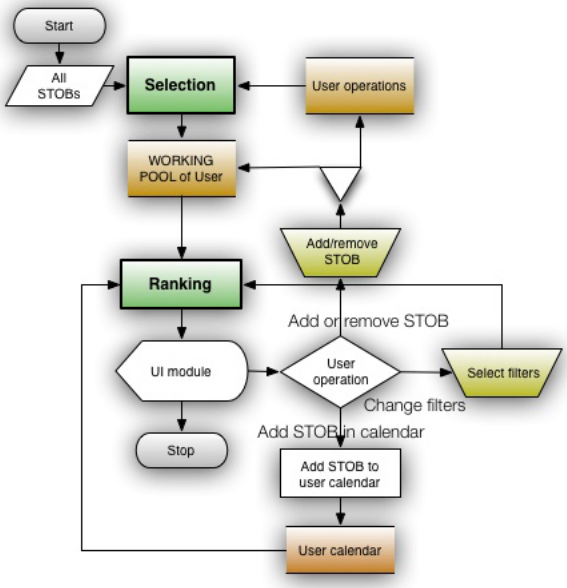

Fig. 1: Selection and Ranking of STOBs

3.2 Ranking

The aim of the ranking phase is to sort the STOB's stored in the user's working pool considering different factors. We can further split it in two sub-phases: the computation of the *individual ranking factors* and their *merge into one or more final rankings*.

Let us first discuss the individual ranking factors, which are all computed for a given user u and a given STOB s. We will introduce the following ranking factors: **thematic interest** ($\mathbb{THI}$), **typology interest** ($\mathbb{TYI}$), **average rating** ($\mathbb{RAT}$), **social interest**

($\mathbb{SOC}$), **feasibility** ($\mathbb{FSB}$), **reachability**($\mathbb{RCH}$). Some are strictly dependent on the user model, some have a social meaning, some depend on spatial-temporal relations.

Each factor $\mathbb{F}$ is computed as a value ranging from 0 to 100, normalizing the result of the computation, if needed[4]. We also compute a qualitative version of the value, that is produced by using two thresholds $0 \leq \theta_{\mathsf{min}} < \theta_{\mathsf{max}} \leq 100$. The thresholds act as landmarks, thus we obtain the three possible qualitative values *low* ($[0, \theta_{\mathsf{min}}]$), *ave* - average ($(\theta_{\mathsf{min}}, \theta_{\mathsf{max}})$), *exc* - exceptional ($[\theta_{\mathsf{max}}, 100]$). We assume that θ_{min} and θ_{max} can be manually adjusted by users, and are initially set to 0 and 100 so as to be rendered ineffective.

Thematic and Typology Interest. The first two ranking factors are the **thematic interest** $\mathbb{THI}_{\mathsf{u}}(\mathsf{s})$ of the user u in the themes of the STOB s and **typology interest** $\mathbb{TYI}_{\mathsf{u}}(\mathsf{s})$ of the user u in the type of the STOB s. These two can be derived from the user model and the STOB description by exploiting classic recommendation techniques, the discussion of which is outside the scope of this paper.

Average rating. We introduce the average rating of a STOB s as:

$$\mathbb{RAT}(\mathsf{s}) = \frac{\sum_{\mathsf{v} \in Raters(\mathsf{s})} \omega_{\mathsf{v}} \cdot \mathsf{RAT}_{\mathsf{v}}(\mathsf{s})}{\sum_{\mathsf{v} \in Raters(\mathsf{s})} \omega_{\mathsf{v}}}$$

where $Raters(\mathsf{s})$ is a set of users who provided a rating for s, ω_{v} is equal to $\mathbb{THI}_{\mathsf{v}}(\mathsf{s})$ if $\mathbb{THI}_{\mathsf{v}}(\mathsf{s})$ is *ave* or *exc*, and equal to 0 if $\mathbb{THI}_{\mathsf{v}}(\mathsf{s})$ is *low*. In other words, we weigh the ratings by the thematic interest of the users expressing them, deeming that uninterested people's opinions are less reliable, and we rule out ratings expressed by people whose thematic interests are too low.

Social Interest. The social interest ranks STOB's for a user u considering how much her friends are interested in participating. As explained in Section 2, a user may add an activity into her calendar specifying if she has the firm intention of participating in it (joined activity) or if she is very interested and wants to be reminded of it, but is still deciding whether to join or not (prospective activity). Therefore, the **social interest** $\mathbb{SOC}_{\mathsf{u}}(\mathsf{s})$ for a user u and a STOB s is defined as follows:

$$\mathbb{SOC}_{\mathsf{u}}(\mathsf{s}) = \frac{2|\mathsf{JND}(\mathsf{u},\mathsf{s})| + |\mathsf{PRS}(\mathsf{u},\mathsf{s})|}{2\,|\mathsf{SN}(\mathsf{u})|} \cdot 100 \tag{1}$$

where $\mathsf{JND}(\mathsf{u},\mathsf{s}), \mathsf{PRS}(\mathsf{u},\mathsf{s})$ are the following two disjoint sets:

- $\mathsf{JND}(\mathsf{u},\mathsf{s}) = \{f \in \mathsf{SN}(\mathsf{u}) | f$ added s as "joined"$\}$ and
- $\mathsf{PRS}(\mathsf{u},\mathsf{s}) = \{f \in \mathsf{SN}(\mathsf{u}) | f$ added s as "prospective"$\}$.

In computing social interest, the number of friends who actually joined the STOB weights two times more than the number of friends who simply showed some interest without making any decision.

Feasibility. $\mathbb{FSB}_{\mathsf{u}}(\mathsf{s})$ represents the possibility of a user u to reach the location where the STOB s takes place and to attend it for a relevant portion, given the appointments stored in her calendar. Let us consider the following situation: the user u is adding the STOB s, with the required time to participate $\mathsf{RQT}(\mathsf{s})$, to her calendar which already contains two STOBs, q and p. The scheduling of q, s and p is such that, taking into account the time needed to move from q to s and from s to p[5], the time left to user u for attending s is $\overline{d}_{\mathsf{u}}(\mathsf{s})$. Our measure of feasibility considers s *not* feasible at

[4] We will not include the normalization in the following formulas for the sake of readability.

[5] We assume an external service (e.g. Google Maps) provides an estimate of the traveling time.

all if $\overline{d}_{\mathsf{u}}(\mathsf{s})$ is less than half of the required time $\mathsf{RQT}(\mathsf{s})$. Conversely, s has maximum feasibility if $\overline{d}_{\mathsf{u}}(\mathsf{s})$ is at least $\mathsf{RQT}(\mathsf{s})$. Therefore, we define feasibility as follows:

$$\mathbb{FSB}_{\mathsf{u}}(\mathsf{s}) = \begin{cases} 0, & \text{if } \overline{d}_{\mathsf{u}}(\mathsf{s}) < \frac{\mathsf{RQT}(\mathsf{s})}{2}. \\ 200\,\frac{\overline{d}_{\mathsf{u}}(\mathsf{s})}{\mathsf{RQT}(\mathsf{s})} - 100, & \text{if } \frac{\mathsf{RQT}(\mathsf{s})}{2} \leq \overline{d}_{\mathsf{u}}(\mathsf{s}) \leq \mathsf{RQT}(\mathsf{s}). \\ 100, & \text{if } \overline{d}_{\mathsf{u}}(\mathsf{s}) > \mathsf{RQT}(\mathsf{s}). \end{cases} \tag{2}$$

Reachability. The last ranking factor we consider is the **reachability** $\mathbb{RCH}_{\mathsf{u}}(\mathsf{s})$ which measures how inclined the user u is to reach the location of the STOB s. As explained in Section 2, the user's position is a circular area defined by its center $c(\mathsf{u})$ and radius $r(\mathsf{u})$. Besides, an additional radius $r_P(u)$ defines her propensity to move, i.e. the distance she is willing to cover to join the STOB. The reachability is maximum (100) when the user's and STOB's centers coincide and it is θ_{min} when the propensity area (whose center is $c(\mathsf{u})$ and whose radius is $r(\mathsf{u}) + r_P(u)$) and the STOB area are tangent. Therefore the reachability is calculated as follows:

$$\mathbb{RCH}_{\mathsf{u}}(\mathsf{s}) = max\{\frac{\theta_{\mathsf{min}} - 100}{r(\mathsf{u}) + r_P(u) + r(\mathsf{s})}\, dist(c(\mathsf{u}), c(\mathsf{s})) + 100, 0\}. \tag{3}$$

Let us now discuss how the ranking factors are used to compute the results that ranking provides to the UI Module for visualization and user interaction. Given a (sub)set of ranking factors $\mathbf{RF}_k = \{\mathbb{F}_1, \ldots, \mathbb{F}_n\}$ a *final ranking* $\mathbf{rank}_k(\mathsf{u}, \mathsf{s})$ for a user u and a STOB s is defined as:

$$\mathbf{rank}_k(\mathsf{u}, \mathsf{s}) = \frac{\sum_{i=1}^{n} \phi_i \cdot \mathbb{F}_i(\mathsf{u}, \mathsf{s})}{\sum_{i=1}^{n} \phi_i}. \tag{4}$$

The values ϕ_i are called *filters* and determine how much each factor contributes to the final ranking. We assume that the filters are initially set to 1 and can be subsequently changed by the user through the UI Module.

In general we assume that the ranking factors can be combined into one or more final rankings, depending on the kind of visualization that the UI Module proposes to the user. This kind of flexibility is useful not only to apply the model to several different systems, but also to allow for contextualization within a single system (for example, we can provide different final rankings for the "proactive" mode and for the "on demand" mode). In our prototype, we compute one final ranking combining all the ranking factors but feasibility. The UI we are designing in fact shows the feasibility factor separately; in a sense it provides a second final ranking on its own.

Besides the final rankings, we use the qualitative values of the ranking factors to compute two sets of STOB's that can prove useful when visualizing the information:

- $\mathsf{LOW}(\mathsf{u}) = \{\mathsf{s} \mid \exists \mathbb{F}_i : \tilde{\mathbb{F}}_i(\mathsf{s}) = \mathit{low}\}$,
- $\mathsf{EXC}(\mathsf{u}) = \{\mathsf{s} \mid \exists \mathbb{F}_i : \tilde{\mathbb{F}}_i(\mathsf{s}) = \mathit{exc}\}$.

$\mathsf{LOW}(\mathsf{u})$ includes all STOB's that score a low value in at least one ranking factor. The UI Module may decide to hide these STOB's whenever there are too many of them to show. $\mathsf{EXC}(\mathsf{u})$ includes all STOB's that score an exceptionally high value in at least one ranking factor, and may be worth to highlight in the UI.

4 Preliminary evaluation

In order to provide answers to the research questions raised in the introduction (Section 1), we performed a user study in which we asked users to express their interest in

participating in a certain activity given different sets of information[6]. We compared the performance of our algorithm when predicting the users' interest in activity participation for two cases: one when the algorithm takes into account only the themes and type of the activity and the other when it considers also the additional factors. Our evaluation focuses on reachability, social interest and average rating because, as mentioned above, feasibility can be considered as a separate factor. Comparing the recommender predictions with users' answers shows that users do take the additional factors into account, if this additional information is available. Consequently, taking into account additional factors improves recommendation accuracy.

4.1 Experimental setting

Our experiment involved 200 users recruited according to an availability sampling strategy[7] and a group of 15 events.

The users were asked to provide the following information:

I1 A score $0 - 10$ expressing their potential interest in each event, *knowing only the event description along with its themes and type*. In the following, these will be referred to as the *init* user scores.

I2 A score $0 - 10$ expressing their potential interest in each event, *being aware of all the factors (themes and types, distance, average rating and social interest)*. These will be referred to as the *full* user scores.

I3 A value $0 - 10$ expressing their personal interest for each theme and type.

I4 A value $0 - 10$ expressing how they think they are affected by distance, average rating and social interest, when deciding if they are interested in an event.

I5 The maximum distance (in kilometers) they would generally travel in order to participate at an event, so that we can compute reachability.

4.2 Results and discussion

We computed the RMSE (*Root Mean Square Error*) between the predictions given by the algorithm and the values provided by users. Tables 1(a) and 1(b) illustrate our experiment results. Notice that in both tables the last column reports the variation of the RMSE expressed as a percentage with respect to the reference value shown in bold face.

Table 1(a) shows the RMSE which compares system results calculated taking into account only themes and type with (i) *init* user scores (in Line 1) or (ii) *full* user scores (in Line 2). Comparing these two cases, it is evident that ignoring additional factors, while the users are aware of them (Line 2), brings to a worse RMSE ($+12.62\%$). This is our first result [R1]: recommendation accuracy worsens if the system does not take into account the additional factors, since the users are aware of them in real-life.

Line 2 of Table 1(a) is repeated as Line 1 in Table 1(b) as reference for the RMSE improvement of the following two lines. Lines 2 and 3 in Table 1(b) report the RMSE obtained when the recommender system takes into account also the additional factors.

[6] The available dataset is available on request, sending a mail to one of the authors.

[7] Much research in social science is based on samples obtained through non-random selection, such as the availability sampling, i.e. a sampling of convenience, based on subjects available to the researcher, often used when the population source is not completely defined.

User scores	System weights THI	TYI	RCH	RAT	SOC	RMSE	$\Delta(\%)$
1 *init*	**0.50**	**0.50**	-	-	-	**2.583**	**-**
2 *full*	0.50	0.50	-	-	-	2.909	$+12.62$

(a) Results of the experiment obtained when the system considers only themes and type.

User scores	System weights THI	TYI	RCH	RAT	SOC	RMSE	$\Delta(\%)$
1 *full*	**0.50**	**0.50**	-	-	-	**2.909**	**-**
2 *full*	0.2	0.2	0.2	0.2	0.2	2.850	-2.03
3 *full*	0.2	0.2	user	user	user	2.795	-3.92

(b) Improvement of the RMSE in case the system takes into account additional factors.

Table 1: Results of the evaluation.

What distinguish those two lines from each other is the adopted set of weights: in Line 2 all the weights are equal while in Line 3 the weights are those provided by the users during the experiment (I4). It is easy to see that the recommender system performs better in these two cases than when it ignores additional factors (Line 1). This is our second result [R2]: the accurate recommender should take the additional factors into account. Moreover, the best result (Line 3) is reached by exploiting user provided weights (-3.92% vs. -2.03%).

5 Related Work

In this section we give a brief overview of the related work in the field of event/activity recommendation, followed by some pointers to works dealing with event ontologies and conclude with some references to works dealing with spatial-temporal representation.

PITTCULT [8] is a collaborative filtering recommender for cultural events, based on trust relations among users for finding similar users. Another collaborative filtering recommender is CUPID [13], which recommends events for the Flemish cultural scene. Similarly to us, they have a pool of events to be recommended, and they apply content and spatial filters after the recommendation process in order to exclude non-compatible events. They use user participation and ratings to find similar users. In [11] the authors propose a collaborative approach to recommendation of future events by starting from a content-based approach, where each user has a parameter vector that relates her preferences to event descriptions and similarity among users is calculated using their past feedback on the events. Only the events within limited distance are considered. A hybrid solution, combining content-based and collaborative filtering techniques, is proposed in [7], which presents an implementation as a Facebook[8] application. Another hybrid approach is described in [3]; in this case an item is recommended to a user if it is similar to the items that this user, or similar users, have liked in the past. In [2] the authors describe *iCity*, an application providing information about cultural resources and events in Torino, Italy. They propose to exploit user's activity to infer her interest about items and therefore to improve content-based recommendations in a Web 2.0 context. Alternatively, location-based recommendation is offered, but it is not combined with interest based recommendation.

[8] `http://www.facebook.com`

All these systems suggest activities ignoring the current user's commitment or her location in the period when the event takes place. Our system offers a relevant improvement, taking into account, in the recommendation phase, the possibility for the user to be present at the event.

In the literature there are already several works on the creation of event ontologies specialized in different domains: [14] used the event ontology for describing music events, [15] is more specialized in the sport domain and [16] considers the events as something that happen. We focus on the general notion of an event intended as a meeting organized earlier, concept applicable across multiple domains and contexts and on which it is possible to think of some recommendations.

Spatial and temporal aspects can be merged in a unique spatial-temporal representation. In this area, there are two main approaches: one considers snapshots of the world at different instants of time [5] and the other uses a 3D or 4D region based representation where time is one of the dimensions [5, 6, 12]. An example of the application of the second approach can be found in [9], where the aim is to define the spatial-temporal location of facts (e.g. actions performed in real life) described by the users and to define if they are co-located.

6 Conclusions and Future Work

Given the growing use of activities in the social networks, such as the *social events* in Facebook[9], a system being able to recommend activities in sophisticated way is more and more relevant in helping users find interesting suggestions without being overwhelmed by a huge mass of items. The solution proposed by our framework is an extension of a classic recommendation system towards spatiotemporal and social aspects. We introduced a semantic model describing three elements (users, STOB's and user's calendar activities) and a process for selecting and ranking new STOB's to be proposed to the user. The final ranking is obtained as a composition of factors, each representing a different aspect of the overall ranking (user's interest, ratings from other users, participation of user's friends, spatiotemporal relationship with other activities and a possibility for the user to reach the location). The opportunity to weight differently the various ranking factors allows for customization of the results according to the taste and needs of the user. Our evaluation shows that recommender systems which take into account additional factors perform better than the ones relying only on type and themes of the events. In addition, allowing users to give weights to these additional factors, brings further improvements into recommendation process. This brings us to a possible future extension of our framework, by adding other ranking factors that can influence the participation. Also, the existing factors can be given various importance in the ranking process and it would be interesting to see how much different factors contribute to the recommendation process. We also plan to investigate how to recommend STOB's to a group of friends, based on their interest in the themes of the STOB or on their participation in similar STOB's.

[9] http://www.facebook.com

References

1. G. Adomavicius and A. Tuzhilin. Context-aware recommender systems. In F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, editors, *Recommender Systems Handbook*, pages 217–253. Springer US, 2011.
2. F. Carmagnola, F. Cena, L. Console, O. Cortassa, C. Gena, A. Goy, I. Torre, A. Toso, and F. Vernero. Tag-based user modeling for social multi-device adaptive guides. *User Modeling and User-Adapted Interaction*, 18:497–538, 2008.
3. C. Cornelis, X. Guo, J. Lu, and G. Zhang. A Fuzzy Relational Approach to Event Recommendation. *Artificial Intelligence*, pages 2231–2242, 2007.
4. C. Cornelis, J. Lu, X. Guo, and G. Zhang. One-and-only item recommendation with fuzzy logic techniques. *Information Sciences*, 177(22):4906–4921, 2007.
5. P. Grenon and B. Smith. SNAP and SPAN: Towards Dynamic Spatial Ontology. *Spatial Cognition & Computation*, 4(1):69–104, Mar. 2004.
6. S. Hazarika and A. Cohn. Abducing Qualitative Spatio-Temporal Histories from Partial Observations. In D. Fensel, F. Giunchiglia, D. L. McGuinness, and M.-A. Williams, editors, *Proc. of the 8th International Conference on Principles and Knowledge Representation and Reasoning, KR-02*, pages 14 – 25. Morgan Kaufmann, 2002.
7. M. Kayaalp, T. Özyer, and S. T. Özyer. A mash-up application utilizing hybridized filtering techniques for recommending events at a social networking site. *Social Network Analysis and Mining*, 1(3):231–239, 2010.
8. D. H. Lee. PITTCULT: trust-based cultural event recommender. In *Proc. of the 2008 ACM conference on Recommender systems*, RecSys '08, pages 311–314. ACM, 2008.
9. S. Likavec, I. Lombardi, A. Nantiat, C. Picardi, and D. Theseider Dupré. Threading facts into a Collective Narrative World. In R. Aylett, M. Y. Lim, S. Louchart, P. Petta, and M. Riedl, editors, *Proc. of the 3rd International Conference on Interactive Digital Storytelling, ICIDS '10*, volume 6432 of *LNCS*, pages 86–97. Springer, 2010.
10. P. Lops, M. De Gemmis, and G. Semeraro. Content-based Recommender Systems: State of the Art and Trends. In F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, editors, *Recommender Systems Handbook*, chapter 3, pages 73–105. Springer US, 2011.
11. E. Minkov, B. Charrow, J. Ledlie, S. Teller, and T. Jaakkola. Collaborative future event recommendation. In *Proc. of the 19th ACM International Conference on Information and Knowledge Management*, CIKM '10, pages 819–828. ACM, 2010.
12. P. Muller. Topological Spatio-Temporal Reasoning and Representation. *Computational Intelligence*, 18(3):420–450, 2002.
13. T. D. Pessemier, S. Coppens, K. Geebelen, C. Vleugels, S. Bannier, E. Mannens, K. Vanhecke, and L. Martens. Collaborative recommendations with content-based filters for cultural activities via a scalable event distribution platform. *Multimedia Tools and Applications*, 58(1):167–213, 2012.
14. Y. Raimond, S. A. Abdallah, M. B. Sandler, and F. Giasson. The music ontology. In S. Dixon, D. Bainbridge, and R. Typke, editors, *Proc. of the 8th International Conference on Music Information Retrieval, ISMIR '07*, pages 417–422. Austrian Computer Society, 2007.
15. J. Rayfield, P. Wilton, and S. Oliver. Sport Ontology. http://www.bbc.co.uk/ontologies/sport/, February 2011.
16. R. Shaw, R. Troncy, and L. Hardman. LODE: Linking open descriptions of events. In A. Gómez-Pérez, Y. Yu, and Y. Ding, editors, *Proc. of the 4th Annual Asian Semantic Web Conference, ASWC 2009*, volume 5926 of *LNCS*, pages 153–167. Springer, 2009.
17. S. Sosnovsky and D. Dicheva. Ontological technologies for user modelling. *International Journal of Metadata Semantic Ontologies*, 5(1):32–71, 2010.