Listening Factors: A Large-Scale Principal Components Analysis of Long-Term Music Listening Histories

Dominikus Baur, Jennifer Büttgen, Andreas Butz
Media Informatics Group,
University of Munich (LMU), Amalienstrasse 17, Munich, Germany
dominikus.baur@ifi.lmu.de, jennybuettgen@web.de, andreas.butz@ifi.lmu.de

ABSTRACT
There are about as many strategies for listening to music as there are music enthusiasts. This makes learning about overarching patterns and similarities difficult. In this paper, we present an empirical analysis of long-term music listening histories from the last.fm web service. It gives insight into the most distinguishing factors in music listening behavior. Our sample contains 310 histories with up to six years duration and 48 associated variables describing various user and music characteristics. Using a principal components analysis, we aggregated these variables into 13 components and found several correlations between them. The analysis especially showed the impact of seasons and a listener's interest in novelty on music choice. Using this information, a sample of a user's listening history or even just demographical data could be used to create personalized interfaces and novel recommendation strategies. We close with derived design considerations for future music interfaces.

Author Keywords
Music; listening history; PCA; latent factors; lifelogging.

ACM Classification Keywords
H.5.5. Information interfaces and presentation (e.g., HCI): Sound and Music Computing.

INTRODUCTION
Ask a thousand people how they are listening to music and you will get a thousand different strategies. While one listener might be an avid playlist builder and manually construct sophisticated song sets or listen to all albums from one artist in a row, another might simply switch on 'shuffle' for the whole library and be done with it. Similarly, more local factors, such as the current time of day, external influences by friends or media, events such as concerts and the current mood all play a role for choosing music. This means that it is hard to judge what factors are important for the user of a music application. Interaction designers have to make assumptions and rely on their own experience when building interfaces. Recommender systems and other personalized systems are also restricted to modeling an abstract listener in a statistical one-size-fits-all approach. People's switch to digital music, however, gives us the chance as researchers to understand how they make their listening decisions. Services like last.fm log music consumption and provide a detailed picture of what music we listened to at what time. These lists of songs together with (partial) demographic information from user profiles give us insights into what factors are important for music lovers and what parameters influence our listening.

In this paper we describe an empirical analysis of long-term music listening histories to determine the order of importance of factors that influence music listening. Using a principal component analysis we identified 13 listening factors that discern one music listener from another and give clues about people's strategies surrounding music. In addition, we describe correlations between the variables that show, e.g., how demographic factors influence music choices. Our research provides answers to how people listen to music on the detailed level of songs, what factors are important for them and how these factors are interconnected. The results can help in designing novel and personalized interfaces or recommender systems.

RELATED WORK
Research in music psychology and sociology provides insight into how people interact with music on a daily basis and what their intentions are [5]. Work by Rentfrow et al. [7], for example, uncovers the role of context in listening to music. Due to the costs of manual data collection, these studies however remain restricted to small samples or short time frames. Automatically collected histories like last.fm's are far more specific but have not been analyzed by music researchers so far. Recommendation based on such histories takes either an item- or user-based approach [1] and can be improved by taking temporal dynamics into account [4]. Actually understanding the users' reasons is not its goal. Recent work by Zheleva et al. [9] describes an adapted version of Latent Dirichlet Allocation for predicting music selection. It works on genre, artist and title information and the social aspects of a music community. The analysis, however, gives only limited information about a single user's intentions. Herrera et al. [3] and Park et al. [6] analyzed online listening histories and found interesting temporal rhythms, but restricted themselves to this aspect of

CHI'12, May 5–10, 2012, Austin, Texas, USA.

the data. A large-scale, song-level analysis for single users with a varied set of attributes has not been performed so far.

METHODOLOGY

Our goal for the study was to classify the most important factors for music listeners. Research has identified many influencing variables (see, e.g., [2,3,5,6,7,9]) but their respective impact is unknown. In order to learn about their importance, we chose an exploratory approach using a principal component analysis (PCA) to extract the most salient characteristics. This being the first such study we wanted to make sure to actually reduce the number of relevant factors and chose PCA as an analysis method.

For the data set we chose music listening histories from last.fm, a web radio that is personalized through minimal user interaction such as 'loving' or 'hating' a song. Users can additionally install so-called 'scrobbling' software that runs in the background and logs tracks that they listen to for at least 30 seconds in their player software or on their mobile devices. The resulting listening histories are accessible via the last.fm webpage or their API. Last.fm's webpage also provides a social network, where users can create profiles, become friends and share music suggestions.

Extracting the Listening History Data Set

Last.fm provides convenient access to almost all of their data through their API. In addition to user profile data, they also host a large library of general metadata about music. The data we acquired fell into four categories: *Demographic information* is an optional type of data that can be entered freely by the users. *Activity* describes a user's engagement on the website and how important the social aspect is to them (e.g., number of friends). The actual *listening histories* are made up of titles and UTC timestamps. Finally, for each title there's a wealth of *music metadata* available, such as the album it appeared on.

We acquired 5,000 user profiles via the last.fm API[1]. They contained over one million unique songs from 130,000 artists and 15 million entries. For each of these songs we then acquired the metadata (album, track number, genre based on user keywords, length, and whether it was available on the last.fm web radio). Last.fm profiles can be set to private thus barring access, but this applied only to 0.92% of our listening histories. Also, to reduce a bias due to the novelty effect and allow for long-term conclusions we only kept histories with at least two years length. We did not filter for profile activity, so the data set also contains inactive profiles. Finally, we adjusted timestamps to the listener's time zone and removed profiles without country information, resulting in 310 suitable profiles.

Representativeness and Limitations

To estimate how representative our data set was, we compared it to the audience demographics for last.fm from the Alexa[2] web traffic service and another data set of 1,000 last.fm listening histories provided by Celma [2]. We compared the (available) demographic information attribute-wise: Our data showed a similar distribution regarding the users' top-5 countries (US, Germany, Poland, UK, Finland; *Alexa*: US, UK, India, Germany, Netherlands; *Celma*: US, UK, Poland, Germany, Norway) and age (M = 24.25, SD = 11.573; *Alexa*: 18-34 year olds are over-represented; *Celma*: M = 25.37, SD = 8.315). Gender distribution had a slight male bias: female-to-male-ratio was about 30% to 70% (*Alexa*: "Similar to the general internet population"; *Celma*: 43% to 57%). The similar demographics in all three data sets indicated that our data base could deliver reliable estimates for last.fm histories.

Listening histories provided by last.fm do not, of course, represent perfect histories with all songs that a person heard in their lives: users can disable the scrobbling process, delete songs from their histories, use non-supported mobile devices, reset their profiles, etc (cf. [8]). They are, however, motivated to avoid such noise and gaps for the quality of their recommendations or the completeness of their personal music logs. Last.fm's web radio also scrobbles songs while listening that are thus not actively chosen by the user. To find such instances, we added a variable (*web radio*, see below) that was increased if a history contained ten songs in a row that were available in the web radio. Finally, taste in music also adheres to societal trends which, however, is a constant factor.

Preprocessing

We could have performed the analysis on the raw (user, timestamp, song title, song artist) variables, but research from music psychology and sociology suggests that there are other factors at play in music listening behavior as well: time of the day or week [3,6,7], for example, or the actual function the listener expects from music [5]. Therefore, we added more variables by preprocessing the raw lists of titles based on their metadata and other factors. The ideas and algorithms for these variables stemmed from the literature and our experience with music. We ended up with 48 variables in 7 different categories:

Demographic info: As providing this information is voluntary for using last.fm, not all user profiles contain it and we only used binary representations if attributes were available for *gender*, *age* and *country*.

Profile and community: Information on whether the user is a paying *subscriber* of the service, the length of their *membership*, the total *play-count* and number of *playlists*, and community information such as the number of *friends* and *shouts* (message posts) on their profile pages.

Music library: The number of *tracks per day*, the *average song length*, the *number of songs in a listening session* and the *duration of listening sessions* (sessions were separated

[1] You can find additional information about the analysis process and the results here: http://do.minik.us/listening-factors

[2] http://www.alexa.com

by at least 30 minutes). We also integrated the ratio of *distinct songs*, *albums* and *artists* (i.e. the number of items that appear only once divided by the number of all items). The *onetimers* attribute is the ratio of songs that were played only once compared to all songs.

Order, shuffle and repetition: This set of variables (all ratios) describes the *in-order listening of an album's songs* and the (chronological) *in-order listening of an artist's albums.* In contrast, we count the number of *shuffled album songs* and *shuffled artist tracks* and check whether the user listened to *complete albums* and how many album songs were *skipped*. We look at the immediate *repetitions of songs, albums* and *artists* within one session and, finally, also whether songs are commonly played as the *intro* to a listening session.

Long-term variables: To determine the fluctuation in listening we calculated the *standard deviations of played tracks for years, weeks* and *days*. 16 variables describe how many genres and artists appear only within *one of the four seasons*, two *half-years* or *halves of a day* (i.e., 8 variables each for genre and artist). The *genre-* and *artist loyalty* describes what percentage of genres and artists stay constant from one year to the next.

Temporal variables: We initially extracted 24 variables for the number of plays per hour of a day and accordingly 12 for months and 7 for days of the week. As these were highly correlated, we were able to condense them into three variables. Those reflect the percentage of plays in one specific timespan compared to the total number of plays: *ratio morning plays* (for the hours from 1AM to 1PM; maximum plays are from 7 to 11PM, minimum 2 to 6AM), *ratio weekend plays* (i.e., Saturday and Sunday) and *ratio first half of the year plays* (January to June).

Web radio and discovery: Our last category contains the ratio of regular to *web radio songs*. It also holds the *artist discovery* variable that counts how often someone starts listening to older songs by an artist after listening to a more recent record.

ANALYSIS AND RESULTS

We first performed Bartlett's test of sphericity for determining whether all variables in the sample are uncorrelated and Kaiser's Measure of psychometric Sampling Adequacy to see if our correlation matrix could be used for the factor analysis. As a second step, we determined the number of components by using the Kaiser criterion and the Cattell scree test. Finally, we rotated the component matrix to make the results easier to distinguish using PASW's VARIMAX algorithm.

Principal Listening Components

Bartlett's test confirmed the uncorrelatedness of the variables ($\chi^2 = 13425.536$, df = 1128, $p < 0.001$). The sampling adequacy was 0.825 according to Kaiser. Also, none of the variables showed a communality value below 0.40 and had to be excluded. Kaiser's criterion suggested 13 principal components, the Cattell scree test 4. After test runs with 4 and 13 components we decided for the latter due to its more distinct results (13 components are able to explain 74.278% of the total variance). The resulting components are sorted by their expressiveness (percent of variance) for distinguishing listening behavior:

C1. Distinct Music in summer, autumn and at night (27.3% of variance): The most important set of variables concern unique genres and artists from summer and autumn but also night-time. Additionally, it contains the *onetimers* variable which is the tendency to listen to songs only once.

C2. Variety (8.3%): This component describes how often songs, albums and artists are repeated, how diverse the library is album- and song-wise, if the user listens to complete albums and the yearly standard deviation of plays.

C3. Distinct Music in winter and spring (7.2%): Similarly to the first factor, unique genres and artists of the first half-year.

C4. Demographic information available (4.9%): As users tend to share this information on an all-or-nothing-basis, this component contains all *available* variables.

C5. User Listening Activity (4.2%): This component contains the *duration of listening sessions*, the *intro* and *tracks per day* variables, the *playcount* and *shuffled artist tracks*.

C6. Loyalty (3.7%): Both *loyalty* attributes, so constancy in listening to artists and genres over the years.

C7. Winter and day music (3.3%): Artists and genres that are unique to listening in winter and during day-time.

C8. General Album Listening (2.9%): This component describes whether a user listens to *shuffled album songs* and *discovered artists*.

C9. Online Activity (2.7%): Describes whether the user is a *subscriber* and uses the *web radio*.

C10. Length of membership and songs (2.6%): The time since becoming a registered last.fm user and the average length of songs.

C11. In-order album and playlist listening (2.5%): Is the user listening to songs in the right *order of the album* and has created *playlists* on the last.fm page?

C12. Social Networking (2.3%): How much is last.fm used like a social network by the user (the number of *friends* and *shouts*).

C13. Skipping and in-order artist listening (2.1%): This last component contains the number of *skipped songs* and whether the user listened to an *artist's albums in order*.

Correlations

Expected correlations exist, for example, between *playcount* and *tracks played per day* ($r(308) = +.927$, $p \leq .01$) and *onetimers* and *distinct songs/albums/artists* (r(308)

= +.926/+.929/+.882, $p \leq .01$). More interesting (because less obvious) correlations can be found between *loyalty genre* and *loyalty artist* ($r(308) = +.923$, $p \leq .01$), which means that there is no large difference between the two concepts. Also, *intro* has a negative correlation with *session length* ($r(308) = -.367$, $p \leq .01$), so the longer the sessions are, the less probable is their starting with the same song.

We performed an additional analysis to investigate the connections between age and gender and listening behavior (we had not enough histories to arrive at useful results for countries). Thus, we took a sample of all users who had entered this demographic information independent of the duration of their membership. We had 546 histories (with an overlap of 196 from the other data set). Regarding *age*, we found that there was a positive correlation with *distinct songs/albums/artists* ($r(544) = +.314/+.341/+.278$, $p \leq .01$), *onetimers* ($r(544) = +.318$, $p \leq .01$), *web radio* ($r(544) = +.259$, $p \leq .01$) and *ratio morning plays* ($r(544) = +.301$, $p \leq .01$). Older listeners use the last.fm web radio, try more new music (which might be a result of the web radio) and seem to get up earlier. For the binary *gender* variable we used an independent samples t-test. Male users had been *members* of the service for a longer time (M=643.7 days, SD=474.2 days, female: M=511.3 days, SD=415.7 days, $t(333)=3.2$, $p=.001$), had a tendency to start their sessions with the same *intro* (M=0.0086, SD=0.0071, female: M=0.0062, SD=0.0043, $t(467)=4.8$, $p=.0001$), listened to more *tracks per day* (M=38.4, SD=68.0, female: M=29.1, SD=36.2, $t(509)=2.1$, $p=.04$) and had a lower *standard deviation of plays for days* (M=0.89, SD=0.19, female: M=0.95, SD=0.18, $t(542)=-2.9$, $p=.004$).

DISCUSSION AND IMPLICATIONS

The results from the analysis provide insights into relevant criteria that discern people's listening behaviors. The temporal context appears to have a strong influence, but goes beyond daily or weekly rhythms [3,6]: Whether people have their own summer/winter artists and genres makes a large difference (C1, C3, C7) and is relevant for the validity of short-term studies about listening (e.g., [5]). There also appears to be a cut between conservative listeners and those more keen on experimenting (C2). C5 describes how extensively users are listening to music and C8 how important song organization via albums is. Unsurprisingly, there are also factors that are less specific for listening to music and more for the last.fm webpage (C4, C9, C12).

Implications of this research apply to the design of recommender systems but also for music user interfaces in general. The order of the identified components shows their relevance. A future recommender system could first rely on external factors such as the current season or time of day but also a user's preference for unknown music to improve results. Automatically determining how many album tracks in the right order are included, if a playlist always starts with the same intro song or how long the resulting playlists are (based on the average session length) can also be taken into account. Interactive media player software could be enriched with additional categories that are not available now, e.g., by adding 'seasons' to filter and sorting functions. Also, the complexity of an interface can be reduced by taking preferences into account: Listeners who only listen to albums or songs by one artist do not have to be confronted with the full song lists. This can especially be beneficial for mobile usage scenarios where complex interfaces are inconvenient. A first level of improvement can be achieved in either case without additional information by simply taking the current date. More sophisticated approaches can use demographic data such as age or gender. Finally, a user's full listening history or snippets of it can be used to extract their preferences regarding the identified dimensions.

SUMMARY AND OUTLOOK

In this paper we represented a principal components analysis of listening histories from the last.fm web service that showed the importance of temporal aspects and character traits. Our research gives an order of importance to listening factors. Next steps are to derive specific listener types and a behavioral framework from this information. A more in-depth analysis of the influence of different tastes in music (classical music versus heavy metal) might also unearth interesting factors. Finally, concrete music applications will show how much taking these factors into account can increase user satisfaction.

ACKNOWLEDGEMENTS

We would like to thank Michael Sedlmair, Raphael Wimmer and our reviewers for their insightful feedback.

REFERENCES

1. Adomavicius, G. and Tuzhilin, A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Trans. Know. and Data Eng. 17*, 6 (2005), 734-749.
2. Celma, O. Music Recommendation and Discovery in the Long Tail. *Springer*, 2010.
3. Herrera, P., Resa, Z., and Sordo, M. Rocking around the clock eight days a week: an exploration of temporal patterns of music listening. *Proc. WOMRAD* (in cnj. ACM RecSys), (2010), 7-10.
4. Koren, Y. Collaborative filtering with temporal dynamics. *Communications of the ACM*, (2010), 89-97.
5. North, A.C., Hargreaves, D.J., and Hargreaves, J.J. Uses of Music in Everyday Life. *Music Perception 22*, 1 (2004), 41-77.
6. Park, C.H. and Kahng, M. Temporal Dynamics in Music Listening Behavior - A Case Study of Online Music Service. *Proc. IEEE/ACIS*, IEEE (2010), 573–578.
7. Rentfrow, P.J. and Gosling, S.D. The do re mi's of everyday life: The structure and personality correlates of music preferences. *Personality and Social Psychology 84*, 6 (2003), 1236-1256.
8. Silfverberg, S., Liikkanen, L.A., and Lampinen, A. I'll press play, but I won't listen: profile work in a music-focused social network service. *Proc. CSCW*, ACM (2011), 207–216.
9. Zheleva, E., Guiver, J., Mendes Rodrigues, E., and Milić-Frayling, N. Statistical models of music-listening sessions in social media. *Proc. WWW '10*, ACM (2010), 1019-1028.