A Comparative Analysis of Personality-Based Music Recommender Systems

Melissa Onori
Department of Engineering
Roma Tre University
Via della Vasca Navale, 79
00146 Rome, Italy
melissa.onori@gmail.com

Alessandro Micarelli
Department of Engineering
Roma Tre University
Via della Vasca Navale, 79
00146 Rome, Italy
micarel@dia.uniroma3.it

Giuseppe Sansonetti
Department of Engineering
Roma Tre University
Via della Vasca Navale, 79
00146 Rome, Italy
gsansone@dia.uniroma3.it

ABSTRACT

This article describes a preliminary study on considering information about the target user's personality in music recommender systems (MRSs). For this purpose, we devised and implemented four MRSs and evaluated them on a sample of real users and real-world datasets. Experimental results show that MRSs that rely on purely users' personality information are able to provide performance comparable with those of a state-of-the-art MRS, even better in terms of the diversity of the suggested items.

Keywords

Personality; music recommendation; evaluation

1. INTRODUCTION

Music plays an important role in entertainment and leisure of human beings. With the advent of Web 2.0, a huge amount of music content has been made available to millions of people around the world. This has provided new opportunities for researchers working on music information with the aim of creating new services that support navigation, discovery, sharing, and the development of online communities among users. Music recommender systems (MRSs) aim to predict what people like to listen to. A recent research field in music recommendation explores the possibility of harnessing information on the target user's personality in the recommendation process.

The goal of the research work described in this paper is to assess the potential benefits of such integration. To this end, we implemented and compared with each other different MRSs, three of them based on users' personality inferred from explicit and implicit feedbacks, and one that does not consider users' personality.

2. RELATED WORK

In the research literature, there exist several works that reveal how information about a user's personality can help infer her music preferences and contribute to a more accurate recommendation process [31]. Therefore, several noteworthy MRSs considering the active user's personality have been proposed. Among others, Ferwerda and Schedl [12] propose an approach where users' personality and emotional states are implicitly extracted by analyzing their microblogs on Twitter. The authors make use of the extraction techniques described by Golbeck [14] and Quercia *et al.* [30], also trying to combine them for better predictions. Hu and Pu [18] compare a personality test-based MRS with a classic rating-based one. The authors point out that users are more inclined to results returned from the former. According to Hu and Pu, the active user perceives less effort and less time to use the personality test-based MRS. They further claim that users show a strong intention to use such MRS again and an unexpected surprise in its results, as they feel that the personality-based approach is able to reveal their hidden preferences, thereby improving the recommendation process. Also Tkalčič *et al.* [34] show that recommenders based on Big Five data can outperform rating-based recommenders. In [19], Hu and Pu consider again their previous results, exploring the use of personality tests for creating psychological profiles of user's friends as well. They enable the MRS to generate recommendations for users and their friends too. They also suggest that personality-based MRSs are preferred by no music connoisseurs, which do not know their music preferences in depth.

3. PERSONALITY

Generally speaking, an individual's personality can be defined as a combination of characteristics and qualities that make up the way she thinks, feels, and behaves in different situations [33]. Personality and emotions shape our everyday life, having a strong influence on our tastes [32], decisions [29], purchases [6], and general behavior [7]. It has been shown that people with similar personalities turn out to have similar preferences [8]. However, giving a more rigorous definition of personality can be challenging, so different theories have been formulated to specifically make easier the comprehension of self and others [9]. Each of these theories differently addresses the problem of representing and characterizing the human personality. We are interested in theories that would allow us to differentiate people from each other through measurable traits. The subject of the psychology of personality traits is the study of the psychological differences between individuals and relies on empirical research.

Initially, it was studied and defined by Gordon W.Allport [1, 2], which specified 17953 specific traits to describe an individual's personality. Then, particular effort was devoted to the attempt of limiting the number of traits that would otherwise be unmanageable. This led to the definition of the well-known Big Five Model [11]. After several revisions, the Big Five factors were finally labeled as follows [24]:

- *Extraversion*;

- *Agreeableness*;

- *Conscientiousness*;

- *Neuroticism*;

- *Openness (to experience)*.

In spite of several criticisms (e.g., [21]), such model is widely adopted in various fields, ranging from Medicine to Business. From the computer science point of view, personality traits include a set of human characteristics that can be modeled and implemented, for example, in personalized services. Prediction of personality traits can be accomplished explicitly (e.g., by administering personality tests), or implicitly (e.g., by monitoring the user's behavior).

3.1 Explicit Acquisition

Nowadays, questionnaires are the most popular method for extracting an individual's personality. They consist of a more or less large number of different questions, which are directly related to the granularity of the traits to be determined. Nunes *et al.* [28] show that the number of items influences the accuracy of measurements of traits. As expected, the higher the number of items, the more accurate the traits extracted. In particular, personality tests based on the Big Five Model are numerous and varied. A reasonable trade-off between accuracy and ease of use is represented by the Big Five Inventory (BFI) [4]. The 44-item BFI has been developed to create a brief questionnaire for efficient and flexible inference of the five factors, without the need to define more individual facets [21].

3.2 Implicit Acquisition

An individual's online behavior has long been the subject of many studies in the social sciences [3, 7, 25]. Results in cognitive psychology show that the general factors of personality can predict the aspects of the Internet use [25]. In fact, personality traits can be reflected in users' actions and ways of surfing the Web [3, 10, 27]. There are also studies that investigate the possibility of inferring the user's personality by user-generated content on social networks such as Facebook and Twitter. For instance, Gao *et al.* derive users' personality traits from their microblogs [13]. Golbeck *et al.* identify users' personality traits by analyzing their Facebook profiles, including peculiarities of language, business, and personal information [15]. Moreover, Golbeck *et al.* [14] and Quercia *et al.* [30] predict users' personality from Twitter, by examining their tweet content and observing their characteristics (e.g., popularity, influential users, etc.). Kosinski *et al.* [23] show that *likes* on Facebook can be used to automatically and accurately predict a set of personal attributes, including personality traits. For instance, the accuracy of prediction of the *Openness* factor is similar to the accuracy that can be obtained through a classic personality test, with

the advantage of not having to force the user to answer a significant number of questions. Along this direction, the authors developed the Apply Magic Sauce (AMS)[1] that allows for the prediction of users' personality from the analysis of their activities on Facebook. Such application, developed at the University of Cambridge Psychometrics Centre, relies on over six million social media profiles and determines personality traits through psychometric evaluations, as described in [23]. The model is based on the dataset of myPersonality project [2].

4. USER STUDY

In this section, we describe the dataset, the setup, and the results of the experimental evaluation.

4.1 Dataset

The experimental tests were performed on the Last.fm [3] music listening data kindly provided by the researchers of myPersonality project [22] and Liam McNamara [26]. From this data, we extracted 1,875 Last.fm users with related information about personality tests and listening histories. The user's preferences were inferred from the *playcount* attribute, which denotes how many times the user listened to that particular song. The final value is obtained by normalizing it between 1 and 5.

4.2 Users

The users who took part in the experimental trials were 65, all of them with an active Facebook account. Their characteristics in terms of gender, age, occupation, and education are illustrated in Tables 1, 2, 3, 4, respectively.

Table 1: Gender

Female	Male
27	38

Table 2: Age

0-18	19-24	25-29	30-35	36-45	46-55	56-65
2	25	26	2	5	4	1

Table 3: Occupation

None	Student	Employee	Professional	Housewife
6	35	21	2	1

Table 4: Education

Primary	Secondary	Bachelor	Master	PhD
6	29	18	10	2

---

[1]http://applymagicsauce.com/

[2]www.mypersonality.org

[3]http://www.last.fm/

4.3 Setup

For presenting the user with the suggested playlists we designed a simple interface that allows for a quick and easy use of the system. Furthermore, we made us of the Spotify APIs [4], which offer a preview of 30 seconds of each song in the playlist. We deemed such time enough for the user to understand whether a given song is to her liking or not. Moreover, listening to the whole playlist is short, thus avoiding that the user will get bored and stop listening to the recommended songs. In this way, the user will be able to express a well-founded opinion.

Each user was required to test all MRSs and evaluate the returned playlists. MRSs were proposed in a random order and with the user completely unaware of their details. Ratings expressed by users in the evaluation phase were related to *novelty*, *serendipity*, *diversity*, *interest*, and *future use*. To this end, each user was asked to provide an assessment in relation to the following five statements:

1. "I found new songs by artists already known to me." *(novelty)*

2. "I found songs by artists that I did not know and, as of now, will begin to listen to." *(serendipity)*

3. "I found songs by artists of different music genres." *(diversity)*

4. "I found the suggested playlist interesting." *(interest)*

5. "I would use this MRS again in the future." *(future use)*

For each of these statements the user could express a numerical value in a Likert 5-point scale (i.e., 1: strongly disagree, 5: strongly agree). In addition, the user could leave a feedback as well.

4.4 Music Recommender Systems

In this section, we introduce the music recommender systems (MRSs) developed as part of our research work.

4.4.1 MRS based on Relations between Explicit Personality and Music Genres

The first MRS acquires information on the target user's personality explicitly, by administering a personality test. We chose the 44-item Big Five Inventory test introduced in Section 3.1, as its length represents an appropriate trade-off between compilation time and results accuracy. Such test is proposed to the target user through a web interface. Once the test is completed, the system analyzes the responses and computes the Big Five factors. In [8], the relations between users' personality types and their preferences in multiple entertainment domains are investigated. The authors derive a set of association rules that connect the Big Five factors with music genres. Based on those rules, this MRS returns the resulting playlist to the user.

4.4.2 MRS based on Explicit Personality and Neighbors

Even the second MRS relies on the user's personality explicitly inferred through the use of the questionnaire. The recommendation mechanism, however, is different. More

---

[4]https://developer.spotify.com/web-api/

---

precisely, this MRS identifies the most similar users to the target one within a dataset containing information related to personality and music habits of a group of Last.fm users. The user u's personality is compared to that of each user v in the dataset by computing the cosine similarity applied to the Big Five factors, which is defined as follows:

$$simp(u,v) = \frac{\sum_{k=1}^{5} p_u^k \times p_v^k}{\sqrt{\sum_{k=1}^{5}(p_u^k)^2}\sqrt{\sum_{k=1}^{5}(p_v^k)^2}} \qquad (1)$$

where p_u^k expresses the value of the Big Five factor k of the user u. Based on such values, the system selects the ten Last.fm users most similar to the user u and generates a playlist from their listening histories.

4.4.3 MRS based on Implicit Personality and Neighbors

The implicit personality acquisition can be carried out by analyzing the user's behavior on the Web, especially on social networks. To this end, we used the APIs of the Apply Magic Sauce (AMS) application introduced in Section 3.2. In order to infer the user's personality, AMS analyzes how she assigns *likes* on Facebook. For such reason, the system allows users to login via Facebook. In this way, AMS enters the user profile, extracts the required information, and returns the predicted information, such as age, intelligence, life satisfaction, interest in specific areas, and her personality traits. Based on such features, the MRS identifies the most similar users to the target one within the Last.fm dataset by computing the similarity function 1. From the information related to the music such users listen to, the MRS builds the personalized playlist for the active user. However, this MRS has a drawback: it is necessary that the user has inserted a sufficient number of *likes* in her profile. Otherwise, the AMS application is not able to predict the user's personality and, as a result, the MRS is not able to deliver any playlist.

4.4.4 MRS based on Music Preferences

This MRS does not exploit information about the user's personality, and has been realized as a baseline to be used in the experimental evaluation. The recommender works as follows. The user is presented with a screenshot of the images of ten songs belonging to the Last.fm top track, and is asked to choose her favorites. Alternatively, the user can enter the title of some of her favorite songs. After that, the system leverages the Last.fm APIs to retrieve songs similar to those chosen by the user and includes them in the suggested playlist. Even though the actual algorithm underlying the Last.fm recommender is unknown, it is reasonable to assume that it mostly relies on collaborative filtering and tagging activity.

4.5 Results

Experimental results are shown in Table 5. In the description of the experimental results, the implemented MRSs are denoted as follows:

I: MRS based on relations between explicit personality and music genres;

II: MRS based on explicit personality and neighbors;

IIII: MRS based on implicit personality and neighbors;

IV: MRS based on music preferences.

Table 5: Results in terms of mean and standard deviation of user ratings

MRS	# of Users	Novelty	Serendipity	Diversity	Interest	Future Use
I	65	2.5 - 1.0	2.5 - 0.8	3.0 - 0.9	3.0 - 0.8	3.4 - 0.8
II	65	2.4 - 0.9	2.6 - 0.8	2.8 - 0.8	3.2 - 0.7	3.3 - 0.8
III	43	2.2 - 0.7	2.2 - 0.6	3.2 - 0.9	2.4 - 0.7	2.8 - 0.9
IV	65	2.9 - 0.8	2.4 - 0.9	1.7 - 0.5	3.5 - 0.6	3.5 - 0.6

The reason for the smaller number of users who experienced the third MRS (i.e., the one based on implicit personality and neighbors) was that not all testers had a sufficient number of *likes* on Facebook to enable the AMS application to predict their personality. It can be noted that the first three systems received very similar assessments, as regards novelty, serendipity, and diversity. Precisely, novelty values are not high, because we asked users if new songs by known artists were in the suggested playlists, not if new artists were in the playlists. Serendipity shows similar values to novelty. Diversity values are quite high, which is obviously positive, since in this way the user can broaden her music knowledge, having a more varied set of possible music listening. The playlist was judged interesting for each system, a bit less for the third one. Users also showed some interest in the reuse of MRSs, a bit less for the third, where the result revealed some skepticism, due to the lower interest in the returned playlist. Different results emerged from the user evaluation of the last system. As expected, the results were higher than the others, except for serendipity (in line with the others) and diversity (lower). In fact, for such MRS the target user directly inserts her preferences. As a result, she was more interested in the suggested playlist and showed higher intention in reusing that recommender. These results may also be related to the difficulty that users appreciate new songs on the first listening and nourish curiosity in music genres different from their usual ones.

5. CONCLUSIONS AND FUTURE WORK

The research work presented here analyzed the effects of integrating the target user's personality in music recommender systems (MRSs). To this end, four different MRSs were developed. Three of them were only personality-based, the fourth did not take into account users' personality at all. The experimental results show that the personality-based ones had performance almost similar to that of a classic MRS. They also prove that the former ones are able to recommend songs with higher diversity than those suggested by the latter one.

This research effort is just beginning, so the possible future developments are manifold. Among others, the extension of the type and number of MRSs to be compared with each other, and the inclusion of music preferences and sentiments extracted from music reviews [16, 17] in the user model. Furthermore, as regards the experimental procedure, we intend to broaden the number of involved users and tested datasets, and to develop a layered evaluation for distinguishing the contributions of the user model from those of the user interface.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] F. H. Allport and G. W. Allport. *Personality traits: their classification and measurement.* Yale University Press, 1921.

[2] G. W. Allport. Concepts of trait and personality. *Psychological Bulletin*, 24:284–293, 1927.

[3] Y. Amichai-Hamburger and G. Vinitzky. Social network use and personality. *Computers in Human Behaviour*, 26(6):1289–1295, 2010.

[4] V. Benet-Martinez and O. E. John. Los cinco grandes across cultures and ethnic groups: multitrait multimethod analyses of the big five in spanish and english. *Journal of Personality and Social Psychology*, pages 729–750, 1998.

[5] S. Berkovsky, E. Herder, P. Lops, and O. C. Santos, editors. *Late-Breaking Results, Project Papers and Workshop Proceedings of the 21st Conference on User Modeling, Adaptation, and Personalization., Rome, Italy, June 10-14, 2013*, volume 997 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2013.

[6] C. Bologna, A. C. De Rosa, A. De Vivo, M. Gaeta, G. Sansonetti, and V. Viserta. Personality-based recommendation in e-commerce. In Berkovsky et al. [5].

[7] B. Caci, M. Cardaci, M. E. Tabacchi, and F. Scrima. Personality variables as predictors of facebook usage. *Psychological reports*, 113(2):528–539, 2014.

[8] I. Cantador, I. Fernández-Tobías, and A. Bellogín. Relating personality types with user preferences in multiple entertainment domains. In Berkovsky et al. [5].

[9] D. S. Cartwright. *Theories and Models of Personality.* W. C. Brown Company, 1979.

[10] F. Celli and L. Polonio. Relationship between personality and interactions in facebook. In X. M. Tu, A. M. White, and N. Lu, editors, *Social Networking: Recent Trends, Emerging Issues and Future Outlook*, chapter 3, pages 41–53. Nova Science Publishers, Inc., 2013.

[11] P. T. Costa and R. R. McCrea. *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI).* Psychological Assessment Resources, Inc., Odessa, Fla. P.O. Box 998, Odessa 33556, 1992.

[12] B. Ferwerda and M. Schedl. Enhancing music recommender systems with personality information and emotional states: A proposal. In I. Cantador, M. Chi, R. Farzan, and R. Jäschke, editors, *Posters, Demos, Late-breaking Results and Workshop Proceedings of the 22nd Conference on User Modeling,*

Adaptation, and Personalization co-located with the 22nd Conference on User Modeling, Adaptation, and Personalization (UMAP2014), Aalborg, Denmark, July 7-11, 2014., volume 1181 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2014.

[13] R. Gao, B. Hao, S. Bai, L. Li, A. Li, and T. Zhu. Improving user profile with personality traits predicted from social media content. In *Proceedings of the 7th ACM Conference on Recommender Systems*, RecSys '13, pages 355–358, New York, NY, USA, 2013. ACM.

[14] J. Golbeck, C. Robles, M. Edmondson, and K. Turner. Predicting personality from twitter. In *PASSAT/SocialCom 2011, Privacy, Security, Risk and Trust (PASSAT), 2011 IEEE Third International Conference on and 2011 IEEE Third International Conference on Social Computing (SocialCom), Boston, MA, USA, 9-11 Oct., 2011* [20], pages 149–156.

[15] J. Golbeck, C. Robles, and K. Turner. Predicting personality with social media. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 253–262, New York, NY, USA, 2011. ACM.

[16] D. F. Gurini, F. Gasparetti, A. Micarelli, and G. Sansonetti. A sentiment-based approach to twitter user recommendation. In B. Mobasher, D. Jannach, W. Geyer, J. Freyne, A. Hotho, S. S. Anand, and I. Guy, editors, *Proceedings of the Fifth ACM RecSys Workshop on Recommender Systems and the Social Web co-located with the 7th ACM Conference on Recommender Systems (RecSys 2013), Hong Kong, China, October 13, 2013.*, volume 1066 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2013.

[17] D. F. Gurini, F. Gasparetti, A. Micarelli, and G. Sansonetti. iscur: Interest and sentiment-based community detection for user recommendation on twitter. In V. Dimitrova, T. Kuflik, D. Chin, F. Ricci, P. Dolog, and G. Houben, editors, *User Modeling, Adaptation, and Personalization - 22nd International Conference, UMAP 2014, Aalborg, Denmark, July 7-11, 2014. Proceedings*, volume 8538 of *Lecture Notes in Computer Science*, pages 314–319. Springer, 2014.

[18] R. Hu and P. Pu. A comparative user study on rating vs. personality quiz based preference elicitation methods. In *Proceedings of the 14th International Conference on Intelligent User Interfaces*, IUI '09, pages 367–372, New York, NY, USA, 2009. ACM.

[19] R. Hu and P. Pu. A study on user perception of personality-based recommender systems. In P. D. Bra, A. Kobsa, and D. N. Chin, editors, *User Modeling, Adaptation, and Personalization, 18th International Conference, UMAP 2010, Big Island, HI, USA, June 20-24, 2010. Proceedings*, volume 6075 of *Lecture Notes in Computer Science*, pages 291–302. Springer, 2010.

[20] IEEE. *PASSAT/SocialCom 2011, Privacy, Security, Risk and Trust (PASSAT), 2011 IEEE Third International Conference on and 2011 IEEE Third International Conference on Social Computing (SocialCom), Boston, MA, USA, 9-11 Oct., 2011.* IEEE, 2011.

[21] O. P. John and S. Srivastava. The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin and O. P. John, editors, *Handbook of Personality: Theory and Research*, pages 102–138. Guilford Press, New York, second edition, 1999.

[22] M. Kosinski, S. C. Matz, S. D. Gosling, V. Popov, and D. Stillwell. Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. *American Psychologist*, 70(6):543–556, Sept. 2015.

[23] M. Kosinski, D. Stillwell, and T. Graepel. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15):5802–5805, Apr. 2013.

[24] R. R. McCrae and O. P. John. An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality*, 60(2):175–215, 1992.

[25] J. C. Mcelroy, A. R. Hendrickson, A. M. Townsend, and S. M. Demarie. Dispositional factors in internet use: Personality versus cognitive style. *MIS Quarterly*, 31(4):809–820, 2007.

[26] L. McNamara, C. Mascolo, and L. Capra. Media sharing based on colocation prediction in urban transport. In *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking*, MobiCom '08, pages 58–69, New York, NY, USA, 2008. ACM.

[27] K. Moore and J. C. McElroy. The influence of personality on facebook usage, wall postings, and regret. *Comput. Hum. Behav.*, 28(1):267–274, Jan. 2012.

[28] M. A. Nunes, J. Bezerra, and A. de Oliveira. Personalityml: A markup language to standardize the user personality in recommender systems. *GEINTEC - Gestão, Inovação e Tecnologias*, 2(3):255–273, 2012.

[29] M. A. S. Nunes and R. Hu. Personality-based recommender systems: An overview. In *Proceedings of the Sixth ACM Conference on Recommender Systems*, RecSys '12, pages 5–6, New York, NY, USA, 2012. ACM.

[30] D. Quercia, M. Kosinski, D. Stillwell, and J. Crowcroft. Our twitter profiles, our selves: Predicting personality with twitter. In *PASSAT/SocialCom 2011, Privacy, Security, Risk and Trust (PASSAT), 2011 IEEE Third International Conference on and 2011 IEEE Third International Conference on Social Computing (SocialCom), Boston, MA, USA, 9-11 Oct., 2011* [20], pages 180–185.

[31] D. Rawlings and V. Ciancarelli. Music Preference and the Five-Factor Model of the NEO Personality Inventory. *Psychology of Music*, 25(2):120–132, 1997.

[32] P. J. Rentfrow and S. D. Gosling. The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, 84(6):1236–1256, 2003.

[33] D. L. Schacter, D. T. Gilbert, D. M. Wegner, and B. M. Hood. *Psychology: Second European Edition.* Worth Publishers. Palgrave Macmillan Limited, 2015.

[34] M. Tkalcic, M. Kunaver, J. Tasic, and A. Košir. Personality based user similarity measure for a collaborative recommender system. In *Proceedings of the 5th Workshop on Emotion in Human-Computer Interaction-Real world challenges*, pages 30–37, 2009.