

Listener-aware Music Recommendation from Sensor and Social Media Data

Markus Schedl

Department of Computational Perception
Johannes Kepler University, Linz, Austria
<http://www.cp.jku.at>

Abstract. Music recommender systems are lately seeing a sharp increase in popularity due to many novel commercial music streaming services. Most systems, however, do not decently take their listeners into account when recommending music items. In this note, we summarize our recent work and report our latest findings on the topics of tailoring music recommendations to individual listeners and to groups of listeners sharing certain characteristics. We focus on two tasks: *context-aware automatic playlist generation* (also known as serial recommendation) using sensor data and *music artist recommendation* using social media data.

1 Introduction

The importance of incorporating user characteristics and contextual aspects into recommender systems has been acknowledged many times [16, 17, 1]. Research that looks into this matter in the domain of music recommendation is scarce, though. Addressing this issue, we summarize our latest work on the tasks of (i) automatic music playlist generation incorporating contextual aspects of the listener and (ii) music artist recommendation tailored according to various user and listening characteristics. Both tasks are highly related to machine learning and data mining. In fact, we approach the former task by gathering a wide variety of listener-centric sensor data from a smart phone app and exploiting machine learning techniques to learn relationships between these features and music metadata (e.g. artist or track name). The latter task is related to data mining as we acquire and analyze huge amounts of listening events produced by users of social media and build recommendation algorithms that consider personal characteristics of the listeners and their listening behavior, also using novel features mined from user-generated data.

2 Automatic Playlist Adaptation based on Sensor Data

Addressing the task of **automatic playlist generation**, we developed an *Android* app for smart devices, dubbed “Mobile Music Genius” (MMG) [8, 6], which collects a variety of user-specific features during playback, ranging from time, location, and weather to ambient noise, light level, and motion. In addition, MMG gathers music metadata (artist, track, mood, and genre) and records player interaction (e.g. play, skip, pause events). Integrating a decision tree algorithm that is trained and retrained on the contextual features using track names as

classes, MMG continuously monitors the context feature values of the listener and uses the classifier to suggest tracks suited to a given context, whenever the changes in context features exceed a threshold. These tracks are subsequently inserted into the playlist after the currently played one.

During a pilot study, we collected 7,628 data points (context features, music metadata, and interaction data) created by 48 students at JKU Linz. Based on this dataset, we investigated a variety of classifiers in cross-validation experiments for the task of **predicting music preferences from context features** [3]. Prediction was effected on four levels: artist, track, mood, and genre. Decision trees and random forests performed best for most prediction tasks, in particular for artist prediction (55% accuracy) and genre prediction (61% accuracy). We further analyzed which features are most important to predict contextual music preferences, and identified running tasks or apps, weather, time, and location, as well as general device properties as the most important ones for all four prediction levels [3]. For the prediction at track level, we found that player-related features, such as activation of repeat mode or the track belonging to a certain playlist, considerably contribute to a good performance, too.

3 Music Recommendation based on Social Media Mining

In our research on music recommenders that are tailored to particular user characteristics, we consider two social media sources to acquire user and listening data: *Last.fm* and *Twitter*. While the former provides convenient API endpoints, we had to elaborate novel information extraction methods to identify listening events from streams of microblogs. In particular, we propose in [4] a hashtag-based filtering approach and a multi-stage rule-based method for matching microblog content and music metadata. Employing these methods yielded several datasets of *Twitter* users' listening events together with time stamps and spatial coordinates [7, 4]. While it can be argued that exploiting social media data biases the results of recommendation experiments towards the user community of the respective platform, we show in [13] that the distribution of *Last.fm* listening events among major music genres is not far away from that reported in official music sales figures provided by the music industry. Exceptions are Classical music, which is underrepresented on *Last.fm* and even more on *Twitter*, and Metal music, which is overrepresented on *Last.fm* [5, 13].

Based on the created social media datasets, in particular a set of 1 million geotagged listening events mined from *Twitter* (MMTD) [4] and a set of 200 million listening events mined from *Last.fm* (LFM-200m) [10], we model listeners by a variety of aspects and subsequently investigate the relationship between these user characteristics and the performance of stand-alone as well as hybrid recommendation algorithms [2, 10, 9]. We further propose new recommendation approaches for geotagged listening data [14, 12].

Our overarching aim is to **identify the optimal music recommendation approach for a given listener** [10, 9]. To this end, we categorize listeners according to age, gender, country, and genre preference [10], or novel listening-related features [2, 9]. For each user group (e.g. listeners in a certain country), we then assess the performance of music artist recommenders, looking into several recommendation algorithms, their variants, and parameter settings (e.g. rating or score aggregation functions for nearest neighbor methods). Algorithms include

user-based collaborative filtering (CF), popularity-based recommendation, and a CF algorithm extended by location-based filtering that we propose in [14]. In addition to these stand-alone recommenders, we create and evaluate hybrid systems, using different score normalization and fusion techniques. We investigate the performance of each algorithmic combination on each user category in cross-validation experiments using dataset LFM-200m. In addition to findings related to algorithmic tuning details, we made interesting observations specific to user categories: (i) recommendations are better when categorizing users according to age and country than according to gender or genre, (ii) younger people seem to be easier to satisfy by recommending overall popular music, whereas from age 41 upwards listeners tend to prefer CF-based recommendations, (iii) popularity-based recommendation works much better for women than for men, (iv) listeners in the investigated countries are served best by CF, except for Russia where mainstream recommendations seem to be more appreciated, and (v) Folk and Blues aficionados prefer our location-enhanced CF recommender.

In addition to the frequently exploited demographic features, we define a set of **novel user features that describe listening behavior**: play count, diversity, and mainstreamness [2] as well as novelty [9]. While *play count* refers to features describing the overall amount of listening events of a user, *diversity* refers to the variety of genres listened to by a user, *mainstreamness* is computed as the share of overall most popular music items among the user’s listening events, and *novelty* as the share of music listened to for the first time by the user in a given time window. In an investigation similar to the one conducted in [10], we find on the MMTD dataset that all recommendation algorithms perform better in terms of F-measure for users with high play count, high diversity, and high mainstreamness [2].

Comparing the results of our investigations of different user groups and recommendation algorithms [10, 9] in terms of precision, recall, and F-measure, it turns out that performance varies strongly between user groups. Focusing on the recommendation algorithms, highest overall precision (6.35%) and F-measure (5.14%) is achieved by the combination of CF and popularity-based recommendation, while highest overall recall (13.85%) is achieved by further integrating content information based on similarity between collaborative music tags.

Our recent work further resulted in several **novel music recommendation approaches**. Exploiting GPS position information of listening events, we propose in [14] a recommendation algorithm that models each user’s geospatial listening distribution via Gaussian mixture models and computes Earth mover’s distances between these models, which are eventually integrated into a CF recommender. Another location-enhanced CF approach exploits geodesic distances between the users’ spatial centroids of listening events, which are either linearly or exponentially weighted [12]. Eventually, we propose in [12] a hybrid music recommender that integrates our previous work on combining audio and web features for music similarity [11] and on rectifying similarity spaces to reduce hubs in recommendations [15].

4 Outlook

We are currently conducting experiments in which we consider a larger variety of user-specific factors, including features related to overall stylistic preferences,

temporal aspects of music consumption, and openness to unknown music. We further plan to investigate the influence of high-level personal characteristics, such as affective states, personality traits, education, intelligence, and music sophistication, on music taste, and to exploit this knowledge to improve music recommendation algorithms.

5 Acknowledgments

This research is supported by the Austrian Science Fund (FWF): P25655 and by the EU-FP7 project no. 601166 (“PHENICX”). The author would further like to thank his colleagues and students who contributed to the work at hand.

References

1. G. Adomavicius and A. Tuzhilin. *Recommender Systems Handbook*, chapter Context-Aware Recommender Systems, pages 217–253. Springer, 2011.
2. K. Farrahi, M. Schedl, A. Vall, D. Hauger, and M. Tkalčič. Impact of Listening Behavior on Music Recommendation. In *Proc. ISMIR*, Oct 2014.
3. M. Gillhofer and M. Schedl. Iron Maiden while jogging, Debussy for dinner? - An analysis of music listening behavior in context. In *Proc. MMM*, Jan 2015.
4. D. Hauger, M. Schedl, A. Košir, and M. Tkalčič. The Million Musical Tweets Dataset: What Can We Learn From Microblogs. In *Proc. ISMIR*, Nov 2013.
5. P. Lamere. Social Tagging and Music Information Retrieval. *New Music Research: Special Issue: From Genres to Tags – Music Information Retrieval in the Age of Social Tagging*, 37(2):101–114, 2008.
6. M. Schedl. Ameliorating Music Recommendation: Integrating Music Content, Music Context, and User Context for Improved Music Retrieval and Recommendation. In *Proc. MoMM*, Dec 2013.
7. M. Schedl. Leveraging Microblogs for Spatiotemporal Music Information Retrieval. In *Proc. ECIR*, Mar 2013.
8. M. Schedl, G. Breitschopf, and B. Ionescu. Mobile Music Genius: Reggae at the Beach, Metal on a Friday Night? In *Proc. ACM ICMR*, Apr 2014.
9. M. Schedl and D. Hauger. Tailoring Music Recommendations to Users by Considering Diversity, Mainstreaminess, and Novelty. In *Proc. ACM SIGIR*, Aug 2015.
10. M. Schedl, D. Hauger, K. Farrahi, and M. Tkalčič. On the Influence of User Characteristics on Music Recommendation. In *Proc. ECIR*, Mar–Apr 2015.
11. M. Schedl and D. Schnitzer. Hybrid Retrieval Approaches to Geospatial Music Recommendation. In *Proc. ACM SIGIR*, Jul–Aug 2013.
12. M. Schedl and D. Schnitzer. Location-Aware Music Artist Recommendation. In *Proc. MMM*, Jan 2014.
13. M. Schedl and M. Tkalčič. Genre-based Analysis of Social Media Data on Music Listening Behavior. In *Proc. ACM Multimedia Workshop: ISMM*, Nov 2014.
14. M. Schedl, A. Vall, and K. Farrahi. User Geospatial Context for Music Recommendation in Microblogs. In *Proc. ACM SIGIR*, Jul 2014.
15. D. Schnitzer, A. Flexer, M. Schedl, and G. Widmer. Local and Global Scaling Reduce Hubs in Space. *Journal of Machine Learning Research*, 13:2871–2902, Oct 2012.
16. Y. Shi, M. Larson, and A. Hanjalic. Collaborative Filtering Beyond the User-Item Matrix: A Survey of the State of the Art and Future Challenges. *ACM Comput. Surv.*, 47(1):3:1–3:45, May 2014.
17. Yuan Cao Zhang, Diarmuid O Seaghdha, Daniele Quercia, Tamas Jambor. Auralist: Introducing Serendipity into Music Recommendation. In *Proc. WSDM*, Feb 2012.