Culture-Aware Music Recommendation

Eva Zangerle
Department of Computer Science, University of Innsbruck, Austria
eva.zangerle@uibk.ac.at

Martin Pichl
Department of Computer Science, University of Innsbruck, Austria
martin.pichl@uibk.ac.at

Markus Schedl
Department of Computational Perception, Johannes Kepler University Linz, Austria
markus.schedl@jku.at

ABSTRACT

Integrating information about the listener's cultural background when building music recommender systems has recently been identified as a means to improve recommendation quality. In this paper, we therefore propose a novel approach to jointly model users by the user's *musical preferences* and his/her *cultural background*. We describe the musical preferences of users by relying on the acoustic features of the songs the users have been listening to and characterize the cultural background of users by cultural and socio-economic features that we infer from the user's country. To evaluate the impact of the proposed user model on recommendation quality, we integrate the model into a culture-aware music recommender system. We show that incorporating both acoustic information of the tracks a user has listened to as well as the cultural background of users in the form of a *music-cultural user model* contributes to improved recommendation performance.

CCS CONCEPTS

• **Information systems → Recommender systems**; *Clustering and classification*; *Music retrieval*;

KEYWORDS

recommender systems, music information retrieval, user modeling

ACM Reference Format:
Eva Zangerle, Martin Pichl, and Markus Schedl. 2018. Culture-Aware Music Recommendation. In *UMAP '18: 26th Conference on User Modeling, Adaptation and Personalization, July 8–11, 2018, Singapore, Singapore.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3209219.3209258

1 INTRODUCTION

Recent advances in recommender systems and music information retrieval have shown that contextual information is vital for highly personalized recommendations. To this end, the geographic location of a user is often exploited as one notion of context. However, location alone does not necessarily serve as a good indicator for the cultural background of a user as geographically close users might have a very different cultural background and cultural aspects may not coincide with political borders [5]. Notably, for recommender systems, the cultural background of a user was found to play a vital role in how recommended items are judged [7]. We hence argue that modeling users based on musical properties of the songs they listen to on the one hand and the user's cultural background on the other hand contribute to capturing music-cultural listening patterns. These patterns particularly describe the complex interrelation between users, their cultural background, and the characteristics of the music they listen to. In this paper, we propose a novel music-cultural user modeling approach to exploit such listening patterns for recommender systems by integrating information about (i) the acoustic qualities of the music users have listened to and (ii) culture-specific information derived from the users' location.

2 MUSIC-CULTURAL USER MODEL

For the proposed music-cultural model, we incorporate a user's musical preferences as well as his/her cultural background.

As for modeling individual *musical preferences*, we gather content-based audio features for each of the tracks in the used LFM-1b dataset [6] by querying the Spotify API[1] (following the lines of [4]). These content features are extracted and aggregated from the audio signal and comprise: danceability (suitability of a track for dancing), energy (perceptual measure of intensity and activity), speechiness (presence of spoken words), acousticness (probability whether a track is acoustic), instrumentalness (signifies whether a track contain vocals), tempo (beats per minute), valence (musical positiveness conveyed) and liveness (presence of audience in the recording).

As for *cultural aspects*, we propose to model these on a country level and integrate two different data sources. First, we use Hofstede's widely accepted model of cultural dimensions [3]. This framework describes a nation's culture and values by the following six dimensions: power distance, individualism, masculinity, uncertainty avoidance, long-term orientation and indulgence. Second, we complement Hofstede's cultural dimensions with socio-economic characteristics of countries. We capture these by figures extracted from the World Happiness Report (WHR) [2]. The WHR provides the following set of measures capturing the perceived happiness of countries: GDP, freedom, healthy life expectancy, generosity, social support, trust, and happiness.

As for the proposed user model, we characterize a user's individual musical preferences along with his/her cultural background in a single feature vector. To capture a user's *individual musical preferences* based on his/her listened tracks, we consider the Spotify audio features. For each of the features, we compute a user's arithmetic mean and standard deviation across all tracks in his/her listening history and add these mean and standard deviation (SD) values to the user's feature vector. We chose to add the standard deviations to mitigate the effects of averaging a large number of features that potentially differ substantially. For the approximation of the *cultural background* of users (or rather, the country they live

ACM ISBN 978-1-4503-5589-6/18/07.
https://doi.org/10.1145/3209219.3209258

[1] https://developer.spotify.com/web-api/get-several-audio-features/.

in) by socio-economic aspects, we rely on the variables of Hofstede's cultural dimensions and the WHR and extract these based on the user's country information. We add these variables to the user's feature vector to find *cultural* listening patterns that presumably reflect cultural similarity better than sole geographic distance. For each of these variables, we perform a linear min-max scaling such that all elements of the vectors are within $[0, 1]$. This eventually yields an 21-dimensional feature vector per user.

EXPERIMENTS AND RESULTS

For the evaluation of the proposed music-cultural user model for music recommendation, we rely on the LFM-1b dataset [6], which provides the full listening histories of 120,322 Last.fm users. Each listening event contains information about the track, artist, album, and user. Since we aim to model musical preferences jointly by individual musical preference and the cultural background of users, we require the data to contain information about the location of the user. For 45.87% of all users within the LFM-1b dataset, country information about the user is available. Therefore, we restrain the dataset to those users (and their tracks). This provides us with a dataset comprising 55,191 users, who have listened to a total of 26,022,625 distinct tracks, which are captured by a total of 807,890,921 listening events. Besides the information contained within the LFM-1b dataset, we also require information about the tracks the users listened to in order to retrieve descriptive content features. For all listening events of users for whom we can obtain country information, we therefore search for the <track, artist, album>-triples extracted from the LFM-1b dataset using the Spotify search API[2] to gather the Spotify URI, which we subsequently use to collect the set of audio features for each track using the audio features API.[3] This results in a dataset of 55,149 users, 394,944,868 listening events and 3,478,399 distinct tracks.The average number of listening events per user is 7,161.

We model the computation of context-aware music recommendations based on the proposed user model as a classification task. Particularly, we utilize the popular XGBoost classifier [1], a scalable end-to-end tree boosting approach. Using XGBoost, we set the learning objective to logistic regression for binary classification. Furthermore, we set the training objective to be the binary classification error rate. For the classification task carried out, we require a rating for each track that allows us to define whether a given track was listened to (and thus considered relevant) for a given user. Hence, we add a binary factor rating to the dataset, that is set to 1 if the user has listened to a track. As our dataset does not contain any implicit feedback of users, we assume tracks the user has not listened to as negative examples. For training the classifier, we split each user's listening history into five folds and use four folds for training an XGBoost model and use the withheld fold as the test set. To evaluate the performance of the proposed contextual user model with respect to recommendation quality, we perform a per-user evaluation. Therefore, we use each user's listening history and perform a *leave-k-out* evaluation per user.

We evaluate the music-cultural model (*Music + Culture*) and also individually evaluate the performance of a model solely relying on musical preferences of users captured by content-based audio features (*Music* model), and analogously a model that describes users and tracks by their cultural background only (*Culture* model). Furthermore, we evaluate an approach that uses each user's listening history and utilizes the country code of each user (e.g., US for users from the United States) as sole contextual information. Here, we aim to evaluate whether the country code may act as a proxy for cultural factors of users (Music + Country).

Model	RMSE
Music + Culture	0.30
Music	0.33
Culture	0.42
Music + Country	0.43

Table 1: Evaluation results of investigated models

The results of this evaluation are depicted in Table 1, where we list the Root Mean Squared Error (RMSE) for each of the evaluated models. As can be seen, our proposed music-cultural model (Music + Culture) outperforms a model that solely describes the musical preferences of a user. Similarly, relying solely on cultural user features also results in a higher RMSE. Notably, the model relying only on the musical preferences of users and their country code achieves the highest RMSE. This shows that the country of a user can not be used as a proxy to accurately describe cultural aspects.

3 CONCLUSIONS AND FUTURE WORK

We have proposed a user model that incorporates a user's musical preferences and his/her cultural background to improve music recommendation systems. Our experiments have shown that such a music-cultural model indeed contributes to improved recommendation quality. Future work includes an analysis of the influence of each of the individual cultural and musical features. Also, we are interested in the differences between countries to obtain a deeper understanding of country-specific listening patterns.

REFERENCES

[1] Tianqi Chen and Carlos Guestrin. 2016. XGBoost: A Scalable Tree Boosting System. In *Proc. of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*. ACM, New York, NY, USA, 785–794. https://doi.org/10.1145/2939672.2939785

[2] John F Helliwell, Richard Layard, and Jeffrey Sachs. 2016. *World Happiness Report.* Sustainable Development Solutions Network.

[3] Geert Hofstede, Gert Jan Hofstede, and Michael Minkov. 1991. *Cultures and organizations: Software of the mind.* Vol. 2. McGraw-Hill.

[4] Martin Pichl, Eva Zangerle, and Günther Specht. 2016. Understanding Playlist Creation on Music Streaming Platforms. In *2016 IEEE International Symposium on Multimedia (ISM)*. IEEE, 475–480.

[5] Martin Pichl, Eva Zangerle, Günther Specht, and Markus Schedl. [n. d.]. Mining Culture-Specific Music Listening Behavior from Social Media Data.

[6] Markus Schedl. 2016. The LFM-1b Dataset for Music Retrieval and Recommendation. In *Proc. of the ACM International Conference on Multimedia Retrieval (ICMR)*. ACM, New York, NY, USA, 103–110.

[7] Markus Schedl and Dominik Schnitzer. 2014. Location-Aware Music Artist Recommendation. In *Proc. of the 20th International Conference on MultiMedia Modeling (MMM 2014)*. Springer Berlin Heidelberg, 205–213.

[2] https://developer.spotify.com/web-api/search-item/
[3] https://developer.spotify.com/web-api/get-several-audio-features/