Exploring Geospatial Music Listening Patterns in Microblog Data

David Hauger and Markus Schedl

Department of Computational Perception,
Johannes Kepler University, Linz
{david.hauger,markus.schedl}@jku.at

Abstract. Microblogs are a steadily growing, valuable, albeit noisy, source of information on interests, preferences, and activities. As music plays an important role in many human lives we aim to leverage microblogs for music listening-related information. Based on this information we present approaches to estimate artist similarity, popularity, and local trends, as well as approaches to cluster artists with respect to additional tag information. Furthermore, we elaborate a novel geo-aware interaction approach that integrates these diverse pieces of information mined from music-related tweets. Including geospatial information at the level of tweets, we also present a web-based user interface to browse the "world of music" as seen by the "Twittersphere".

Keywords: microblogs, geospatial music taste, music listening patterns

1 Introduction

Due to their continuously growing importance and usage, social media provide a valuable source of user-generated and user-related information. Especially microblogs – due to their nature of being less conversational and providing means to share activities, opinions, experience, and information [21] – are well-suited for discovering breaking news and for user-centric information retrieval [32], [25], [34].

Since its advent in 2006, `Twitter`'s [8] popularity has been continuously growing, resulting in being today's most popular microblogging service. According to `Twitter`'s last official announcements in March 2011 they claimed to have more than 200 million registered users creating a billion posts per week [1]. Given this remarkable user base, it is no surprise that `Twitter` has already been used for various information retrieval and datamining tasks, including analyzing the spread of diseases [23], detecting earthquakes [26] and hot topics [30], recommendation of information sources [9] and ranking tweets according to the relevance of the user [14], [33]. There have also been attempts to identify spam users based on the temporal entropy of tweets containing URLs [31].

One of the many types of information posted via tweets, i.e. messages on the `Twitter` platform limited to 140 chararacters, is information on the music a user is currently listening to. This information may be provided either manually (e.g. included in personal comments) or automatically by plugins for music players or music portals [7]. This research aims at identifying geospatial music listening patterns of the music-tweeting community (although these users are not necessarily representative for the total population).

Section 2 presents related work on microblog mining and geospatial visualization of musical information. In Section 3 we present a novel approach to mine `Twitter` posts for music listening-related information. Additionally, we suggest the use of genre-based clustering and propose a method to co-occurrence-based similarity estimation to organize and visualize the extracted information. Section 4 illustrates how geospatial music listening data may be supportive for various tasks. We present a number of use cases and the user interface of a visualization framework to interactively browse and dynamically explore the world of tweeted music listening events.

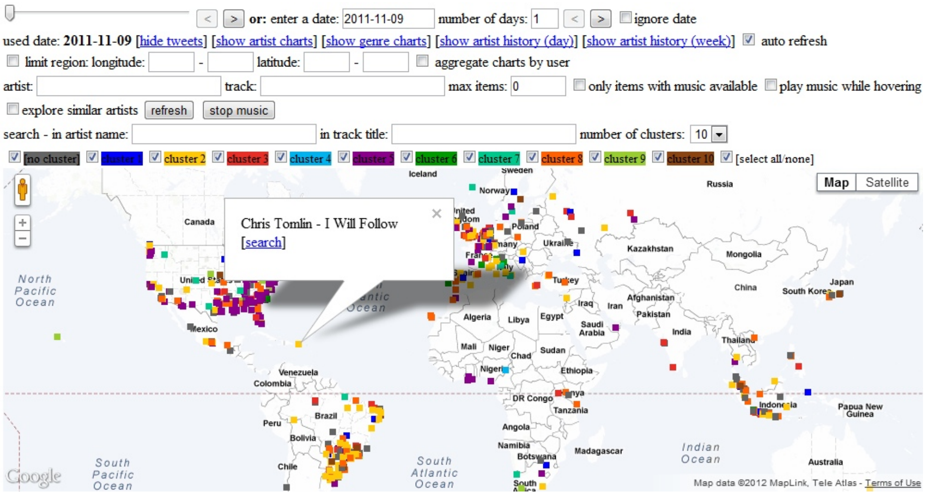


Fig. 1. Visualization of all tweets with the mouse hovering one tweet in Puerto Rico. Map image provided by `Google Maps` [4], ©Google 2012.

2 Related Work

The work at hand, as far as we are aware of, is the first to provide a framework to explore the `Twitter` "world of music" and to visualize geospatial music listening patterns in an interactively explorable environment.

Related work may be categorized into work related to mining microblog data and the geospatial visualization of musical information.

cluster	assigned genre tags (top 20)
1	Electronic, House, Electronica, Dance, Techno, Electro, Trance, Downtempo, Synthpop, Minimal techno, Progressive House, Deep house, Tech house, Drum and bass, Breakbeat, Electropop, Dub, Dubstep, Electro house, Electroclash
2	Rock, Alternative, Alternative rock, Hard rock, Progressive rock, Classic rock, Heavy metal, Psychedelic rock, Grunge, Garage rock, Christian, Alternative metal, Progressive, Stoner rock, Nu metal, Christian rock, Post-grunge, Rock and roll, Southern rock, Modern rock
3	Indie, Indie rock, Indie pop, Post-punk, Lo-fi, Emo, Britpop, Dream pop, Math rock, Power pop, Indietronica, Indiepop, Noise pop, Chamber Pop, Piano rock, Twee pop, Dance-punk, Neo-Psychedelia, Hamburger Schule, Jangle pop
4	Experimental, Ambient, Noise, Psychedelic, Dark ambient, Drone, IDM, Industrial, Post-rock, Avant-garde, Instrumental, Glitch, New Age, Noise rock, Contemporary classical, Breakcore, Space rock, Electroacoustic, Darkwave, Krautrock
5	Hip-Hop, Rap, hip hop, Underground hip hop, Underground, Gangsta rap, Reggae, Dirty South, Turntablism, Southern rap, Grime, Dancehall, G-funk, Horrorcore, Crunk, Ragga, Reggaeton, Memphis rap, Chicano rap, Experimental hip hop
6	punk, Punk rock, Pop punk, Ska, Street punk, Ska punk, Garage punk, Garage, Anarcho-punk, Skate punk, Folk punk, Streetpunk, Psychobilly, Skacore, Horror punk, Riot Grrrl, Melodic, Celtic punk, Deathrock, Christian punk
7	Folk, Singer-songwriter, Acoustic, Celtic, Folk rock, Country, Americana, World, Irish, Indie folk, Traditional, Bluegrass, Neofolk, Medieval, Ethnic, Freak folk, New Weird America, Trad, Folk metal, Acoustic rock
8	Pop, Rnb, Pop rock, New Wave, J-pop, Disco, Eurodance, Soft rock, Turkish, Anime, Latin pop, K-pop, Europop, Ballad, Russian pop, Turkish pop, C-pop, Asian, Gospel, Teen pop
9	Jazz, Funk, Soul, Fusion, Blues, Lounge, Piano, Acid jazz, Free jazz, Swing, Smooth jazz, Nu jazz, Jazz fusion, Soundtrack, Contemporary Jazz, Easy listening, Vocal jazz, Bossa nova, Classical, Big band
10	Hardcore, Metalcore, Metal, Hardcore punk, Death metal, Post-hardcore, Thrash metal, Screamo, Gabber, Black metal, Grindcore, Melodic hardcore, Straight edge, Deathcore, Melodic death metal, Progressive metal, Hardcore techno, Mathcore, Thrashcore, Power metal

Table 1. Groups of genre tags using NNMF for 10 clusters.

2.1 Mining Microblog Data

Hardly any research has been conducted on the intersection between microblog mining and music information retrieval (MIR). Among the few works, Schedl et al. [29] analyze artist popularity on the country level, using term frequencies of `Twitter` posts as one source of information. Zangerle et al. [15] compute inverse document frequency on a fulltext index to map tweets to artists and tracks. The authors propose a co-occurrence-based approach to construct a song recommender system. Schedl and Hauger [28] use microblog data from all cities with more than 500,000 inhabitants in order to calculate deviations of musical taste from the mainstream on country and city level.

General work on microblog mining includes the following: Java et al. [17] analyze microblogs from `Twitter`, `Jaiku`, and `Pownce` in order to study network properties and friendship relations as well as intentions of using those systems. Moreover, they report on geographical distributions of `Twitter` users and the growth of the network. Furthermore, Java et al. aim to identify trends and communities based on keywords. Kwak et al. [18] extend Java et al.'s approach to trend detection by gathering tweets mentioning `Google`'s most frequently used search terms and analyzing the re-tweeting behavior. The authors particularly stress the recentness as one of the major advantages of this source of information.

There is a wide field of different applications that exploit information shared via microblogs. Exploiting geospatial data, De Longueville et al. [13] used data from `Twitter` for forest fire detection. Lee et al. [19] mined `Twitter` for information on earthquakes and plotted them on a world map. As most of those tweets had no information on geo-coordinates attached, they used city names to define positions of tweets. As mentioned in their paper, geo-coordinates are hardly available as they require GPS-enabled devices – which is one of the reasons why they have not been exploited earlier. Bollen et al. [11] mined `Twitter` for emotion-related terms in order to calculate the "public mood", which was then linked to the emergence of stock markets trying to predict future trends.

2.2 Geographic Visualization of Musical Information

Most visualization approaches for musical information are based on various types of content- or context-based features (or similarity measures). These features are mapped to visual aspects such as position, color, distance, or font size. Geographic information is usually not taken into account. However, Raimond et al. [24] combine information from different sources to retrieve geospatial information on artists in order to be able to locate them on a map. Similarly, Govaerts and Duval [16] aim to detect artist origin and plot the results on a map. Another possibility to link music to geographical information is presented by Byklum [12], who searches lyrics for geographical content like names of cities or countries.

A different approach for combining music and geospatial information is presented by Park et al. [22]. They started from geospatial positions and tried to generate music matching the selected environment, based on ambient noise, surroundings, traffic, etc.

As far as we know, geospatial information has not yet been scientifically used to visualize listening patterns, which is most probably due to the fact that this is a relatively new type of information available.

3 Methodology

3.1 Data Acquisition & Processing

For the work reported in this paper we used the `Twitter` Streaming API to retrieve tweets with geospatial coordinates available (preliminary analysis showed that this applies to less than 3% of the tweets). Between September 2011 and August 2012 we crawled `Twitter` for potentially music-related hashtags, e.g. `#nowplaying`, `#np`, `#itunes`, `#musicmonday` and `#thisismyjam`. The most frequently used music-related hashtag `#nowplaying` and its abbreviation `#np` have already been proven successful to determine music listening-related tweets [27]. During these nine months we retrieved 2,337,489 tweets including both one of the hashtags mentioned above and geospatial information.

However, microblog data is not standardized, neither in terms of the content nor concerning the usage of hashtags. For instance, `#nowplaying` is also used to refer to activities other than music listening (among others, sports events, movies, or games), to a much smaller extent though. Moreover, some tweets are music-related, but contain no information that could be used for our purposes (e.g. "#nowplaying my favorite songs again and again...").

Having obtained the tweets, our goal was to parse and analyze the content to extract artist information. Dictionary-based text matching algorithms and word stemming [10] are not suited to process this type of data, as artist names may match common speech terms. This results in "I", "You", "Me", and "Love" as the most popular, often erroneously detected, artists in our tests, using a list of artists from `freebase` [6]). Artist names that are part of other artist names also pose a serious problem.

In order to overcome these difficulties, we elaborated an alternative approach. Preliminary observations revealed that music-related tweets often contain patterns, such as:

- *song title* by *artist name* [on *some platform*]
- *artist name*: "*song title*"
- *song title #artist name*
- *song title – artist name*
- *artist name – song title*

Therefore, we decided to adopt a multi-level, pattern-based approach, matching only potential artist names against the artist dictionary. Starting with the specific patterns listed above and continuing the search with more general ones (e.g. any term separated by special characters) in case the mentioned ones could not be applied, we were able to eliminate erroneous detections of common speech terms and account for the problems with artist names occurring as substrings in other artist names.

However, relying exclusively on artist information and ignoring song titles still left us with some remaining ambiguity. For instance, the tweet "#np Lena – Satellite" matches the patterns "*artist name – song title*" and "*song title – artist name*", with both "Lena" and "Satellite" being valid potential artist names [28].

Consequently, we decided to add track information. For the approach described in this paper we used the `musicbrainz` database [3] as knowledge base for artist names and related song titles.

Applying the approach just described, we were able to map 697,614 of the retrieved tweets (29.8%) to 97,515 unique tracks by 20,567 unique artists ("Drake" being the most popular one with 12,998 tweets).

In the following, we present different approaches to facilitate exploration of music collections, which we implemented in the proposed UI.

3.2 Genre-Based Clustering

Aiming to visualize geospatial music listening activities, we had to come up with a meaningful color-mapping. The first approach presented in this paper organizes tweets in a number of clusters, where a cluster may represent, e.g., genre, mood, country, or language and each cluster is assigned a specific color. As genre classification is the most traditional way of organizing music, our default clustering is based on genres.

Earlier work made use of `allmusic`'s [5] 18 major genres to categorize music [28]. Since `allmusic` over-emphasizes the "Pop and Rock" genre (with more than 60% of the artists being assigned to it), using these genre labels would result in one big heterogeneous cluster encompassing many different styles, which might be not very helpful to the users.

Therefore, we decided to employ tag-based clustering. For each artist we gathered the available tags from `last.fm` [2]. In order to group artists by genre we filtered the tags using a list of 1,944 known genres from `freebase` [6]). Applying non-negative matrix factorization (NNMF) [20], we split the artists (and genre tags) into k clusters, k ranging from 10 to 20 in our experiments, which seemed a reasonable range. The top-20 genre tags for 10 clusters are listed in Table 1. A higher number of clusters increases their homogeneity, but results in a higher number of necessary colors, increasing visual clutter. To chose a tradeoff between granularity and diversity of colors we allow users to set the number of clusters manually.

3.3 Similarity Estimation

In addition to the approach for static color mapping using a clustering algorithm described above, we also implemented a dynamic visualization approach. One possibility to explore music collections is to find songs by artists similar to a seed artist. Therefore, we calculate similar artists to display and apply a color mapping expressing the similarity scores with respect to a selected seed artist.

To calculate the similarity between two artists i and j we used the co-occurrence-based similarity function

$$sim(i,j) = \frac{cooc_{i,j}}{\sqrt{occ_i \cdot occ_j}}$$

with occ_i being the total number of occurrences of artist i, and $cooc_{i,j}$ being the number of co-occurrences of artists i and j. The co-occurences of i and j are defined as the number of users twittering about songs by artist i as well as about songs by artist j. This similarity function has already been proven successful [28].

3.4 Visualization & User Interaction

To visualize the geographic distribution of tweets the coordinates can be mapped to a 2D representation of a world map (cf. Figure 1). In Sections 3.2 and 3.3 two different approaches for similarity-based color-mapping have been proposed. Interaction possibilities can be categorized as follows:

- Interactions with the visualized tweets: Each tweet is represented by a small square on the map. Hovering it with the mouse opens an information window (see Figure 1) presenting information on artist and song title. Further, it is possible to apply a variety of filters (e.g. date, genre, artist name, track title)
- Interactions with the underlying map: This includes basic navigation and zooming as well as opportunities to geographic filtering.
- Interactions for statistical purposes: To facilitate analysis we offer tools to calculate play counts on different levels (song, artist, and genre) along with the mentioned filters and temporal aggregation.

3.5 Implementation

To retrieve the tweets provided by `Twitter`'s Streaming API we use the command line tool `curl` requesting all tweets with geospatial information. After filtering these tweets for music related hashtags the textual information of the tweets is analyzed according to the patterns mentioned in Section 3.1. Potential artist names and according song titles are automatically matched against our `musicbrainz` server, i.e. a `Postgres` server hourly updating to provide a current copy of the `musicbrainz` database. If a match for a certain tweet is found, this information is written to a `MySQL` database. Old tweets that could not be mapped to known songs are regularly checked again, as they may refer to songs contained in a later update of the `musicbrainz` database.

For visualization and the interactive user interface as described in Section 3.4, we decided to create an overlay visualization for `Google Maps`, using the navigation functionality provided by the `Google Maps` API [4]. For the client-side part we relied on web technologies including HTML5 and AJaX.

4 Exemplary Use Cases

In order to illustrate how users might want to explore the "world of music" using geospatial information and the concepts described above, the current section presents a number of use cases as well as approaches to achieve these goals.

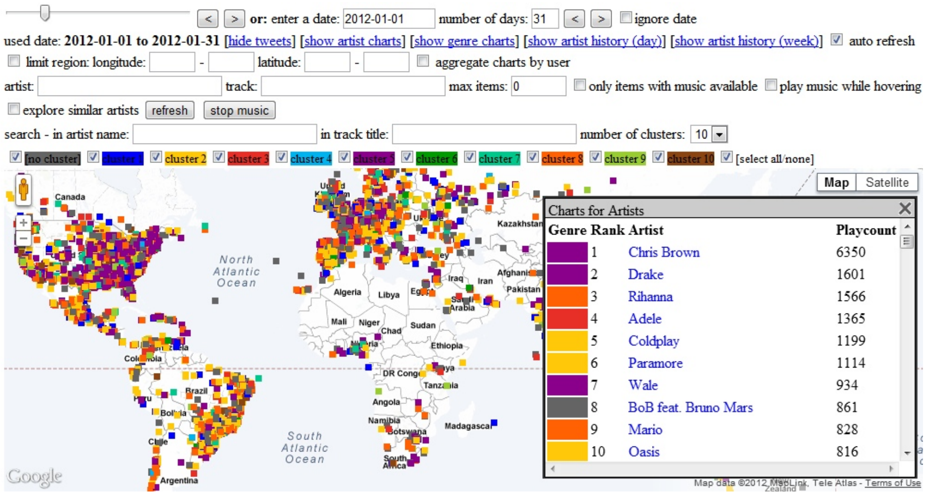


Fig. 2. Visualization of all music tweets and play counts aggregated on artist level for January 2012. Map image provided by `Google Maps` [4], ©Google 2012.

4.1 Acoustically Exploring the World of Music

As the most natural way of exploring music is listening, the proposed system aims to provide short mp3-snippets for the tracks referred to by the tweets. To this end, we matched the tweets to a collection of 2.3 million tracks, resulting in available snippets for 12,070 of the 60,651 identified tracks. To facilitate aural exploration, a "play" button is displayed in the respective information windows. Additionally, the user interface offers a mode in which snippets are automatically played when hovering the corresponding item. Furthermore, it is possible to set a filter to omit tweets without a snippet attached.

4.2 Detecting Globally Popular Artists

As shown in Figure 2 it is possible to display the play counts for all artists. Alternatively, to reduce the effect of very active `Twitter` users promoting their favorite artists, the charts may be aggregated on user level, i.e. the charts refer to users twittering about these artists instead of particular play counts.

To explore temporal dynamics, charts may be generated for customizable time windows, which enables, for instance, daily or weekly charts. Moreover, the filters mentioned in Section 3.4 can be applied.

4.3 Detecting Local Trends

In addition to global popularity estimations, analysis may be restricted to tweets of a certain geographic area. The current version of the system allows to set a

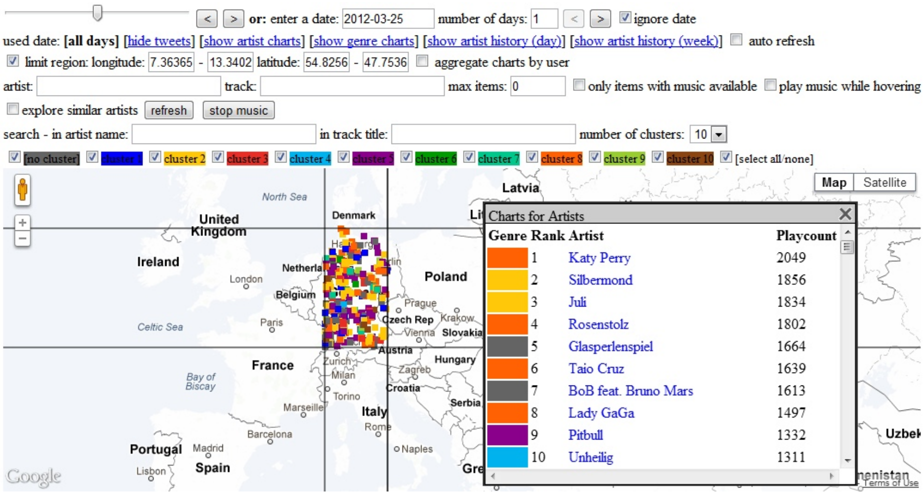


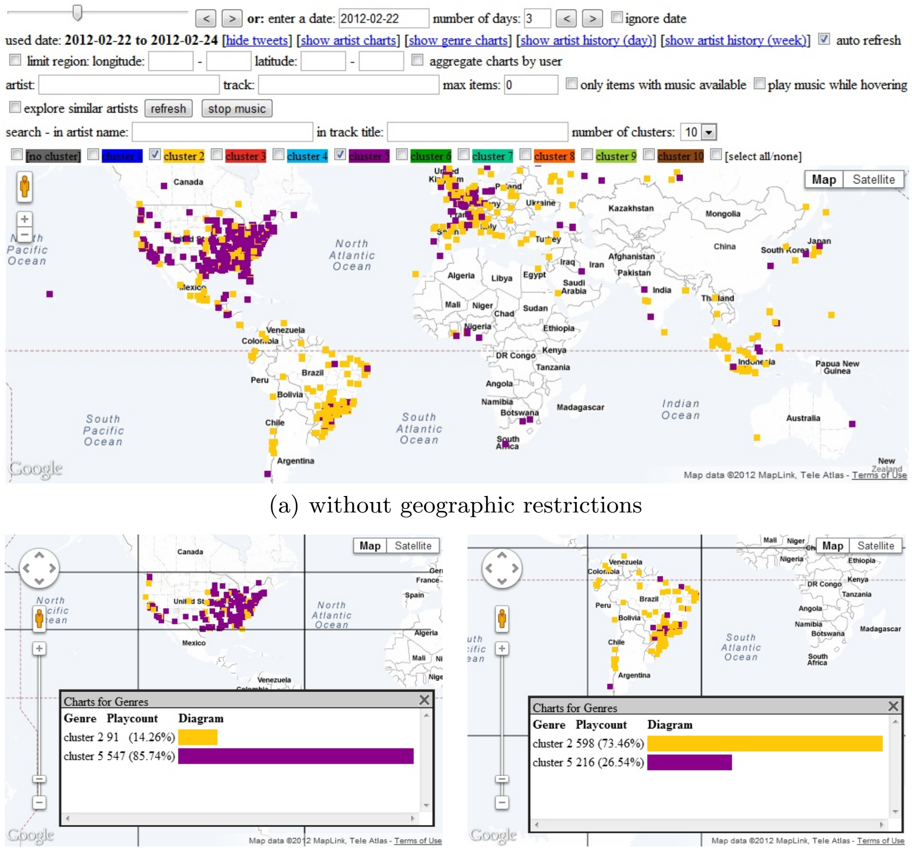
Fig. 3. Visualization of music tweets and play counts for a geographic region (roughly corresponding to Germany). Map image provided by `Google Maps` [4], ©Google 2012.

rectangular bounding box as shown in Figure 3, which allows to calculate local charts. This may help to identify local trends as well as popular local artists.

When visually exploring the map, some trends might be surmised. For instance, the first overview already gives the impression that the cluster consisting of Hip-Hop, Rap, etc. is relatively wide-spread in the United States (cluster 5 in our examples), whereas South America shows a strong preference for the Rock cluster (cluster 2 in our examples). As the user interface of our framework allows to (de-)select single clusters, we can compare those two clusters directly to each other as shown in Figure 4. Here we can see an arbitrarily selected period of three days where we can observe the previously mentioned pattern. Selecting these two areas and comparing their genre charts to each other (see Figures 4(b) and (c)) reveals that for the given period of time, cluster 2 is indeed three times as popular as cluster 5 in South America, but cluster 5 is 5.6 times as popular as cluster 2 in the United States of America. This pattern remains consistent for the whole period of observation. Further investigation reveals that also France shows a relatively high occurrence of Hip-Hop/Rap, whereas Spain and Italy (like South America) have a much stronger Rock cluster.

4.4 Exploring an Artist's Popularity

Having compared artist or genre distributions, one might be interested in detailed information on a specific artist. In addition to filtering tweets, it is possible to display the play counts for the different tracks by an artist. Figure 5 displays the play counts for songs by Madonna and shows how popularity changes with new releases. In this case, the release of the album "MDNA" (a popular track of which is "Girl Gone Wild") in March 2012 and the pre-release in February



(a) without geographic restrictions

(b) restricted to US

(c) restricted to South America

Fig. 4. Visualization of music tweets of two genre clusters for a period of three days. Map image provided by `Google Maps` [4], ©Google 2012.

2012 can be seen well in the resulting charts. Optionally, these charts can be restricted to evaluate only tweets from within a geographic region.

4.5 Retrieving Similar Artists

Another means of music exploration is by similar artists. The proposed system offers a "similar artist mode", where users can enter a seed artist (tweets of this seed are displayed in black on the world map). According to the co-occurrence-based similarity function $sim(i,j) = \frac{cooc_{i,j}}{\sqrt{occ_i \cdot occ_j}}$ (see Section 3.3), the 50 most similar artists are calculated. The similarity scores are mapped to the range $[0, 255]$. The resulting values are subsequently mapped to the RGB color space using the red channel only.

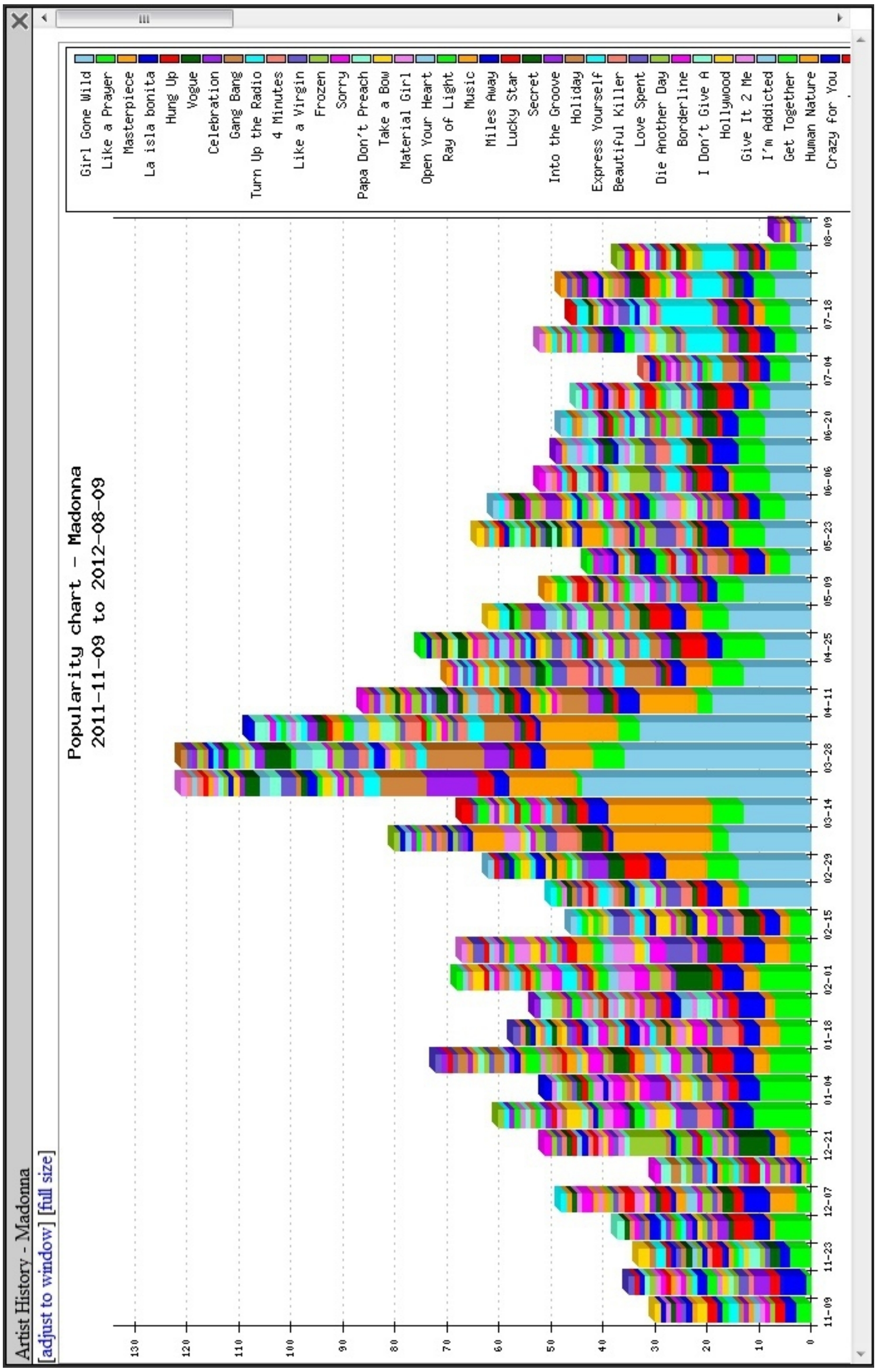


Fig. 5. Play counts for a single artist (“Madonna”) on track level, aggregated by week

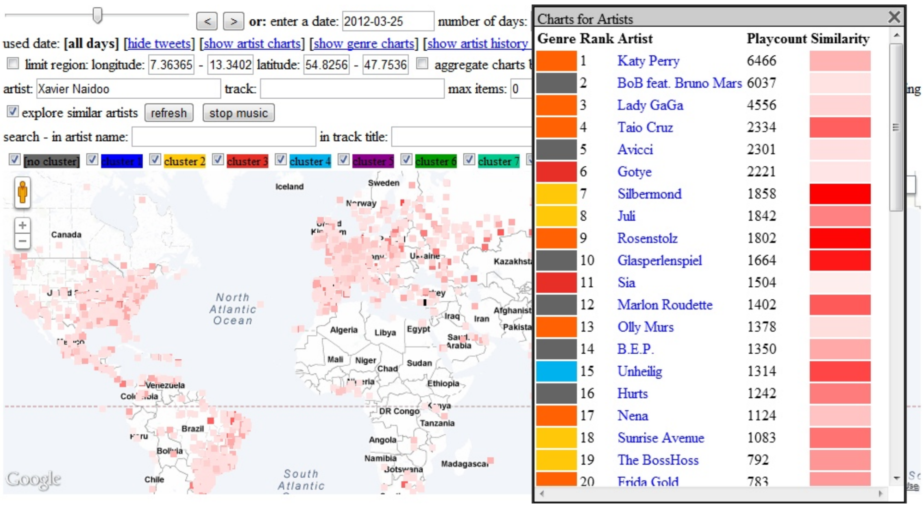


Fig. 6. Map and play counts for a seed artist ("Xavier Naidoo", black) and the 50 most similar artists, ranging from red (most similar) to white (least similar). Map image provided by `Google Maps` [4], ©Google 2012.

Figure 6 shows similar artists as well as a popularity chart among these similar artists. Clicking on artist names results in a new query using this artist as the new seed artist. This offers a multimodal data view, combining popularity and similarity information.

5 Summary and Future Work

We proposed a pattern-based approach to extract music listening activities from microblogs. Applying this approach to a data set covering nine months of microblogging activity gathered via `Twitter`'s Streaming API, we indexed tweets that offer geospatial information. In addition, we presented a framework to visualize this information and elaborated a user interface for interactively exploring world-wide music listening histories and detecting listening patterns. Using tag information we implemented genre-based clustering and used these clusters as source of information for the graphical representation. Alternatively, to detect similar artists, we implemented a visualization of the artists most similar to a selected seed artist using co-occurrence-based similarity measures. This approach could additionally be used or extended by various other types of similarity measures (e.g., based on term weight or on features obtained via signal-based audio processing), and might serve as an alternative way of proposing artists and/or tracks in dynamic playlist generation.

In order to be able to test users' hypotheses on observable listening patterns we provide possibilities to filter the data set by geographic coordinates. As a possible extension we could use information on the geographic boundaries of political regions to perform evaluations on country level. As already mentioned, genre tags are only one of many ways of clustering music – so we are exploring a variety of different clustering features and algorithms. Furthermore, we could make use of URLs or other links contained in the tweets. Via real-time processing of tweets, we could relate this information to album releases and concert tours, and further analyze temporal dynamics of artist popularity. As part of future work we will also look into building personalized music retrieval models, for which geolocalized information on music consumption might serve to incorporate cultural specifics in listening activity.

6 Acknowledgments

This research is supported by the Austrian Science Funds (FWF): P22856-N23 and Z159.

References

1. http://blog.twitter.com/2011/03/numbers.html (access: August 2012).
2. http://last.fm (access: August 2012).
3. http://musicbrainz.org (access: August 2012).
4. https://developers.google.com/maps/ (access: August 2012).
5. http://www.allmusic.com (access: August 2012).
6. http://www.freebase.com (access: August 2012).
7. http://www.spotify.com (access: August 2012).
8. http://www.twitter.com (access: August 2012).
9. M. Armentano, D. Godoy, and A. Amandi. Recommending Information Sources to Information Seekers in Twitter. In *International Workshop on Social Web Mining, Co-located with IJCAI 2011*, 2011.
10. R. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. Addison Wesley, 1999.
11. J. Bollen, H. Mao, and X. Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1):1–8, 2011.
12. D. Byklum. Geography and Music: Making the Connection. *Journal of Geography*, 93(6):274–278, 1994.
13. B. De Longueville, R. S. Smith, and G. Luraschi. "omg, from here, i can see the flames!": a use case of mining location based social networks to acquire spatio-temporal data on forest fires. In *Proceedings of the 2009 International Workshop on Location Based Social Networks*, LBSN '09, pages 73–80, New York, NY, USA, 2009. ACM.
14. Y. Duan, L. Jiang, T. Qin, M. Zhou, and H.-Y. Shum. An Empirical Study on Learning to Rank of Tweets. In C.-R. Huang and D. Jurafsky, editors, *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, pages 295–303. Tsinghua University Press, August 2010.
15. E. T. C. K. for Music Recommendations. Eva zangerle and wolfgang gassler and günther specht. In *Making Sense of Microposts (#MSM2012)*, pages 14–17, 2012.
16. S. Govaerts and E. Duval. A Web-based Approach to Determine the Origin of an Artist.. In K. Hirata, G. Tzanetakis, and K. Yoshii, editors, *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2010)*, pages 261–266. International Society for Music Information Retrieval, 2009.

17. A. Java, X. Song, T. Finin, and B. Tseng. Why We Twitter: Understanding Microblogging Usage and Communities. In *Proc. WebKDD and SNA-KDD*, San Jose, CA, USA, Aug 2007.
18. H. Kwak, C. Lee, H. Park, and S. Moon. What is Twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World wide web*, WWW '10, pages 591–600, New York, NY, USA, 2010. ACM.
19. C.-H. Lee, H.-C. Yang, T.-F. Chien, and W.-S. Wen. A Novel Approach for Event Detection by Mining Spatio-temporal Information on Microblogs. In *International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2011)*, pages 254–259, July 2011.
20. D. D. Lee and H. S. Seung. Learning the Parts of Objects by Non-negative Matrix Factorization. *Nature*, 401(6755):788–791, 1999.
21. A. Oulasvirta, E. Lehtonen, E. Kurvinen, and M. Raento. Making the ordinary visible in microblogs. *Personal Ubiquitous Comput.*, 14(3):237–249, Apr. 2010.
22. S. Park, S. Kim, S. Lee, and Woon Seung Yeo. Online Map Interface for Creative and Interactive MusicMaking. In *Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME 2010)*, pages 331–334, Sydney, Australia, 2010.
23. M. J. Paul and M. Dredze. You Are What You Tweet : Analyzing Twitter for Public Health. *Artificial Intelligence*, pages 265–272, 2011.
24. Y. Raimond, C. Sutton, and M. Sandler. Automatic Interlinking of Music Datasets on the Semantic Web. In *Linked Data on the Web (LDOW2008)*, 2008.
25. D. M. Romero, B. Meeder, and J. Kleinberg. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In *Proceedings of the 20th International Conference on World Wide Web (WWW 2011)*, WWW '11, pages 695–704, New York, NY, USA, 2011. ACM.
26. T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake Shakes Twitter Users: Real-Time Event Detection by Social Sensors. In *Proceedings of the 19th International Conference on World Wide Web (WWW 2010)*, May 2010.
27. M. Schedl. Analyzing the Potential of Microblogs for Spatio-Temporal Popularity Estimation of Music Artists. In *Proc. IJCAI: International Workshop on Social Web Mining*, Barcelona, Spain, July 2011.
28. M. Schedl and D. Hauger. Mining Microblogs to Infer Music Artist Similarity and Cultural Listening Patterns. In *Proceedings of the 21st International World Wide Web Conference (WWW 2012): 4th International Workshop on Advances in Music Information Research: "The Web of Music" (AdMIRe 2012)*, Lyon, France, 2012.
29. M. Schedl, T. Pohle, N. Koenigstein, and P. Knees. What's Hot? Estimating Country-Specific Artist Popularity. In *Proceedings of the 11th Internat. Society for Music Information Retrieval Conference (ISMIR 2010)*, Utrecht, Netherlands, August 2010.
30. B. Sharifi, M.-A. Hutton, and J. Kalita. Summarizing Microblogs Automatically. In *Proceedings of NAACL HLT*, June 2010.
31. G. V. Steeg. Information theoretic tools for social media. In *Making Sense of Microposts (#MSM2012)*, pages 1–1, 2012.
32. J. Teevan, D. Ramage, and M. R. Morris. #TwitterSearch: A Comparison of Microblog Search and Web Search. In *Proceedings of the 4th ACM International Conference on Web Search and Data Mining (WSDM'11)*, Hong Kong, China, Feb 2011.
33. J. Weng, E.-P. Lim, J. Jiang, and Q. He. TwitterRank: Finding Topic-sensitive Influential Twitterers. In *Proceedings of the Third ACM International Conference*

on Web Search and Data Mining, WSDM '10, pages 261–270, New York, NY, USA, 2010. ACM.

34. S. Wu, J. M. Hofman, W. A. Mason, and D. J. Watts. Who Says What to Whom on Twitter. In *Proceedings of the 20th International Conference on World Wide Web (WWW 2011)*, pages 705–714, New York, NY, USA, 2011. ACM.