

Identifying time zones in a large dataset of music listening logs

Gabriel Vigliensoni
CIRMMT
McGill University
Montréal, Canada
gabriel@music.mcgill.ca

Ichiro Fujinaga
CIRMMT
McGill University
Montréal, Canada
ich@music.mcgill.ca

ABSTRACT

Knowing where listeners are is an important contextual dimension that can be used in context-aware music recommendation systems to improve their performance. This paper presents our research on identifying the time zone where listeners are by analysing their weekly aggregated music listening profiles. We collected a large dataset of full music listening histories (N=594K) of users of the Last.fm's *scrobbler* service from all around the globe, and formulated six approaches for identifying the time zone where these listening profiles have been generated based on their listeners' behaviour. The performance of these approaches was compared with a manually labelled dataset of listening profiles' time zones. We found that the best method was based on the assumption that people, in general, sleep during night time and submit fewer music logs. This approach, implemented by estimating the local minima of people's weekly aggregated listening profile, resulted in a 75 percent correctly identified time zones with a tolerance of ± 1 hour.

Categories and Subject Descriptors

H.5.5 [Information Interfaces and Presentation (e.g., HCI)]: Sound and Music Computing—*Methodologies and techniques*; J.4 [Computer Applications]: Social and Behavioural Sciences; I.5.3 [Pattern Recognition]: Applications—*Waveform Analysis*

General Terms

User modelling and personalization, Social network analysis, Information extraction from social media, Context-aware retrieval and recommendation

Keywords

Music information retrieval, listening histories, listening context, music recommendation, big data

1. INTRODUCTION

The context of music listening has been the object of study of a growing number of publications, particularly coming from the Music Psychology field. It has been suggested that the act of music listening long time ago left the spaces devoted exclusively to music enjoyment and music nowadays is listened to in a wide variety of contexts [4]. As music now accompanies our everyday life activities, the act of music listening does not only have music and listener as determinant factors, but the context of listening has appeared as another variable that influences, and is influenced, by the other two factors [3]. It has been also established that people consciously understand these interactions [2], and use them when choosing music for doing activities with non-musical goals [11].

Traditional music recommendation systems—content-based, collaborative filtering, and hybrid systems—have not typically considered the listening context to make recommendations. However, *context-aware* music recommendations systems profit from findings of previous research on musical preferences and listening context to improve the performance of their recommendations, tailoring them to the listeners' everyday life activities and context. In fact, the context of listeners has been extracted by using environmental features such as time, date, weather, noise conditions, and listener's activity and location [1, 7, 8], and context-aware music recommendation systems have outperformed traditional music recommendation strategies [5, 9, 12].

Our long-term goal is to extract contextual features of everyday music listening by analysing a large dataset of listening profiles. We hypothesize that there is enough information in these profiles to infer people's location and activities when listening to music. However, these profiles have to be normalized in time—as if all listeners would be in the same location—in order to perform a straightforward comparison. In particular, in this paper we are interested in identifying the time zones where a large dataset of music listening histories have been generated, because knowing this time shift will allow us to normalize the listening profiles afterwards. Calculating time zones from listening histories is not trivial because people travel and people's behaviour change in time, and so aggregation techniques were implemented in order to minimise these changes. Time-zone information could be extracted from the listeners' self-declared country, however this information is not always available or, if available, it does not work for countries with multiple time zones.

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.

SoMeRA'14, July 6–11, 2014, Gold Coast, QLD, Australia.

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM 978-1-4503-3022-0/14/07\$15.00.

<http://dx.doi.org/10.1145/2632188.2632203>

Section 2 of this paper presents the characteristics of our dataset and how we collected it, and Section 3 provides details about the approaches we implemented for inferring the time zones from listeners’ music listening patterns, and also about the design and results of an experimental comparison of these approaches.

2. DATA COLLECTION

We are interested in identifying the time zones where a large amount of people have been listening to music, given their music listening histories. Listening profiles can be extracted from several datasets that allow music listening behaviour research: (i) the *Yahoo! Music dataset*: provides ratings for songs collected by the interaction of users with Yahoo! Music services and an online survey conducted by Yahoo! research¹; (ii) the *Million Song Dataset, Taste Profile Subset*: this dataset gives `<user, song, play count>` triplets for more than 1M users²; (iii) the UPF MTG’s *Last.fm Dataset-1K*: offers full listening histories for 1K listeners in the form of `<user, timestamp, artist, song>` tuples with data collected from Last.fm³; (iv) the UPF MTG’s *Last.fm Dataset-360K*: supplies a dataset of `<user, artist, play count>` for 360K listeners with data collected from Last.fm⁴; and (v) the *EMI One Million Interview dataset*: in-construction database of interviews about the interests, attitudes, behaviour, and familiarity of 1M listeners with different aspects of music.⁵ Although the UPF MTG’s *Last.fm Dataset-1K* provided full listening histories we needed for our analysis, to be able to have a large enough sample to infer statistically significant patterns from the global population we decided to collect the data directly from Last.fm.

Last.fm stands out from most online digital music services because it does not only record music logs between the tracks offered by the system and its users, but also between songs and listeners of a wide range of third-party music and media players by means of their *scrobbler* service. This service “scrobble” (i.e., to automatically submit the tracks a listener play to her profile in the Last.fm database) music logs from more than 600 digital music services, browsers, media players, and devices, such as Android, Google Chrome extension, Firefox, Grooveshark, iOS, Pandora, Rdio, Spotify, iTunes, Winamp, Squeezebox, and many others.⁶ The Last.fm database has been recording music logs since October 2002, and has more than 70 million registered users.⁷

We now present the criteria and acquisition methods used to collect music listening histories for a large number of listeners (N=594K) that have used the Last.fm’s *scrobbler* service.

¹Available at <http://webscope.sandbox.yahoo.com/catalog.php?datatype=r>

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

³Available at <http://www.dtic.upf.edu/~ocelma/MusicRecommendationDataset/lastfm-1K.html>

⁴Available at <http://www.dtic.upf.edu/~ocelma/MusicRecommendationDataset/lastfm-360K.html>

⁵Part of the data available at <http://musicdatascience.com/emi-million-interview-dataset/>

⁶Non-exhaustive list of scrobblers available at <http://build.last.fm/category/Scrobblers>

⁷Data retrieved using the Last.fm API on April 21, 2014

2.1 Data criteria and acquisition

Aggregating people’s music listening histories implies collapsing their music logs into several periods of time. In order to obtain even data across aggregated weeks, months, seasons, or years, we looked for listeners with at least two years activity submitting music logs since their first scrobble was submitted. Also, as people sometimes register for a service, try it, and never used it again, we looked for listeners with an arbitrary average of, at least, ten scrobbles per day in order to ensure they have been actively submitting music logs. With only these two restrictions we forced all listeners in our dataset to have a minimum of 7300 music logs submitted to the Last.fm database.

Data acquisition was performed by means of using several machines calling the Last.fm API during a period of one year in order to gently comply with their terms of service.⁸ Most interactions with the Last.fm web services require knowing listeners’ usernames in advance, and so in order to obtain a large number of them, we started this research project by sampling the “Recently Active Users” Last.fm webpage⁹ periodically, and used the retrieved usernames as seeds to get more usernames. But with a hint from Eugenio Tacchini, we dramatically increased the acquisition of usernames by using Last.fm ID (*lfid*) numbers, which increased sequentially, when calling the Last.fm API.

We collected our data by using the Last.fm’s API method `user.getRecentTracks()`. With this call we obtained full listening histories of listeners that we stored in the form of `<username, timestamp, artist-MBID, album-MBID, track-MBID>` tuples. Along with this data, we also stored part of the metadata available for each listeners: their mandates *lfid*, *user type* (i.e., their user status in Last.fm, such as subscriber or user), and *registration time* in coordinated universal time (UTC); and the optionals, self-declared *age*, *country*, and *gender*. We also recorded the timestamp of the first and last music logs, and the average number of scrobles per day.

2.2 Data cleaning

Once we started acquiring data we noted that some of the listeners were “super users”: they had far more music logs than the average of listeners. After close inspection, we realized that there were two issues in some music listening histories: (i) there were listeners with many duplicated music logs (i.e., same timestamp and music items IDs); and (ii) some people had scrobbles that were too close in time (i.e., less than 30 seconds apart, which is the minimum that Last.fm requires to consider a played track as a valid scrobble). We hypothesised these issues are artifacts produced by the interaction of scrobblers and the Last.fm servers and database that have not been fixed. We decided to filter out all duplicated logs and also the scrobbles which were less than 30 seconds apart in time. There were some other extra issues (e.g., listeners with the same *lfid* but two versions of their username, or listeners with strange timestamps in their listening history), but these were less common in the dataset and so we did not filter these logs or listeners.

⁸Last.fm legal terms and policies are available at <http://www.last.fm/legal>

⁹<http://www.last.fm/community/users/active?page=1>

2.3 Data integration

In order to centralize all metadata acquired from Last.fm into one machine and facilitating the data inspection we implemented a Solr search server instance.¹⁰ However, as the amount of music listening logs we collected grew rapidly (i.e., more than 70M music logs per day), it was not feasible to perform fast calculations in stand alone computers or small servers, and so we decided to move our data into a high-performance computing (HPC) facility. In order to improve the I/O performance and the manageability of the data we created compressed TAR files of a couple of Gigabytes for storing the raw listening logs, and we used the HDF5 data model for storing the features extracted from the data. A map-reduce approach based on the message-passing interface (MPI) was implemented to process the data.¹¹

2.4 Data demographics

Our current dataset consists of 27 billion music logs taken from 594K users’ music listening histories. In this massive repository there were more than 555K different artists, 900K albums, and 7 million tracks. Table 1 summarizes some features of the data and metadata in our dataset.

| Dataset | Listeners | Logs | Artists | Albums | Tracks | |
|-----------------------|-----------|-----------|----------|------------|--------------|------|
| | 594K | 27MM | 555K | 900K | 7M | |
| Listener's | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| Age (years) | 0 | 21 | 24 | 25.4 | 27 | 113 |
| Number of logs | 7K | 24K | 37K | 49K | 60K | 998K |
| Scrobbling age (days) | 731 | 1192 | 1653 | 1721 | 2188 | 3929 |
| Gender | Declared | Non-decl. | | Female | Male | |
| (%) | 81.6 | 18.4 | | 28.7 | 71.3 | |
| User type | Alumni | Moderator | Staff | Subscriber | User | |
| (number and %) | 70 (~0%) | 21 (~0%) | 33 (~0%) | 14K (2.4%) | 580K (97.6%) | |

Table 1: Dataset summary

In terms of age, 72 percent of the listeners declared their age. The median was 24, and the mean 25.4. 98 percent of the listeners with age declared have a self-declared age in the range [15, 54] years old, and among listeners in the extra two percent there are some that declared more than 100 or less than 5 years old, which should correspond to “liars”. In order to reduce noise, most of the analysis we carried on listening behaviour in relation to the listener’s age excluded listeners outside of the [15, 54] age range.

In terms of gender, more than 80 percent of people declared a gender. Among them, the percentage of listeners that declared to submit music logs as *male* was more than double the percentage of listeners that chose *female*. Also, Last.fm assigned their users a “user type” according to their involvement with the service: *subscribers* were those that paid a monthly instalment to Last.FM for getting unlimited streaming tracks, *users* were people using the service with no special privileges, *alumni*, *moderator*, and *staff* were statuses for people that formerly had worked or were actually working for Last.fm. We considered that *gender* and *user type* were relevant for analysing listening histories because

¹⁰Solr search platform available at <https://lucene.apache.org/solr/>

¹¹MPI Standard available at <http://www.mpi-forum.org/docs/mpi-3.0/mpi30-report.pdf>

they could provide insights about the correlation of these features with music behaviour and preferences, but it is out of the scope of this paper.

In terms of location, 82 percent of the listeners in our dataset reported a country. Among these countries, the United States had by far the largest percentage of listeners (21 percent), and 19 countries had at least one percent of the total amount of listeners in the dataset. These “top countries” combined accounted for more than 85 percent of the total number of listeners in the dataset. Fig. 1 shows a world map with the relative number of listeners per country, normalized by the corresponding number of inhabitants in each country.¹² The colour palette is based on vigintiles (i.e., 20 quantiles) of the data, where red indicates the highest vigintile and blue the lowest one. This map could be interpreted as the degree of Last.fm market penetration by country. By looking at the higher vigintiles—red and orange colours—we can see that listeners from mostly all time zones were represented in our dataset. Moreover, while Northern European and Australasian countries had the largest proportion of listeners submitting music logs to Last.fm, the United States was no longer the first ranked country. People in Africa, South Asia, and Far East Asia did not extensively use Last.fm.

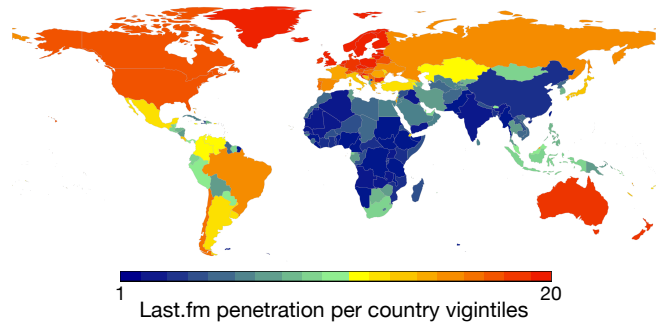


Figure 1: Relative number of listeners per country, normalized by the corresponding number of inhabitants in each country. Red and blue colours indicate highest and lowest vigintile.

3. IDENTIFYING TIME ZONES IN MUSIC LISTENING PATTERNS

Last.fm records UTC timestamped logs using UNIX time stamp format for all logs submitted by listeners regardless of where they actually are. In other words, all music logs have the same temporal point of reference, but there is no data about the city, country, or time zone where these logs were generated. Also, UTC does not change with seasons, and so changes in local national time for countries following daylight saving time are not stored. For the sake of our project—identifying time zones from listeners’ listening patterns—these two issues were a problem because the aggregated listening patterns from people in different time zones were shifted depending on where they were, and so we could not directly compare their patterns to obtain some knowledge from them.

¹²Population data for the year 2012 taken from <http://data.worldbank.org/indicator/SP.POP.TOTL>

The next subsection describes the approaches we developed to normalize the time zone of listeners, and provides the details and results of an experiment we carried to determine the best time-zone normalization approach.

3.1 Normalizing the time zone of listeners

In order to compare aggregated music listening patterns, we had to find a way to shift them in time as if all listeners were in the same time zone. For coming up with different approaches we relied on two assumptions: (i) listeners, in general, share an overall music listening pattern during the day time which resembles a normal distribution of music listening logs; and (ii) listeners, in general, sleep during night time and submit fewer music logs. For testing the former assumption we designed an implementation based on finding the time lag to obtain maximum cross correlation with a sample population of listeners located in time zone 0; and finding the local minima of the weekly aggregated listening patterns for testing the latter. In order to test if the performance improved we also formulated variants of these two implementations, adding up to a total of six approaches for time-zone identification.

3.1.1 Time zone 0 cross correlation

Our first approach relied on the idea that listeners, in general, share a similar listening pattern profile. The time zone 0 cross correlation (*TZ0_XCORR*) approach calculated the lag value k which returned the maximum correlation between $x[t + k]$ and $y[t]$ given a cross correlation function $ccf(x, y)$. The k value was the estimation of how many hours of difference there were between any two listeners. As a fixed control time series $y[t]$ we chose to use the aggregated listening profiles for listeners with self-declared country in time zone GMT +0. To accomplish this, we took all listeners in the dataset which had declared as their country the United Kingdom, and we expected this population would be large enough to minimize the effect of people that submitted music logs while travelling, or that might have lied about their self-declared country. We calculated their weekly aggregated listening mean, and we used it as the control time series. Fig. 2 shows mean and 95% CI error bars of the control time series. It can be seen that there is no statistical difference to claim that there were differences in music listening behaviour on weekdays and weekends, in opposition to what has been previously found [6, 10]. Also, there is no overlap between the CI error bars for peak and valleys in day and night, and so we found statistically significance support to our second assumption which stated that listeners scrobble less music logs songs during the night time that during the day.

3.1.2 Local minima approach

Based on the assumption that the chances of being scrobbling by night are less than in the day, we looked for the multi local minima within a week (i.e., the local minima indexes in the weekly aggregated listening profile with a moving neighbourhood time period of 12 hours). However, as we also expected that listeners change their behaviour on weekends—being different than on weekdays—we focused our analysis in extracting the local minima in weekdays only. We implemented this approach by using the “WMTSA” R package¹³,

¹³ Available at <http://CRAN.R-project.org/package=wmtsa>

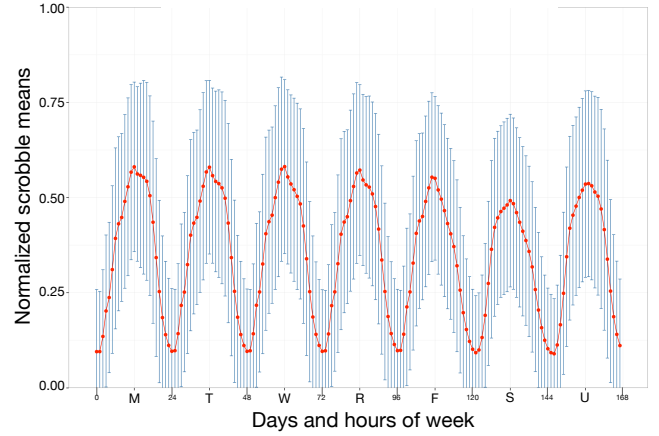


Figure 2: Control time series mean and 95% CI error bars for the *TZ0_XCORR* approach computed with the weekly aggregated listening profile from 42K listeners self-declared to be in the United Kingdom.

which gave us back a matrix with the local minima indexes and values. We then normalized the indexes values into the range $[-12, 12]$, we averaged them, and we finally rounded this value. This integer number corresponded to the estimated time zone for each listener. Averaging and rounding the multi local minima points facilitated us to overcome the problems of: (i) points that were not retrieved, because the effectively retrieved points were used for the final estimation; and (ii) false positive points, because their value was averaged and their effect minimized.

After implementing this approach, we realized that the local minima indexes of listening profiles with “flat zones” (i.e., valleys in a listening profile with no substantial change for a period of time) were pointing to the beginning rather than at the middle of the “flat zones”. In order to address this issue, we implemented a variant of the local minima approach that computed the local minima for the original as well as the reversed version of the time series of weekdays. The indexes obtained with the backward version were reversed, and were averaged with those retrieved by the forward version. By this method we expected to more accurately point the local minima index to the middle area of the “flat zones”. The approach that calculated the time zone based on only one forward pass was named forward local minima (*FF_LM*), and the approach that calculated the time zone based on two passes, one on each direction, was named forward-backward local minima (*FB_LM*).

3.1.3 Seasonal decomposition

In addition to the previous experimental factors for testing the best approach to identify time zones, we also employed time series decomposition to isolate the cyclic seasonal data from any trend and noise in the weekly aggregated music listening profiles. We implemented this approach to verify if by just using the seasonal decomposed data—removing the noise and trend of the profile—the performance of any of the previously explained approaches improved.

Hence, we ended up with *raw* and *seasonal* variants for each one of our approaches based on the time zone 0 cross correlation and the local minima. Hence, the six approaches we tested were: *TZ0_XCORR*, *FF_LM*, *FB_LM*, *SEAS_TZ0_XCORR*, *SEAS_FF_LM*, and *SEAS_FB_LM*.

3.2 Experimental procedure

The purpose of our experiment was to compare the performance of two approaches and four variants in identifying the time zones where weekly aggregated music listening patterns were generated. To accomplish our goal, we randomly drew listening histories for 384 listeners from our dataset, aggregated their data into a week, and manually labelled each one of these profiles in a time zone within the range $[-12, 11]$. We named this subset the *control dataset*. Although we expected that labelling listeners' time zones would be easy, we realized that their annotation was difficult. Most people in the control dataset had cyclic patterns but some of them had slight changes in their weekly patterns, and so choosing one specific time zone was not obvious. Furthermore, a few listeners did not have a clear cyclic pattern at all. Fig. 3 shows the weekly aggregated listening profile of six listeners in different time zones. While it seems easy to estimate the time zone of listeners in the upper row, for the ones in the bottom row, the annotation is problematic.

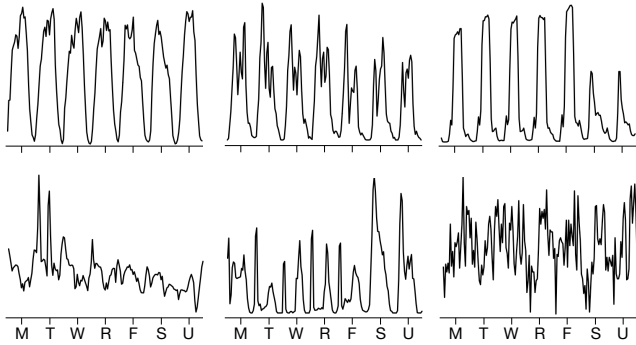


Figure 3: Weekly aggregated music listening profiles of six listeners in our control dataset. While the time zones for the profiles in the upper row can be easily estimated, the lower row profiles time zones are problematic.

After we created the control dataset of time zones per listening profile, we proceeded to compute the time zones for 1,000 populations replicated from the original sample of 384 listening profiles using the bootstrap technique with the six aforementioned approaches. We wanted to estimate the performance of each method not only by measuring the percentage of perfect matches, but also by how close the time zones were estimated, and so we quantified the performance of each approach by their time difference in hours. For example, if the computed and the manually labelled listening profile had the same time zone, their time difference was zero, but if the computed profile was shifted was lagged two hours to the left, their time difference was -2 hours.

Fig. 4 shows the performance of the six approaches in identifying time zones of listening profiles. Bars and colours indicate the time differences in hours, ranging from $[-12, 11]$,

where a zero-hour difference is orange. 95 percent CI error bars show upper and lower limits for 1,000 populations replicated from the original sample of 384 listening histories using bootstrap at $\alpha = 0.05$. It can be seen that although small, the largest percentage of correctly computed time zones (i.e., time difference was zero) was achieved by both *TZ0_XCORR* and *SEAS_TZ0_XCORR*, with statistically significant differences with the other approaches. However, when analysing the performance of all methods with a tolerance of ± 1 hour, the *FB_LM* and *FF_LM* approaches—methods based on the assumption that people scrobble less in the night, had a much better performance. In fact, these methods computed appropriately the time zones, with a one-hour window tolerance, for 75 and 70 percent of the dataset respectively. However, there was no statistically significant difference to determine which one of these two methods was better, and so the approach we designed to overcome the problem of listeners with “flat zones” was not effective. Finally, the *seasonally* decomposed versions of the local minima approaches had a poorer performance than their *raw* counterpart, which implied that the decomposition was also not relevant.

4. CONCLUSIONS AND FUTURE WORK

This paper presented our research on identifying the time zones where a dataset of listening histories have been generated. We started by summarising the currently available datasets for music listening research and explained why we chose to collect our own data, providing details about the collection method, the filtering of data, and some of its demographic features. We then stated the assumptions we made for formulating and implementing our approaches for time-zone extraction from listening patterns: (i) listeners, in general, share an overall music listening pattern during the day time; and (ii) listeners, in general, sleep during night time and submit fewer music logs. We finally detailed an experiment that compared the performance of two different approaches and its variants with a control dataset of manually labelled listening profiles' time zones. Overall, the best approach was based on finding the local minima of weekly aggregated listening profiles, which was based on the assumption that people share moments of sleeping at night. This approach recognized correctly 75 percent of the time zones with ± 1 hour of tolerance.

Some drawbacks may be raised concerning our experimental design and analysis. First, it can be said that the time zone of listeners is not fixed, especially considering long listening histories. In fact, people may travel for vacations or work while still being submitting music logs; or may move to a different time zone indefinitely. Although the aggregation of long listening histories minimizes the former issue, it can not cope with the latter. It would be interesting, though, to investigate the actual percentage of people migrating to different countries, and see if this point could have an effect in our analysis. Second, in spite of the fact that individual schedule variation (e.g., morning people as opposed to night people) could possibly exceed small differences in time zones, the approaches we have presented in this paper for identifying time zones in people's listening profiles can still be used to know the shift of these cyclic listening patterns in time, allowing a more straightforward comparison between them.

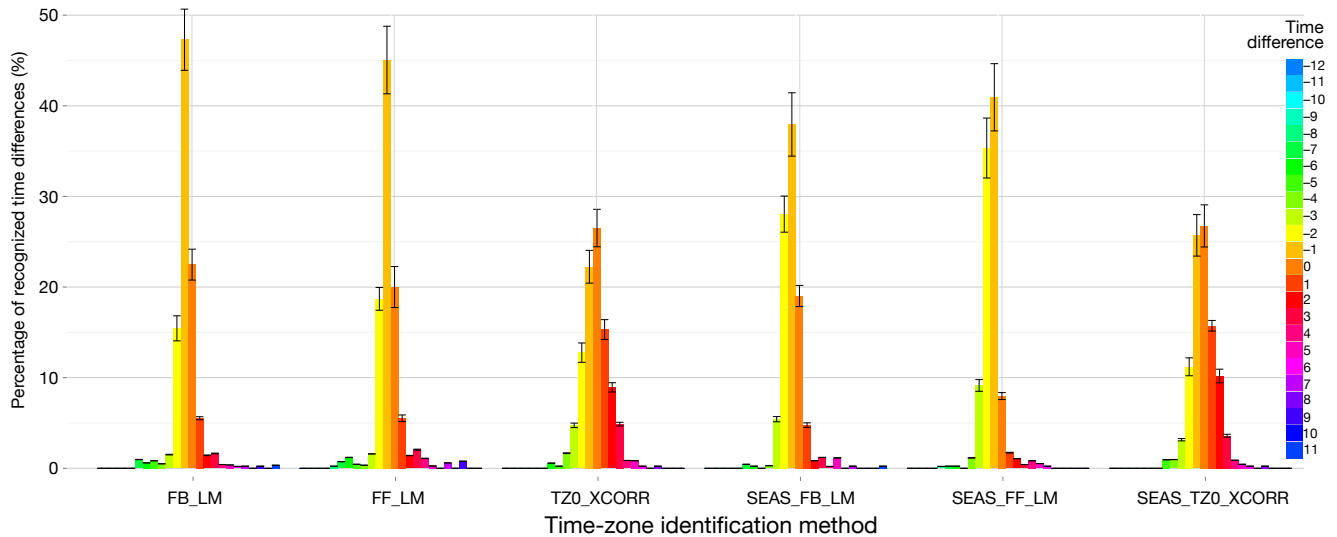


Figure 4: Performance of six approaches in identifying time zones of listening profiles. The plot shows percentage and 95% CI error bars for each time difference between the manually labelled and computed time zones for 1,000 populations taken with replacement from a sample of 384 random listening histories.

As ideas for further research, we believe that time-zone normalization could be improved by paying attention to the national public holidays for each country. We hypothesise that listeners’ listening behaviour in holidays is different in comparison with working days, and so this could be used to find the specific time zone and country where they have been submitting music logs. Furthermore, the daylight saving time shift could be used to find out the hemisphere where people is. Its application in the summer time implies a shift of one hour, but in opposite directions for each hemisphere. These changes in people’s routines could be detected in their music listening profiles, and could be used to determine where they are.

5. ACKNOWLEDGMENTS

This research has been supported by BecasChile Bicentenario, CONICYT, Gobierno de Chile, and the Social Sciences and Humanities Research Council of Canada. Important part of this work was made using ComputeCanada’s High Performance Computing resources. The authors also would like to thanks Eugenio Tacchini for his idea about calling the Last.fm API using IDs instead of usernames.

6. REFERENCES

- [1] S. Cunningham, S. Caulder, and V. Grout. Saturday night or fever? Context aware music playlists. In *Proceedings of the Audio Mostly Conference*, pages 1–8, Piteå, 2008.
- [2] T. DeNora. *Music in everyday life*. Cambridge University Press, 2000.
- [3] D. J. Hargreaves, R. MacDonald, and D. Miell. How do people communicate using music. *Musical communication*, pages 1–25, 2005.
- [4] V. J. Konečni. Social interaction and musical preference. In D. Deutsch, editor, *The Psychology of Music*, pages 497–516. New York: Academic Press, 1982.
- [5] D. Lee, S. E. Park, M. Kahng, S. Lee, and S. Lee. Exploiting contextual information from event logs for personalized recommendation. In R. Lee, editor, *Computer and Information Science*, pages 121–39. Springer-Verlag, 2010.
- [6] A. C. North, D. J. Hargreaves, and J. J. Hargreaves. Uses of music in everyday life. *Music Perception*, 22(1):41–77, 2004.
- [7] G. Reynolds, D. Barry, T. Burke, and E. Coyle. Towards a personal automatic music playlist generation algorithm: The need for contextual information. In *Proceedings of the 2nd Audio Mostly Conference*, pages 84–9, 2007.
- [8] G. Reynolds, D. Barry, T. Burke, and E. Coyle. Interacting with large music collections: Towards the use of environmental metadata. In *2008 IEEE International Conference on Multimedia and Expo*, pages 989–92, 2008.
- [9] D. Shin, J. Lee, J. Yeon, and S.-g. Lee. Context-aware recommendation by aggregating user context. In *2009 IEEE Conference on Commerce and Enterprise Computing*, pages 423–30, 2009.
- [10] J. Sloboda, S. O’Neill, and A. Ivaldi. Functions of music in everyday life: An exploratory study using the experience sampling method. *Musicae Scientiae*, 5(1):9–32, 2001.
- [11] J. A. Sloboda, A. Lamont, and A. Greasley. Choosing to hear music: Motivation, process and effect. In S. Hallam, I. Cross, and M. Thaut, editors, *The Oxford handbook of music psychology*, pages 431–40. Oxford University Press, Oxford, 2009.
- [12] J. Su, H. Yeh, P. Yu, and V. Tseng. Music recommendation using content and context information mining. *Intelligent Systems, IEEE*, 25(1):16–26, 2010.