Recommending Music for Places of Interest in a Mobile Travel Guide

Matthias Braunhofer
Free University of Bolzano
Piazza Domenicani 3
Bolzano, Italy
mbraunhofer@unibz.it

Marius Kaminskas
Free University of Bolzano
Piazza Domenicani 3
Bolzano, Italy
mkaminskas@unibz.it

Francesco Ricci
Free University of Bolzano
Piazza Domenicani 3
Bolzano, Italy
fricci@unibz.it

ABSTRACT

Context-aware music recommender systems suggest music items taking into consideration contextual conditions, such as the user mood or location, that may influence the user preferences at a particular moment. In this paper we consider a particular kind of context-aware recommendation task: selecting music suited for a place of interest (POI), which the user is visiting, and that is illustrated in a mobile travel guide. We have designed an approach for this novel recommendation task by matching music to POIs using emotional tags. In order to test our approach, we have developed a mobile application that suggests an itinerary and plays recommended music for each visited POI. The results of the study show that users judge the recommended music suited for the POIs, and the music is rated higher when it is played in this usage scenario.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*information filtering*

General Terms

Design, Experimentation, Human Factors

Keywords

Music recommender systems, context awareness, mobile services, tags

1. INTRODUCTION

Music recommender systems are decision support tools that reduce the information overload by retrieving relevant music items based on a user's profile. The recommendation process can be content-based, i.e., using features of the music liked by the user to predict what else the target user may like [2], or collaborative-based, which finds users with similar music preferences and recommends to the target user

items liked by these users [5]. However, most of the available music recommender systems suggest music regardless of the contextual conditions which can be important to predict the user's preferences at a particular moment. In fact, a study on users' musical information needs [4] has shown that people often seek music for a contextual situation like an occasion, an event or an emotional state rather than by artist or song information. To cope with these needs, recently there has been an emerging interest in contextual, or situational music selection [6, 1]. The idea of such music selection is to recommend music depending on the user's actual situation, emotional state, or any other contextual condition that might be relevant to increase the user's satisfaction for the selected music.

In this line of research, we are considering the problem of retrieving music that fits the surroundings of the user, and more specifically, we are looking for effective ways to select musical content suited for the place of interest (POI) that the user is visiting or browsing using an information service. The intuition is that a Vivaldi's concerto sounds as a better choice for a "calle" (a narrow street) in Venice rather than a Bach's organ fugue, which may better suit an old gothic church. Being able to select music for a place can be used for creating new engaging location-aware music delivery services. In particular, in this paper we are considering a scenario where a tourist is sightseeing a city (Bolzano, Italy) using a mobile city guide. The guide recommends a walking itinerary and, while the user is visiting the suggested POIs, it plays music that suits the visited POIs. The goal is to enhance the user's experience, to create a more engaging travel guide tool, and to recommend music that could be better evaluated by the user.

In a previous paper [3] we have illustrated our approach to match music tracks to POIs; it exploits user-assigned emotional tags to both music tracks and POIs. The approach was evaluated in an online experiment where the users were required to evaluate the appropriateness of the music selected by the system for POIs. The experiment was carried out using a web application where the POI descriptions were shown and music played in the background. Since such "simulated" environment can not fully reflect the real-world settings of a visited POI (i.e., the surroundings, weather conditions, other people around) it was crucial to further evaluate our approach in real-world settings. Therefore, we have implemented the above mentioned mobile city guide for Android phones and conducted a live user study. The main goals of this study were to test the following hypotheses: a) users agree with the music-to-POI match produced by our
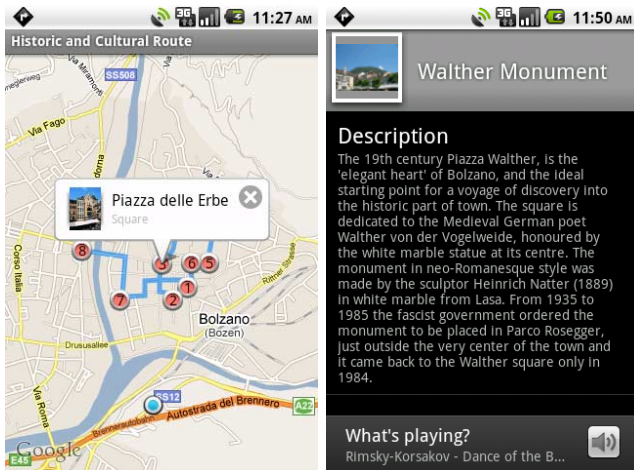
Figure 1: Sample screenshots of the application

approach, and b) users tend to rate the selected music tracks higher in this mobile, in-context, usage scenario, compared with the rating situation where the context is not defined.

The rest of the paper is structured as follows: in section 2 we describe the design of the mobile guide, section 3 describes the evaluation procedure. The results of the study are presented in section 4. Finally, conclusions are drawn and future work directions defined in section 5.

2. ADAPTING MUSIC TO POI

This section briefly describes the usage scenario and the technologies used to develop PlayingGuide: a travel guide that illustrates the POI the user is close to and plays music suited for that POI. Users of PlayingGuide are tourists, new to Bolzano, interested in exploring some of the city's POIs. After the user has launched this Android application she may choose a travel itinerary that is displayed on a map indicating the user's current GPS position and the locations of the POIs in the itinerary (see Figure 1, left). Then, every time the user is nearby to a POI (in the selected itinerary or not), she receives a notification alert conveying information about the POI. While the user is reading this information, the system plays a music track that suits the POI (see Figure 1, right). For example, the user might hear "Bach – Air" while visiting the Cathedral of The Assumption of Our Lady and "Rimsky-Korsakov – Dance of the Bumble Bee" during a visit to the Walther Monument.

PlayingGuide has been implemented in a fat client architecture, i.e., the entire application runs locally on the mobile device and allows to synchronize local data changes with a remote server. The guide recommends music for a POI in the following two steps:

Step 1. Given a POI and a set of music tracks as input, the first step to generate a music recommendation is to compute the similarity between this POI and the available music tracks.

Our dataset consists of 75 music tracks (classical music and movie soundtracks) and 32 POIs in Bolzano. Both POIs and music tracks have been tagged using a restricted tag vocabulary in a separate user study [3]. The tag vocabulary consists of 33 emotion adjectives from the Geneva Emotional Music Scale [9] and 13 adjectives describing the physical

qualities (color, temperature, etc.): 46 adjectives in total. On average, a POI is annotated with 16.34 tags and a music track with 13.67 tags.

In order to compute the music-to-POI similarity, both POIs and music tracks are represented as vectors, with one component corresponding to each tag in our tag vocabulary, together with a weight for each component. The weight of tag t with respect to a POI (or music track) d is given by:

$$w_{t,d} = \begin{cases} -\log p_t & \text{if } tf_{t,d} > 0 \\ 0 & \text{otherwise} \end{cases}, \qquad (1)$$

where $tf_{t,d}$ is the number of times d has been annotated with t, and p_t denotes the fraction of all POIs (or music tracks) annotated with t. Then, the similarity between a POI d_1 and a music track d_2 is computed as the weighted Jaccard similarity of their vector representations $\vec{V}(d_1)$ and $\vec{V}(d_2)$:

$$sim(d_1, d_2) = \frac{\sum_{i=1}^{M} \min(\vec{V}_i(d_1), \vec{V}_i(d_2))}{\sum_{i=1}^{M} \max(\vec{V}_i(d_1), \vec{V}_i(d_2))}. \qquad (2)$$

The major reason for choosing the weighted Jaccard similarity metric [8] (with the weights as shown in Equation 1) is that it has been shown to provide good matching music when compared with other similarity metrics [3]. In these experiments, users were presented, in a web application, with a sequence of POIs and for each POI a set of music tracks was suggested by some alternative similarity metrics and could be played. By measuring the user selections for the musics that in the user's opinion suit the presented POI, we could identify the best similarity metric, i.e., the metric that produced the largest proportion of music tracks actually selected by the users.

Step 2. Given the music-to-POI similarity scores, the final step before delivering a music track recommendation for a POI, is to sort the music tracks by decreasing music-to-POI similarity score and then to randomly pick out one of the top N (in our case 3) music tracks. The motivation for not always choosing the top-scoring music track for each POI is to avoid, or at least minimize, the probability that the same music tracks are played for POIs that have been annotated with similar sets of tags, and therefore to ultimately suggest more diverse music tracks while the user is visiting an itinerary.

3. USER STUDY EVALUATION

In order to evaluate the proposed music-to-POI matching approach, we compared the performance of PlayingGuide with an alternative system variant having the same user interface but not matching the music with the POI; it rather suggests music tracks that, according to our similarity metric, have low similarity with the POI. We call the original PlayingGuide variant MATCH, and this second variant MUSIC. We adopted a between-groups design, involving 26 subjects (researchers or students from our faculty). Subjects were randomly assigned to the MATCH and MUSIC variant in a random way (13 each). We note that the outcome of this comparison was not evident at all, since with a superficial evaluation, even the less similar tracks could be considered suited, and there are not large differences among these tracks (all of them being popular classical or movie soundtrack music).

The subjects were instructed in: the purpose of the experiment; their task and the procedure of the experiment; and in the usage of our test device, a Google Nexus One mobile phone. Following this introductory phase, each subject was given a phone with earphones, and was asked to complete the "Historic and Cultural Route" to visit various POIs in Bolzano. This route required subjects to walk approximately 45mins in the center of Bolzano. Whenever a subject was approaching a POI, either belonging to the route or not, a notification invited the user to inspect the POI's details and listen to the recommended music track. If the recommended music track was perceived as unsuited, subjects could also pick an alternative music track from a shuffled list of four possible alternatives: two randomly generated and two with high music-to-POI similarity scores.

Immediately after viewing the details of a POI and listening to the accompanying music, subjects were asked, in a feedback dialog, to answer to three questions related to the POI and the recommended music, namely: a) "How much did you like the place of interest?", b) "How much did you like the music?", and c) "Was it a good music for that place of interest?". The first two questions were to be rated on a five-star rating scale (with 1 star being the lowest score and 5 stars being the highest score), whereas the third question required a simple "Yes" or "No" answer.

After the "Historic and Cultural Route" had been completed, subjects were asked to fill out a paper-and-pencil questionnaire based on the Computer System Usability Questionnaire (CSUQ) [7] to assess the overall usability and effectiveness of the system. The subjects rated various statements on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree).

Finally, three months later, the subjects who participated to this evaluation were asked to rate again the music tracks that were recommended for the various POIs. The new ratings were collected through a simple web interface where the music tracks were played one by one. Since the web interface presented the music tracks without any reference to POIs, it enabled us to collect the subjects' ratings without any influence produced by the match between the POI and the music or the contextual situation of the visit.

4. EVALUATION RESULTS

Analyzing the logging data collected in the experiment, we found that the mean listening time (in seconds), to the music recommended for each POI, was slightly higher in the MATCH condition ($M = 38.43$, $SD = 24.56$) than in the MUSIC condition ($M = 36.91$, $SD = 23.40$). However, this difference is not statistically significant ($p = .58$ in a t-test).

By means of the feedback dialog on the phone, a total of 308 responses regarding the various visited POIs and their recommended music tracks were obtained: 157 (51%) from subjects in the MATCH group, and 151 (49%) from subjects in the MUSIC group. The mean ratings to the question "How much did you like the place of interest?" were similar across both conditions, being somewhat higher in the MUSIC condition ($M = 3.93$, $SD = 0.99$) than in the MATCH condition ($M = 3.78$, $SD = 1.08$). There is, however, no significant difference between the conditions: $p = .21$ in a t-test. This is as expected, since MATCH and MUSIC subjects visited almost the same POIs and they were asked to rate each POI independently from the recommended music track.

The mean rating for the question "How much did you like the music?" (posed by the feedback dialog on the phone as well) was 3.82 ($SD = 1.02$) in the MATCH condition and 3.53 ($SD = 1.17$) in the MUSIC condition. The observed difference is statistically significant: $p = .023$ in a t-test. This result seems to support the hypothesis that users like the suggested music more when matching the visited POIs. In fact, this could not be validated, since considering the ratings for the same music tracks collected via the web interface (see Table 1), i.e, where the users could rate the music tracks without any reference to POIs, we again found a significantly larger mean rating for the music tracks suggested by the MATCH variant compared to those suggested by the MUSIC variant. Table 1 shows the mean ratings acquired for the tracks in MATCH and MUSIC conditions via the feedback dialog on the phone (i.e., in context) and the mean ratings for the same tracks acquired via the web interface three months later (i.e., without context).

Rating acquisition condition	Group of music tracks	
	MATCH	MUSIC
In context (mobile)	3.78	3.34
Without context (web)	3.22	2.95

Table 1: Mean ratings for the music tracks in MATCH and MUSIC groups

Hence, the higher ratings given to music tracks by the subjects in the MATCH condition could also be determined by the fact that, in general, the users liked more these tracks. However, this data validate the hypothesis that listening to a music track on the mobile device, in this particular situation, has the effect of increasing the rating for the music track. In fact, in both the MATCH and MUSIC groups a two-tailed, paired, t-test shows a significant difference between the mean ratings acquired with the mobile phone (in context) and on the Web interface (without context): for the MATCH group $p < .001$; and for the MUSIC group $p = .03$.

In the final, and more important, evaluation of the effectiveness of our proposed music-to-POI matching approach, we measured the proportion of "Yes" answers to the question "Was it a good music for that place of interest?". This was substantially higher in the MATCH condition (0.77) than in the MUSIC condition (0.60). This difference in proportions is statistically significant, $\chi^2(1, N = 308) = 10.89$, $p < .001$. We can conclude that users evaluate the music tracks recommended with our proposed method to better suit the POIs than the music tracks suggested in the control setting.

Moreover, as an additional indication of the benefit of matching music to POI, when the subjects, possibly unsatisfied with recommended music, freely selected an alternative track from the suggested list of four options (two matching and two random), they strongly preferred music matching with the POI. Out of 77 manual music selections, 58 (75%) were chosen from the tracks matching to the POI and 19 (25%) from the randomly suggested tracks, i.e., the probability that a user selects a matched music track is about three times higher than that of selecting a random music track. This preference for matched music tracks is also statistically significant, $\chi^2(1, N = 77) = 19.75$, $p < .001$, which proves our hypothesis that users prefer tracks for POIs that are generated by our music-to-POI matching approach.

Finally, Table 2 illustrates the ratings given by the subjects to each statement in the usability questionnaire. Both variants received very positive responses. In general, the differences between the groups' mean ratings are not statistically significant. This is not surprising, since both groups tested the same system but with different music recommendation approaches. There is, however, a marginally significant difference in statement 11 ("The music was correctly selected for each POI."), $p = .051$ in a t-test, indicating again that MATCH subjects perceived the music recommended and played for the various POIs as more appropriate compared to MUSIC subjects.

Statement	MATCH mean (SD)	MUSIC mean (SD)
1. It was simple to use this system.	6.08 (0.51)	6.46 (0.66)
2. I am able to complete my work quickly using this system.	5.58 (1.00)	5.54 (1.27)
3. I feel comfortable using this system.	6.00 (0.60)	5.92 (1.12)
4. It was easy to learn to use this system.	6.17 (0.83)	6.54 (0.78)
5. Whenever I make a mistake using the system, I recover easily and quickly.	5.60 (1.07)	5.20 (1.62)
6. The information provided with this system is clear.	5.90 (1.29)	5.92 (1.04)
7. It is easy to find the information I needed.	6.00 (1.18)	5.77 (0.93)
8. The organization of information on the system screens is clear.	6.08 (1.24)	6.31 (1.11)
9. The interface of this system is pleasant.	6.25 (0.62)	6.69 (0.63)
10. I like using the interface of this system.	6.17 (0.83)	6.38 (0.65)
11. The music was correctly selected for each POI.	**5.00** (0.74)	**4.08** (1.38)
12. I liked the music played for each POI.	5.08 (0.67)	4.38 (1.98)
13. I would recommend it to a friend.	6.00 (0.74)	5.92 (1.19)
14. Overall, I am satisfied with this system.	6.00 (0.74)	6.00 (0.82)

Table 2: Usability questionnaire ratings

5. CONCLUSIONS AND FUTURE WORK

In this paper we have presented PlayingGuide, a novel mobile context-aware recommender system, that suggests and plays music tracks while users are visiting POIs. PlayingGuide matches music tracks to POIs, based on tags that users have assigned to them. To achieve this, we have chosen a similarity metric between tagged resources, and used it to compute the matching between music tracks and POIs. For a given POI, a music track with a high music-to-POI similarity score is finally provided as a recommendation to the user in the context of the visit to the POI.

We have formulated the following two experimental hypotheses: a) users agree with the music recommendations generated by our music-to-POI matching approach, and b) users consider the selected music tracks as more appealing if they are suggested and played on the implemented mobile application in the context of the visit to a POI. In a live user study, we were able to confirm our hypotheses.

The topic of matching music to POIs is relatively new, and there are several research questions that deserve future work, e.g., what is the impact of alternative similarity measures on the quality of music recommendations, how can our approach be extended to other music genres than just classical and movie soundtracks, what alternative techniques other than tagging can be used to match music to POIs, and how other contextual factors in addition to the POI must influence the music selection process. Nevertheless, our results already demonstrate that the proposed approach could be used to create new and appealing music retrieval services.

6. REFERENCES

[1] R. Cai, C. Zhang, C. Wang, L. Zhang, and W.-Y. Ma. Musicsense: contextual music recommendation using emotional allocation modeling. In *Proceedings of the 15th International Conference on Multimedia*, pages 553–556, New York, NY, USA, 2007. ACM.

[2] K. Hoashi, K. Matsumoto, and N. Inoue. Personalization of user profiles for content-based music retrieval based on relevance feedback. In *Proceedings of the 11th ACM international conference on Multimedia*, pages 110–119, New York, NY, USA, 2003. ACM.

[3] M. Kaminskas and F. Ricci. Location-adapted music recommendation using tags. In *UMAP 2011 - 19th International Conference on User Modeling, Adaptation, and Personalization*, 2011.

[4] J. Y. Kim and N. J. Belkin. Categories of music description and search terms and phrases used by non-music experts. In *Proceedings of the 3rd International Conference on Music Information Retrieval*, Paris, France, 2002.

[5] I. Konstas, V. Stathopoulos, and J. M. Jose. On social networks and collaborative recommendation. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 195–202, New York, NY, USA, 2009. ACM.

[6] J. Lee and J. Lee. Context awareness by case-based reasoning in a music recommendation system. In *Ubiquitous Computing Systems*, pages 45–58. Springer, 2007.

[7] J. Lewis. IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. *International Journal of Human-Computer Interaction*, 7(1):57–78, 1995.

[8] B. Markines, C. Cattuto, F. Menczer, D. Benz, A. Hotho, and G. Stumme. Evaluating similarity measures for emergent semantics of social tagging. In *Proceedings of the 18th international conference on World Wide Web*, pages 641–650. ACM, 2009.

[9] M. Zentner, D. Grandjean, and K. R. Scherer. Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, 8(4):494–521, August 2008.