Myusic: a Content-based Music Recommender System based on eVSM and Social Media

Cataldo Musto[1], Fedelucio Narducci[2], Giovanni Semeraro[1], Pasquale Lops[1], and Marco de Gemmis[1]

[1] Department of Computer Science
University of Bari Aldo Moro, Italy
name.surname@uniba.it

[2] Department of Information Science, Systems Theory, and Communication
University of Milano-Bicocca, Italy
narducci@disco.unimib.it

Abstract. This paper presents Myusic, a platform that leverages social media to produce content-based music recommendations. The design of the platform is based on the insight that user preferences in music can be extracted by mining Facebook profiles, thus providing a novel and effective way to sift in large music databases and overcome the cold-start problem as well. The content-based recommendation model implemented in Myusic is eVSM [4], an enhanced version of the vector space model based on distributional models, Random Indexing and Quantum Negation. The effectiveness of the platform is evaluated through a preliminary user study performed on a sample of 50 persons. The results showed that 74% of users actually prefer recommendations computed by social media-based profiles with respect to those computed by a simple heuristic based on the popularity of artists, and confirmed the usefulness of performing user studies because of the different outcomes they can provide with respect to offline experiments.

1 Introduction and Related Work

One of the main issues of the so-called *personalization pipeline* is preference acquisition and elicitation. That step has always been considered the *bottleneck* in recommendation process since classical approaches for gathering user preferences are usually time consuming or intrusive. The widespread diffusion of social networks in the age of Web 2.0 offers a new interesting chance to overcome that problem, since users spend 22% of their time on social networks[3] and 30 billion pieces of content are shared on Facebook every month [3]. In this scenario, to harvest social media is a recent trend in the area of Recommender Systems (RSs): it can merge the un-intrusiveness of implicit user modeling with the accuracy of explicit techniques, since the information left by users is freely provided and actually reflects real preferences.

[3] http://blog.nielsen.com/nielsenwire/social/

This paper presents Myusic, a tool that provides users with music recommendations. The goal of the system is to catch user preferences in music and filter the huge amount of data stored in platforms such as iTunes or Amazon in order to produce personalized suggestions about artists users could like. The filtering model behind Myusic is eVSM, an enhanced extension of VSM based on distributional models, Random Indexing and Quantum Negation. As introduced in [4], eVSM provides a lightweight semantic representation based on distributional models, where each artist (and the user profile, as well) is modeled as a vector in a semantic vector space, according to the tags used to describe her and the co-occurrences between the tags themselves. The model is based on the assumption that a user profile can be built by combining the tag-based representation (obtained by crawling Last.fm platform) of the artists she is interested in. Next, classical similarity measures can be exploited to match item descriptions with content-based user profiles. A prototype version of Myusic was made available online for two months in order to design a user study and evaluate the effectiveness of the model as well as its impact on real users.

Generally speaking, this work concerns to the area of music recommendation. The commonly used technique for providing recommendations is collaborative filtering, implemented in very well known services, such as MyStrands[4], Last.fm[5] or iTunes Genius. An early attempt to recommend music using collaborative filtering was done by Shardanand [8]. Another trend is to use content-based recommendation strategies, which analyze diverse sets of low-level features (e.g. harmony, rhythm, melody), or high-level features (metadata or content-based data available in social media) [2] to provide recommendations. The use of Linked Data for music recommendation is investigated in [6]. Recently, Bu et al. [1] followed the recent trend of harvesting information coming from social media for personalization tasks and proposed its application for music recommendation. Finally, Wang et al. [9] showed the usefulness of tags with respect to other content-based sources.

The paper is organized as follows: the architecture of the systems is sketched in Section 2; Section 3 focuses on the results of a preliminary experimental evaluation and finally Section 4 contains conclusions and directions for future research.

2 Myusic: content-based music recommendations

The general architecture of Myusic is sketched in Figure 1. We can identify four main components:

Crawler. The CRAWLER module queries Last.fm through its public APIs to build a corpus of available artists. For each artist, the name, a picture, the title of the most popular tracks, their playcount and a set of tags that describe that artist are crawled. All the crawled data are locally stored.

[4] http://www.mystrands.com
[5] http://www.last.fm

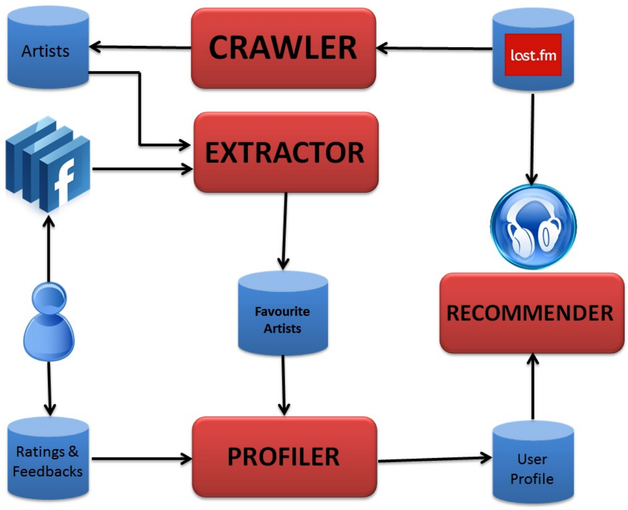

Fig. 1. Myusic architecture

Extractor. The EXTRACTOR module connects to Facebook, extracts artists the user likes (*Favourite Music* section in the Facebook profile, see Figure 2), and maps them to the data gathered from Last.fm in order to build a preliminary set of artists the user likes. This information is locally modeled in her own profile to let her receive recommendations even in her first interaction with Myusic, thus avoiding the cold-start. Implicit information coming from the links posted by the user and the events she attended are extracted, as well.

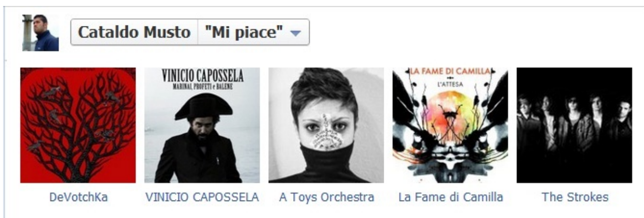

Fig. 2. User Preferences from a Facebook profile

Profiler. The process of building user profiles is performed in two steps. First, a weight is assigned to each artist returned by the EXTRACTOR. The weight of a specific artist is defined according to a simple heuristic: if a user posted a song, that information can be considered as a *light* evidence of her preference for that artist, while the fact that she explicitly clicked on "Like" on her Facebook page can be considered as a *strong* evidence. For example, on a 5-point Likert scale, a score equal to 3 is assigned to the artists whose name appear among the links posted by the user, while a score equal to 4 is assigned to those occurring in her favorite Facebook pages. If an artist occurs in both lists (that is to say, the user likes it and posted a song, as well), 5 out of 5 is assigned as score. Next, a profiling model has to be chosen. The eVSM framework provides four different profiling models [5]: a basic profile (referred to as *RI*), a simple variant that exploits negative user feedbacks (called *QN*), and two weighted counterparts which give greater weight to the artists a user liked the most (respectively, *W-RI* and W-QN). Regardless the profiling model, in eVSM user profiles are defined in eVSM by means of two vectors, $\boldsymbol{p}_{+u}$ and $\boldsymbol{p}_{-u}$, which represent user preferences and negative feedbacks, respectively. They are defined as follows:

$$\boldsymbol{p}_{+u} = \sum_{i=1}^{|I_u^+|} \boldsymbol{a}_i * r(u, a_i) \tag{1}$$

$$\boldsymbol{p}_{-u} = \sum_{i=1}^{|I_u^-|} \boldsymbol{a}_i * (MAX - r(u, a_i)) \tag{2}$$

where I_u^+ is the set of user favorite artists, I_u^- is the set of artists the user dislikes, MAX is the highest rating that can be assigned to an item, $r(u, a_i)$ is the score assigned to the artist a_i and $\boldsymbol{a}_i$ is the vector space representation of the artist. Since each artist is described through a set of tags $t_1 \dots t_n$ extracted from Last.fm, the vector space representation is a weighted vector $\boldsymbol{a}_i = (w_{t_1}, \dots, w_{t_n})$ where w_{t_i} is the weight of the tag t_i. Generally speaking, W-QN model combines $\boldsymbol{p}_{+u}$ with $\boldsymbol{p}_{-u}$ through a Quantum Negation operator implemented in eVSM framework, while W-RI model exploits only the information coming from $\boldsymbol{p}_{+u}$ and does not take into account negative feedback. Finally, RI and QN follow the same insight of their weighted counterpart with the difference that they do not exploit the user rating $r(u, a_i)$, thus a uniform weight is given to each artist.

Recommender. Given a semantic vector space representation based on distributional models for both artists and user profiles, through similarity measures it is possibile to produce as output a ranked list of suggested artists. The cosine similarity for all the possible couples $(\boldsymbol{p}_u, \boldsymbol{a})$ is computed, where $\boldsymbol{p}_u$ is the vector space representation of user u, while $\boldsymbol{a}$ is the vector describing the artist a. Figure 3 shows an example of recommendation list. The platform allows the user to express feedbacks on recommendations. Positive and negative feedbacks are used to respectively update positive and negative profile vectors and to trigger the recommendation process again.

3 Experimental Evaluation

The goal of the experimental evaluation is to validate the design of the platform by carrying out a user study whose goal is to analyze the impact and the effective-

Fig. 3. An example of recommendation list in Myusic platform

ness of the different configurations of eVSM implemented in Myusic. Specifically, a user study involving 50 users under 30, heterogeneously distributed by sex, education and musical knowledge (according to the availability sampling strategy) has been performed. They interacted for two months with the online version of Myusic. A crawl of Last.fm was performed at the end of November, 2011 and data about 228,878 artists were extracted. Each user explicitly granted the access to her Facebook profile to extract data about favourite artists. At the end of the Extraction step, a set of 980 different artists the 50 users like were extracted from Facebook pages. Generally speaking, 1,720 feedbacks were collected: 1,495 of them came from Facebook profiles, while 225 were explicitly provided by the users (for example, expressing a feedback on their recommendations). The collected feedbacks were highly unbalanced since only 116 (6.71%) on 1,720 were negative. Last.fm APIs were exploited to extract the most popular tags associated to each artist. The less expressive and meaningful ones (such as *seenlive*, *cool*, and so on) were considered as noisy and filtered out. The design of the user study was oriented to answer to the following questions:

- **Experiment 1**: Does the cold-start problem can be mitigated by modeling user profiles which integrate information coming from social media?
- **Experiment 2**: Do the users actually perceive the utility of adopting weighting schemes and negation when user profiles are represented?
- **Experiment 3**: How does the platform perform in terms of novelty, serendipity and diversity of the proposed recommendations?

In the first experiment, users were asked to login and to extract their data from their own Facebook page. Next, a user profile was built according to a profiling model *randomly* chosen among the 4 described above and a preliminary set of recommendations was proposed to the target user. In order to evaluate the effectiveness of the EXTRACTOR we compared the recommendation list generated through eVSM to a baseline represented by a list produced by simply ranking the most popular artists. Next, we asked users to tell which list they preferred. Obviously, they were not aware about which list was the baseline and which one was built through eVSM. A plot that summarizes users' answers is provided in Figure 4-a. It is straightforward to note that users actually prefer social media-based recommendations, since 74% of them preferred that strategy with respect to a simple heuristic based on popularity of the artists stored in database. However, even if the results gained by this profiling technique were outstanding, it is necessary to understand why 26% of the users simply preferred the most popular artists. Probably, there is a correlation between users' knowledge in music and the list they choose. It is likely that users with very *generic* tastes prefer a list of popular singers. Similarly, it is likely that users with a poor knowledge in music might prefer a list of well-known singers with respect to a list where most of the artists, even if related to their tastes, were unknown. A larger evaluation with users, split according to their musical knowledge, may be helpful to understand the dynamics behind users' choices. Similarly, it would be good to investigate the impact of the amount of the information extracted from Facebook profiles with the accuracy of the recommendations. The second experiment was performed in two steps. In the first step users were asked to login and to extract their data from their own Facebook page, as in Experiment 1. Next, two profiles were built by following the RI and the W-RI profiling models, respectively. Finally, recommendations were generated from both profiles, and users were asked to choose the configuration they preferred. As in Experiment 1, they were not aware about which recommendations were generated by exploiting their weighted profile and which ones were produced through its unweighted counterpart. Results of this experiments are shown in Figure 4-b. Differently from the results obtained from an *in-vitro* experiment performed in a scenario of movie recommendation [5], users did not perceive as useful the introduction of a weighting scheme designed to give higher significance to the artists the user likes the most. On the contrary, the RI profiling model was the preferred one for 70% of the users involved in the experiment. Similarly, in the second step of the experiment the RI profiling model was compared to the QN one, in order to evaluate the impact on user perception of modeling negative preferences. Also in this case the results were conflicting with the outcomes that emerged from the *in-vitro* experiment since 65% of the users preferred the recommendations generated through the profiling technique that does not model negative preferences. Even if the results of Experiment 2 did not confirmed the outcomes of the offline evaluation of eVSM they are actually interesting. First, they confirmed the usefulness of combining offline experiments with user studies thanks to the different outcomes they can provide. Indeed, in user-centered applications such as content-based recommender systems, user perception and user feedbacks play a central role and these factors need to be taken into account. In general, further investigation is needed because most of these results may be due to a specific *bias* of the designed experiment. As stated above, the extraction of data from Facebook pages crawls information about what a specific user likes, so very few negative feedback were collected (less than 7%). Consequently, the negative part of the user profile was very poor and this might justify the results. It is likely that collecting more negative feedbacks would be enough to confirm the usefulness of negative information. Finally, in Experiment 3 users were asked to express their preference on the recommendations produced through the RI profiling model (since it emerged as the best one from the previous experiment) in terms of novelty, accuracy and diversity. The results of this experiment are sketched in Figure 4-c. In general, the results are encouraging since most of the users expressed a positive opinion about the system. Specifically, Myusic has a positive impact on final users in terms of trust, since the opinion of 92% of the users ranges from *Good* to *Very Good*. This is likely due to the good accuracy of the recommendations produced by the system. Indeed, more than 80% of the users considered as accurate or very accurate the suggestions of the system. Similarly, also the outcomes concerning diversity were positive, since more than 60% labeled the level of diversity among the recommendations as *Very Good*. The only aspect that needs improvements regards the novelty of recommendations since 34% of the users labeled as not novel the suggestions produced by the system. This outcome was somehow expected since overspecialization it is a typical problem of content-based recommender systems (CBRS). However, even if these results lead us to carry on this research, they have to be considered as preliminary since this evaluation needs to be extended by comparing results of eVSM with other state of the art models, such as LSI, VSM or collaborative filtering.

4 Conclusions and Future Directions

In this paper we proposed Myusic, a music recommendation platform. It implements a content-based recommender system based on eVSM, an enhanced version of classical VSM. The most distinguishing aspect of Myusic is the exploitation of Facebook profiles for acquiring user preferences. An experimental evaluation carried out by involving real users demonstrated that leveraging social media is an effective way for overcoming the cold-start problem of CBRS. On the other hand, the exploitation of relevance feedback and user ratings generally did not improve the predictive accuracy of Myusic. Users showed to trust the system, and Myusic also achieved good results in terms of accuracy and diversity of recommendations. Those results encouraged keeping on this research. In the future we will investigate the adoption of recommendation strategies tailored on the music background of each user, even by learning accurate interaction models in order to classify users [7]. Furthermore, we will try to introduce more unexpected suggestions. Experiments showed that novelty needs to be improved.

References

1. S. Bu, J.and Tan, C. Chen, C. Wang, H. Wu, L. Zhang, and X. He. Music recommendation by unified hypergraph: combining social media information and music

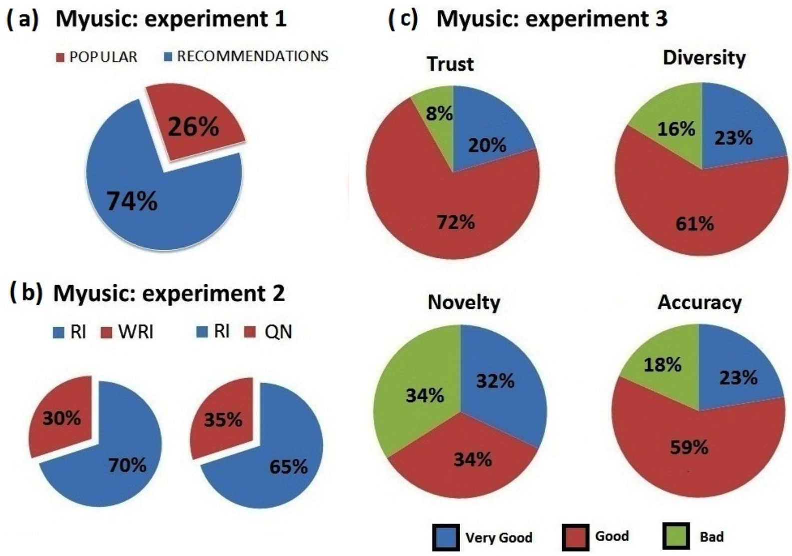

Fig. 4. Results of Experiments

content. In *Proceedings of the international conference on Multimedia*, MM '10, pages 391–400, New York, NY, USA, 2010. ACM.
2. C. Hahn, S. Turlier, T. Liebig, S. Gebhardt, and C. Roelle. Metadata Aggregation for Personalized Music Playlists. *HCI in Work and Learning, Life and Leisure*, pages 427–442, 2010.
3. J. Manyka, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers. Big data: The next frontier for innovation, competition, and productivity. Technical report, McKinsey Global Institute, 2011.
4. C. Musto. Enhanced vector space models for content-based recommender systems. In *Proceedings of the fourth ACM conference on Recommender systems*, RecSys '10, pages 361–364. ACM, 2010.
5. C. Musto, G. Semeraro, P. Lops, and M. de Gemmis. Random indexing and negative user preferences for enhancing content-based recommender systems. In *EC-Web*, pages 270–281, 2011.
6. A. Passant and Y. Raimond. Combining Social Music and Semantic Web for Music-Related Recommender Systems. In *Social Data on the Web, Workshop of the 7th International Semantic Web Conference*, Karlsruhe, Deutschland, Oktober 2008.
7. Giovanni Semeraro, Stefano Ferilli, Nicola Fanizzi, and Fabio Abbattista. Learning interaction models in a digital library service. In Mathias Bauer, Piotr J. Gmytrasiewicz, and Julita Vassileva, editors, *User Modeling*, volume 2109 of *Lecture Notes in Computer Science*, pages 44–53. Springer, 2001.
8. U. Shardanand. Social information filtering for music recommendation. Bachelor thesis, Massachusetts Institute of Technology, Massachusetts, 1994.
9. D. Wang, T. Li, and M. Ogihara. Are tags better than audio? the effect of joint use of tags and audio content features for artistic style clustering. In *ISMIR*, pages 57–62, 2010.