

User Geospatial Context for Music Recommendation in Microblogs

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

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

Katayoun Farrahi
Department of Computing
Goldsmiths, University of
London, UK
k.farrahi@gold.ac.uk

ABSTRACT

Music information retrieval and music recommendation are seeing a paradigm shift towards methods that incorporate user context aspects. However, structured experiments on a standardized music dataset to investigate the effects of doing so are scarce. In this paper, we compare performance of various combinations of collaborative filtering and geospatial as well as cultural user models for the task of music recommendation. To this end, we propose a geospatial model that uses GPS coordinates and a cultural model that uses semantic locations (continent, country, and state of the user). We conduct experiments on a novel standardized music collection, the “Million Musical Tweets Dataset” of listing events extracted from microblogs. Overall, we find that modeling listeners’ location via Gaussian mixture models and computing similarities from these outperforms both cultural user models and collaborative filtering.

Categories and Subject Descriptors

Information systems [Information retrieval]: Music recommendation; Human-centered computing [Collaborative and social computing]: Social media

General Terms

Algorithms, Evaluation

Keywords

Music Information Retrieval; Music Recommendation; User Geospatial Context; Personalization; Evaluation

1. INTRODUCTION

Accounting for contextual aspects of the listener in music retrieval and recommendation systems has recently been identified as one of the most important challenges in these

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
SIGIR '14, July 6–11, 2014, Gold Coast, Queensland, Australia.
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-2257-7/14/07 ...\$15.00.
<http://dx.doi.org/10.1145/2600428.2609491>.

fields [6]. Such contextual aspects include time, location, activity, weather, mood, or people with whom the listener is currently together [11]. Given today’s availability of (mobile) devices that are equipped with sensors, it becomes feasible to gather corresponding data on a large scale.

In recent work, it has been shown that incorporating information about a listener’s position may help improve music recommendation [12]. However, Schedl and Schnitzer follow a simplistic approach to model a user’s position, by computing the centroid of the GPS coordinates of listening events. In particular for users whose geospatial listening distribution is centered around more than one location this model is problematic as the centroid will likely correspond to a location that is not representative for the user’s main places of music consumption.

Addressing this shortcoming, we propose here two location-based approaches to derive similarities between users from GPS coordinates in microblog data about music listening, and we combine these similarities with user-based collaborative filtering for the task of music recommendation. The user context in our work is hence addressed by taking into account *musical preference* and *geospatial data*, using a standardized collection of listening behavior mined from microblog data [4]. Our contextual similarity models either represent the listeners’ location by modeling the distribution of her GPS coordinates via Gaussian mixture models (the corresponding model is further called GEO) or by assigning political or cultural regions to the user (CULT).

In the remainder of the paper we first discuss related work (Section 2), outline the dataset used (Section 3), and present the proposed music recommendation models (Section 4). Subsequently, we detail the setup of the experiments and discuss results (Section 5). Conclusions and possible extensions round off this paper (Section 6).

2. RELATED WORK

Literature on location-based music recommendation is relatively sparse. This can be explained by the fact that geospatially annotated music listening data is barely publicly available for academic research, exceptions being the recently published datasets “MusicMicro” [10] and “Million Musical Tweets Dataset” [4]. Among the few works, Park et al. [7] use geospatial data to recommend music that matches a selected environment, based on aspects such as ambient noise, surrounding, or traffic. Kaminskis et al. [5] propose various approaches (exploiting a knowledge base, human annota-

Level	Amount	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Users	27,778	5	7	10	27.69	17	89,320
Artists	21,397	1	1	2	35.95	9	11,850
Tracks	108,676	1	1	1	7.08	4	2,753
Continents	7	9	4,506	101,400	109,900.00	142,200	374,300
Countries	166	1	12	71	4,633.00	555	151,600
States	871	1	7	40	814.60	194	148,900

Table 1: Data set statistics.

tions, and annotations automatically predicted from audio signals) to recommend music for a given point of interest, such as a particular monument. Byklum [2] explore the use of song lyrics to link music to cities and countries. Raymond et al. [8] combine information from different sources to derive geospatial information on artists, aiming at locating them on a map. Perhaps closest to the work at hand is [12], in which Schedl and Schnitzer incorporate position information from GPS coordinates into a music recommendation model. Their model is, however, not suited to deal with users having several geographical clusters of listening activity. Furthermore, the authors only consider user distance based on pure GPS coordinates, ignoring cultural or political regions, which may better reflect differences in music preference¹.

From a higher point of view, this work relates to context-based and hybrid recommendation systems, a detailed review of which is unfortunately beyond the scope of the paper. A decent literature overview and comprehensive discussion are given in [9].

3. MICROBLOG MUSIC DATA

Publicly available, standardized datasets of music listening behavior are rare. In particular when looking for datasets of real-world, commercial size, to the best of the authors’ knowledge the “Million Song Dataset”² [1] and the “Yahoo! Music Dataset”³ [3] used in the KDD Cup 2011⁴ are the only ones. While the former comes with tens of millions of $\langle \text{user}, \text{song}, \text{playcount} \rangle$ triples and the latter with hundreds of millions of $\langle \text{user}, \text{song}, \text{rating} \rangle$ instances, none of these collections include listener location.

Recently, a dataset of music listening activities inferred from microblogs has been released [4]. This so-called “Million Musical Tweets Dataset”⁵ (MMTD) is an extension of the “MusicMicro” set [10], is freely available, and comprises almost 1,100,000 listening events of 215,000 users listening to a total of 134,000 unique songs by 25,000 unique artists. Since this dataset includes longitude and latitude coordinates for each listening event, it is well-suited for location-aware music recommendation. We thus use it in our experiments.

Since the MMTD dataset contains a lot of users with too few listening events to make reliable recommendations, we consider only users with at least five listening events by different artists. Furthermore, we map GPS coordinates to

¹Consider, for example, North Korea vs. South Korea, which are geographically close, but may not necessarily share similar music preference.

²<http://labrosa.ee.columbia.edu/millionsong>

³<http://music.yahoo.com>

⁴<http://www.sigkdd.org/kdd2011/kddcup.shtml>

⁵<http://www.cp.jku.at/datasets/MMTD>

semantic locations, such as continent, country, or state, to allow for cultural user modeling. In this way, we obtain a subset of the MMTD set with basic statistics shown in Table 1. The second column shows the total amount of the entities in the corresponding first row, whereas the right-most six columns show principal statistics based on the number of tweets.

4. MUSIC RECOMMENDATION MODELS

In the considered music recommendation models, each user $u \in U$ is represented by a set of artists listened to $A(u)$. All approaches determine for a given seed user u a number K of most similar neighbors $V_K(u)$, and recommend the artists listened to by these $V_K(u)$, excluding the artists $A(u)$ already known by u . Formally, the recommended artists $R(u)$ for user u are computed as $R(u) = \bigcup_{v \in V_K(u)} A(v) \setminus A(u)$ and $V_K(u) = \operatorname{argmax}_v^K \operatorname{sim}(u, v) \forall v \in (U \setminus u)$. In considering geographical information for user-context models, we investigate the following approaches, which differ in the way this similarity term $\operatorname{sim}(u, v)$ is computed.

GEO: Our geospatial approach models each user’s location with a Gaussian mixture model (GMM) of up to 3 components. For each user, the set of latitude and longitudes for each tweet are taken. If at least 3 unique locations exist in the data, then a GMM with 3 components is used. Otherwise, if 2 unique tweet locations exist, a model with 2 components is used, otherwise a GMM with a single component is used. Once each user is represented as a GMM of their tweet locations, a distance is computed between each pair of users. The earth mover’s distance is considered by sampling from each user’s GMM and computing the distance between their probability distributions. The inverse distances are used for determining user similarity $\operatorname{sim}(u, v)$.

CULT: In the cultural approach, we select the neighbors for the seed user only according to a geographical similarity computed by means of the Jaccard index on listening distributions over semantic locations. We consider as such semantic categories continent, country, and state. For each user, we obtain the relevant locations by computing the relative frequencies of his listening events over all locations. To exclude continents, countries, and states that are unlikely to contribute to the user’s cultural circle, we retain only locations at which the user has listened to music with a frequency above his own average⁶. On the corresponding listening vectors over locations of two users u and v , we compute the Jaccard index to obtain $\operatorname{sim}(u, v)$. Depending on the location category user similarities are computed on, we distinguish CULT_continent, CULT_country, and CULT_state.

CF: We also consider a user-based collaborative filtering

⁶This way we exclude, for instance, locations where the user might have spent only a few days during vacation.

approach. Given the artist playcounts of seed user u as a vector $\vec{P}(u)$ over all artists in the corpus, we first omit the artists that occur in the test set (i.e. we set to 0 the playcount values for artists we want our algorithm to predict). We then normalize $\vec{P}(u)$ so that its Euclidean norm equals 1 and compute similarities $sim(u, v)$ as the inner product between $\vec{P}(u)$ and $\vec{P}(v)$.

CF+GEO: This approach is a combination of CF and GEO. To this end, we construct a new user-user-similarity matrix as the point-wise product of the similarity matrices given by the CF and by the GEO approaches, to incorporate both music preference and geospatial information. The similarity $sim(u, v)$ between two users u and v is then given by this new matrix.

CF+CULT: This approach works by combining the CF similarity matrix with the CULT similarity matrix in the same way as the CF+GEO approach does.

RB: For comparison, we further implemented a random baseline model that randomly picks K users and recommends the artists they listened to. The similarity function can thus be considered $sim(u, v) = rand[0,1]$.

5. EVALUATION

5.1 Experimental Setup

To assess the quality of the different approaches, we perform 5-fold cross validation experiments on the user level, i.e. we predict 20% of each user’s artists based on the remaining 80% (varying the artists in the 20:80 split five times per user). We compute performance measures (precision, recall, and accuracy) by averaging over all users in the corpus. To compare performance between different approaches, we use a parameter N for the number of recommended artists, and adapt dynamically the number of neighbors K to be considered for the seed user u . This is necessary as (i) we do not know a priori the number of artists listened to by the K neighbors and (ii) this number varies between approaches. Using a fixed K will thus not yield the same amount of recommended artists for different approaches. To determine a suited value of K for a given recommendation approach and a given N , we start the respective approach with $K = 1$ and iteratively increase K until the number of recommended artists equals or exceeds N . In the latter case, we sort the found artists according to their overall popularity among the K neighbors and consider as recommendations the top N .

As we could compute Gaussian mixture models for the GEO approach only for the 767 most active users in the dataset due to data sparsity, the following results are reported for this subset. However, Pearson’s correlation coefficient for precision and recall obtained on this subset and on the full dataset is, respectively, 0.9987 and 0.9547 (computed as mean correlation for the models other than GEO). Results can thus be considered highly comparable.

5.2 Results and Discussion

Figure 1 shows a precision/recall-plot for all investigated approaches, depicting precision/recall-pairs resulting from different values of N in the range [1, 500]. Figure 2 further illustrates accuracy over various values of N .

As can be seen, all approaches outperform the random baseline. The CF+GEO approach performs best on average. Combinations of CF+CULT perform inferior to CF

(except for the most general continent setting), but outperform CULT alone. For small values of N , CF+GEO performs very similar to CF and to the best combination of CF and CULT. We can also see that adding the GEO component to the CF approach improves CF’s performance, though only marginally.

Another interesting finding is that modeling users as distributions over coordinates and determining neighbors by computing similarities between these distributions (GEO) performs quite similar to defining user similarity based on whether users typically reside on the same continent, in the same country, or in the same state (CULT). However, the former performance is more stable than the latter. The unstable results for approaches involving a CULT component for higher values of N are likely due to fact that the limited number of closest neighbors does not allow to recommend a considerable number of N artists, in particular for narrow geographical levels such as state. This also holds for the combination with CF as the CF weights are suppressed via the point-wise multiplication of CF and CULT. Other fusion techniques will be investigated in future work.

Highest recall for almost any precision value is achieved by CF+GEO. This is due to the fact that considering location helps to better diversify recommendations, yielding to higher recall for users with music tastes further away from the mainstream. Compared to CF approaches that commonly suffer from a popularity bias (i.e. popular items are recommended more frequently than lesser known ones), geospatial models seem capable of alleviating this issue.

6. CONCLUSIONS AND OUTLOOK

We proposed two models to define and incorporate geospatial information into music recommendation approaches: one based on statistical modeling of users’ listening positions via Gaussian mixture models (GEO), the other on semantic user locations (CULT) at the levels of continent, country, and state. We further combined GEO and CULT with collaborative filtering techniques. Evaluation conducted on a standardized dataset of geolocated listening events showed that music recommendations based on a combination of collaborative filtering (CF) and GEO outperform both CULT and CF as well as their combinations.

Future work includes investigating other aspects of the user context, for instance, listening time, demographics, or language. We further aim at integrating content-based music descriptors, such as timbre, rhythm, lyrics, or tags. Learning to rank and fusion techniques other than weighting similarities will be assessed, too. Finally, analyzing whether recommendation performance for users in the “long tail”, i.e. users with particular music taste, is different from that of mainstream music listeners is likely to yield clues on how to tailor music recommendations to different user groups.

Acknowledgments

This research is supported by the European Union Seventh Framework Programme FP7 / 2007-2013 through the PHENICX project (no. 601166) and by the Austrian Science Funds (FWF): P22856 and P25655.

7. REFERENCES

- [1] T. Bertin-Mahieux, D. P. Ellis, B. Whitman, and P. Lamere. The million song dataset. In *Proc. ISMIR*, Miami, USA, 2011.

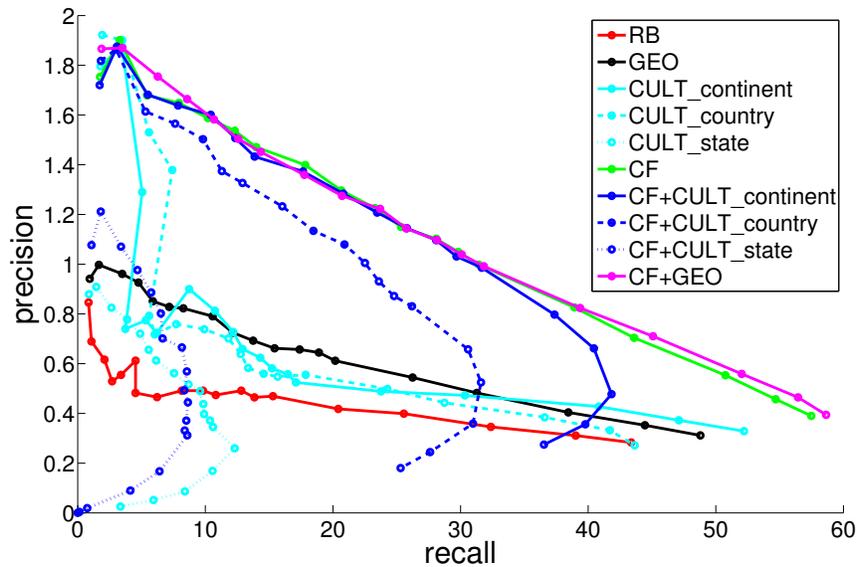


Figure 1: Precision/recall plot.

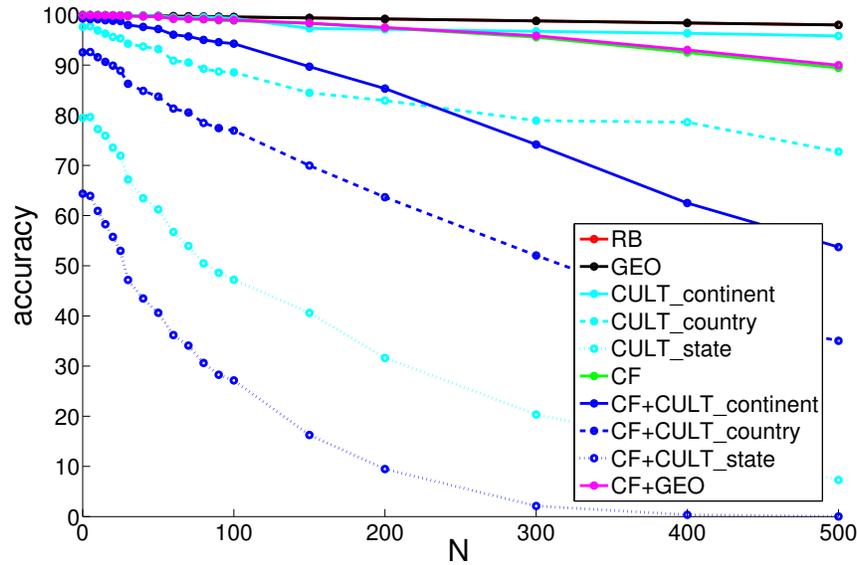


Figure 2: Accuracy for different numbers N of recommended artists.

[2] D. Byklum. Geography and Music: Making the Connection. *Journal of Geography*, 93(6):274–278, 1994.

[3] G. Dror, N. Koenigstein, Y. Koren, and M. Weimer. The Yahoo! Music Dataset and KDD-Cup’11. *Journal of Machine Learning Research: Proceedings of KDD-Cup 2011 competition*, 18:3–18, October 2012.

[4] D. Hauger, M. Schedl, A. Košir, and M. Tkalčič. The Million Musical Tweets Dataset: What Can We Learn From Microblogs. In *Proc. ISMIR*, Curitiba, Brazil, 2013.

[5] M. Kaminskas, F. Ricci, and M. Schedl. Location-aware Music Recommendation Using Auto-Tagging and Hybrid Matching. In *Proc. RecSys*, Hong Kong, China, 2013.

[6] C. Liem, M. Müller, D. Eck, G. Tzanetakis, and A. Hanjalic. The Need for Music Information Retrieval with User-centered and Multimodal Strategies. In *Proc. MIRUM*, Scottsdale, AZ, USA, 2011.

[7] S. Park, S. Kim, S. Lee, and W. S. Yeo. Online Map Interface for Creative and Interactive MusicMaking. In *Proc. NIME*, Sydney, Australia, 2010.

[8] Y. Raimond, C. Sutton, and M. Sandler. Automatic Interlinking of Music Datasets on the Semantic Web. In *Proc. WWW: LDOW Workshop*, Beijing, China, 2008.

[9] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, editors. *Recommender Systems Handbook*. Springer, 2011.

[10] M. Schedl. Leveraging Microblogs for Spatiotemporal Music Information Retrieval. In *Proc. ECIR*, Moscow, Russia, 2013.

[11] M. Schedl, A. Flexer, and J. Urbano. The neglected user in music information retrieval research. *Journal of Intelligent Information Systems*, 41:523–539, Dec 2013.

[12] M. Schedl and D. Schnitzer. Hybrid Retrieval Approaches to Geospatial Music Recommendation. In *Proc. SIGIR*, Dublin, Ireland, 2012.