Towards time-aware contextual music recommendation: an exploration of temporal patterns of music listening using Circular Statistics

Zuriñe Resa

Master Thesis MTG - UPF / 2010

Master in Sound and Music Computing

Master Thesis Supervisor:

Perfecto Herrera

Department of Information and Communication Technologies

Universitat Pompeu Fabra, Barcelona

Abstract

Music is present in many situations of our daily life, it's a way of showing our personality traits, a way of reinforcing our social identity or even a way of inducing an emotional state or mood. These factors that are highly linked to our music consumptions and preferences, are highly changeable, they do vary over time and they can be assessed to be linked to physiological or natural rhythms. For this reason, it seems reasonable to explore the influence of these rhythms in music listening activity. This thesis is an attempt to explore temporal patterns in relation with the time of the day or the day of the week which can be observed when tracking specific artists or music genres. The final goal is to characterize music listeners based on the information extracted regarding the time of reproducing each artist and genre, assessing temporal patterns and thus, cyclic rhythms.

So as todetect these temporal patterns or rhythms, a circular statistical analysis is performed over a data-set containing the listening habits of almost a thousand users of an online radio - music recommender system. This analytical approach offers a wide range of strategies to examine "circular data", data where the period of measurement is rotationally invariant(as the daily hours which range form 0 to 24, 24 being the same as 0). This way temporal patterns are identified regarding the time of the day or the day of the week (respectively, a period of 24 hours and of 7 days). We show that for certain users, respectively for artist and genres, 20% and 40% of the found temporal patterns can be used to predict music listening selections with above-chance accuracy.

This finding has as a possible application personalized playlist generation and music recommendation based on providing user-specific suggestions at the "right" moment.

Acknowledgments

First of all, I would like to thank Dr. Xavier Serra, for all his guidance, support and for being so patient with me along these two years of master. Thank you for giving me the opportunity to carry out this master, which in some point has change my life, having the chance to know more interesting people than I could imagine and growing as a person and researcher.

Then, I'd like to sincerely thank my supervisor Perfecto Herrera, who has been immensely involved in this project and has teach me all that enthusiasm that is behind an idea and its consequence research. I'm proud that I chose this topic, as he said "far away from making music or sounds", because it has been a challenge to find the way and being sure enough of the work we were doing. Also, I'd like to express my gratitude to Emilia Gomez, who has been really involve in the development of the whole Master, following each of our steps. Thank you for your guidance and advices.

Moreover, I thank all the researches of the MTG, for teaching some lessons or courses, and thus being an incomparable "transfer" of knowledge, or just for sharing a conversation, comments or opinions. Specially, I would like to thank Mohamed Sordo, for all his invaluable "technical" support and as we say here, for "giving without willing anything back".

Furthermore, I'd like to thank all the master students I have had the pleasure to share this great experience of the master degree with along these two years. Those last year "parties" at the room 55.312 of the MTG won't be never forgotten. Particularly, I'd like to personally thank Laia, Fran, Pablo and Frederic, who have been my fellow mates during these two years, who have trust me whenever I was doubting and have become my little family here.

Finally, I'd like to thank my parents and friends for their support. As someone in the field said some years ago, " ... I hope they finally understood what I was working on". I extend a thank to my flat mates, who had support me whenever I was down, trying to understand what I was saying.

And to Jon, for being a tireless source of inspiration, confident of my work and the best live partner I could ever imagine.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this chapter the context and motivation of the thesis are stated, as well as the goals and expected results. Briefly, the structure of the thesis is also presented.

1.1 Motivation

"When do people listen to music? Is there any temporal pattern that describes music consumption and preferences of each individual?" People asked to answer these kind of questions would give an overview of similar responses such as: "Depending on the time of the day", "The situation or context I'm involved", "If I'm working or at home", "Surely on weekends I listen to different styles of music than on weekdays"... As said, when answering the first proposed question almost all the people agree that there are times of the day, week or month in which they do listen to different type of music genres, music moods or even artists. Indeed, music listening habits are stated to be one of the most descriptive aspects of our personality, highly linked with the way we are [32]. There are in fact, many associations between individuals' music preferences or tastes and many aspects of our lifestyles, that vary across different cultures, which do reflect our personality traits, our aim on identity building and our social self view [30]. In the same way, the observation that human behavior is modulated by rhythms of hormonal releases, exposure to light, weather conditions, moods and also the activity we are engaged into[2],[14] paves the way to our main hypothesis: there are musical decisions that reflect the influence of those rhythms and therefore show temporal patterns of occurrence. On one side, the existing links between music and mood and between music and activity on the other side, lead us to clarify the connection between natural rhythms and music selection. In both cases, music has functional values either as a mood regulator[33] or as an activity regulator [16]. Therefore, as mood and activity are subject to rhythmic patterns and cycles, music selection expressed in playlists could somehow reflect that kind of patterning[33], [37].

In order to identify individuals' music preferences and the link between them and time of the day or day of the week, we use a data-set containing the playcounts of users from Last.fm [1]. Last.fm is a popular Internet radio site for streaming music, founded in the United Kingdom in 2002. The site offers numerous social networking features and can recommend and play artists similar to the user's favourites. The playcounts from user's of Last.fm are accessible via the some methods of the open API. So as to characterize the mentioned rhythmic patterns or cycles by means users' playlists, we use circular statistics analysis. This technique allows us to detect daily or weekly rhythm by observing the time of reproduction of an specific artist or genre. Therefore, for each user playlist we perform the analysis in two different ways; by tracking the data regarding a specific artist selection time and regarding a specific genre one.

1.2 Goals and expected results

The goals for this thesis is to characterize different individuals music listening behavior, by means of identifying temporal patterns in music listening activity. The

[1] http://www.lastfm.es/home

core of the process relies on the circular statistical analysis of a data-set coming from online radio users' playlists. So as to perform the analysis of the data, the next steps are followed:

1. Data acquisition
2. Data cleaning and organizing
3. Analysis using Circular Statistics(for daily and weekly rhythm)

   Circular features extraction/selection

   Temporal patterns detection(Hypothesis testing)
4. Evaluation of the performed analysis

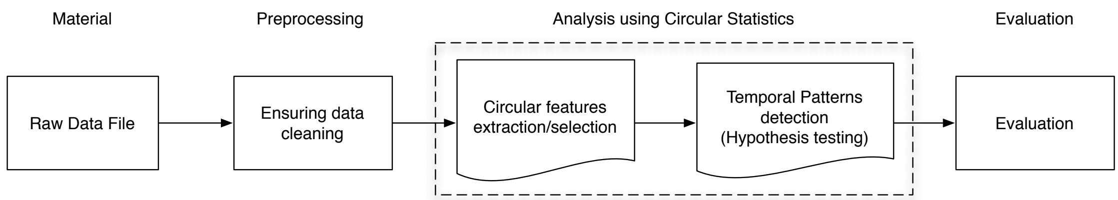

Figure 1.1: General process block diagram

Once the above steps are performed, the expectation at the end of this research is that we have a set of artists/genres for each individual who does follow a specific daily or weekly rhythm, and thus, are relevant regarding their music listening activity in the mentioned online radio system. We also expect to confirm that each individual's listening behavior is far away from the global average one, which indeed would mean that, personality or emotional state is reflected at the time we listen to music in a personal way.

1.3 Structure of the thesis

About the structure of this document, chapter one states the motivation and goals of this thesis. Chapter two talks about the state of the art of each of the different fields that are involved in the development of this research. Chapter three explains the methodology used for performing the adequate analysis of the acquired data-set, results of which are presented in the fourth chapter. Finally, the fifth chapter discusses the whole process.

Chapter 2

State of the art

2.1 Music relates with personality, behaviour and emotions

Music is a ubiquitous social phenomenon, it is the center of many social activities and is part of our individual every day life. Music listening activity is present in almost every situation of our daily life, at home, in the office, in the car, whenever you go to a shop and in every restaurant or club. At this very moment, there are people that no matter which situation their are in, they are listening to music. Moreover, the music isn't a merely entertainment activity but is deeply related to community and identity building. Just as individuals shape their social and physical environments to reinforce their dispositions or self-views, the music they select does serve as a similar function. Very often, listeners do identify themselves with their musical choices, equating their music collection with their personality [21]. Who hasn't heard or said: "My personality goes into my iPod". Cattell who was among the first ones to theorize about how music could contribute to understanding personality, believed that preferences for certain types of music reveal important information about unconscious aspects of personality [7]. So as to this, many researchers on the field have regarded music preferences as a manifestation of more explicit personality traits [32], [15]. Indeed, some studies about music sharing among communities and networks discovered that when individuals have access to other's playlists, we are able to make judgements more about the personality than about the type of music they do listen to [38]. Due to these findings about communities (i.e. groups of people working in the same factories, departments...), by thinking that if you have some musical taste in common with other person, it won't be difficult to have also some personal views or ways of acting similar. Actually, as we do with many other facets of our lives, we tend to listen to music that our friends or colleges listen to, which helps us to define our social identity, personality and musical tastes or preferences. There are many reasons beyond the question "why do you listen to music?", and as said, personality traits are so linked to it. Musical taste or preferences, are so tightly linked to personality but also they do depend on context and situations in which we are involved. It is believed that we all have "lifesoundtracks", trying to explain that the music we listen to inspires, motivates, calms, excites, and generally moves along the actions of our everyday life [25]. People do often employ music as a way to motivate themselves to accomplish certain tasks or as mood induction. Music enhances our emotional states, varies or calms them. Many brain centers are activated when we do listen to music; as they are the ones related to the cognitive centers or the frontal lobe circuits, which are involved in planing, motivation or expectations building. In fact, is easy to realize that for some music styles the heart rate or the blood pressure changes, and so, each one of us is an expert in selecting the right music so as to achieve a certain emotional state[34]. As an example, after doing exercise people tend to listen to low arousal music (related to low tempo music) as a way to moderate their arousal, even they probably had prefer to listen to high arousal music while doing exercise [18]. So, as emotional states do, they vary over time, they do change according to particular contexts and situations; our musical preferences or taste will work in the same way.

2.2 Chronobiology, or how time influences natural rhythms

Chronobiology is the discipline that deals with the time and rhythm in living organisms. The influence of circadian rhythms (those showing a repetition pattern every 24 hours approximately, usually linked to the day-night alternation), but also of ultradian rhythms (those recurring in a temporal lag larger than one day like the alternation of work and leisure or the seasons) has been demonstrated on different levels of organization of many living creatures, and it is critical that some biological cycles are maintained in order to keep an optimum health [26]. Many researches in the field agree on the importance of tuning this rhythmicity in our bodies, in appreciating them as a balance indicator [14].

Below an interdisciplinary point of view of this issue, their interests nowadays point to the use of music for health purposes. Trying to examine the influence of music on the rhythmical system of humans, could lead to ascertain if the harmonic laws of music are able to normalize unbalanced rhythmical functions and human internal organisms coordination [2]. In fact, the results of the studies for this modern branch at medium path between medicine and biology, called chronomedicine and chronobiology, have shown that the human organism has not only a complicated interior design in terms of space structure, but also referring to the time design [24].

There are two main groups of rhythms that appear as relevant from any point of view of any living organism, the long wave and the short wave rhythms. On one hand, the longer wave rhythms referred to those rhythmic functions that do have a relation with the environment in terms, that the outside environmental rhythms do have an effect on those human functions related to yearly or daily rhythms [2]. These rhythms are related to the year, month, week or day, by means of their periodic duration. Another issue to take into account is the environment where we places ourselves, which has an effect on synchronizing the entire time structures of our organism, as a time keeper that assures their correct environmental classification. As an example, this is the effect of the circadian rhythm, or the day-night rhythm, which corresponds to the 24 hour cycle rotation of the earth. Many of our processes of our body are dependent of this cycle, and in fact, we are synchronized to it by means of the daily structure of waking up with the sunrise, going to work, having lunch in the middle of the day when the sun is on the middle of its trajectory and going to bed when the sun is down [14]. However, even these different length rhythms may appear as independent, they are all related from the longer to the shorter ones. This means that the daily rhythm(and all the rhythms that are underneath it) is affected by the annual, monthly and weekly rhythm.

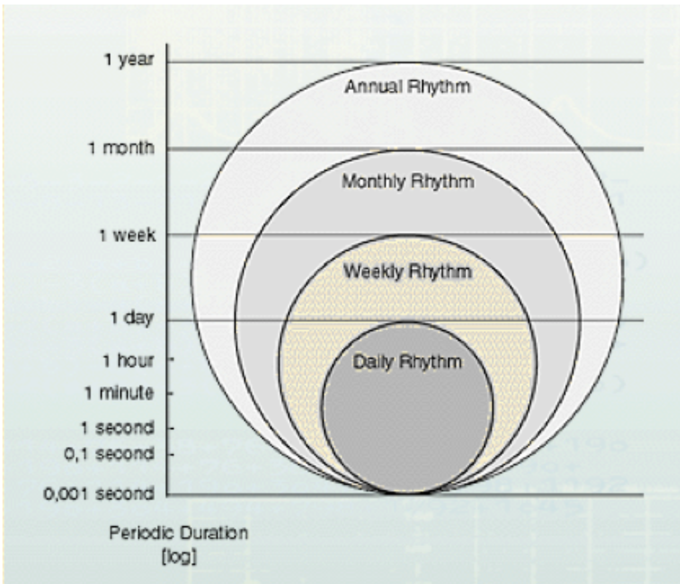

Figure 2.1: Biological long wave rhythms from [17]

On the other hand, we have the shorter wave rhythms in the human organism spectrum, which do not have any direct link with the environment [2]. These are mainly endogenous autonomic functional variations, in which temporal periodicities are placed underneath the daily rhythm. Those autonomic functional variations are referring to those internal processes of the information transfer system(the nervous system), the transportation and distribution system(breathing, circulation...) and the metabolic system(relaxation, nutrition and digestion). Even if they are not immediately connected with the environmental rhythms, they are linked by means of a balance within the daily rhythm constrain. As said, these three major internal systems work in different tempos but yet in synchrony. In the middle of these three organism internal structures is placed the functional behavior of breathing and heart beating, which produces a balance that forms an organic foundation for rhythmic-musical sensations and actions.

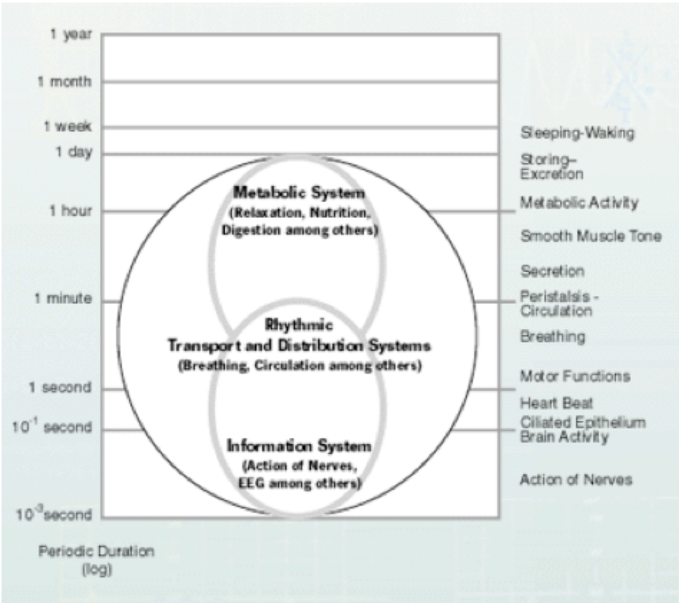

Figure 2.2: Biological short wave rhythms from [17]

It has been proved in laboratory studies that the music listening action can change (alter or relax) heart rate, respiration, blood pressure, brain waves and many neurochemical levels such as adrenalin,dopamine, etc [34]. In everyday life, we do use music also for this purpose, trying to asses a balance for our personal clock, inducing a higher beat rate when we feel upset or relaxing it when we are stressed. Somehow, we find listening to music a way to stabilize our "internal clock", as a conductor that brings out the fine tempo that leads us to find the harmonic melody that best fits in those specific moments of our lives [31], [17].

2.3 Temporal variability in music listening and preferences

As mentioned in the point before, our entire organism is based on rhythms, being internal or coming from the environment that surrounds us. Indeed, these organic rhythms can be balanced or altered by the context or situations we are standing in, as they are not only related with the intrinsic functionality of our organism, but they do also relate to our senses and thus, to our emotional states. Emotions or moods are tightly connected to our physiology, and thus, they do vary over time. So as to this, individuals personalities have also some temporal dynamics, or rhythms, which at the first years of growth, are more changeable than when becoming adults. Context and situations are also related to all these individuals' concerns, as we are not standing alone in the universe where we live. People surround us, situations, as well as physiological states, are underlying this personality and emotion variability. Obviously, we don't feel in the same way where we are stressed at work than when the weekend time arrives. We do not also feel the same comfortability degree in some social situations than in others. Also our physiology tends to be altered while we make a change of situation, such as going on holidays or even while moving house. Since music is tightly connected to those brain centers that modulate emotions, humans have become experts in finding the music that enhances the particular emotional state we are standing in on every kind of different situation [25]. Therefore, we have stated a relation between emotion, personality, physiology and music which is being reinforced by means of the temporal variable [22]. While physiology includes some rhythmical patterns, temporal variability or changes over the time are derived from it. At the same time, these concepts are applied by means of temporality to emotions and personality, and thus to music listening activity. If music can enhance or relax emotion sensation and personality traits reinforcement, as they change, music does it in the same way. Furthermore, this temporal variability, as linked to physiological rhythms can be converted and identified as temporal patterns.

Once we have stated these reasonable links, it could be of interest to lead an analysis of the temporal patterns of different individuals that can be an intrinsic characteristic of him/her. As many other behaviors, we have assessed that the time variable could be a relevant analysis facet for music listening activity, and trying to extract some temporal patterns a potential tool for describing people's behaviors. So as to this, while discrete observations of an individual's idiosyncratic behavior can appear almost random, typically and as we have stated, there are always repeating and easily identifiable routines in every person's life, as were in physiological processes [11]. These patterns become more apparent when the behavior is temporally, spatially, and socially contextualized. Because of this contextualization matter, studies about trying to identify some kind of routines in people's behavior have established that many people consider as ground truth or universal patterns weekdays vs. weekends, working hours vs. home hours, or even hot season vs. cold season [11].
Regarding these statements, it seems reasonable to hypothesize that also temporal patterns could be observed while listening to music. By observing and analyzing temporal structures or periodicities in people's playlists, we could extract some behavioral patterns which surely will be related to personality, emotions and thus, to natural rhythms.

2.4 Retrieving temporal patterns from music listening activity

Following the explanation in the preceding section, our path for achieving temporal patterns from music listening activity would be by analyzing and extracting temporal structures or patterns from listeners' playlists. It is known that self-composed playlists are looked upon as a tool for self-expression and identity portrayal, and thus it seems plausible to state that our living rhythmicity will be also stamped in them [37]. Trying to observe people's behavior around music listening implies having a large database with listeners' playlists, which means to work with somehow "active people" by means of listening to music. As our main goal is to find these periodic temporal patterns, we will be trying to retrieve them from a database of an online radio station, Last.fm. Indeed, Last.fm is a website that describes itself as a "social music revolution." Started in 2002, it is now in over twelve languages, more than 240 countries, and has over 40 million users. The premise of the website is to provide users with a profile page, enable them to track and post listening histories through the "audioscrobbler" [2] software on their profile, and facilitate community through online interaction such as forming groups, posting comments to profile pages, using discussion forums, and messaging users. Hence, the site serves as a way to reflect on a user's self and personality through his/her music consumption as music taste signifies aspects of his/her identity. Some work has been done around this profiling of the website, how people interact with the system and how the system in fact, constructs a "music identity" of the user by which the system recognizes him/her. In this respect, quantifying listening habits is seen as a certain external objectification of self and personality [13]. In the same way, the website also includes a recommendation system based on this "user profile" and "user music identity", from which it already "knows" the user music taste and preferences and makes recommendations

[2]Scrobbling is the process in which software syncs the name of the songs users listen to from their media player or MP3 player to their profile on the site.

based on them. But, as we can see, there's a lot of information of each of the users in their profiles, but also in their playlist. Indeed, we will be working with the data extracted from the users' playlist, outlining these temporal patterns from the temporal information stamped in them.

Trying to retrieve temporal patterns or identify temporal structure from human behavior, no matter which activity we are talking about, has to do with applying a specific method or approach for the analysis of the data from which the most relevant facet will be the time. When working on retrieving routines form a behavioral data-set, one simple and so valid approach is to describe each individuals' data-set in some characteristic behaviors or structures by means of "eigenbehaviors" or "eigenvectors", what has been called eigendecomposition [11]. These characteristic behaviors (as could be leaving home early, going to work, breaking for lunch and returning home in the evening) are computed from the principal components of individuals behavioral data. In fact, a linear combination of an individual's eigenbehaviors can accurately reconstruct the behavior from each day in the data. By providing this type of behavioral caricature, it is possible for the primary eigenbehaviors to be used to accurately predict an individuals subsequent behavior.

However, even though we are trying to achieve or retrieve behaviors or routines from the users playlists or in these cases, listeners' data-sets, what our research question concerns are processes which do have somehow a rhythmic oscillation or pattern. So as to perform this kind of analysis, Circular statistics or the Spectral Analysis of time series data are believed to be the best analytic tools [22]. Circular Statistics, a technique developed forty years ago and largely used in biological and physical sciences have also been exploited in personality research in order to study the temporal evolution and the recurrent patterns of mood [22], [20]. The theory of circular statistics was developed to analyze data where angles do have a meaning, as for example; data such as time of the day, phase of the moon or day of year that exhibit a different periodicity. Moreover, circular statistics is an alternative to common methods or procedures for identifying cyclic variations or patterns, which include spectral analysis of time-series data or time-domain based strategies[20]. Although these approaches are frequently used, their prerequisites are seldom met. Firstly, they assume that time-series are at least interval scaled (for discrete data). Secondly, conventional time series analysis requires regularly spaced data or equidistant measurements in time. Although our data is rounded to the nearest hour, it could happen that the results obtained are not rounded in this way (so as to get greater degree of accuracy). Another prerequisite, which maybe is the one that more influences our decision about not working with these types of methods, is that time series data follows a Gaussian distribution- whereas hourly counts or weekday counts clearly dont follow such distribution.

All these assumptions and prerequisites are one of the reasons why we did choose Circular Statistics as the best tool to analyze our dataset. Another reason, as important as the first one, comes from the thought that time is considered to be cyclic in nature (e.g daily hours, with 0:00 being the same as 24:00, and days of the week which also follow a circular pattern). Also circular statistics techniques have been designed for analyzing variables measured on a periodic scale (specifically angular or directional data), temporal data such as time of the day, day of the week, the month or the year can also be subjective to this kind of statistical analysis. Thus, when trying to retrieve these periodical patterns from the Last.fm users' playlists, circular statistics are believed to be the tool that best fits with the type of analysis and data we will be working with. Nevertheless, to our knowledge, it is the first time that it is used in the analysis of music-related behavior, though some applications have been previously reported [4], [10].

2.5 Applications: Time as a personalized music recommendation facet

The ubiquitous availability of our personal music files by means of portable devices or personalized radio-like services (the mentioned Last.fm or many others as Spotify, iTunes...) and the digitalization of music makes possible to enhance new music discoveries, personalization and recommendation services [37]. A music recommendation system which aims on proposing to the user unknown and interesting artists or tracks to the users, we can point to, not only delivering the right music but, delivering it at the right moment in time. This allows to consider the context of listening as a relevant variable in any user model intended to be exploited for music recommendation. As a starting point of this issue, we will present the basic functionalities of the recommender systems, how the contextual information about users has been improving the recommendation and which would be the application of our findings in them.

A recommender system is a personalized system that helps users to find items of interest based on some information about their historical preferences. For this purpose, there exist some recommendation strategies usually based on collaborative filtering, content based filtering or a combination of both of them. Collaborative filtering recommends items based on similarity of the preferences of a group of users, named as neighbors. In this case, the basis is that if a group of users did have the same interests in the past, they will have similar ones in the future. In this way, recommendations can be proposed based on the information of the whole group. Content based recommenders, in contrast, have their root in information retrieval and filtering. Their approach is based on the analysis of pre-rated items by the users and thus, creating a user profile based on them. While these approaches do maintain the personalization of the user, they do miss the fact that the users interact with the systems within a particular "context" [27]. Furthermore, several studies have maintained that a change in the context makes behavior of a user change and in fact, when applying contextual modeling of the users, the performance of the recommendation system improves both in terms of predictive accuracy and true positive ratings [36], [8]. However, it is difficult to identify a contextual variable that affects purchasing the behavior of costumers in some way. Approaches in finding the contextual variable that could fit for successful recommendations have the genre, age, geographic location or even occupation as a clue for building users' personalized profiles and thus, they do improve in their recommendation ratings of the users. Nevertheless, it has only recently happened that time has been included as context information [6], [23].

However, with the rising of new technologies, these approaches have changed while using these technologies feedback information for a contextualization purpose. As an example, new devices included in mobile phones such as sensors or bluetooth had been used for implicit user feedback information acquiring, and thus making successful recommendations based on them [36], [8]. Nevertheless, there have been also approaches that do not need to use the feedback of any kind of new technology for creating a contextual framework of the user from which personalized recommendations could be offered. Based on the implicit user feedback coming from the time facet of the user's behavior with the system has been used as a contextual variable that improves the recommendations [1]. The approach is based on building user's microprofiles according to some non-overlapping time partitions. These microprofiles do serve to have a more precise model of the user, which in essence means to know more certainly which are the users' tastes and needs, in this case in an specific time slot. For this purpose a fundamental problem is how to discover meaningful time partitions for each user, which surely would have a base on time cycles. Regarding this work, time partitions had been made according to some basic statements, such as the morning time slot would be considered from 6 am to 9 am, which could differ from one user to another. So, most of that reported work was, though, on finding optimal temporal partitions.

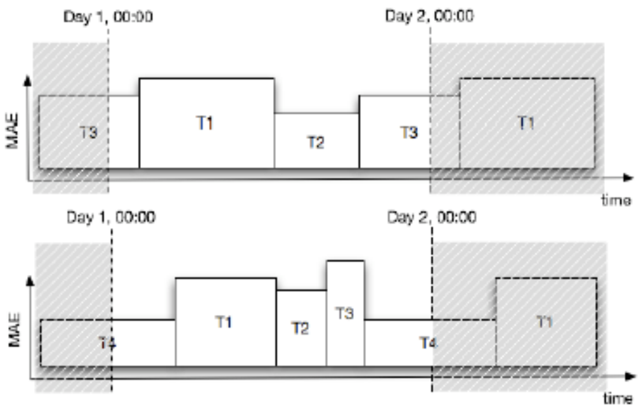

Figure 2.3: Examples of the partitioning T for a day from [1]

Our work would precisely lie on finding those moments in time, like the hour of the day or the day of the week, in a personalized way, working with the listening habits of an amount of users from a specific music recommendation system. In fact, what we are presenting in this research is that there are other options, maybe complementary to this idea of time partitions, that keep the temporal dimension as a continuous and circular one (e.g. Circular Statistics).

As said, a music recommendation system, does propose to the user interesting and unknown artist or tracks, based on user's musical taste. But, as we have commented above, musical taste or preferences do underlie many other behavioral, emotional and personality related aspects. In fact, listeners' mood determines their music selection, which in fact is tightly connected with the environment we are in [33]. Indeed, our research seems to be well aligned with a generic trend concerned with detecting hidden patterns of human behavior at the individual level thanks, mainly, to the spread of portable communication and geolocation technologies [3], [29]. We have been talking on how natural and physiological rhythms do affect to these musical preferences, in the sense that these could change according to some specific contextual rhythms or periodicities which can be observed in the behavioral aspect of each individual. Regarding the microprofiling approach, and that the main problem is to find meaningful time partitions, a solution for it could be to apply these type of rhythms to the music listening and recommendation experience. Here we do propose the use of this contextual information (i.e. time, basically the hour of the day or the weekday) of each individual as a characterization of their user profiles for a context aware music recommendation system. Taking into account that the music listening experience is tightly connected to many time based rhythms, we do hypothesize that the combination of personal aspects, such as in this case time, and musical preferences or taste could lead to a more personalized contextual characterization of users, and thus, in an improvement of the provided musical recommendations.

Chapter 3

Methodology

This chapter explains the methodology used for analyzing the data. First, we will make a description of the material we will be working with, how we managed to acquire it and how is it structured. After, the preprocessing stage will be explained, emphasized on the noise filtering step which allows us to get the clean working data-set. In this case, clean data-set means to avoid noise produced by sporadic reproductions of artists or genres by users. Once we have performed this data cleaning step, we point out the splitting of this clean data-set in both the prediction and validation data-sets, which would be used in the next analysis steps and for the Evaluation one. Finally, Circular Statistics methods and techniques are presented, which will be used for analyzing the data-set and extracting the concerning results.

3.1 Material

3.1.1 Data Collection

Getting access to the musical choices made by a large amount of people who are listening to music during several years, even in our current digital lifestyle, is not an easy task yet. Many music players store individual users' records, but they are not publicly available. As a workable solution, we have taken advantage of Last.fm

API, which makes possible to get the playcounts and related metadata of their users. In a preliminary step of the analysis, we used a smaller data-set (50 users) than the one whose results are presented here. As raw data, we will be working with the full listening history of 992 unique users, expressed as 19,150,868 text lines and spanning variable length listening histories from 2005 to 2009 (Thanks to scar Celma who kindly shared the Last.fm data file [3]). The file contains ¡user-id, timestamp, artist-mbid, artist-name, song-mbid, song-title¿ tuples collected using the user.getRecentTracks() method. The data is formatted one entry per line as follows (tab separated, "\t"):

userid-timestamp-artid-artname-traid-traname.tsv

```
userid "\t" timestamp "\t" musicbrainz-artist-id "\t" artist-name "\t"
musicbrainz-track-id "\t" track-name
```

Example:

```
user-000639 2009-04-08T01:57:47Z MBID The Dogs D'Amour MBID Fall in Love Again?
```

3.1.2 Genre tagging

Due to the two different appoaches that we plan to perform over the data, we need to assign the genre tag each track entry. For this purpose, we use again the Last.fm API, but in this case, we use the track.gettoptags() method. This method returns a list of tags for each track, which have been assigned by the users of Last.fm. At the same time, each tag has been assigned by the users a "weight"[4], which is a sign of its reliability. This list of tags, however, may relate to different aspects of music (e.g. genre, mood, instrumentation, decades, ...). Since in our case we need a single genre per track, we proceed as follows:

- Tags are cleaned in order to remove special characters or any other undesirable

[3] http://www.dtic.upf.edu/ ocelma/MusicRecommendationDataset/lastfm-1K.html

[4] Last.fm relevance weight of tag t to track x , ranging from 0 to 100

characters, such as spaces, hyphens, underscores, etc. (e.g. hip-hop -> hiphop, hip hop -> hiphop)

- Irrelevant tags (i.e., those having a low weight) are removed.
- Tags are matched against a predefined list of musical genres, which consist of 272 unique musical genres/styles gathered from Wikipedia and Wordnet.
- From the matched genres in the previous step, the highest-weight genre is selected for the given artist. If there is more than a single tag with the highest weight assigned to a track, we select that with less popularity (i.e. popularity is an attribute of each genre tag we computed according to the number of occurrences in our data set). We proceed this way in order not to biased the data against unpopular artists.

In our data-set case, we did match the 80% of the original tracks, obtaining an amount of 249 unique genres. In order to have an overview of our data, we search for those genres that have more repetitions. The figure 3.1 shows the distribution of what we have called the "Top10" of genres, a *ranking* made from the overall repetitions of each genre in the whole data-set.

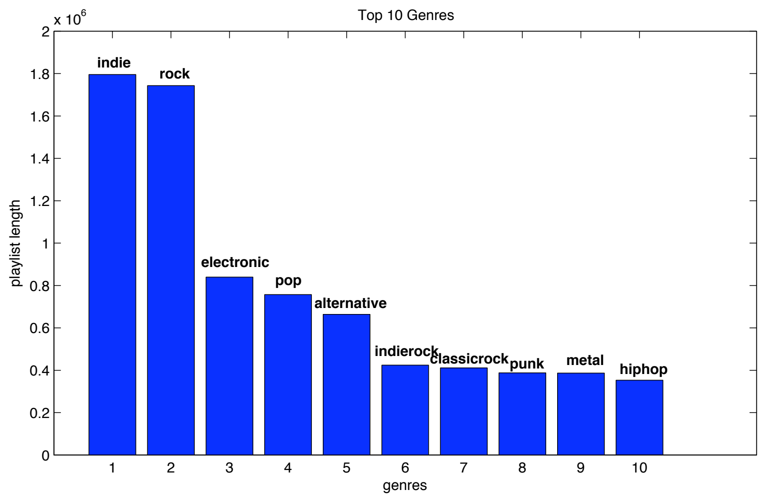

Figure 3.1: Top10 genre distribution

3.2 Data Cleaning

Data coming from Last.fm contains playcounts that cannot be attributable to specific listening decisions on the side of users. If they select radio-stations based on other users, tags or similar artists there are chances that songs, artists and genres will not recur in a specific user's profile. In general, even if having data coming from personal players which in fact, do obey solely to the user's will, we would discard (i) users that do not provide data to be processed, and (ii) artists and genres that only appear occasionally. We prefer to sacrifice a big amount of raw data, and keep those that help us to identify a few clearly recurring patterns, maybe for just a bunch of users, artists or genres.
So as to achieve the mentioned cleaning goals we follow the next procedure:

1. For each user, we compute the average length of each artist/genre in his/her playlist. Indeed, for each user's playcounts, we take each artist/genre repetitions and we compute the average of all of them.

2. Again, for each user's data-set, we filter out all those artists/genres which playlist length is below the user's overall playlist length.

3. In order to get rid of low frequency playing users, we compute the median value of the percentage of artists/genres left after the last filtering step. Those users whose percentage of artists/genres is above this median value, will be considered as "valid" users, and will be the ones chosen for the next analysis steps. Users whose percentage of artists/genres is below the median percentage are discarded.

4. So as to verify that the selected users present a high degree of variability in their listening habits (i.e. that they don't listen at the same day or at the same time only), we check each user playlist exhaustively. Those "valid" users that present this kind of behavior won't be useful for further analysis steps, so we will filter them.

The whole process since we acquired the raw data file till we get the "clean" data files is presented in the diagram 3.2:

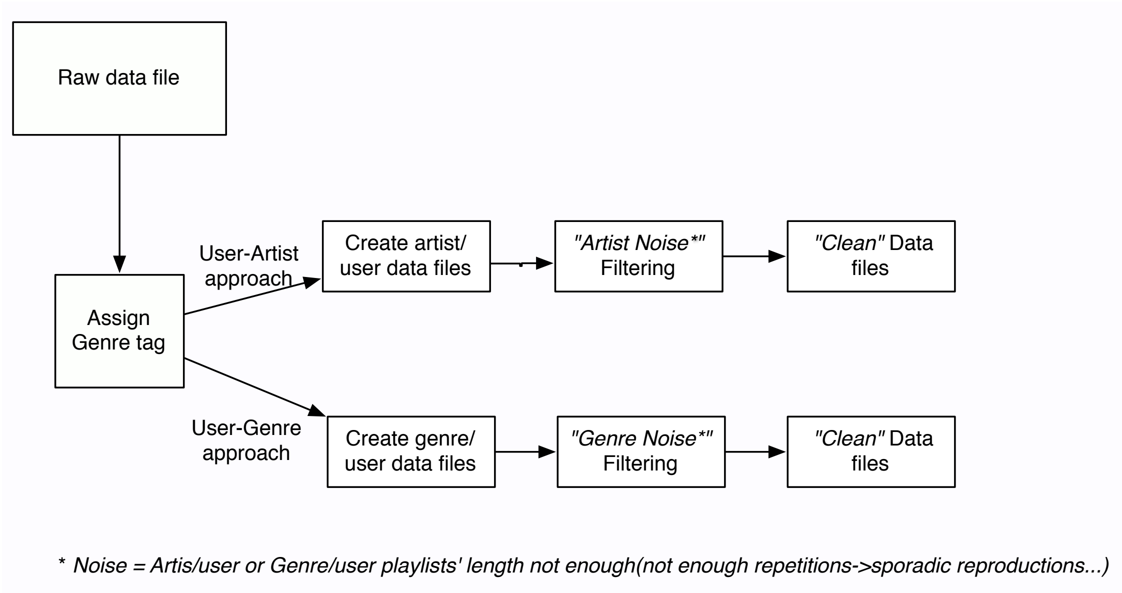

Figure 3.2: Data Cleaning process block diagram

3.2.1 User/Artist based approach: Data overview

Regarding the first point of this cleaning process where we compute for each user the average length of each listened artist and genre, overall, we got that almost the 95% of the users' playlists' average were below the 200 reproductions per artist. The histogram 3.3 shows this distribution:

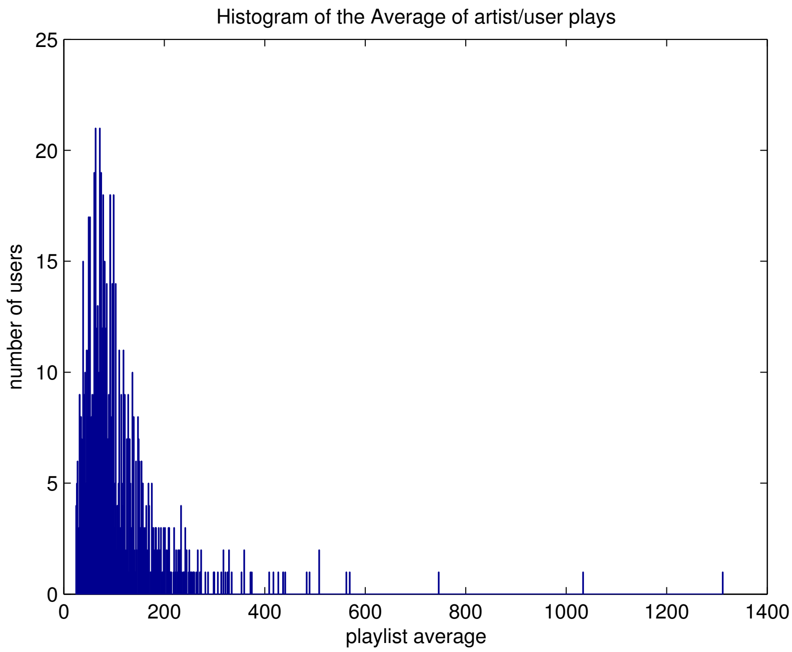

Figure 3.3: User-Artist average playlist length histogram

As an example, for a specific user case, figure 3.4 shows that over the 70% of the listened artists were reproduced less than 20 times. This fact is a clear example of the "long-tailed" artists reproductions, where few artists are frequently reproduced and many of the total amount of artists are sporadically reproduced. As said, these last ones are not considered as useful for a further analysis, as regarding the scope of the study, we want to explore repetitive behaviors, cyclic ones, which do need of enough repetitions for considering them as personal user's habits.

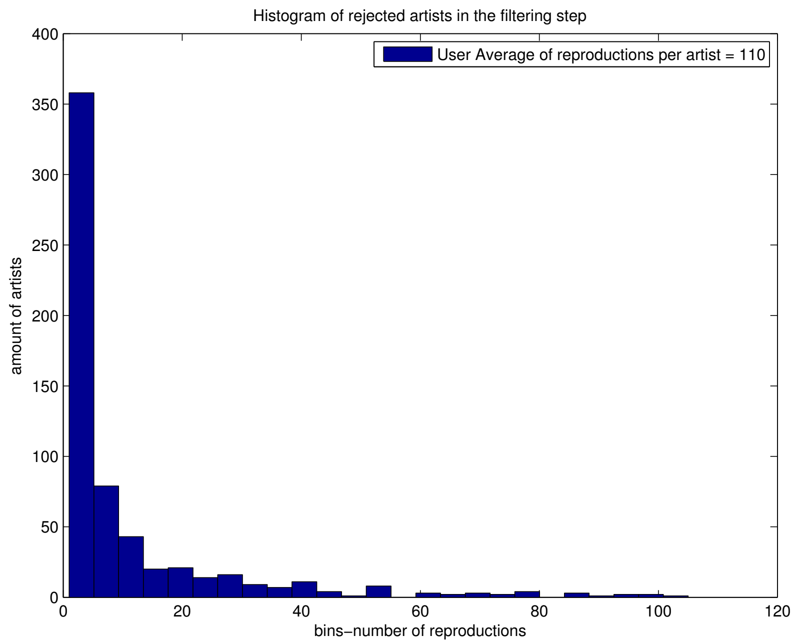

Figure 3.4: Histogram of filtered-out artists for an specific user

3.2.2 User/Genre based approach: Data overview

As said, we follow the same methodology used in the user/artist based approach. We also compute the average playlist length for each user regarding each listened genre playlist length. For the user/genre approach, almost the 88.56% of users' playlists' average were below 1000 reproductions. Although this value is greater that in the artists case, it seems reasonable concerning the point that we are considering the genre's playlist and thus, there will be a larger amount of artists for the same genre, which implies a larger playlist length for each of the listened genres. As in the artist case, here a long-tail distributions of the amount of listened genres is observed.

We can observe the average playlists' length regarding the amount of users in the histogram 3.5:

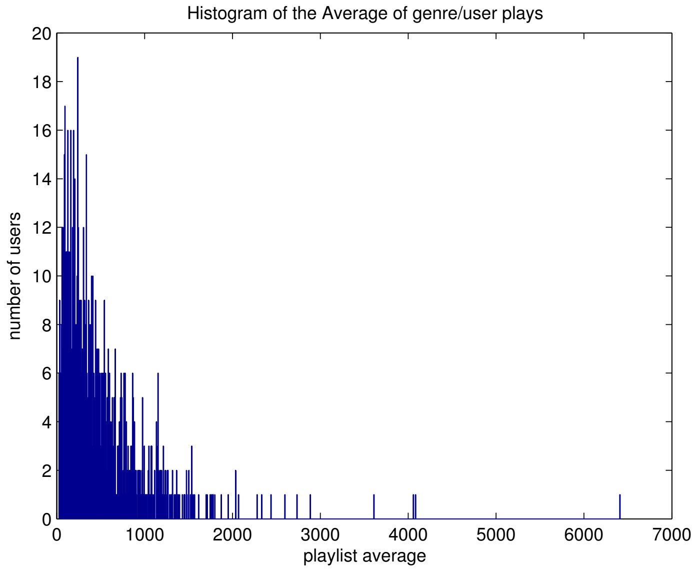

Figure 3.5: User-Genre average playlist length Histogram

3.3 Prediction and Validation Data Sets

Once we get rid of all the suspected noise, we will model listener's behavior by characterizing their listening tendencies. When the temporal preferences of each listener are modeled we will evaluate their robustness by testing them as they were predictions of the user's behavior. In fact, that's one of the purposes of user modeling, to infer unobservable information about a user from observable information about his/her actions, preferences, goals... So as to build these user's models, we follow a content-based approach where user's past behavior is a reliable indicator of his/her future behavior. This approach requires large amounts of data (which is not a problem in this case) and is particularly suitable for situations where users tend to exhibit a characteristic behavior. In order to evaluate our prediction model we split our clean data-set in two groups. One will be used to generate the temporal predictions while the other one will be used to test them, so we follow an approach simpler than a linear prediction model. The test set contains all the data in the last year of listening activity for a given subject (i.e. the "future behavior"). The prediction-generation set contains the data coming from two years of listening previous to the year used in the test set (i.e. the "past behavior").

3.4 Analysis using Circular Statistics

Common methods for identifying cyclic variations or patterns include spectral analysis of time-series data or time-domain based strategies. Although these kind of approaches are commonly used, the prerequisites of such sophisticated procedures are seldom met, such as interval-scaled time-series variables. As an alternative to these kind of procedures, in this thesis we wanted to explore the effectivity of methods from a different field of statistics, directional or circular statistics. These methods do offer another opportunity for the detection of patterns in time, where fewer prerequisites have to be met.

Under the circular statistics framework, variables or data considered to be cyclic in nature are meant to have a period of measurement that is rotationally invariant. In our case, this period is referred to the daily hours and the days of the week. Therefore, for the analysis, the rotationally invariant period would be reflected as daily hours ranging from 0 to 24, considering that 24 hours is the same as 0 hours. Regarding the weekly rhythm, Monday at 0 hours would be considered the same as Sunday at 24 hours.

Concerning the scope and goals of our research, for the analysis stage we will be dealing with some descriptive statistics features (mean direction and resultant mean vector length). For the validation and the characterization of temporal patterns, aimed to detect the corresponding temporal patterns, we will apply some methods from inferential statistics methods (Rayleigh's and Omnibus test).

3.4.1 Feature extraction using Circular Statistics

Willing to characterize Last.fm users by finding temporal patterns behind their entries registered by this online radio service, we first approach the problem by extracting the most descriptive parameters from the data-set using circular statistics methods.

For this purpose, we did choose some of the basic methods and functions that are used for exploring and summarizing important properties of a sample of angular data such as central tendency, spread, symmetry or peakedness. As we have two different approaches, for both of them we follow the same kind of processes, though in the user/genre approach some modifications were done to some of the steps. At the same time, the whole analysis was repeated regarding the time of the day and the day of the week. Due to this, we'll have a parallel overview of the data, having the information both for the circadian rhythm(daily rhythm) and circaseptan rhythm(weekly rhythm). In fact, as our raw data is not measured in angles or directions, but in hours (from 0 to 24) and weekdays (from 0 to 6, 0 corresponding to Monday and 6 to Sunday), the first step in circular analysis corresponds to the conversion to a common angular scale. We chose the angular scale in radians, and thus we apply the next conversion to our data-set:

$$\alpha = \frac{2\pi x}{k} \tag{3.1}$$

where x is the representation of the data in the original scale, α is its angular direction (in radians) and k is the total number of steps on the scale where x is measured. For the daily hour case, x would have values between 0 and 24 and k = 24. Alternatively, for the weekday analysis, x would have a scale range from 0 to 6 (0 referring to a Monday listening event and 6 to a Sunday listening one) and thus, k = 6. As noted, the effect of this conversion can be easily transformed back to the original scale. Once our data is placed in the right scale we can then compute the circular statistics needed. Trying to have a complete overview of the split data-sets, for the central tendency we compute the circular mean listening direction for each listened artist/genre per user. This mean direction cannot be computed by simply averaging the data points. Indeed, the data is first converted to directions defined by:

$$C = \sum_{i=1}^{n} cos\alpha_i, S = \sum_{i=1}^{n} sin\alpha_i, \tag{3.2}$$

$$R = C^2 + S^2 \tag{3.3}$$

thus, according to Fisher[12], the mean direction is given by

$$\overline{\alpha} = \begin{cases} \tan^{-1}\left(\frac{S}{C}\right) & S > 0, C > 0 \\ \tan^{-1}\left(\frac{S}{C}\right) + \pi & C < 0 \\ \tan^{-1}\left(\frac{S}{C}\right) + 2\pi & S < 0, C > 0 \end{cases} \tag{3.4}$$

In the case of the sample spread, we compute the circular variance, circular standard deviation and circular concentration. This last one is also known as the mean resultant vector length, and it would be a quite representative property of the each sub data-set. In fact, the above quantity R is named as the mean resultant vector, greatly associated with the mean direction $\overline{\alpha}$, whose length is given by:

$$\overline{R} = R/n \tag{3.5}$$

The length of the mean resultant vector is one of the most crucial measurement of the circular spread of the data sample and can be used for hypothesis testing.

In the next section we will talk about which of this properties are determinant in further analysis steps, which would be performed for specific or personalized cyclic patterns detections.

3.4.2 Selection of features extracted using Circular Statistics

From the whole set of circular statistical descriptors, we had to make a decision on which of them are of more interest or relevancy. For our research problem the mean direction will tell us the average hour of the day or the average day of the week the user listens to a specific artist or genre.This average hour and day may reflect the listening trend of each user regarding the artist or genre he or she listens to. At the same time, this mean direction value could be used for testing the significance of the conclusions that can be extracted from it, as for example if the mean direction for two or more groups is the same one. One way of assessing this significancy is by applying the Watson-Williams test [39].

The mean resultant vector length, the value of which ranges between 0 and 1, describes the spread of the data around the circle. In fact, values for $\overline{R}$ near 0, are referred to data that follow a uniform circular distribution (i.e. events occur uniformly in time). For values around 1, the data is said to be concentrated around the mean direction. So, this parameter is strongly linked with the mean direction in a data set and the distribution of the data sample around it, thus, it gives us interesting information about the "strength" of the mean direction for each data sample. Moreover, this statistic is also useful because its statistical significance can be assesed using the Rayleigh's test [12] or the Omnibus' test [40].

3.4.3 Detection of temporal patterns

Once we have established that our analysis framework would be based on applying circular statistics to our data-set and we have selected the relevant statistical features, the next step is to establish the path we follow for the detection of the temporal patterns or rhythms. As commented in the last point, inferential statistics methods let us to determine if the results obtained from the analysis are significant enough to make some judgements upon them. Therefore, our goal is to state if the mean directions obtained for each sample are representative enough to consider them as a tendency or pattern of each user behavior. We perfom the Rayleigh and Omnibus test over the results obtained from the analysis regarding the daily and weekly data for both approaches, the artist listening and the genre listening based one. When performing these tests we can assess if any artist or genre for a specific user has a strong listening tendency regarding an hour of the day or a day of the week. Both tests are performed over the mentioned R statistic and share a common null hypothesis:

H_0: The population is uniformly distributed around the circle

with the alternative hypothesis

H_A: The population is not uniformly distributed around the circle

The Rayleigh test asks how large the resultant vector length R must be to indicate a non-uniform distribution [12]. It is particularly suited for detecting unimodal deviations from uniformity. If the data indeed is unimodal, it is the most powerful test for assesing it. The approximate p-value under H_0 is computed according to equation 27.4 in [40] :

$$p = exp\left(\sqrt{(1 + 4N + 4\,(N^2 - R_n^2))} - (1 + 2N)\right), \tag{3.6}$$

where a small p indicates a significant departure from uniformity and provides evidence to reject the null hypothesis. For N as small as 10, the R_n value can be approximated as $R \cdot N$ up to three decimal places.

The Omnibus or Hodges-Ajne test [40] for circular uniformity is an alternative to the Rayleigh test that works well for unimodal, bimodal and multimodal distributions. To conduct the test, the smallest number m that occurs within a range of 180 degrees is computed. Under the null hypothesis, the probability of observing an m this small or smaller is:

$$p = \frac{1}{2^{N-1}} (N - 2m) \binom{N}{m} \tag{3.7}$$

which can for $N > 50$ be approximated by

$$p \simeq \frac{\sqrt{2\pi}}{A} exp(-\pi^2 / \left(8A^2\right)), \tag{3.8}$$

where $A = \frac{\pi\sqrt{N}}{2(N-2m)}$

Regarding the fact that our goal is to establish that users listening activity shows some tendencies or modally distributed music listening behaviors, this test seems to be suitable for achieving these goals since it has been specially designed for detecting general deviations from uniformity at the price of some statistical power. Also, it works well for many distributions shapes. Therefore, we used this test in order to detect deviations from the uniformity and assess modally distributed data.
Figure 3.6 shows an example of the circular representation of a specific user listening behavior for a specific artist along 24 hours. We can observe how the data is modally distributed (i.e. shows a distribution with a more frequent direction direction). The left side diagram shows the daily distribution of listening and the right one the circular histogram. The blue lines in left diagram are a representation of the data trajectories, samples of data that appear one after the other. The red line represents the mean vector direction and length in both cases. As explained in the section 3.4, the mean direction refers to the mean listening hour or day, using this circular representation, and the length, to the concentration of the data around than mean direction.

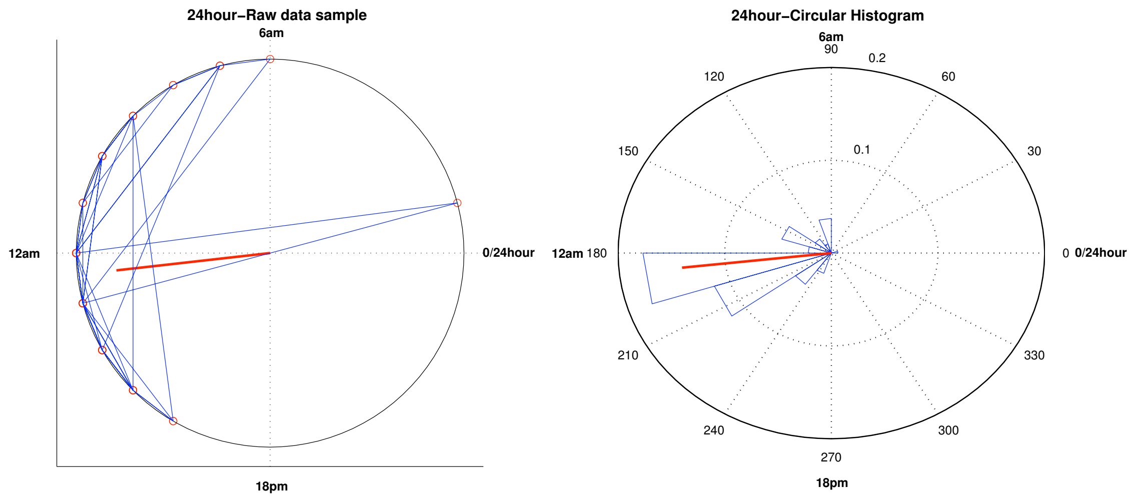

Figure 3.6: Circular representation of Artist *"Y"*daily rhythm for an specific user

In figure 3.7 we observe the circular representation of a specific user genre pattern during 24 hours. This one even it seems that the data is more uniformly distributed, we can identify those directions (hours) for which the listener has a more frequent listening behavior. Due to this fact, the strength of the tendency is not as high as in the last example (see the length of the red line, the mean direction vector length).

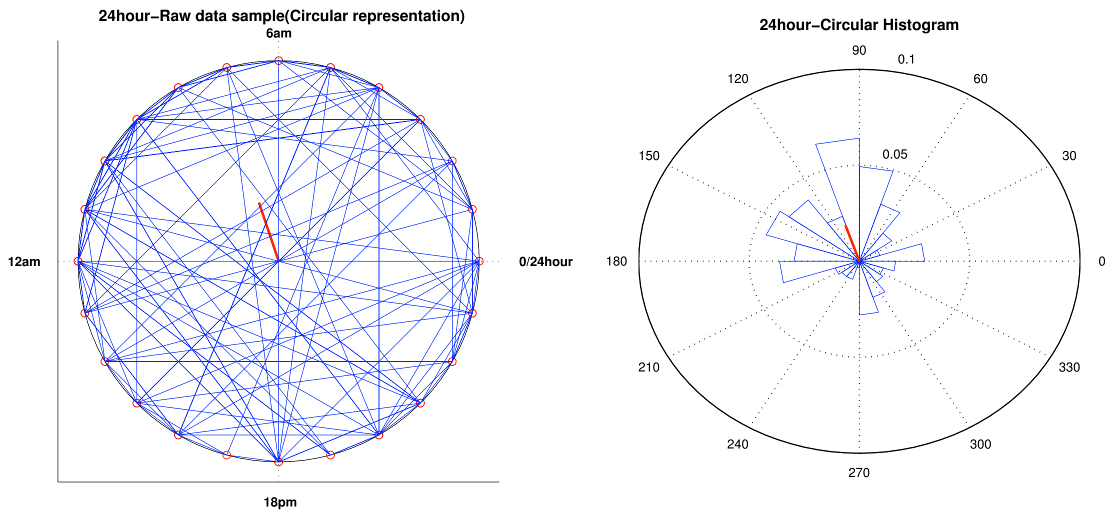

Figure 3.7: Circular representation of Genre *"Y"*daily rhythm

So as to exploit the information given by the genre tag, we perform another kind of test over the data coming from the listening activity of each genre for each user and the data of each genre for all the "valid" users. In this point, we want to investigate if the listening trend of each genre for each user is the same one as the global one. Regarding the personalized scope of the project, we will reject those genres listened to by each user that show a listening trend similar to the global one. Furthermore, we want to state that the genres listened by an specific user show a tendency which is personal, unique, and that in fact, don't have anything in common with the rest of the population. Therefore, we perform a specific type of test concerning the mean direction. The Watson-Williams two- or multi-sample test of the null hypothesis is a circular analogue of the two sample t-test or the one factor ANOVA (i.e. analysis of variance). Thus it assesses the question whether the mean directions of two or more groups are identical or not.

H_0: All of s share a common mean direction, i.e., $\overline{\alpha}_1 = \cdots = \overline{\alpha}_s$

H_A: Not all s have a common mean direction.

The test statistic is calculated via [39], [35]

$$F = K \frac{(N-s)\left(\sum_{j=1}^{s} \overline{R}_j - \overline{R}\right)}{(s-1)\left(N - \sum_{j=1}^{s} \overline{R}_j\right)} \tag{3.9}$$

where $\overline{R}$ is the mean resultant vector length when all samples are pooled and the $\overline{R}_j$ the mean resultant vector length computed on the jth group alone (similar to total variance within group variance in the ANOVA setting). The correction factor K is computed from

$$K = 1 + \frac{3}{8k} \tag{3.10}$$

where k is the maximum likelihood estimate of the concentration parameter of a von Mises distribution with resultant vector length $\overline{r}_w$. This k is computed using the approximation given by [12] (section 4.5.5). Here, $\overline{r}_w$ is the mean resultant vector length of the s resultant vectors r_j computed for each group individually. The obtained value of the test statistic is then compared to a critical value at the δ level obtained from $F_{\delta(1),1,N-2}$.
The Watson-Williams test assumes underlying von Misses distributions with equal concentration parameter, but has proven to be fairly robust against deviations from these assumptions [40]. The sample size for applying the test should be at least 5 for each individual sample. If binned data is used, bin widths should be no larger than 10 deg. Rejecting the null hypothesis only provides evidence that not all of the s groups come from a population with equal mean direction, not if all groups have pairwise differing mean directions or evidence of which of the groups differ.
Nevertheless, this statement is what we were searching for when performing this test, reject the hypothesis that a listener listens to a specific genre at the same moment or day that all the population that listens to that genre do.

In order to consider the results from any of these tests significant so as to reject the null hypothesis, we set a significancy level of 0.05. So as to consider any result from the test as significant (null hypothesis rejection) the obtained value from the test would have to be $p < 0.05$.

All the circular statistics analyses have been performed with the CircStat for Matlab [5]. This toolbox is intended to provide the users of Matlab, a set of functions and solutions for most common problems in descriptive and inferential statistics for circular data. There are other comparable toolboxes for other language environments such as the CircStats package for the R programming environment [28], another one for Stata [9] and Oriana, which is a commercially available program that offers the basic functionality for the analysis of circular data [19].

3.4.4 Validation of detected patterns

Once we have detected the temporal patterns by applying the inferential methods explained in the section before, the validation of them is performed using the prediction and validation data-sets. In the evaluation step, we will test the accuracy of our prediction model, by calculating the percentage of times the event that actually occurs in the test set was predicted in the model within an error range.

Chapter 4

Results

In this section we will present the results obtained from the analysis, thus the temporal patterns that have been detected and the evaluation performed over them.

4.1 Data Cleaning output

As a consequence of the data cleaning process, in the artist selection based approach, our working data set now contains data from 466 users. The cleaning process has kept the 62% of their total playcounts, as shown in the diagram 4.1. This 62%

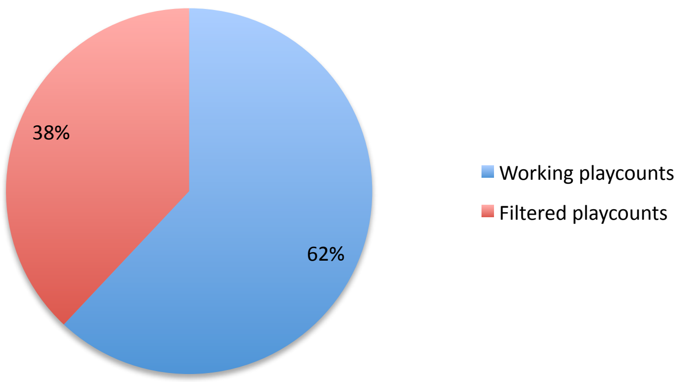

Figure 4.1: Overall percentages of playcounts corresponding to the working data set regarding the Artist Selection

corresponds to the 4.5% ($\sigma = 3.3$) of the initial number of artists. This dramatic reduction of the number of artists should not be surprising as many listening records show a "long-tail" distribution, with just a few frequently played (i.e. repeated) artists, and many of them seldom played as showed in the figure 3.4.
After we get rid of these *noisy* artists, we verify the listening variability of our 466 "valid" users. None of the users had a playlist which only contained reproductions from a specific day of the week or a specific hour, so our clean data-set did manage to have a high degree of variability upon which we could perform the circular statistical analysis.
On the other hand, when focusing on musical genre listening, the working data set includes 515 users, from which the 78% of their playcounts was kept as we can observe in the diagram 4.2:

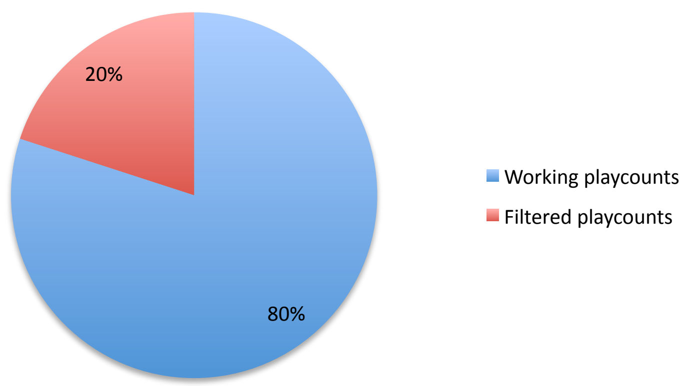

Figure 4.2: Overall percentages of playcounts corresponding to the working data set regarding the Genre Selection

These 78% of playcounts comprise the 8.6% ($\sigma = 2.8$) of the total number of genres. Again, overall, a "long-tail" distribution of the listened genres is present in the data set. As in the artist case, we verify the listening variability of our 515 "valid" users and again, none of these users show a playlist which only contained a unique listening day or hour. Thus, we conclude that regarding the genre based approach, our working data-set is suitable enough for performing the following analysis steps.

4.2 Temporal patterns of Artist Selection

Once we have cleaned our data-set, as mentioned, we compute the mean circular direction and the mean resultant vector length for each artist. Therefore, these values can be considered as a description of the listening tendencies for each artist by each user. Both parameters were calculated for the daily and for the weekly data.

So as to asses the relevance of each user's listening trends, we tested that the distribution of the playcounts was different from uniform, and that it was modally distributed (i.e. showing a tendency around several hours or around a day of the week) and discarded those that were not fulfilling these requirements (a significance level of 0.05 -null hypothesis rejection probability $p<0.05$- was set). Regarding the daily hour approach, for each listener clean data-set, on average almost the 93% (σ = 13) of the artists passed the uniformity test, so for the artist percentage listening trend is meant to be concentrated around a specific hour. However, regarding the original data-set, on average for each user only the 7% (σ = 3,2) of the artists show a listening hour tendency. For the weekly approach, the per user average when performing the test over the clean data-set is of 99,8% (σ = 0,8), so almost a 100% of the artists conforming the clean data-set show significantly a tendency for a listening hour. As in the daily case, from the original data-set, on average a 7,5% (σ = 3,2) of the artists passed the significance test.

4.3 Temporal patterns of Genre Selection

Following the same structure, in this section we will overview the results obtained from the whole analysis regarding the genre listening approach.

As in the artist case, we perform the same uniformity test in a will to detect modally distributed data (i.e. showing some kind of tendency). For the daily hour case, on average almost the 97% ($\sigma = 9.25$) of the genres belonging to each user clean data set passed the uniformity test. But regarding each user original data set, the percentage falls to a 11% ($\sigma = 2.8$). In the weekly approach, we got that 100% of the genres in each user clean data set show a tendency or a modal distribution of the data around a day or an hour, passing by far the uniformity test. Anyway, as in the daily approach, when looking at the whole original data set of each user, on average only a 11.3% ($\sigma = 2.69$) of the listened genres passed the significance test.
As mentioned in the section 3.4.3, as a way to exploit the information given by the genre tag we perform the Watson-Williams test, which as said, is the circular equivalent for the two sample t-test. It's aim is to find if two groups of samples do have the same mean direction, and thus, we use it for assessing if the genre listening activity of each user is the same one as the global one (i.e. the listening activity for all the users for each genre). Since performing this test had a very high computational cost we randomly chose a 10% of the users' clean data to perform this test. Nevertheless, we found that for this portion of the users, their listening trends regarding the genre selection was different to the global in almost a 87% ($\sigma = 14$) for the daily listening approach and for a 70% ($\sigma = 24$) for the weekly case. So, we can state that a high percent of the listening trends where personal, not global ones, reinforcing our findings that the user's behavior where intrinsic and have to say with their personality and preferences.

4.4 Evaluation

According to the analysis and circular statistics presented before, temporal patterns of music selection have been observed in a modest but not negligible percent of the studied listeners. These patterns can be observed at the artist level (i.e., certain artists tend to be played at certain days of the week or at certain hours of the day), and also at the genre level, meaning that those listeners tend to select music of certain genres at specific days of the week or at specific hours of the day. In order to assess the robustness of these patterns, we tested the listening trends using a sample of playcounts extracted from a time period posterior to the one used for building the model, and spanning one year (e.g., for user X, data to build the model belonged to years 2006 and 2007 and the model was tested on data belonging to year 2008). For each user and artist we computed a hit if the difference between the day of playing in the train and test conditions, expressed as a circular mean value in radians, was less than 0.45, which is the equivalent to a half-a-day error (an alternative, but more restrictive index could also be used when considering the exact day instead of the radians).

At the artist level, data from 466 users, including 7820 different songs and a grand total of 23669 playcounts were used in the validation of the temporal listening patterns of artist. When predicting the day of listening, an overall 32.4% of hits was found for the songs in the test collection, which exceeds by far the chance expectations ($1/7 = 14.28\%$). An additional analysis can be done focusing on each listener as the final goal of the model is providing user-specific contextual recommendation. In this case we computed the per-user average, which was of 34.5%, with a standard deviation of 17.8.

Identical data treatment was done with the time of the day. In this case the threshold to consider a hit was an absolute difference between the predicted and the observed time of less than 0.13 radians (in this case, is the equivalent to a half-an-hour error). An overall 17.1% of hits was found using the test set data (here the chance expectation baseline was 4.1%). The per-user average was of 20.5 with a standard deviation of 16.4.

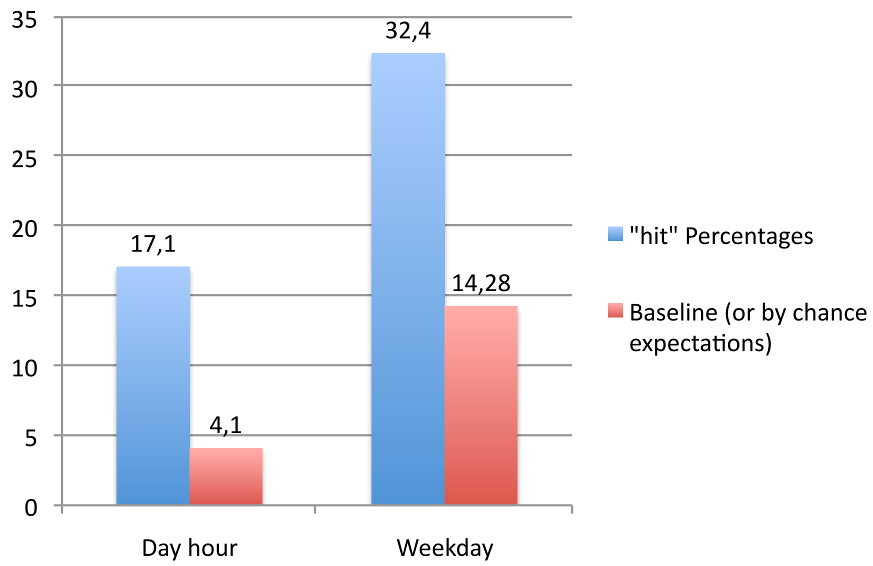

Figure 4.3: Overall Artist based evaluation percentages

At the genre level, data from 456 users, including 5100 songs, and 117 different genres were used in the validation of the genre data. In order to consider a "hit" in the prediction of listening time and day for a given genre we set the same thresholds than for evaluating the artist prediction. For the time of the day, an overall 22.6% of hits was found, while the per-user average was of 23.2%. It is interesting to note that relaxing the required accuracy of the prediction to plus/minus one hour error we reached 39.9% of average hits and per-user average 41% ($\sigma = 28.4$). For the day of the week, the overall hit percent found was 40.9%, while the per-genre average and the per-user average were, respectively, 40.7% ($\sigma = 24.1$) and 41.7% ($\sigma = 26.3$). These values were also consistently observed for the most frequent genres, while the dispersion increased in less-frequent genres.

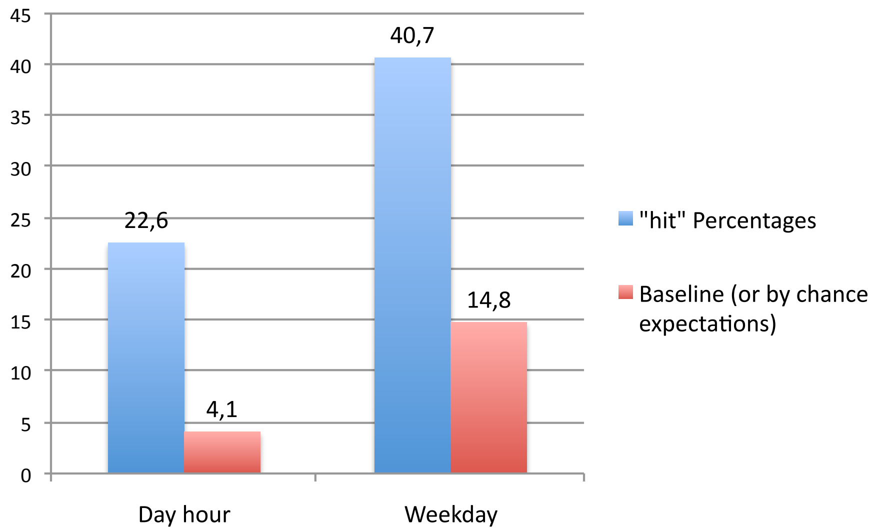

Figure 4.4: Overall Genre based evaluation percentages

4.5 Case Example

Regarding that the scope of the project is focused in finding personal preferences when listening to music, in this section we will present some examples of the obtained results for specific subjects. Indeed, an example of a user which listening trends have been detected regarding the genre selection, from which successful predictions have been done resulting in a high percent of "hits" is presented in the Example A. In the Example B we present a subject for which the detected temporal patterns have not been so successful for building the prediction model, ending up in a low "hit" rate.

4.5.1 Example A

As said, in this example we can observe the detected temporal patterns of a specific user regarding the genre selection according to the day of the week. The table 4.1 presents the obtained circular mean direction value in radians and days [5], both for

[5] 0 representing monday and 6 sunday

the prediction and the validation data sets. This values referred to the most frequently listened day of the week, regarding the weekly rhythm as analysis period. In this example we can observe that more than the 50% of the listened genres where able to generate good predictions models and compute a high number of "hits" (in the weekly case, we compute a hit if the difference between the day of playing in the prediction and validation data sets, expressed as a circular mean value in radians, was less than 0.45 radians equivalent to half a day error). The figures 4.5 and 4.6 reflect the circular representation of the prediction and validation data respectively for the "indie" genre which predictions where consider to be successful. Observing the circular histogram, we can note that the distribution of the data follows almost the same pattern in both cases.

Genre	Circular mean value (radians)(prediction)	Circular mean value (days) (prediction)	Circular mean value (radians) (validation)	Circular mean value (days) (validation)
altcountry	0,307	0	-0,19312	6
alternative	0,044007	0	1,0472	1
classicrock	-0,44918	6	-0,83294	5
hardcore	0,10246	0	-1,9713	4
indie	0,20287	0	-0,035333	6
indierock	0,644	1	1,4482	1
punk	-0,32306	6	-0,56331	5
rock	-0,17663	6	-0,58282	5
ska	-0,25477	6	-0,54638	5

Table 4.1: Example of the results obtained for a specific user X regarding the genre selection by means of the weekly rhythm

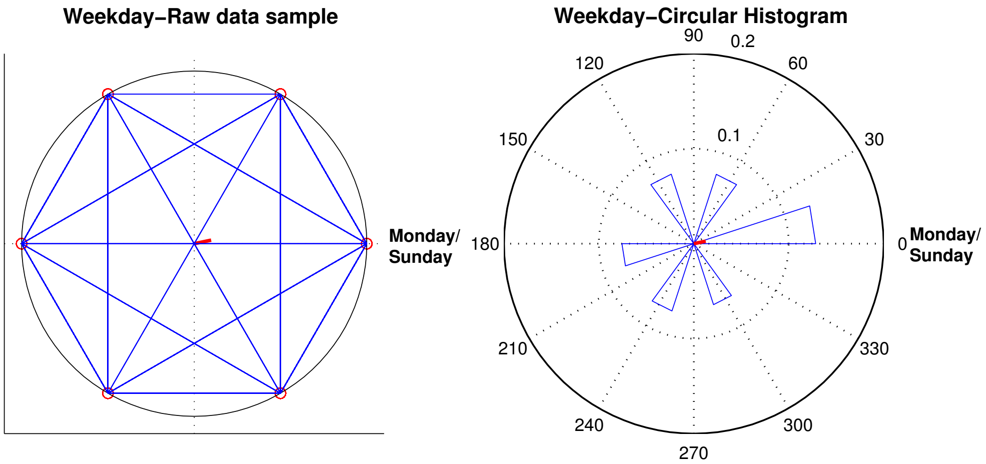

Figure 4.5: Circular data distribution of the indie genre used for building the prediction

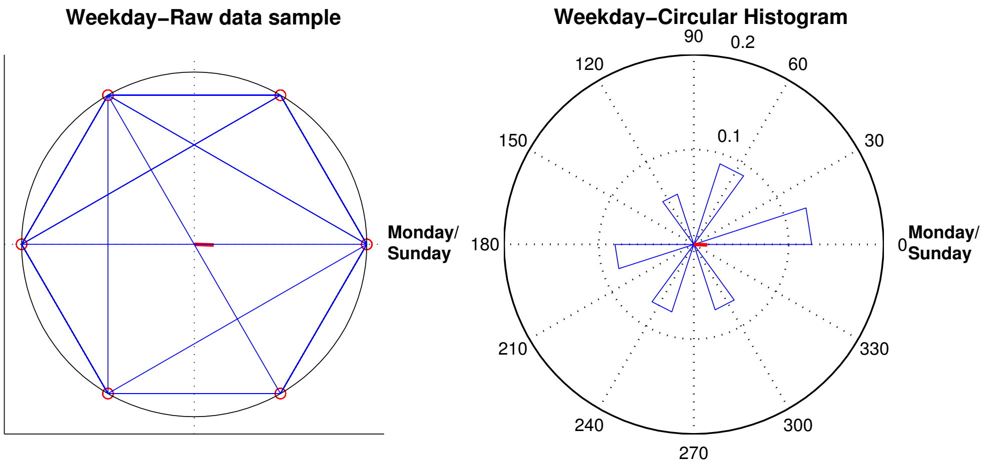

Figure 4.6: Circular data distribution of the indie genre used for the validation data set

4.5.2 Example B

As mentioned before, in this example we show a user for which his/her temporal patterns weren't clear enough for building good predictions. Only one of the listened genres was consider for generating a good prediction, with the same considerations than in the case before. As we can observe in the figures 4.7 and 4.8, especially looking at the circular histograms, the distribution of the data for "alternative" genre that didnt generate a "hit", doesn't follow the same distribution at all.

Genre	Circular mean value (radians)	Circular mean value (days)	Circular mean value (radians) (validation)	Circular mean value (days) (validation)
acoustic	-0,10153	6	-0,80463	5
alternative	0,076154	0	1,2521	1
alternativerock	0,11248	0	0,60589	1
classicrock	0,21714	0	-0,33347	6
electro	0,10844	0	0,98614	1
electronic	2,8296	3	0,94155	1
emo	-0,14516	6	-0,034627	6
hiphop	0,1165	0	-0,60589	5
indie	0,16336	0	-2,6821	3
indierock	0,50904	0	-2,733	3
punk	0,45443	0	1,2898	1
rock	0,13832	0	-1,2694	5
ska	0,19808	0	-0,33347	6
trance	-0,46976	6	-1,8518	4

Table 4.2: Example of the results obtained for a specific user Y regarding the genre selection by means of the weekly rhythm

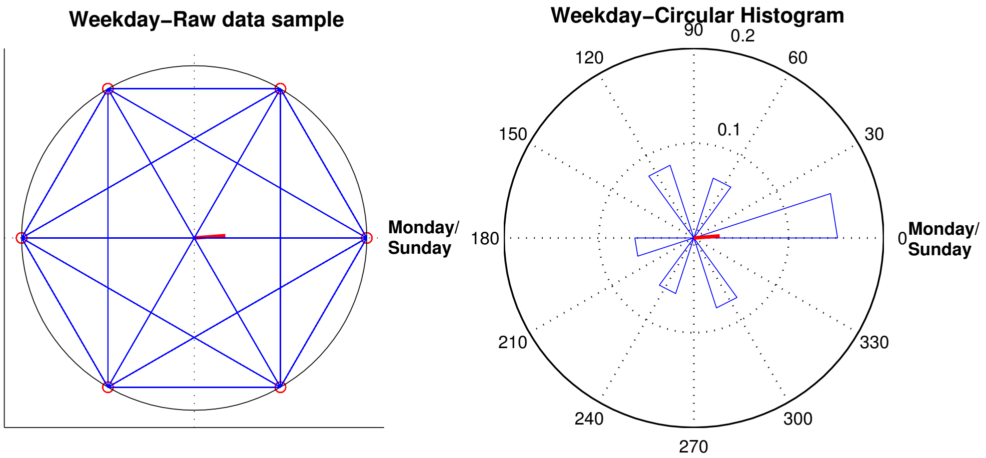

Figure 4.7: Circular data distribution of the alternative genre used for building the prediction

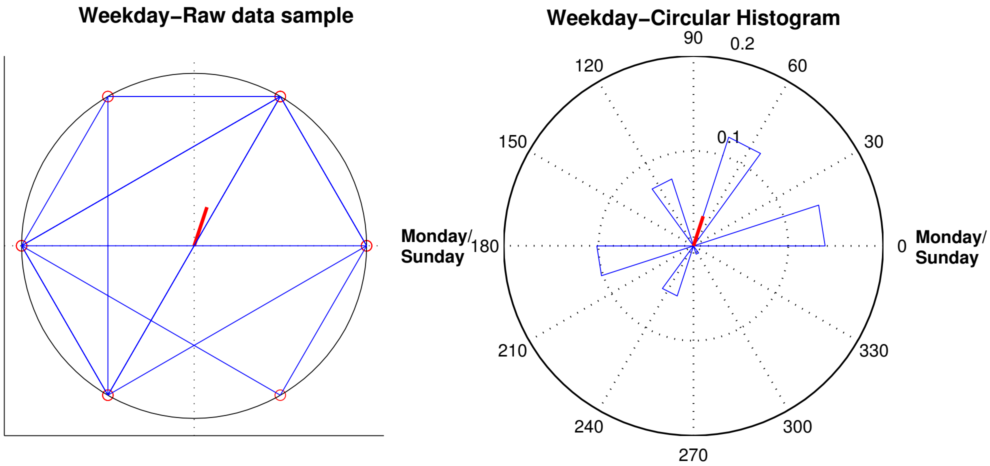

Figure 4.8: Circular data distribution of the alternative genre used for the validation data set

Chapter 5

Conclusions and Future Work

In this final chapter of the document a summary of the contributions of this thesis are presented as well as some discussion and conclusions among them. Also, some plans or suggestions for future work are studied, where improvements and applications of the work are explained.

5.1 Summary of contributions

As stated in the Introduction 1.2, the goals for this thesis where:

- Characterize different individuals music listening behavior by means of identifying temporal patterns in music listening activity.
- Obtain a set of artists/genres for each individual which do follow a specific daily or weekly rhythm.
- Confirm that each individual listening behavior is far away from the global one.

After performing a quite in depth review of the literature about music preferences and contextualization of them, the present study is, as far as we know, the first one inquiring the possibility that our music listening behavior may follow some detectable circadian and ultradian patterns, at least under certain circumstances. In this case, the circumstances are referred to the preprocess the raw data coming from logs of users from the online radio station Last.fm, which aim to filter out all that data that wasn't useful for the study. So as to detect temporal patterns of music listening behavior, the need of a "robust" (robust by means of playlists with enough number of repetitions) data set lead us to filter out sporadical or not frequent reproductions. However, although we got rid of all that "noise", we managed to keep 62% of the total playcounts in the artist filtering process and 80% in the genre filtering one.

Working with this "clean" data set and using different circular statistics tools, we have discovered that a non-negligible amount of listeners tend to prefer to listen certain artists and genres at specific moments of the day and/or at certain days of the week. In fact, from the working data set, more than the 90% of the artists and genres do present a tendency to be listened regarding the weekly and daily rhythm. Nevertheless, this 90% of the artists and genres, correspond to the 4.5% in the artist case and 8.6% in genre case of the total amount of artists and genres before the filtering process. So, this small set of artists and genres that are being listened to following a circadian or ultradian rhythm do characterize each listener behavior. This last statement is reinforced with the results obtained for the genre selection, where more than the 70% of user's listening habits are far away from the global ones, which consolidates the idea that the listening activity has so much to say about self personality and preferences. In fact we have also observed that between 20% and 40% of these trends are robust enough to allow predictions to be successful for contextual music recommendation regarding each user's listening behavior.

Another thing to be pointed is the usefulness of the Circular Statistics to face up the problem. The few requirements that have to be met , the simplicity of understanding the results and the clear way of representing them have become it an essential tool for achieving the goals of this thesis.

Concerning that the scope of the project is highly linked with the contextual recommendation based systems, we wrote a short paper where we describe the idea of the project, as well as the methodology we follow and the results and conclusions we got. This paper has been accepted for the *WOMRAD 2010* , a Workshop on Music Recommendation and Discovery part of the RecSys 2010, the 4th ACM conference on Recommender Systems.

5.2 Future Work

As future work, regarding the implementation part of the work, we believe that more sophisticated prediction models could be tested such as Markov Models or even Bayesian Networks [41]. The use of Markov models based on the thought that the occurrence of the next event depends only on a fixed number of previous events, could be approached for anticipating listeners behavior. However, when thinking about modeling listeners behavior, an implicit working assumption of our and many predictive models is that users maintain their behavior or interests over time. Users have interests and behaviors that change over time, and prediction models should be able to adapt to these changes. A solution for this issue would be to retrain a prediction model whenever its performance deteriorates beyond a certain threshold. Another way of solving this problem would be using the mentioned Bayesian Networks, which show an incremental adaptive behavior.

Concerning the analysis part of the work, there are still some questions that can be answered performing a deeper and more general analysis of the data. As an example, temporal patterns could differ between women and men, which is a common question when observing patterns of behavior or communication. In a more general approach we could think of investigating if there are artists or genres that clearly show a greater degree of cyclic rhythmicity (e.g the genre chill-out, could be that happens to be listened almost at nights), or others that don't show it at all. Although this last question appear to be a bit far from the scope of the project (we have been working in a personalize way, trying not to classify nor to create communities), could be a interesting information of the data-set, characterizing it as a whole.

Another future plan, in this case trying to find a workable place for this study, would be to embed all the processes as part of an application. We can think about ways to implement it into existing systems or as an added functionality of a web-based recommender service could be explored. The main contribution of our study is that better music recommendations could be done regarding the time of the day or the day of the week, based on the rhythmicity that our results have shown. So, when thinking about implementing or embedding it into existing systems, the preferred purpose of the application would be to give to the user personalized recommendations according to the day and the hour of the query. One option so as to fulfill the temporal patterns detection and to give similarity based recommendations could be to use existing systems such as the one given by the Last.fm API. Another option could be to simply use the locally registered reproductions of the user in his/her music player to compute the temporal patterns behind them. In fact, our temporal patterns detection system could be seen as a plug-in for a music player which would add a personalized functionality to it. In the following lines we would describe the process we may follow for development of this kind of plugins. As an example, we will explain how this could be implemented using Last.fm API's methods. First of all, the system would make a query to last.fm so as to acquire each user's listening habits or history. This information would be parse and given to temporal patterns detection process. This process would only have to be carried out one time, identifying those artists or genres which show a characteristic cyclic rhythm for that specific user. Once this temporality has been detected for a set of artists and genres, we could make new queries to last.fm for similar artists or genres regarding the day of the week the query is being done and the hour of the day. Another similar approach to exploit the temporality of the process, would be instead of querying last.fm and presenting the results from the query to the user, by means of some pop-ups show the recommendation to the user, without making it sound and so, letting the user choose if he/she wants to get the recommendation or continue listening to whatever he/she was listening. The diagram below explains graphically the process:

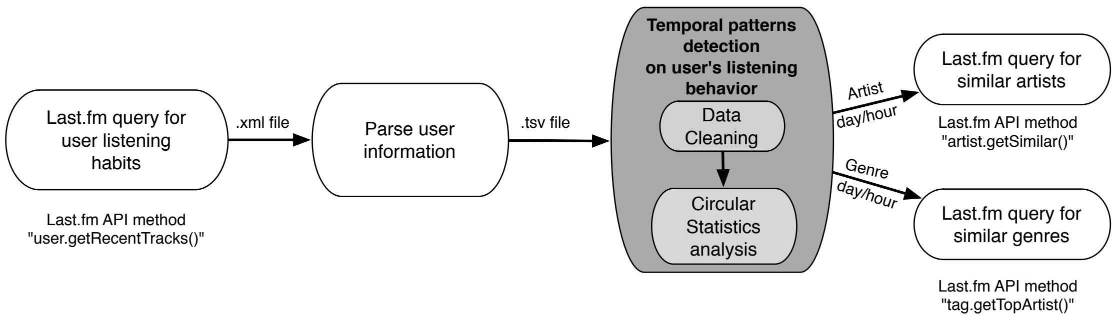

Figure 5.1: Block diagram of a possible application

So as to exploit the analysis of the found temporal patterns, another possible application would be as an added functionality of visualization of the music listening habits of the users for music player. Implementing a visualization based on the circular representation of the habits of each user would be useful for the listeners, so as to have an overview of the listening hours and days of a specific artist or genre. This type of information is valuable in the sense that normally the listeners do not have a time aware annotation of their records, and so, could be very informative to visualize them.

Last, but not least, as the results from the evaluation confirm, better recommendations could be assesed based on the results obtained from the temporal patterns detection regarding the artist or genre selection. So, recommendations according to the time of the day or the day of the week could result on a increase of the satisfaction of the users of music recommenders systems.

Appendix A

Short Article accepted for *WOMRAD 2010* Workshop on Music Recommendation and Discovery (ACM RecSys 2010)

Rocking around the clock eight days a week: an exploration of temporal patterns of music listening

Perfecto Herrera
perfecto.herrera@upf.edu

Zuriñe Resa
Music Technology Group
Department of Technology
Universitat Pompeu Fabra
zuri_resa@hotmail.com

Mohamed Sordo
mohamed.sordo@upf.edu

ABSTRACT
Music listening patterns can be influenced by contextual factors such as the activity a listener is involved in, the place one is located or physiological constants. As a consequence, musical listening choices might show some recurrent temporal patterns. Here we address the hypothesis that for some listeners, the selection of artists and genres could show a preference for certain moments of the day or for certain days of the week. With the help of circular statistics we analyze playcounts from Last.fm and detect the existence of that kind of patterns. Once temporal preference is modeled for each listener, we test the robustness of that using the listener's playcount from a posterior temporal period. We show that for certain users, artists and genres, temporal patterns of listening can be used to predict music listening selections with above-chance accuracy. This finding could be exploited in music recommendation and playlist generation in order to provide user-specific music suggestions at the "right" moment.

Categories and Subject Descriptors
H.5.5 Sound and Music Computing – methodologies and techniques, modeling.

General Terms
Measurement, Experimentation, Human Factors.

Keywords
Music context analysis, Playlist generation, User modeling, Music metadata, Temporal patterns, Music preference.

1. INTRODUCTION
Among the requirements of good music recommenders we can point to, not only delivering the right music but, delivering it at the right moment. This amounts to consider the context of listening as a relevant variable in any user model for music recommendation. As existing technologies also make it possible to track the listening activity every time and everywhere it is happening, it seems pertinent to ask ourselves how this tracking can be converted into usable knowledge for our recommendation systems. Music listening decisions might seem expressions of free will but they are in fact influenced by interlinked social, environmental, cognitive and biological factors [21][22].

WOMRAD 2010 Workshop on Music Recommendation and Discovery, colocated with ACM RecSys 2010 (Barcelona, SPAIN).

Chronobiology is the discipline that deals with time and rhythm in living organisms. The influence of circadian rhythms (those showing a repetition pattern every 24 hours approximately, usually linked to the day-night alternation), but also of ultradian rhythms (those recurring in a temporal lag larger than one day like the alternation of work and leisure or the seasons), has been demonstrated on different levels of organization of many living creatures, and preserving some biological cycles is critical to keep an optimum health [18]. The observation that human behavior is modulated by rhythms of hormonal releases, exposure to light, weather conditions, moods, and also by the activity we are engaged into [12][3] paves the way to our main hypothesis: there are music listening decisions that reflect the influence of those rhythms and therefore show temporal patterns of occurrence. The connection would be possible because of the existing links between music and mood on one side, and between music and activity on the other side. In both cases, music has functional values either as mood regulator [23] or as an activity regulator [13]. Therefore, as mood and activity are subject to rhythmic patterns and cycles, music selection expressed in playlists could somehow reflect that kind of patterning [26][23]. More specifically, in this paper we inquire on the possibility of detecting that, for a specific user, certain artists or musical genres are preferentially listened to at certain periods of the day or on specific days of the week. The practical side of any finding on this track would be the exploitation of this knowledge for a better contextualized music recommendation. Our research is aligned with a generic trend on detecting hidden patterns of human behavior at the individual level thanks, mainly, to the spread of portable communication and geolocation technologies [4][20].

2. RELATED RESEARCH
While recommendations based on content analysis or on collaborative filtering may achieve a certain degree of personalization, they do miss the fact that the users interact with the systems in a particular context [19]. Furthermore, several studies have shown that a change in contextual variables induces changes in user's behaviors and, in fact, when applying contextual modelling of the users (i.e., considering the time of the day, the performed activity, or the lighting conditions), the performance of recommendation systems improves both in terms of predictive accuracy and true positive ratings [8][25]. Although context-based music recommenders were available since 2003 [1], time information is a recently-added contextual feature [7][17].

A generic approach to the characterization of temporal trends in everyday behavior has been presented in [10], where the concept of "eigenbehavior" is introduced. Eigenbehaviors are characteristic behaviors (such as leaving early home, going to work, breaking for lunch and returning home in the evening) computed from the principal components of any individual's behavioral data. It is an open research issue if Eigenbehaviors could provide a suitable framework for analyzing music listening patterns. A model tracking the time-changing behavior of users and also of recommendable items throughout the life span of the data was developed for the Netfix movie collection [14]. This allowed the author to detect concept drifts and the temporal evolution of preferences, and to improve the recommendation over a long time span.

Although research on behavioral rhythms has a long and solid tradition, we are not aware of many studies about their influence on music listening activities. The exception is a recent paper [2] where users' *micro-profiles* were built according to predefined non-overlapping temporal partitions of the day (e.g., "morning time slot",). The goal of the authors was to build a time-aware music recommender and their evaluation of the computed micro-profiles showed their potential to increase the quality of recommendations based on collaborative filtering. Most of that reported work was, though, on finding optimal temporal partitions. As we will see, there are other feasible, maybe complementary, options that keep the temporal dimension as a continuous and circular one by taking advantage of circular statistics. Developed forty years ago and largely used in biological and physical sciences, circular statistics has also been exploited in personality research for studying temporal patterns of mood [15][16]. To our knowledge, it is the first time they are used in the analysis of music-related behavior, though applications to music have been previously reported [5][9].

3. METHODOLOGY

3.1 Data Collection

Getting access to yearly logs of the musical choices made by a large amount of listeners is not an easy task. Many music playing programs store individual users' records of that, but they are not publicly accessible. As a workable solution, we have taken advantage of Last.fm API, which makes possible to get the playcounts and related metadata of their users. As raw data we have started with the full listening history of 992 unique users, expressed as 19,150,868 text lines and spanning variable length listening histories from 2005 to 2009. The data contained a user identifier, a timestamp, Musicbrainz identifiers for the artist and track, and a text name for the listened track.

The artist genre information was gathered from Last.fm using the Last.fm API method *track.getTopTags()*, which returns a list of tags and their corresponding weight[1].This list of tags, however, may relate to different aspects of music (e.g. genre, mood, instrumentation, decades...). Since in our case we need a single genre per track, we first clean tags in order to remove special characters or any other undesirable characters, such as spaces, hyphens, underscores, etc. Then irrelevant tags (i.e., those having a low weight) are removed and the remaining ones are matched against a predefined list of 272 unique musical genres/styles gathered from Wikipedia and Wordnet. From the genre tags we obtained for each song, we select the one with the highest weight. If there are several tags with the highest weight, we select the one with the least popularity (popularity is computed as the number of occurrences of a specific genre in our data-set).

3.2 Data cleaning

Data coming from Lastfm.com contain playcounts that cannot be attributable to specific listening decisions on the side of users. If they select radio-stations based on other users, on tags or on similar artists there are chances that songs, artists and genres will not recur in a specific user's profile. In general, even in the case of having data coming from personal players obeying solely to the user's will, we should discard (i) users that do not provide enough data to be processed, and (ii) artists and genres that only appear occasionally. We prefer to sacrifice a big amount of raw data provided those we keep help to identify a few of clearly recurring patterns, even if it is only for a few users, artists or genres.

In order to achieve the above-mentioned cleaning goals we first compute, for each user, the average frequency of each artist/genre in his/her playlist. Then, for each user's dataset, we filter out all those artists/genres for which the playlist length is below the user's overall average playlist length. Finally, in order to get rid of low-frequency playing users, we compute the median value of the number of artists/genres left after the last filtering step, which we will name as "valid" artists/genres. Those users whose number of "valid" artists/genres is below the median percentage value are discarded.

3.3 Prediction and Validation Data Sets

Once we get rid of all the suspected noise, we split our dataset in two groups. One will be used to generate the temporal predictions while the other one will be used to test them. The test set contains all the data in the last year of listening for a given subject. The prediction-generation set contains the data coming from two years of listening previous to the year used in the test set.

3.4 Circular Statistics

Circular statistics are aimed to analyze data on circles where angles have a meaning, which is the case when dealing with daily or weekly cycles. In fact, circular statistics is an alternative to common methods or procedures for identifying cyclic variations or patterns, which include spectral analysis of time-series data or time-domain based strategies [15]. Although these approaches are frequently used, their prerequisites (e.g., interval scaling, regularly spaced data, Gaussianity) are seldom met and, as we mentioned above, these techniques have rarely been used to analyze music-related data and therefore we wanted to explore its potential.

Under the circular statistics framework, variables or data considered to be cyclic in nature are meant to have a period of measurement that is rotationally invariant. In our case this period is referred to the daily hours and the days of the week. Therefore, taking into account the rotationally invariant period of analysis this would be reflected as daily hours that range from 0 to 24, where 24 is considered to be the same as 0. Regarding to the weekly rhythm, Monday at 0h would be considered to be the same as Sunday at 24h.

[1] Last.fm relevance weight of tag *t* to artist *a*, ranging from 0 to 100.

The first step in circular analysis is converting raw data to a common angular scale. We chose the angular scale in radians, and thus we apply the following conversion to our dataset:

$$\alpha = \frac{2\pi x}{k}$$

where x represents raw data in the original scale, α is its angular direction (in radians) and k is the total number of steps on the scale where x is measured. In fact, we denote α as a vector of N directional observations α_i (i ranging from 1 to N). For the daily hour case, x would have values between 0 and 24, and k = 24. Alternatively, for the weekday analysis, x would have a scale from 0 (Monday) to 6 (Sunday) and thus, k = 6. As noted, the effect of this conversion can be easily transformed back to the original scale. Once we have converted our data to angular scale, we compute the *mean direction* (a central tendency measure) by transforming raw data into unit vectors in the two-dimensional plane by

$$r_i = \begin{pmatrix} \cos\alpha_i \\ \sin\alpha_i \end{pmatrix}$$

After this transformation, vectors r_i are vector-averaged by

$$\bar{r} = \frac{1}{N}\sum_i r_i$$

The quantity $\bar{r}$ is the *mean resultant vector* associated to the mean direction, and its length $\bar{R}$ describes the spread of the data around the circle. For events occurring uniformly in time $\bar{R}$ values approach 0 (uniform circular distribution) whereas events concentrated around the mean direction yield values close to 1 (see figure 1 for an example). A null hypothesis (e.g., uniformity) about the distribution of data can be assessed using Rayleigh's [11] or Omnibus (Hodges-Ajne) tests [27], the latter working well for many distribution shapes. Once we have detected significantly modally distributed data by means of both tests, we verify that it wasn't completely pointing to a single day or hour. All the circular statistics analyses presented here have been performed with the CircStat toolbox for Matlab [6].

4. RESULTS

4.1 Data cleaning

As a consequence of the cleaning process, our working dataset now contains data from 466 valid users. The cleaning process has kept 62% of their total playcounts, which corresponds to 4,5% of the initial amount of artists. This dramatic reduction of the artists should not be surprising as many listening records show a "long-tail" distribution, with just a few of frequently played artists, and many of them seldom played. On the other hand, when focusing on musical genre listening, the working dataset includes 515 users, from which 78% of their playcounts has been kept. These playcounts comprise 8.6% of the total number of genres. Again, a long-tail distribution of the amount of listened genres is observed.

4.2 Temporal Patterns of Artist Selection

Once we have cleaned our dataset, we compute the mean circular direction and the mean resultant vector length for each artist and user. Therefore, these values can be considered as a description of the listening tendencies for each artist by each user. Both parameters were calculated for the daily and for the weekly data.

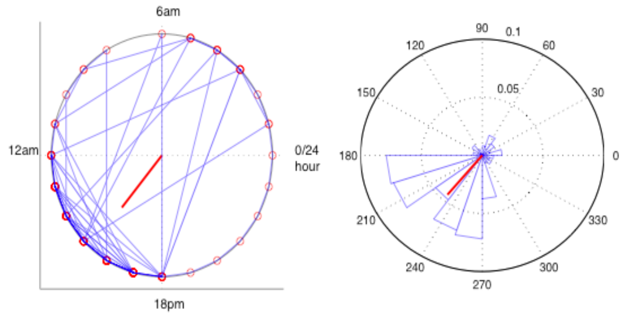

Figure 1. Circular representation of a specific user listening behavior for a specific artist along 24 hours. The left side diagram shows the daily distribution of listening, and the right one the circular histogram. The red line represents the mean vector direction and length in both cases.

In order to assess the relevance of these listening trends, we tested that the distribution of playcounts was different from uniform, and that it was modally distributed (i.e, showing a tendency around an hour or around a day of the week) and discarded those that were not fulfilling these requirements (a null hypothesis rejection probability $p<0.05$ was set for the tests).

In the hour prediction problem, for each listener's clean dataset almost 93% (σ=13) of the artists passed on average the uniformity test (i.e., listening to them is meant to be concentrated around a specific hour). However, considering the raw dataset, only a per-user average of 7% (σ=3.2) of the artists show a listening hour tendency. For the weekly approach, the per-user average in the clean dataset is 99.8% (σ=0.8), indicating that there are some artists showing a clear tendency towards a preferred listening day. Considering the original raw dataset, they correspond to a 7.5% (σ= 3.2) of all the played artists.

Data from 466 users, including 7820 different songs and a grand total of 23669 playcounts were used in the validation of the temporal listening patterns of artists. For each user and artist we computed a "hit" if the absolute difference between the playing day in the prediction and test conditions, expressed as a circular mean value in radians, was less than 0.45 (the equivalent to a half-a-day error). For the time of the day a half-an-hour error was accepted, corresponding to a difference between the predicted and the observed time of less than 0.13 radians.

When predicting the day of listening, an overall 32.4% of hits was found for the songs in the test collection, which exceeds by far the chance expectations (1/7=14.28%). As the final goal of the model is providing user-specific contextual recommendation, an additional per-user analysis yielded 34.5% of hits (σ=17.8). Identical data treatment was done with the time of the day yielding an overall 17.1% of hits (chance expectation baseline: 1/24=4.1%) and a per-user hit rate of 20.5% (σ=16.4).

4.3 Temporal Patterns of Genre Selection

Data from 456 users, including more than 5100 songs and 117 genres, were used for the validation of the genre-related patterns. In order to consider a "hit" in the prediction of listening time and day for a given genre, we set the same thresholds than for evaluating the artist prediction. For the time of the day an overall 22.6% (and per-user 23.2%) of accurate predictions was found. It is interesting to note that relaxing the required accuracy of the prediction to plus/minus one hour error we reached 39.9% of average hits and per-user average 41% (σ=28.4). For the day of the week, the overall hit percent was 40.9%, while the per-genre average and the per-user average were, respectively, 40.7% (σ=24.1) and 41.7% (σ=26.3). It is interesting to note that among the best predictable genres we find many of infrequent ones but also many of the most frequent ones.

5. CONCLUSIONS

The present study is, as far as we know, the first one inquiring the possibility that our music listening behavior may follow some detectable circadian and ultradian patterns, at least under certain circumstances. We have discovered that a non-negligible amount of listeners tend to prefer to listen to certain artists and genres at specific moments of the day and/or at certain days of the week. We have also observed that, respectively for artists and for genres, 20% and 40% time-contextualized music recommendations can be successful. In our future work agenda, more sophisticated prediction models will be tested, and also ways to implement them into existing music recommenders.

6. ACKNOWLEDGMENTS

Our thanks to Òscar Celma who kindly shared the Last.fm data file, accessible from this URL: http://www.dtic.upf.edu/~ocelma/MusicRecommendationDataset/lastfm-1K.html

7. REFERENCES

[1] Anderson, M., Ball, M., Boley, H., Greene, S., Howse, N., Lemire, D, and McGrath, S. 2003. Racofi: A rule-applying collaborative filtering system. In Proc. of COLA'03.

[2] Baltrunas, L. and Amatriain, X. 2009. Towards Time-Dependant recommendation based on implicit feedback. RecSys09 Workshop on Context-aware Recommender Systems (CARS-2009).

[3] Balzer, H.U. 2009. Chronobiology as a foundation for and an approach to a new understanding of the influence of music. In R. Haas and V. Brandes (Eds.), *Music that Works*. Wien/New York: Springer Verlag.

[4] Barabasi, A.L. 2010. *Bursts: The Hidden Pattern Behind Everything We Do*. New York: Dutton Books.

[5] Beran, J. 2004. *Statistics in Musicology*, Boca Raton: CRC.

[6] Berens P., 2009, CircStat, a Matlab Toolbox for Circular Statistics, *Journal of Statistical Software*, 31, 10.

[7] Boström, F. 2008. AndroMedia - Towards a Context-aware Mobile Music Recommender. Master's thesis, University of Helsinki, Faculty of Science, Department of Computer Science. https://oa.doria.fi/handle/10024/39142.

[8] Coppola, P., Della Mea, V., Di Gaspero, L., Menegon, D., Mischis, D., Mizzaro, S., Scagnetto, I. and Vassena, L. 2009. The context-aware browser. IEEE Intelligent Systems, 25,1, 38-47.

[9] Dressler, K. and Streich, S. 2007. Tuning Frequency Estimation Using Circular Statistics. 8th Int. Conf. on Music Information Retrieval (ISMIR-2007), 357-360.

[10] Eagle, N. and Pentland, A.S. 2009. Eigenbehaviors: Identifying structure in routine. *Behavioral Ecology and Sociobiology*, 63, 7, 1057-1066.

[11] Fisher N.I., 1993, *Statistical Analysis of circular data*, Cambridge: Cambridge University Press.

[12] Foster, R.G., and Kreitzman, L. 2005. *Rhythms of Life: The Biological Clocks that Control the Daily Lives of Every Living Thing*. Yale: Yale University Press.

[13] Hargreaves, D. J. and North, A. C. 1999. *The functions of music in everyday life*: Redefining the social in music psychology. Psychology of Music 27, 71-83.

[14] Koren, Y. 2009. Collaborative filtering with temporal dynamics, New York, NY, USA, 447-456.

[15] Kubiak, T. and Jonas, C. 2007. Applying circular statistics to the analysis of monitoring data: Patterns of social interactions and mood. *European Journal of Personality Assessment*, 23, 227-237.

[16] Larsen, R.J., Augustine, A.A., and Prizmic, Z. 2009. A process approach to emotion and personality: Using time as a facet of data. *Cognition and Emotion*, 23, 7, 1407-1426.

[17] Lee, J.S. and Lee, J.C. 2008. Context awareness by case-based reasoning in a music recommendation system. *4th Int. Conf. on Ubiquitous Computing Systems*, 45-58.

[18] Lloyd, D., and Rossi, E. 2008. *Ultradian Rhythms from Molecules to Mind: a new vision of life*. New York: Springer.

[19] Lombardi, S., Anand, S. and Gorgoglione, M. 2009. Context and Customer Behavior in Recommendation. RecSys09 Workshop on Context-aware Recommender Systems.

[20] Neuhaus, F., 2010. *Cycles in Urban Environments: Investigating Temporal Rhythms*. Saarbrücken: LAP.

[21] Radocy, R.E. and Boyle, J.D. 1988. *Psychological Foundations of Musical Behavior* (2nd ed.) Springfield, IL: Charles C. Thomas.

[22] Rentfrow, P.J. and Gosling, S.D. 2003. The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, 84, 6, 1236-1256.

[23] Reynolds, G., Barry, D., Burke, T., and Coyle, E. 2008. Interacting with large music collections: towards the use of environmental metadata. *IEEE International Conference on Multimedia and Expo*, 989-992.

[24] Saarikallio, S., and Erkkilä, J. 2007. The role of music in adolescents' mood regulation. *Psych. of Music*, 35, 1, 88-109.

[25] Su, J.H., Yeh, H.H., Yu, P.S., Tseng, V. 2010. Music recommendation using content and context information mining. *IEEE Intelligent Systems*, 25, 16-26.

[26] Valcheva, M. 2009. Playlistism: a means of identity expression and self-representation. Technical Report, Intermedia, University of Oslo. http://www.intermedia.uio.no/download/attachments/43516460/vit-ass-mariya_valcheva.pdf?version=1

[27] Zar J.H. 1999, *Biostatistical Analysis* (4th edition), Upper Saddle River, NJ: Prentice Hall.

Bibliography

[1] X. Amatriain. Towards time-dependant recommendation based on towards time dependant recomenadation based on implicit feedback. RecSys09, Workshop on Context-aware Recommender Systems (CARS-2009), 2009.

[2] H.U. Balzer. Chronobiologyas a foundation for and an approach to a new understanding of the influence of music. *Music that works*, pages 25–82, 2009.

[3] A.L. Barabási and A. Gelman. Bursts: The hidden pattern behind everything we do. *Physics Today*, 63(5):46, 2010.

[4] J. Beran. *Statistics in musicology.* Boca Raton, FLA: CRC Press, 2004.

[5] P. Berens and V. MJ. The circular statistics toolbox for matlab. Technical report, Technical Report 184, Max Planck Institute for Biological Cybernetics. URL http://www. kyb. tuebingen. mpg. de/publication. html, 2009.

[6] F. Boström. Andromedia - towards a context-aware mobile music recommender. Master's thesis, University of Helsinky, Faculty of science, Department of Computer Science., https://oa.doria.fi/handle/10024/39142., 2008.

[7] R.B. Cattell and D.R. Saunders. Musical preferences and personality diagnosis: A factorization of one hundred and twenty themes. *Journal of Social Psychology*, 39:3–24, 1954.

[8] P. Coppola, V. Della Mea, L. Di Gaspero, D. Menegon, D. Mischis, S. Mizzaro, I. Scagnetto, and L. Vassena. Context-aware browser. *IEEE Intelligent Systems*, 2009.

[9] N.J. Cox. Circstat: Stata modules to calculate circular statistics. *Statistical Software Components*, 1998.

[10] K. Dressler, S. Streich, and F. Publica. Tuning frequency estimation using circular statistics. In *Proceedings of the International Conference on Music Information Retrieval ISMIR*, 2007.

[11] N. Eagle and A.S. Pentland. Eigenbehaviors: Identifying structure in routine. *Behavioral Ecology and Sociobiology*, 63(7):1057–1066, 2009.

[12] NI Fisher. *Statistical analysis of circular data*. Cambridge Univ Pr, 1996.

[13] C. Fitzpatrick. Scrobbling identity: Impression management on last. fm. *Technomusicology: A Sandbox Journal*, 1(2), 2008.

[14] R.G. Foster and L. Kreitzman. *Rhythms of life: The biological clocks that control the daily lives of every living thing*. Profile Books, London, 2004.

[15] L. Geoff, S. Saarikallio, and P. Toiviainen. Personality traits correlate with characteristics of music-induced movement. *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009) Jyväskylä, Finland*, 2009.

[16] D.J. Hargreaves and A.C. North. The functions of music in everyday life: redefining the social in music psychology. *Psychology of Music*, 27(1):71, 1999.

[17] G. Hildebrandt. Biological rhythms and their counterparts in music, 2009.

[18] P.N. Juslin and J.A. Sloboda. *Music and emotion: Theory and research*. Oxford University Press Oxford, England, 2001.

[19] WL Kovach. Oriana for windows. *Kovach Computing Services Inc., Anglesey, Wales, United Kingdom*, 1994.

[20] T. Kubiak and C. Jonas. Applying circular statistics to the analysis of monitoring data: Patterns of social interactions and mood. *European Journal of Psychological Assessment*, 23(4):227, 2007.

[21] R. Lambiotte and M. Ausloos. Uncovering collective listening habits and music genres in bipartite networks (11 pages). *Physical Review- Series E-*, 72(6):66107, 2005.

[22] R.J. Larsen, A.A. Augustine, and Z. Prizmic. A process approach to emotion and personality: Using time as a facet of data. *Cognition & Emotion*, 23(7):1407–1426, 2009.

[23] J.S. Lee and J.C. Lee. Context awareness by case-based reasoning in a music recommendation system. In *Proceedings of the 4th international conference on Ubiquitous computing systems*, pages 45–58. Springer-Verlag, 2007.

[24] B. Lemmer. *Discoveries of Rhythms in Human Biological Functions: A Historical Review*, volume 26, pages 1019–1068. Informa Healthcare New York, USA, 2009.

[25] D.J. Levitin and J. McGill. Life soundtracks: The uses of music in everyday life. *unpublished, Dept. of Psychology McGill University*, 2007.

[26] D. Lloyd and E.L. Rossi. Ultradian rhythms from molecules to mind: a new vision of life. 2008.

[27] S. Lombardi, SS Anand, and M. Gorgoglione. Context and customer behavior in recommendation. RecSys09, Workshop on Context-aware Recommender Systems (CARS-2009), 2009.

[28] U. Lund, C. Agostinelli, M.C. Agostinelli, and GNU License. The circular package. 2006.

[29] F. Neuhaus. *Cycles in Urban Environments: Investigating Temporal Rhythms.* Saarbrücken: Lambert Academic Publishing, 2010.

[30] A.C. North and D.J. Hargreaves. Lifestyle correlates of musical preference: 2. media, leisure time and music. *Psychology of music*, 2007.

[31] Nigel Osborne. *Comunicative Musicality:Exploring the basis of human companionship*, volume 01, inbook Towards a chronobiology on music, page 456. Oxford University Press Oxford, England, 01 edition, 01 2008.

[32] P.J. Rentfrow and S.D. Gosling. The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, 84(6):1236–1256, 2003.

[33] 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 989–992, 2008.

[34] J.A. Sloboda. Music structure and emotional response: Some empirical findings. *Psychology of music*, 19(2):110–120, 1991.

[35] MA Stephens. Multi-sample tests for the fisher distribution for directions. *Biometrika*, 56(1):169–181, 1969.

[36] Hsin-Ho; Yu Philip S.; Tseng Vincent S.; Su, Ja-Hwung; Yeh. Music recommendation using content and context information mining. *IEEE Intelligent Systems*, 25:16–26, 2010.

[37] M. Valcheva. Playlistism: a means of identity expression and self-representation. Technical report, Technical Report, Intermedia, University of Oslo, 2009.

[38] A. Voida, R.E. Grinter, N. Ducheneaut, W.K. Edwards, and M.W. Newman. Listening in: practices surrounding itunes music sharing. In *Proceedings of the*

SIGCHI conference on Human factors in computing systems, pages 191–200. ACM New York, NY, USA, 2005.

[39] GS Watson and EJ Williams. On the construction of significance tests on the circle and the sphere. *Biometrika*, 43(3-4):344, 1956.

[40] J.H. Zar. *Biostatistical analysis*. Prentice-Hall Englewood Cliffs, NJ, 1974.

[41] I. Zukerman and D.W. Albrecht. Predictive statistical models for user modeling. *User Modeling and User-Adapted Interaction*, 11(1):5–18, 2001.