

Ameliorating Music Recommendation

Integrating Music Content, Music Context, and User Context for Improved Music Retrieval and Recommendation

Markus Schedl

Department of Computational Perception
Johannes Kepler University Linz, Austria

Linz, Austria
markus.schedl@jku.at

ABSTRACT

Successful music recommendation systems need to incorporate information on at least three levels: the music content, the music context, and the user context. The former refers to features derived from the audio signal; the second refers to aspects of the music or artist not encoded in the audio, nevertheless important to human music perception; the third refers to contextual aspects of the user which change dynamically.

In this paper, we briefly review the well-researched categories of music content and music context features, before focusing on user-centric models, which have been neglected for a long time in music retrieval and recommendation approaches. In particular, we address the following tasks: (i) geospatial music recommendation from microblog data, (ii) user-aware music playlist generation on smart phones, and (iii) matching places of interest and music.

The approaches presented for task (i) rely on large-scale data inferred from microblogs, motivated by the fact that social media represent an unprecedented source of information about every topic of our daily lives. Information about music items and artists is thus found in abundance in user-generated data. The questions of how to infer information relevant to music recommendation from microblogs and what to learn from them are discussed. So are different ways of incorporating this kind of information into state-of-the-art music recommendation algorithms.

The presented approaches targeted at tasks (ii) and (iii) model the user in a more comprehensive way than just using information about her location and music listening habits. We report results of a user study aiming at investigating the relationship between music listening activity and a large set of contextual user features. Based on these, an intelligent mobile music player that automatically adapts the current playlist to the user context is presented.

Eventually, we discuss different methods to solve task (iii), i.e., to determine music that suits a given place of inter-

est, for instance, a major monument. In particular, we look into knowledge-based and tag-based methods to match music and places.

Keywords

user-centric music retrieval, mobile music player, adaptive playlist generation, hybrid music recommendation

Categories and Subject Descriptors

Information systems [Information search and retrieval]:
Music recommendation

1. COMPUTATIONAL FEATURES FOR MUSIC DESCRIPTION

Music is a highly multimodal object. It may be represented by a score sheet, listened to by playing a digital audio file, described by a song's lyrics, and visualized by an album cover or a video clip. During the last two decades of research in Music Information Retrieval (MIR), various computational approaches have been proposed in order to infer semantic information about music items and artists from many data sources [2, 8]. The availability of such descriptors enables a wide variety of exciting applications, such as automated playlist generation systems [18], music recommendation systems [3], intelligent browsing interfaces for music collections [9], or semantic music search engines [7].

In [15] a broad categorization of the respective music descriptors is presented. As depicted in Figure 1, these categories are *music content*, *music context*, *user properties*, and *user context*. Features describing the music content are inferred directly from the audio signal representation of a music piece (e.g. timbre, rhythm, melody, and harmony). However, there also exists aspects influencing our music perception which are not encoded in the audio signal. Such contextual aspects that relate to the music item or artist are referred to as music context (e.g. political background of a songwriter or semantic labels used to tag a music piece). These two categories of computational features have been addressed in a vast amount of literature on MIR.

The perception of music is highly subjective though, in turn depends on user-specific factors too. Aspects that belong to the two categories of user properties and user context thus need to be considered when building user-aware music retrieval and recommendation systems. Whereas the user

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. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MoMM2013, 2-4 December, 2013, Vienna, Austria.

Copyright 2013 ACM 978-1-4503-2106-8/13/12 ...\$15.00.

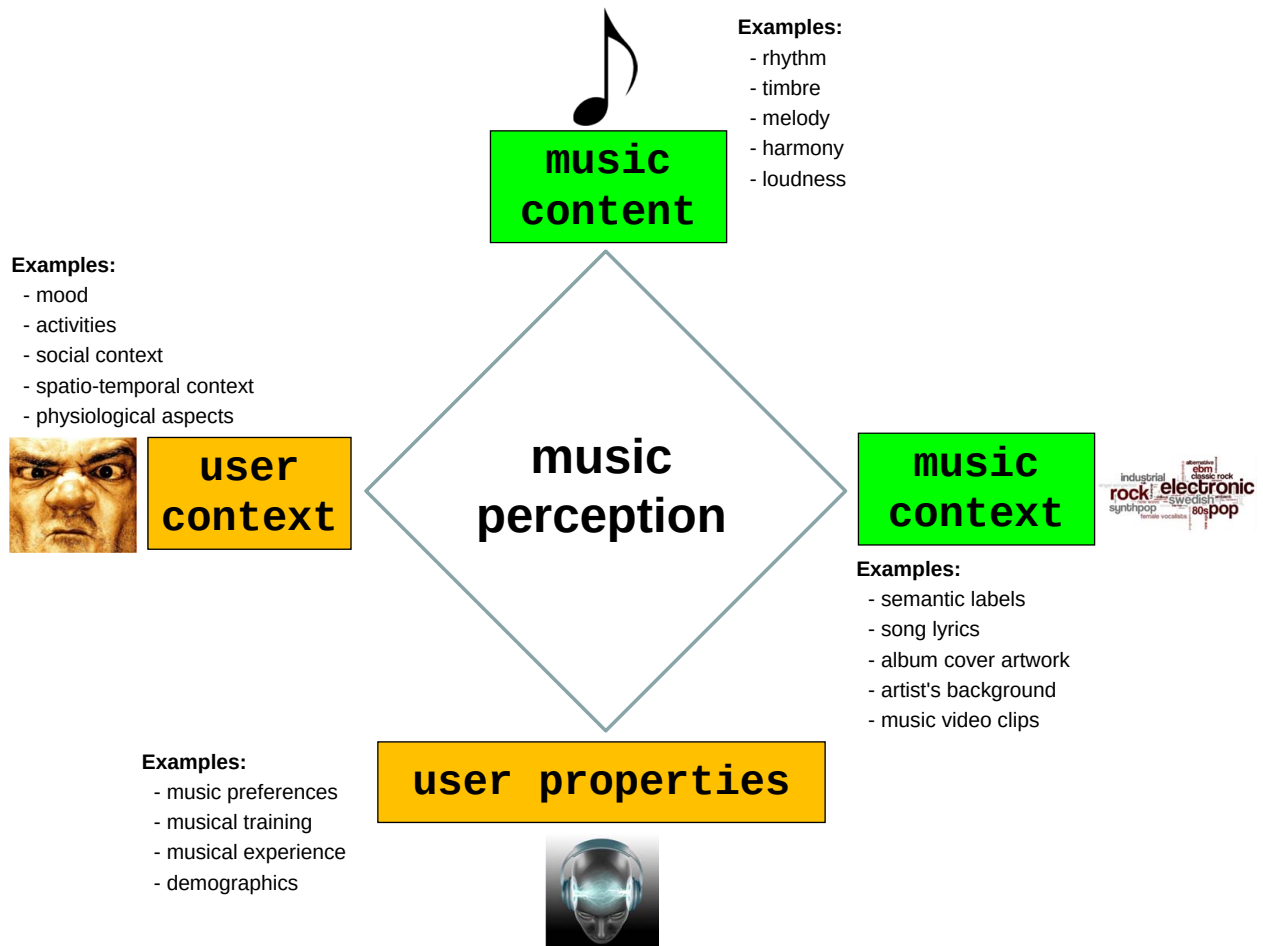


Figure 1: Categorization of music and user descriptors.

properties encompass static or only slowly changing characteristics of the user (e.g. age, musical education, or genre preference), the user context refers to dynamic user characteristics (e.g. her mood, current activities, or surrounding people).

In the following, we present some of our work on user-centric music retrieval and recommendation, which takes into account aspects of user properties and user context, in addition to music-related features. We first show how to extract music listening events from microblog posts and how to use this data to build location-aware music recommendation systems (Section 2). Subsequently, we present our intelligent mobile music player (“Mobile Music Genius”) that automatically adapts the music playlist to the state and context of the user. We also report on preliminary experiments conducted to predict music a user is likely to prefer given a particular context (Section 3). Eventually, we address the topic of recommending music that suits particularly interesting places, such as monuments (Section 4).

2. GEOSPATIAL MUSIC RECOMMENDATION

Microblogs offer a wealth of information on users’ music listening habits, which can in turn be exploited to build user-aware music retrieval and recommendation systems. We

hence crawled **Twitter** over a period of almost two years, in an effort to identify tweets (i) reporting on listening activity and (ii) having attached location information. Making use of **Twitter**’s Streaming API allows to continuously gather 1–2% of all posted tweets. We first filter this stream, excluding all tweets that do not come with location information or do not include listening-related hashtags, such as *#nowplaying* or *#music*. Subsequently, we apply a pipeline of pattern matching methods to the remaining tweets. Using a data base of artist names and song titles (**MusicBrainz**), this results in data items containing temporal and spatial information about listening activity, the latter being defined by artist and song. We further enrich the data items by mapping the position information given as GPS coordinates to actual countries and cities (where available). The resulting final data set (“**MusicMicro**”) is presented in detail in [14], an extension to it (“**Million Musical Tweets Dataset**”) in [5]. These data sets can be used, for instance, to explore music listening activity based on time (cf. Figure 2) or location (cf. Figure 3). In Figure 2, the popularity of songs by “Madonna” on the microblogosphere is illustrated, by aggregating the respective listening events into 2-week-bins. It can also be clearly seen when new albums and songs were released. Figure 2 gives an example of the spatial music listening distribution. Please note that we also employ

a machine learning technique to predict the genre of each song and subsequently map genres to different colors. Our corresponding user interface to browse the microblogosphere of music is called “Music Tweet Map”¹.

In addition to these tools for visual analysis of music listening patterns on the microblogosphere, the enriched microblog data enables to build hybrid music recommender systems. To this end, we developed approaches that integrate state-of-the-art techniques for music content- and music context-based similarity computation and ameliorate these by simple location-based user models [16]. More precisely, given the “MusicMicro” collection [14], we first compute a linear combination of similarity estimates based on the PS09 features [11] (for audio) and on $tf \cdot idf$ features [13] (derived from artist-related web pages), yielding a joint music similarity measure. We experimented with different coefficients for the linear weighting and found that adding even only a small component of the complementary feature boosts performance.

Based on these finding, we elaborated a method to integrate user context data, in this case location, into the joint similarity measure. More precisely, we first compute for each user u the geospatial centroid of her listening activity $\mu(u)$. In order to recommend music to a user u , we then use the geodesic distance between $\mu(u)$ and $\mu(v)$, computed for all potential target users v , to weight other distance measures based on music-related features. Incorporating this method into a standard collaborative filtering approach, thus giving higher weight to nearby users than to users far away when computing music-related similarities between users, we show in [17, 16] that this location-specific adaptation of similarities can outperform standard collaborative filtering and content-based approaches.

3. USER-AWARE MUSIC PLAYLIST GENERATION ON SMART PHONES

The importance of taking into account the contextual aspects of the user when creating music recommenders or music playlist generators is underlined by several scientific works [1, 10, 4]. We present in this paper for the first time our “Mobile Music Genius”² (MMG) player, which is an intelligent mobile music player for the **Android** platform. The player aims at dynamically and seamlessly adapt the music playlist according to the music preference of the user in a given context. To this end, MMG continuously monitors a wide variety of user context data while the user interacts with the player or just enjoys the music. From (i) the contextual user data, (ii) implicit user feedback (play, pause, stop, skip events), and (iii) meta-data about the music itself (artist, album, track names), MMG learns relationships between (i) and (iii), i.e. which kind of music she prefers in which situation. The underlying assumption is that music preference changes with user context. A user might, for instance, want to listen to an agitating rock song when doing outdoor sports, but might prefer some relaxing reggae music when being at the beach at a sunny and hot day.

Table 1 lists some examples for user context attributes that are continuously monitored. In addition to these unobtrusively gathered data, we ask the user for her activity and mood each time a new track is played, presuming that both

¹<http://www.cp.jku.at/projects/MusicTweetMap>

²<http://www.cp.jku.at/projects/MMG>

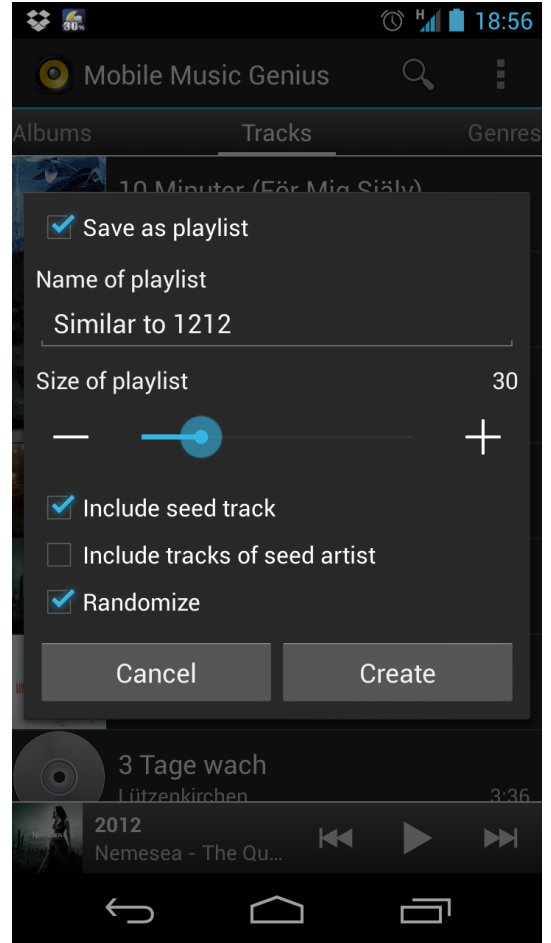


Figure 4: Automatic music playlist generation with “Mobile Music Genius”.

strongly influence music taste, but are not easy to derive with high accuracy from the aspects listed in Table 1.

The methods used to create and continuously adapt the playlist in MMG work as follows. In principle, creating a playlist can either be performed manually, like in a standard mobile music player, or automatically. In the latter case, the user selects a seed song and is then given the options shown in Figure 4: she can decide on the number of songs in the playlist, whether the seed track or tracks by the seed artist should be included in the playlist, and whether she wants her playlist shuffled, i.e. the nearest neighbors to the seed track randomly inserted into the playlist, instead of ordered by their similarity to the seed.

This automatic creation of a playlist does not yet take into account the user context. Instead, it relies on collaboratively generated tags downloaded from **Last.fm**. To this end, MMG gathers for each piece in the user’s music collection tags on the level of artist, album, and track. Subsequently, tag weight vectors based on the importance each tag is attributed according to **Last.fm** are computed (similar to $tf \cdot idf$ vectors). The individual vectors on the three levels are merged into one overall vector describing each song.

If the user now decides to create a playlist based on a selected seed song s , the cosine similarity between s and all other songs in the collection is computed and the songs clos-

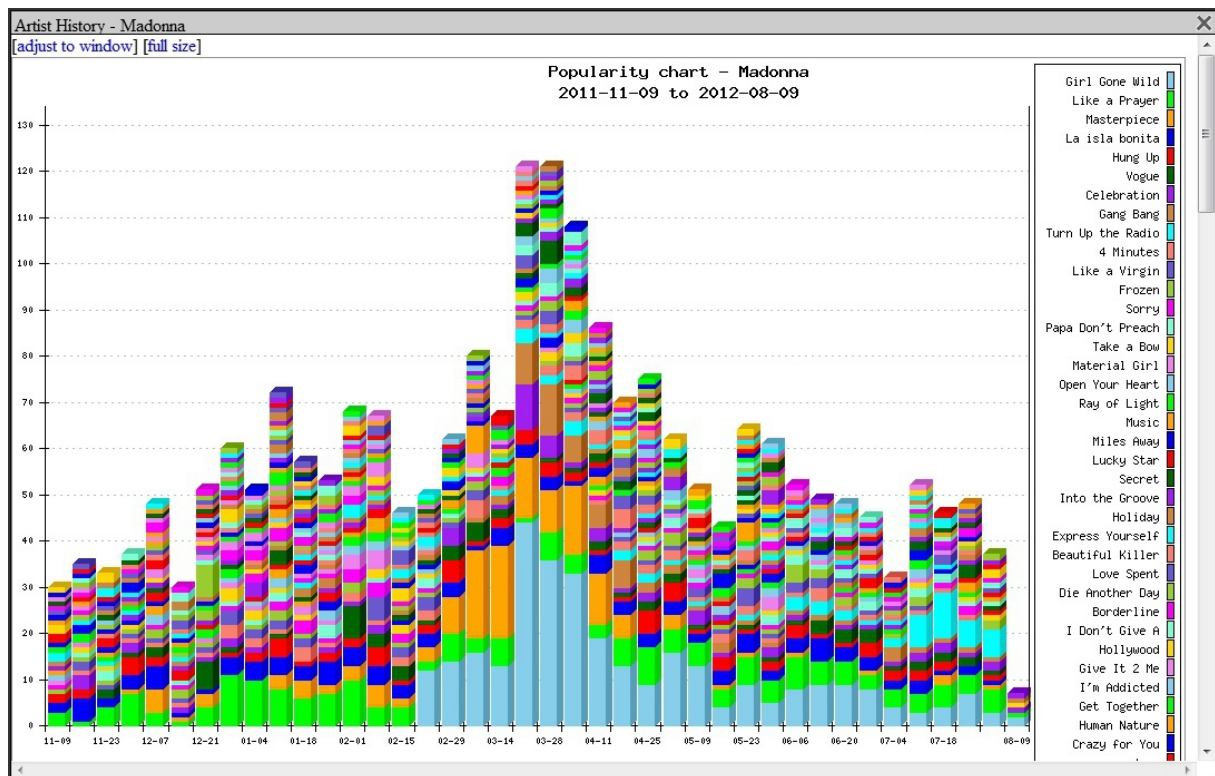


Figure 2: Temporal music listening distribution of songs by “Madonna”.

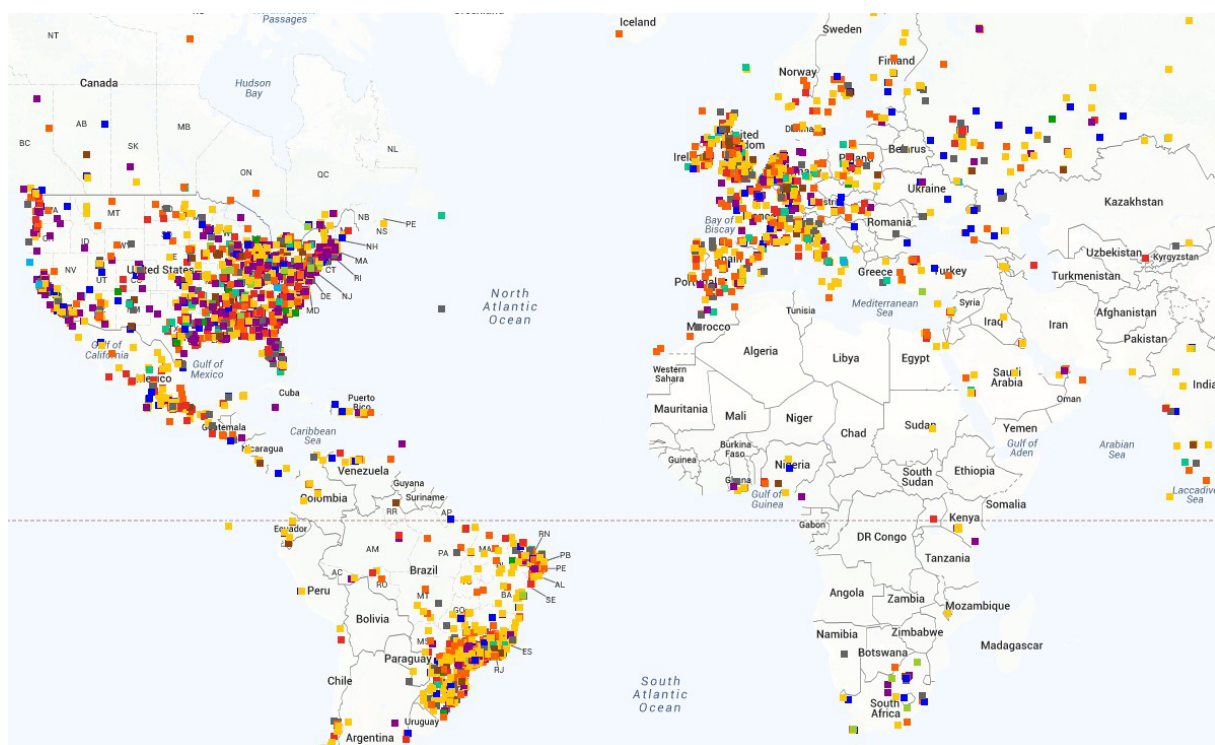


Figure 3: Spatial music listening distribution of music genres (depicted in different colors).

Category	Exemplary Attributes
Time	day of week, hour of day
Location	provider, latitude, longitude, accuracy, altitude, nearest relevant city, nearest populated place
Meteorological	wind direction and speed, clouds, temperature, dew point, humidity, air pressure, weather condition
Ambient	light, proximity, noise
Physical activity	acceleration, orientation of user, orientation of device
Task activity	screen state (on/off), docking mode, recently used tasks
Phone state	operator, state of data connection, network type
Connectivity	mobile network: available, connected, roaming; WiFi: SSID, IP, MAC, link speed, networks available; Bluetooth: enabled, MAC, local name, available devices, bonded devices
Device	battery status, available internal/external storage, available memory, volume settings, headset plugged
Player state	playlist type, repeat mode, shuffle mode

Table 1: Some user context attributes monitored by MMG.

est to s are inserted into the playlist. Of course, the constraints specified by the user (cf. Figure 4) are taken into account as well.

As for automatically adapting the playlist, the user can enable the respective option during playback. In this case, MMG continuously compares the current user context vector c_t , which is made up of the attributes listed in Table 1 (and some more), with the previous context vector c_{t-1} , and triggers a playlist update in case $|c_t - c_{t-1}| > \rho$, where ρ is a sensitivity threshold that can be adapted by the user. If such an update is triggered, the system first compares c_t with already learned relations between user contexts and songs. It then inserts into the playlist, after the currently played song, tracks that were listened to in similar contexts. Since the classifier used to select the songs for integration into the playlist is continuously fed relations between user context and music taste, the system dynamically improves while the user is listening to music.

In order to assess how well music preference can be predicted from the user context, we first built a data set by harvesting data about users of the MMG player over a period of two months, foremost students of the Johannes Kepler University Linz³. This yielded about 8,000 single listening events (defined by artist and track name) and the corresponding user context vectors. We subsequently experimented with different classifiers. This is still work in progress, but preliminary results are quite encouraging. Indeed, when predicting music artists from user contexts, the instance-based k-nearest neighbors classifier reached 42% accuracy, a rule learner (“JRip”) 51%, and a decision tree learner (“J48”) 55%, while a simple baseline majority voter (“ZeroR”) that always predicts the most frequent class only achieved 15% accuracy. Experiments have been conducted using the Weka⁴ data mining software.

4. MATCHING PLACES OF INTEREST AND MUSIC

Selecting music that not only fits some arbitrary position in the world, but is tailored to a meaningful place, such as a monument, is the objective of our work presented in [6]. We hence propose five approaches to music recommendation for places of interest: (i) a knowledge-based approach, (ii) a user tag-based approach, (iii) an approach based on music

auto-tagging, (iv) an approach that combines (i) and (iii), and (v) a simple personalized baseline.

Since all approaches, except for (v), require training data, we first collected user annotations for 25 places of interest and for 123 music tracks. To this end, we used the web interface depicted in Figure 5, to let users decide which tags from an emotion-based dictionary fit a given music piece. Similar interfaces were used to gather additional human-generated data used as input to the computational methods, in particular image tags and information about the relatedness of music pieces and places of interests.

Based on these data, the knowledge-based approach makes use of the DBPedia⁵ ontology and knowledge base. More precisely, the likelihood of a music piece to relate to a place of interest is approximated by computing from the ontology the graph-based distance between the musician node on the one hand and the node representing the place of interest on the other hand.

The tag-based approach makes use of the human annotations gathered in the initial tagging experiment. To estimate the relatedness of a music piece m to a place of interest p , the Jaccard index between m ’s and p ’s tag profiles is computed, which effectively calculates the overlap between the sets of tags assigned to m and assigned to p .

Music auto-tagging is performed using a state-of-the-art auto-tagger [19]. Again the Jaccard index between m ’s and p ’s tag profiles is computed to estimate the relatedness of p to m . The benefit over the human based-tagging approach is that there is no need to gather tags in expensive human annotation experiments, instead a large set of tags for unknown music pieces can be inferred from a much smaller set of annotated training data. To this end, we use a Random Forest classifier.

We further propose a hybrid approach that aggregates the recommendations produced by the knowledge-based and the auto-tagging based approaches, employing a rank aggregation technique.

Finally, a simple personalized approach always recommends music of the genres the user indicated to like in the beginning of the experiment, irrespective of the place of interest. More details on all approaches are provided in [6].

Evaluation was carried out in a user study via a web interface (cf. Figure 6), involving 58 users who rated the suitability of music pieces for places of interest in a total of 564 sessions. A session corresponds to the process of viewing im-

³www.jku.at

⁴www.cs.waikato.ac.nz/ml/weka

⁵www.dbpedia.org

Tag:

☐ Melancholic	☐ Bright
☐ Heavy	☐ Animated
☒ Tender	☐ Energetic
☐ Cold	☐ Spiritual
☒ Modern	☒ Serene
☐ Ancient	☐ Calm
☐ Affectionate	☐ Sad
☒ Dark	☐ Strong
☒ Lightweight	☐ Colorful
☒ Open	☐ Thrilling
☐ Warm	☐ Agitated
☐ Sentimental	☐ Bouncy

Submit

Fritz Kreisler - Liebesfreud
http://en.wikipedia.org/wiki/Fritz_Kreisler

00:08 00:31 FMP3

"Friedrich 'Fritz' Kreisler (February 2, 1875 – January 29, 1962) was an Austrian-born violinist and composer. One of the most famous violin masters of his or any other day, he was known for his sweet tone and expressive phrasing. Like many great violinists of his generation, he produced a characteristic sound which was immediately recognizable as his own. Although he derived in many respects from the Franco-Belgian school, his style is nonetheless reminiscent of the gemütlich (cozy) lifestyle of pre-war Vienna."

Skip this item

Figure 5: Web interface used to tag music pieces and to investigate the quality of the tags predicted by the music auto-tagger.

ages and text descriptions for the place of interest, listening to the pooled music recommendation results given by each approach, and rating the quality of the recommendations. To measure recommendation quality, we computed the likelihood that a music piece marked as well-suited was recommended by each approach, averaged over all sessions. Summarizing the main findings, all context-aware approaches (i)–(iv) significantly outperformed the simple personalized approach (v) based only on users’ affinity to particular genres. The auto-tagging approach (iii) outperformed both the human tags (ii) and the knowledge-based approach (i), although just slightly. Superior performance was achieved with the hybrid approach (v) that incorporates complementary sources of information.

5. CONCLUSION

As demonstrated by the examples given in the paper, combining user-centric information with features derived from the content or context of music items or artists can considerably improve music recommendation and retrieval, both in terms of common quantitative performance measures and user satisfaction. In the future, we are likely to see many more algorithms and systems that actively take user-centric aspects into account and intelligently react to them. In particular in the music domain, novel recommendation algorithms that address cognitive and affective states of the users, such as serendipity and emotion, are emerging [12, 20].

6. ACKNOWLEDGMENTS

This research is supported by the Austrian Science Fund (FWF): P22856, P25655, and the FP7 project PHENICX: 601166. The author would also like to thank David Hauger, Marius Kaminskas, and Francesco Ricci for the fruitful collaborations on extracting and analyzing listening patterns from microblogs (David) and on music recommendation for places of interest (Marius and Francesco), which led to the

work at hand. Special thanks go to Georg Breitschopf who spent a lot of time elaborating and evaluating user context-aware playlist generation algorithms and a corresponding mobile music player.

7. REFERENCES

- [1] J. T. Biehl, P. D. Adamczyk, and B. P. Bailey. DJogger: A Mobile Dynamic Music Device. In *CHI 2006: Extended Abstracts on Human Factors in Computing Systems*, pages 556–561, Montréal, Québec, Canada, 2006.
- [2] M. A. Casey, R. Veltkamp, M. Goto, M. Leman, C. Rhodes, and M. Slaney. Content-Based Music Information Retrieval: Current Directions and Future Challenges. *Proceedings of the IEEE*, 96:668–696, April 2008.
- [3] O. Celma. *Music Recommendation and Discovery – The Long Tail, Long Fail, and Long Play in the Digital Music Space*. Springer, Berlin, Heidelberg, Germany, 2010.
- [4] S. Cunningham, S. Caulder, and V. Grout. Saturday Night or Fever? Context-Aware Music Playlists. In *Proceedings of the 3rd International Audio Mostly Conference of Sound in Motion*, October 2008.
- [5] D. Hauger, M. Schedl, A. Košir, and M. Tkalčič. The Million Musical Tweets Dataset: What Can We Learn From Microblogs. In *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR)*, Curitiba, Brazil, November 2013.
- [6] M. Kaminskas, F. Ricci, and M. Schedl. Location-aware Music Recommendation Using Auto-Tagging and Hybrid Matching. In *Proceedings of the 7th ACM Conference on Recommender Systems (RecSys)*, Hong Kong, China, October 2013.
- [7] P. Knees, T. Pohle, M. Schedl, and G. Widmer. A Music Search Engine Built upon Audio-based and Web-based Similarity Measures. In *Proceedings of the*

La Scala, Milan, Italy
http://en.wikipedia.org/wiki/La_Scala



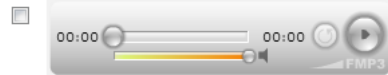
La Scala is a world renowned opera house in Milan, Italy. The theatre was inaugurated on 3 August 1778 and was originally known as the New Royal-Ducal Theatre at La Scala. The premiere performance was Antonio Salieri's 'Europa riconosciuta'. Most of Italy's greatest operatic artists, and many of the finest singers from around the world, have appeared at La Scala during the past 200 years.

Session 1 out of 10:

Listen to the tracks and select those that in your opinion are **suited** for the described place:

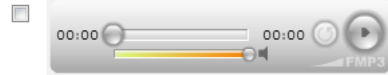
Reincidentes - Ay Dolores

<http://en.wikipedia.org/wiki/Reincidentes>



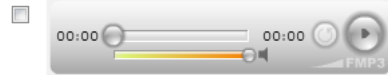
Vincenzo Pucitta - La Vestale, Opera seria 1st act

http://en.wikipedia.org/wiki/Vincenzo_Pucitta



The Shower Scene - This Is The Call Out

http://en.wikipedia.org/wiki/The_Shower_Scene



Duchess Maria Antonia of Bavaria - Pallid' ombra che d'intorno

http://en.wikipedia.org/wiki/Duchess_Maria_Antonia_of_Bavaria

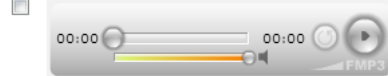


Figure 6: Web interface to determine which music pieces fit a given place of interest.

30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), Amsterdam, the Netherlands, July 2007.

- [8] P. Knees and M. Schedl. A Survey of Music Similarity and Recommendation from Music Context Data. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)*, 10(1), 2013.
- [9] P. Knees, M. Schedl, T. Pohle, and G. Widmer. An Innovative Three-Dimensional User Interface for Exploring Music Collections Enriched with Meta-Information from the Web. In *Proceedings of the 14th ACM International Conference on Multimedia*, Santa Barbara, California, USA, October 2006.
- [10] B. Moens, L. van Noorden, and M. Leman. D-Jogger: Syncing Music with Walking. In *Proceedings of the 7th Sound and Music Computing Conference (SMC)*, pages 451–456, Barcelona, Spain, 2010.
- [11] T. Pohle, D. Schnitzer, M. Schedl, P. Knees, and G. Widmer. On Rhythm and General Music Similarity. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR)*, Kobe, Japan, October 2009.
- [12] S. Sasaki, T. Hirai, H. Ohya, and S. Morishima. Affective Music Recommendation System Using Input Images. In *ACM SIGGRAPH 2013 Posters*, Anaheim, CA, USA, 2013.
- [13] M. Schedl. #nowplaying Madonna: A Large-Scale Evaluation on Estimating Similarities Between Music Artists and Between Movies from Microblogs. *Information Retrieval*, 15:183–217, June 2012.
- [14] M. Schedl. Leveraging Microblogs for Spatiotemporal Music Information Retrieval. In *Proceedings of the 35th European Conference on Information Retrieval (ECIR)*, Moscow, Russia, March 24–27 2013.
- [15] M. Schedl, A. Flexer, and J. Urbano. The neglected user in music information retrieval research. *Journal of Intelligent Information Systems*, July 2013.
- [16] M. Schedl and D. Schnitzer. Hybrid Retrieval Approaches to Geospatial Music Recommendation. In *Proceedings of the 35th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, Dublin, Ireland, July 31–August 1 2013.
- [17] M. Schedl and D. Schnitzer. Location-Aware Music Artist Recommendation. In *Proceedings of the 20th International Conference on MultiMedia Modeling (MMM)*, Dublin, Ireland, January 2014.
- [18] D. Schnitzer, T. Pohle, P. Knees, and G. Widmer. One-Touch Access to Music on Mobile Devices. In *Proceedings of the 6th International Conference on Mobile and Ubiquitous Multimedia (MUM)*, Oulu, Finland, December 12–14 2007.
- [19] K. Seyerlehner, M. Schedl, P. Knees, and R. Sonnleitner. A Refined Block-level Feature Set for Classification, Similarity and Tag Prediction. In *7th Annual Music Information Retrieval Evaluation eXchange (MIREX)*, Miami, FL, USA, October 2011.
- [20] Yuan Cao Zhang, Diarmuid O Seaghdha, Daniele Quercia, Tamas Jambor. Auralist: Introducing Serendipity into Music Recommendation. In *Proceedings of the 5th ACM Int'l Conference on Web Search and Data Mining (WSDM)*, Seattle, WA, USA, July 2012.