i-ETC: ISEL Academic Journal of Electronics,
Telecommunications and Computers
CETC2011 Issue, Vol. 2, n. 1 (2013) ID-3

Social Mobility Advisor

André Costa, David Antunes and Porfírio Filipe
Área Departamental de Engenharia de Electrónica e Telecomunicações e de Computadores
Instituto Superior de Engenharia de Lisboa, Lisbon, Portugal
{31250, 31212}.alunos.isel.pt, pfilipe@deetc.isel.ipl.pt

Keywords: Sustainable Mobility, Social Network, Reputation Systems, Recommender Systems, Incremental Collaborative Filtering, Scalability.

Abstract: Seamless access to updated information about mobility options is a key factor for sustainable economic development. In this context, the adoption of information systems, with innovative features, meet these needs by stimulating the interest and curiosity of end users. There are well known difficulties, even for regular travelers, to get the most suitable mobility options combining available transport modes. Also, the sustainable mobility has become the new imperative for transport policy. In this paper we present a collaborative platform that supports a Social Network, in which users can search and share sustainable mobility experiences. This platform includes a reputation and a recommender systems specially designed to deal with mobility options.

1 INTRODUCTION

We live in a society of information. Sharing and accessing that information has become an acquired right. The Internet has become a great ally in the dissemination of information process, partly due to the popularity of Social Networks and its collaborative based interaction. Currently, there are successful examples of websites that operate either under the social network basis or by taking advantage of other social networks (e.g., authentication, friend list invitations, etc.). *Trip Advisor* (www.tripadvisor.com), *Rail Europe* (http://www raileurope.com) and *Momondo* (http://www momondo.pt/) are some examples in the mobility context, which try to provide mobility options. However, these services are unable to indicate alternative means of transportation; also, the results do not reflect the personal tastes of those who use the service. Therefore, we identify the need for a system that is able to provide multimodal mobility options, and is built under each user's tastes and interests in order to generate relevant search results and recommendations.

In this paper we present a collaborative solution that aims to assist travelers to share their experiences and find the best suitable mobility alternatives, bearing in mind the best sustainable options. Those actions are supported by the collaborative community interactions, in other words, under the social network concept. This approach leads to a credibility problem which we had into account by assuring both users and information are cataloged by confidence levels. We also propose an automated moderation mechanism that tries to reduce the number of required staff to administer the system.

The Social Mobility Advisor (SMA) project aims to contribute for a more accessible, social and sustainable mobility across the Europe, dealing with reputation and recommender systems in order to build a social network.

2 CONCEPT

Social networks (SN) have evolved over time as a way of information exchange, even before the existence of the Internet. In fact, already in ancient Egypt there were groups of people with common interests who came together to share knowledge. With the emerging of Internet, information share and access has become more simple, fast and effective. Therefore, SN naturally shifted to this environment, grouping a greater number of people, breaking down physical, geographical and cultural barriers.

SMA takes advantage of collaborative approach of social networks, by supplying a way for users to share their mobility experiences where the participating entities collaborate actively to improve the quality of information, updating information items, for example, targets, gates, and operators / transport authorities. Thus, we can incrementally gather and provide mobility experiences, decreasing information disparity problem on this matter, that daily harms regular travelers.

Due to the system's collaborative way, on which information deeply depends on users' interactions, we identify the lack of confidence as a major problem. This problem is commonly addressed by commercial companies through the use of reputation systems. We managed to combine existing reputation systems tendencies (e.g., *Amazon*, *eBay*, *Slashdot*) in order to minimize this problem. With the aim of minimize the need for dedicated human resources, especially to maintenance, the SMA social network tends to be self-managed by the participants which are dynamically assigned, with more or less privileges, depending on their current reputation.

So far we have described a process that can be used to populate the system with filtered information. Despite that, a user may ask: "Which information is relevant for me?". We consider that this question cannot be left unanswered. Our goal here is to provide relevant information for each user, rather than dump random (and possible undesirable) information for every users. Recommender systems help to increase the effectiveness and capability of recommendation generation, which will also highly increase the computation weight. The underlying techniques we present try to overcome the scalability problem, by computing heavy calculations into simpler incremental summations.

The information available in the SMA social network is the result of the integration of information from internal sources (internal repository with mobility options maintained by the participants) as well as information provided by external sources, such as *Wikipedia* (http://www.wikipedia.org/), *Google Maps*, *Google Places* (http://code.google.com), *Facebook* (http://www facebook.com/) and *Brighter Planet* (http://www.brighterplanet.com/).

3 REPUTATION SYSTEM

The traditional cues of trust and reputation that we are used to observe and depend on in physical world are missing in online environments, so that electronic substitutes are needed. A reputation system collects, distributes, and aggregates feedback about participants' past behavior, helping users to decide whom to trust and encouraging trustworthy behavior [1, 3].

We assumed that such mechanism can be adopted in our context in order to overcome the confidence problem. The proposed reputation system is based on the combination of existing ones, such as *Amazon* (http://www.amazon.com/), *Stackoverflow* (http://stackoverflow.com/) and *eBay* (http://www.ebay.com/). Also, we grouped users by credibility, assigning them reputation levels.

3.1 Reputation Level

The reputation levels system is inspired in videogames like *World of Warcraft* (http://eu.battle net), which needs exponential requirements to level up. Equation 1 illustrates the adopted formula to generate reputation level requirements for the next level, supplying the current level.

$$RN = (8 * NC) * (45 + (5 * NC)$$

Equation 1: Reputation level requirements.

The reputation level is an important issue in our reputation system, since it will define the weight of the user's ratings, described below, as well as their role/privileges in Social Network.

3.2 Ratings

Our system provides an input for users to express their opinion related to item's information quality by rating it as being helpful or not helpful (+1, -1), as showed in Illustration 1. This range permits a less ambiguous evaluation, since it will influence the reputation of those who submitted the item.

Illustration 1: Binary rating representation.

Using this schema, is possible to filter and organize information by its quality and, at the same time, indirectly rate the user who submitted that information. This approach aims to decrease public reputation of those who submit poor information and, inversely, reward who submit relevant information.

Item's information reputation is determined using the Equation 2, which consists on a summation of all positive and negative ratings weighted by rater's reputation level. Rater's reputation level is used to avoid unfair ratings [1], one of reputation systems known problem, in assumption that users with better reputation provide more reliable ratings.

$$A_i = a * b - (c' * d') + c * d$$

Equation 2: Binary rating calculation.

a – Item i current rating (in cache);

b – Summation of item i raters reputation (in cache);
c – New rating for item i of active user;
d – Item i active user reputation level (in cache);

Reevaluation factors:

c' – Old rating for item i of active user;
d' – Reputation level of active user when submitted the old evaluation (in cache);

In order to calculate the mobility item's quality itself (e.g., how good is a gate, operator, transport, etc.), our system also provides an input for users to express their mobility experience quality by a quantitative rating from 1 to 5 stars, as illustrated in Illustration 2, plus comment.

Illustration 2: Star rating representation.

Item's quality reputation is calculated by the Equation 3 and it consists in a weighted average between evaluations and raters reputation. Rater's reputation level is used to avoid unfair ratings as mentioned before.

$$A_i = \frac{a * b - (c' * d') + c * d}{b - d' + d}$$

Equation 3: Start rating calculation.

a – Item i current rating (in cache);
b – Summation of item i raters reputation (in cache);
c – New rating for item i of active user;
d – Item i active user reputation level (in cache);

Reevaluation factors:

c' – Old rating for item i of active user;
d' – Reputation level of active user when submitted the old evaluation (in cache);

3.3 User's reputation

As mentioned before, users indirectly rate others by rating their submitted items as being helpful or not helpful, according to the Equation 4.

$$R = -(a' * r') + a * r$$

Equation 4: User's reputation adjustment.

R – Reputation increment;
a – Rater reputation level;
r – New rating (+1, -1);

Reevaluation factors:

a' – Old rater reputation level;
r' – Old rating;

For example, when a user A submits an item and a second user B on level 2 rates that item as useful (+1), the user A gets two points of reputation as presented in Illustration 3.

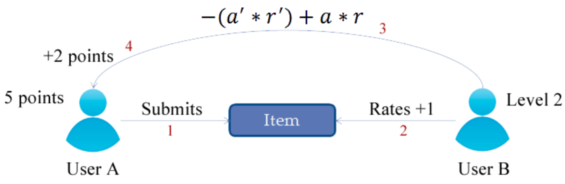

Illustration 3: Evaluation example.

3.3 Interaction Incentives

Due to the lack of incentive for users to provide their mobility experiences and ratings, we included incentive mechanisms in our reputation system. Users are rewarded for data submission, where the importance of the submitted item defines the assigned reputation points. Table 1 shows the possible configuration values of rewarded points.

Action	Reputation
Binary rating	+2
Star rating	+1
Mobility item submission, except route	+4
Route submission	+6
Comment submission	+1
Collaboration action	+3
Revaluation	0

Table 1: Possible configuration for incentives.

For example, someone that submits a new gate will be rewarded with 4 reputation points.

We also adopted a badge reward system, used by systems like *Stackoverflow* (http://www.stackoverflow.com/) and *Foursquare* (https://www foursquare.com/) with great acceptance by the community, where users are rewarded by reaching certain objectives.

In our context, we reward the users when they reach, for example, a pre-defined number of done routes or rated items.

3.4 Moderation

In our approach, the information is strongly dependent on user's collaborative interactions consequently, moderation actions are needed. To reduce the amount of necessary staff to keep the system, we included a role based moderation mechanism, inspired on Slashdot, which takes advantage of user's reputation to delegate them moderation actions. This way, most reliable users are assigned to perform some moderation actions. Our mechanism has two moderation roles, where the first, called "collaborator", are assigned to manage regular users by editing/hiding comments and item's information. The second role, called "moderator", was introduced to reduce the number of unfair collaborators, as described by Slashdot [1].

4 RECOMMENDER SYSTEM

In everyday life, we are confronted with recommendations from other people, television, ads, news reports from news media, and so on. In a Web environment, recommender systems can help us sift through all the available information to find which is most valuable for us, accordingly with our tastes and interests. E-commerce companies rely on recommender systems to ensure that customers are recommended the right products, improving sales. [4] It is natural to assume that these mechanisms can be introduced in order to assist travelers to find relevant mobility alternatives.

In our study, collaborative filtering (CF) is described as one of the most successfully technique to generate relevant recommendations. The fundamental assumption of CF is that if users X and Y rate *n* items similarly, or have similar behaviors (e.g., buying, watching, listening), and hence will rate or act on other items similarly.

The CF algorithms can be divided in three main parts, listed below:

- ***Representation of input data***: deals with the submitted ratings data, which is represented by a *user-item* matrix.
- ***Neighborhood formation***: is the most important step. It focuses on the problem of how to identify the most similar users (neighbors) with active user - u_x. The most commonly used technique is the Pearson's correlation, that measures the extent to which two variables linearly relate with each other [4], described in Equation 5

$$sim_{u_x,u_y} = \frac{\sum_{h=1}^{n'}\left(r_{u_x i_h} - \overline{r_{u_x}}\right)\left(r_{u_y i_h} - \overline{r_{u_y}}\right)}{\sqrt{\sum_{h=1}^{n'}\left(r_{u_x i_h} - \overline{r_{u_x}}\right)^2}\sqrt{\sum_{h=1}^{n'}\left(r_{u_y,i_h} - \overline{r_{u_y}}\right)^2}}$$

Equation 5: Pearson's correlation calculation.

where the $h \in H$ summations are over the items that both the users u_x and u_y have rated and $\overline{r_u}$ is the average rating of the co-rated items of the uth user.

- ***Recommendation generation:*** deals with the problem of finding the top-N recommended products from the neighborhood users.

A CF algorithm should be both accurate (the recommended objects should subsequently receive high ratings) and efficient in terms of computational complexity. Therefore, accomplishing that is quite an engineering challenge. CF fails to scale up its computation with the growth of both the number of users and items in the database. [2] This is mainly caused by the *neighborhood formation* task, which performs heavy computations in order to calculate the similarity for each pair of users.

We studied "classic" collaborative filtering techniques and realized that real-time recommendations, as expected in a web application, would tend to generate unacceptable performance values. In order to address the scalability challenge, we included in our project an Incremental Collaborative Filtering (ICF) based algorithm which allows generating recommendations in a reasonable time.

4.1 Incremental CF

To address scalability problem, we adopted the proposal presented by Papagelis *et al* [2]. It consists on an incremental user-based CF method, based on incremental updates of the user-to-user similarity matrix, defined in step two (*neighborhood formation*). The adopted mechanism to calculate the user similarity is the Pearson's correlation, described in Equation 5.

In order to update users' similarity, based on previous values, we adopted the notation described on Equation 6.

$$A = \frac{B}{\sqrt{C}\sqrt{D}} \rightarrow A = sim_{u_x,u_y}, B = \sum_{h=1}^{n'} \left(r_{u_x i_h} - \overline{r_{u_x}}\right)\left(r_{u_y,i_h} - \overline{r_{u_y}}\right),$$

$$\mathrm{C} = \sum_{h=1}^{n'} \left(r_{u_x,i_h} - \overline{r_{u_x}}\right)^2, \mathrm{D} = \sum_{h=1}^{n'} \left(r_{u_y,i_h} - \overline{r_{u_y}}\right)^2$$

Equation 6: A, B, C and D definition factors.

A, B, C and D are updated every time a user u_a submits a new rating or updates an existing one. The new similarity calculation is based on previous factor's values.

$$\mathrm{A}' = \frac{\mathrm{B}'}{\sqrt{\mathrm{C}'}\sqrt{\mathrm{D}'}} \rightarrow \mathrm{A}' = \frac{\mathrm{B} + \mathrm{e}}{\sqrt{\mathrm{C} + \mathrm{f}}\sqrt{\mathrm{D} + \mathrm{g}}}, \mathrm{B}' = \mathrm{B} + \mathrm{e},$$

$$\mathrm{C}' = \mathrm{C} + \mathrm{f}, \mathrm{D}' = \mathrm{D} + \mathrm{g}$$

Equation 7: New similarity calculation formula.

The new similarity value A' is calculated based on A', B', C and D' factors generated by adding e, g and g increments on A, B, C and D values, as described in Equation 7. The formulas to calculate these increments are described in Table 2.

		Submission of a new rating		Update of an existing rating	
u_y had rated i_a	e	$e = (r_{u_a,i_a} - \overline{r_{u_a}'})(r_{u_y,i_a} - \overline{r_{u_y}}) - \sum_{h=1}^{n'} d\overline{r_{u_a}}(r_{u_y,i_h} - \overline{r_{u_y}})$	(2)	$e = dr_{u_a,i_a}(r_{u_y,i_a} - \overline{r_{u_y}}) - \sum_{h=1}^{n'} d\overline{r_{u_a}}(r_{u_y,i_h} - \overline{r_{u_y}})$	(8)
	f	$f = (r_{u_a,i_a} - \overline{r_{u_a}'})^2 + \sum_{h=1}^{n'} d\overline{r_{u_a}}^2 - 2\sum_{h=1}^{n'} d\overline{r_{u_a}}(r_{u_a,i_h} - \overline{r_{u_a}})$	(3)	$f = dr_{u_a,i_a}^{\ 2} + 2dr_{u_a,i_a}(r_{u_a,i_a} - \overline{r_{u_a}'}) + \sum_{h=1}^{n'} d\overline{r_{u_a}'}^2 - 2\sum_{h=1}^{n'} d\overline{r_{u_a}'}(r_{u_a,i_h} - \overline{r_{u_a}})$	(9)
	g	$g = (r_{u_y,i_a} - \overline{r_{u_y}})^2$	(4)	$g = 0$	(10)
u_y had not rated i_a	e	$e = -\sum_{h=1}^{n'} d\overline{r_{u_a}}(r_{u_y,i_h} - \overline{r_{u_y}})$	(5)	$e = -\sum_{h=1}^{n'} d\overline{r_{u_a}}(r_{u_y,i_h} - \overline{r_{u_y}})$	(11)
	f	$f = \sum_{h=1}^{n'} d\overline{r_{u_a}}^2 - 2\sum_{h=1}^{n'} d\overline{r_{u_a}}(r_{u_a,i_h} - \overline{r_{u_a}})$	(6)	$f = \sum_{h=1}^{n'} d\overline{r_{u_a}'}^2 - 2\sum_{h=1}^{n'} d\overline{r_{u_a}'}(r_{u_a,i_h} - \overline{r_{u_a}})$	(12)
	g	$g = 0$	(7)	$g = 0$	(13)

Table 2: e, f and g calculation formulas.

We managed to build a recommendation algorithm which is able to operate under an online environment. The incremental calculation used in neighbourhood formation returns the same values as the "classical" technique.

4 CONCLUSION

This paper describes a collaborative platform for sharing and getting sustainable mobility options by taking advantage of emerging concepts, like Social Network.

We developed a platform that offers to its users the possibility of information sharing focused on user's needs. In this perspective, users can interact and collaborate within a virtual community supported by the World Wide Web. Therefore, users are involved in a media dialogue, as both producers and consumers of information, which contrasts with the access through a portal (website) where they are limited to being passive consumers of content specifically created for them.

The implemented system manages to organize information by its credibility and also assign to most reputable users moderation actions, through the adopted reputation system. The submitted information is weighted in order to organize it by its relevance and the users are cataloged by reputation levels. Therefore, we can decrease information disparity.

Currently, the SMA social network has exceeded the testing phase, which counted with more than one hundred users. The present database includes more than ten thousand items (gates, transports, operators), covering mobility information in a worldwide context, and about one thousand rates and comments.

As future work, we consider relevant to embrace mechanisms that allow transportation authorities to interact directly with the system in order to keep updated information about mobility items provided by them, ticket pricing and departure/arrival schedules, converging our solution in a door-to-door scenario. Also, these authorities could be part of incentive mechanism, offering real prizes (discounts, vouchers) for those who best contribute to system's information quality.

Everyone can obtain more information and join the social network accessing http://start.isel.pt.

REFERENCES

[1] Audun Jøsang, Roslan Ismail and Colin Boyd. *A survey of Trust and Reputation Systems for Online Service Provision.* (The Netherlands 2007), Elsevier Science Publishers B. V. Amsterdam.

[2] M. Papagelis, I. Rousidis, D. Plexousakis, E. Theoharopoulos, *Incremental collaborative filtering for highly-scalable recommendation algorithms*, M.-S. Hacid, N. V. Murray, Z. W. Ras, and S. Tsumoto, editors, ISMIS, volume 3488, Lecture Notes in Computer Science, 2005.

[3] P. Resnick, R. Zeckhauser, E. Friedman and Ko Kuwabara, *Reputation Systems: Facilitating Trust in Internet Interactions*, Communications of the ACM, 43, 2000.

[4] B. M. Sarwar, G. Karypis, J. A. Konstan and J. Riedl. *Analysis of recommendation algorithms for e-commerce*, New York, USA, 2000, ACM.