From Sensors to Songs: A learning-free novel music recommendation system using contextual sensor data

Abhishek Sen, Martha Larson
Intelligent Systems
Delft University of Technology, Netherlands
a.sen@student.tudelft.nl, m.a.larson@tudelft.nl

ABSTRACT

Traditional approaches for music recommender systems face the known challenges of providing new recommendations that users perceive as novel and serendipitous discoveries. Even with all the music content available on the web and commercial music streaming services, discovering new music remains a time consuming and taxing activity for the average user. The goal for our proposed system is to provide novel music recommendations based on contextual sensor information. For example, contextual place information can be inferred with intelligent use of techniques such as geo-fencing and using lightweight sensors like accelerometers and compass to monitor location. The inspiration behind our system is that music is not in the past, neither in the future, but rather enjoyed in the present. For this reason, the system does not rely on learning the user's listening history. Raw sensor data is fused with information from the web, passed through a cascade of Fuzzy Logic models to infer the user's context, which is then used to recommend music from an online music streaming service (SoundCloud) after filtering out songs based on genre preferences that the user dislikes. This paper motivates and describes the design for a mobile application along with a description of tests that will be carried out for validation.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search, Retrieval and Filtering

Keywords

Context-aware, Music recommender systems, Fuzzy logic, Sensor data fusion

1. INTRODUCTION

Music plays a central role in the daily lives of many people. Today, music streams are readily available through services such as YouTube, Spotify and Apple Music. However, the process of discovering music remains a tedious and unpleasant experience. With a general trend towards streaming music as opposed to downloading music, many music services aim to learn user music tastes and listening behaviors to provide personalized music recommendations. The assumption of the learning approach, namely, that past behavior is a good predictor of future behavior, is certainly not ill founded. Such services will certainly satisfy users looking for a highly predictable music experience. However, users interested in expanding their music horizons will not be satisfied by algorithms that rely on previous listening history or preferences (artist/genre), since they do not support new music discovery. Such algorithms fail to provide the serendipity that is extremely important for users to discover music that is new, but is also not completely alien to them. Instead, to design the system presented here, we make a new assumption. We consider music listening to be independent of the past (history) or the future (prediction) and instead consider it as a function of the present (current context). We use the term *context* to refer to the sum of a user's experience at a given moment, including place, surroundings, activities that the user is currently pursuing and atmospheric effects on the user's mood. We assume that listeners have similar expectations of which music fits a particular context. We rely on the idea that this collective conception of 'music that fits a moment' will provide users with a sense that the recommendations of our system fit their current needs, and at the same time allow them to discover music that they would not have otherwise found themselves.

2. RELATED WORK

There is a large volume of prior research in the field of context-aware music recommender systems (e.g., [2], [**?**], [10], [7]). Bonnin and Jannach present a comprehensive literature survey on automated playlist generation and categorize existing approaches in [1]. They mention the importance of context in automatic playlist generation and also how similarity-based algorithms are an obvious approach when the system's goal is to maximize the homogeneity of the playlist. As a downside, serendipity and diversity are negatively affected since most songs recommended will be of a similar type, i.e., the same with respect to artist or genre. One of their core recommendations for future research is to assess multiple criteria at the same time and explore the trade-offs between homogeneity and diversity of playlists. Our system, explained further in Section 5, addresses these recommendations by balancing diversity and homogeneity and does not rely on learning the user's past listening be-

LocalRec'15, September 19, 2015, Vienna, Austria.

haviors.

In [10], Wang et. al propose a system that is context-aware, probabilistic and learns the user's listening habits over time for better recommendations. Their system utilizes contextual sensor data and integrates this information with music content analysis to provide relevant music recommendations per context. However, the study requires the musical signal of the songs to be pre-analyzed by music analysis and was also evaluated with offline music. In the version presented in this paper, our system instead focuses on music metadata that is directly available and does not rely on learning the user's listening behaviors. Okada et. al present a system in [8] that focuses on the user interface aspects of context-aware music recommender systems, an area often ignored by researchers. One of their core objectives is to explore how context plays a key role in a user's listening behavior and how this information can be conveyed to the user. In the next sections, we will see how this prior work inspired key design choices in our system.

3. DESIGN CONCEPT

Our main design concept is—as the title states—from sensors to songs. We want to recommend novel music to users by inferring their context from sensory data. To achieve the desired surprise and delight factor, the system should not have to learn the user's music tastes and listening behavior. We believe this non-learning characteristic of the system to be, currently, a quite radical approach to music recommendation. It allows users to discover new music continually without any impediments, such as the need to interact frequently with the system. Through inference of user preferences based on collection-wide user experiences of context, we think the system will achieve a level of personalization that is ideal for music recommender systems—without the need to learn everything about the user's listening history.

We are aware that user music preferences are also highly personal. However, instead of making the assumption that music recommendation is "all about personalization", our system strives to integrate "minimum necessary personalization". We do this in two ways. First, we rely on the idea of the context as mentioned above. The situations in which users find themselves can be expected to reflect their lifestyles and overall music preferences for places and activities. A system like ours that relies on collective music preferences of users for specific contexts, is actually providing a level of personalization, albeit indirectly. Second, we allow users a minimum degree of control, e.g., in excluding songs from genres that the user dislikes.

4. DESIGN METHODOLOGY

Inspired by the design concept, our system focuses on providing novel music recommendations with an emphasis on incorporating contextual user information. Our design methodology aims to inform the possibilities for a sensor-based music recommender, with a user centered approach. The goal of sensors embedded within any device is to 'sense' the environment for information such as temperature, acceleration etc. This inherent capability of sensors makes them an ideal choice for use in interpreting user context, especially since most users carry 'smart' devices such as smartphones close to them at all times. This allows the system to respond to major context changes implicitly without requiring any user action. It is also important for our system to be lightweight and run efficiently and not drain the device's battery during normal usage.

Given that the context inference might not be perfect due to 'noisy' sensory data, we want to give the user a choice of playlists. As discussed in our design concept, to exploit the communal behaviors of music listening across different contexts, we will generate contextual tags to retrieve music from SoundCloud[1]. Knees and Schedl [5] discuss tags as a form of text-based approach given their community-based characteristics. SoundCloud has a music database of over 100 million songs, which are richly annotated with tags. Tags of a track that are related to the context provide us with evidence that listeners generally associate the track with that context. Contextual tag-based queries then allow us to retrieve songs from SoundCloud that both, fit contexts and allow users to discover new music.

We conducted an intensive focus group study with 6 Master's students from different faculties at the Delft University of Technology and the feedback gave us key insights for our design process. All of them described music discovery as a tedious and challenging activity even with all the music available on the web. They described their ideal music recommender system would know which song to play for any given situation and not just based on their past history.

One of their main complaints about current music recommender systems was that most systems tend to repeat the same type of songs unless the user has explicitly made a different selection. They were also of the opinion that even though such a system might provide 'bad' recommendations at times, they would simply move on to the next song and continue listening. This insight suggested that our system does not have to infer the user context perfectly and that we could hedge our predictions by providing the user a choice of playlists for the most likely contexts. The group also mentioned that they all had different music tastes and each had their own music preferences for different contexts—this led us to include a genre preferences block as shown in Figure 1 so that in addition to knowing what the user enjoyed listening to of late, more importantly, the system "knows" the kind of music the user really does not enjoy hearing.

5. PROPOSED SYSTEM

The proposed system architecture as shown in Figure 1 is the materialization of our design concept, methodology, and the focus group feedback. The system is divided into three main components: context inference, music retrieval/analysis and music recommendation.

5.1 Context Inference

Context inference as shown in Figure 1 is done by fusing sensor data and passing it through fuzzy logic models.

5.1.1 Sensors

Table 1 shows a list of contextual information categories and the sensors used for their inference. All the sensors used in the system are embedded inside most smartphones and this trend is likely to continue with future 'smart' devices such as smartwatches and other wearables. The system is scalable and additional sensors can be easily integrated to further improve the context inference process.

[1]https://developers.soundcloud.com/docs/api/reference

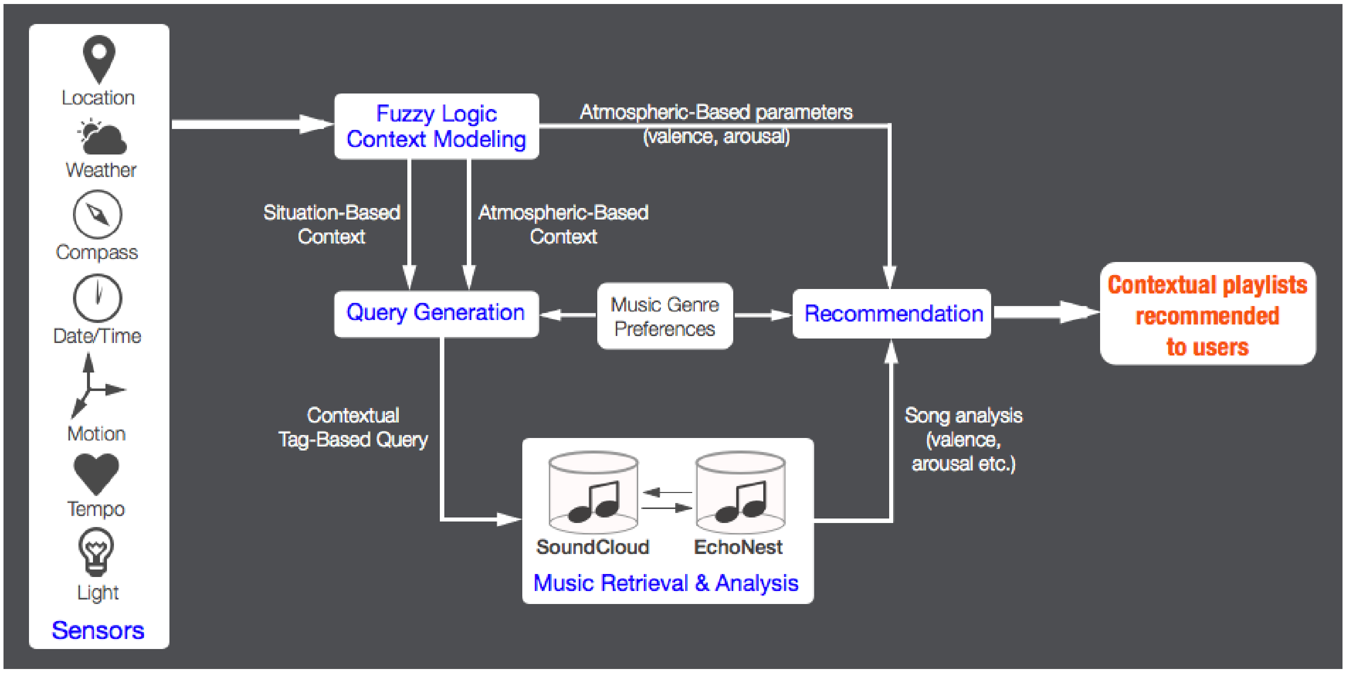

Figure 1: Proposed playlist generation system architecture: Sensor-based novel music recommender

Table 1: Contextual Information

Category	Sensors
Location	Wifi, GPS, Accelerometer, Compass, Cellular
Indoor/Outdoor	Compass, Light
Activity	Accelerometer, Gyroscope
Date/Time	System Clock
Weather	Temperature, Humidity, Pressure, Sunshine

5.1.2 *Fuzzy Logic Context Modeling*

Using fuzzy logic for context inference makes the system extremely flexible and easy-to-understand, and allows it to process imprecise sensor data with ease. Motivated by our non-learning design concept, fuzzy logic makes it possible to translate user-supplied human language rules into mathematical values that can be used for making decisions, thus making the system logic easily understandable. Given the computational challenges of fusing multi-modal sensor data, fuzzy logic provides an extremely light-weight and efficient technique. The Fuzzy Logic Context Modeling block comprises two main internal models as shown in Table 2.

Table 2: Fuzzy Logic Context Models

Category	Inputs
Atmospheric-Based	Temperature, Humidity, Pressure, Sunshine
Situation-Based	Activity, Day of week, Time of day, Indoor/Outdoor, Place

5.1.3 *Atmospheric-Based Context Model*

The Atmospheric-based model generates values for valence and arousal based on prior psychology research on the impacts of different weather factors on people's mood (e.g., [3], [4]). Weather information is integrated into the system using Yahoo API[2]. Mood is a very difficult characteristic to judge on a personal level—especially since everyone's mood could be influenced by a multitude of factors. For this reason, we decided to use the most important weather condition factors that are thought to most universally affect people in a certain geographic area to get a rough estimate of which quadrant of Russell's widely accepted circumplex model of affect the user might be in [9]. The objective here is not to accurately determine the user's mood but to get a general idea depending on the impacts of weather on their mood.

5.1.4 *Situation-Based Context Model*

For the situation-based context model, the focus group results informed us of the most common situations in which participants listen to music and we chose to pick the top 7 for our system: waking up, commuting, working/studying, exercising, relaxing, housework and sleeping. To determine the situation, we use fuzzy rules such as the following:

IF *Activity* IS *Stationary* AND *DayOfWeek* IS *Weekday* AND *TimeOfDay* IS *Afternoon* AND *Indoor/Outdoor* IS *Indoor* AND *Place* IS *Office* THEN *Context* IS *Working or Studying*

The activity states that our system identifies are stationary, walking, running and driving. These activity states are provided by the iOS platform. To accurately distinguish between the stationary and driving state, we utilize GPS to get the user's speed and make a decision accordingly. The indoor/outdoor sensor inputs to this model determine whether the user is indoors or outdoors using sensors such as light and compass and is adapted from Zhou et. al's proposed system in [11]—we do not use cellular signal strength in our system due to lack of development support on iOS.

[2]https://developer.yahoo.com/weather/

Our system is currently able to identify five general place categories for users—home, office, library, gym and other. These areas are recognized without the user having to explicitly enter information. The system monitors significant location updates and marks any visited locations as possible candidates for any of the above five places in a two-step process.

First, using the Foursquare Venues API we reverse geocode the location's coordinates to the library or gym place categories. If no results are returned, the visit information is then passed through an internal fuzzy model to determine the home and office place categories based on fuzzy rules. Once a place has been annotated with a category (not always), the system sets up a geofence around it for a specified radius. From this point on, any time the user enters or leaves this place, a place context change event is triggered and the user's context is recomputed by processing all the other sensory inputs as shown in the Situation-based Context Model in Table 2. If a change in user context is detected, a new contextual song query is formulated to request a new set of songs from SoundCloud. The proposed technique of monitoring places ensures that we do not drain the smartphone's battery by only using the GPS when needed.

5.2 Music Retrieval & Recommendation

Once the user's context has been analyzed, the next step is to retrieve songs from SoundCloud based on this information. The system performs query expansion using Last.fm APIs[3] to translate the fuzzy model's output into query tags. For example, for the 'Exercising' tag, Last.FM returns a set of similar tags such as 'fitness', 'workout' and 'motivation'. These results are aggregated and the top-ten tags are used by the system for this context. For the weather-to-mood tag generation process, social mood tags from [6] were used as tag seeds for the query expansion. Through this process of query expansion and the ever-evolving music community on SoundCloud, the chances of retrieving novel music is very high.

Post song retrieval, the system then performs music content analysis using the EchoNest APIs[4] for each of the retrieved songs by looking at parameters such as energy and valence. However, since each song is analyzed at runtime by uploading the tracks from SoundCloud to EchoNest, this process takes some time and increases as the song duration increases. For this reason, this analysis is done in the background while the user is listening to music. After the songs have been retrieved from SoundCloud and/or EchoNest, the system then generates situational-based and atmospheric-based playlists. It is important at this step that the user identifies the subtle differences between the recommended playlists and we plan on using visual imagery and colors to convey the contextual differences to the user.

6. OUTLOOK & CONCLUSIONS

The system is currently under development as a mobile application on the iOS platform on iPhone models 5S and newer running iOS 8[5]. The evaluation plan is to have about 20-25 participants use the application for a week and answer the following research questions: "What information shown on the user interface would make the music recommendations transparent for users?", "Which modality (mood-based-context or activity-based-context) influences the user more?", "Is the content-based re-ranking for song relevancy necessary for recommendations?".

Future work in this topic includes a number of challenges such as removing the hard-coding of contextual tags and making the tag generation process dynamic. Other alternatives would be to include playlist titles and tracks within the recommendations for playlists. Our design concept and motivations for this system however remain the same—to expand the musical horizons of users while making the music discovery process less tedious and more serendipitous.

[3] http://www.last.fm/api

[4] http://developer.echonest.com/docs/v4

[5] https://developer.apple.com/library/ios/navigation/

7. ACKNOWLEDGEMENTS

The contribution of the second author was funded in part by CrowdRec (EC FP7 Project 610594).

8. REFERENCES

[1] G. Bonnin and D. Jannach. Automated generation of music playlists: Survey and experiments. *ACM Comput. Surv.*, 47(2):26:1–26:35, 2014.

[2] D. Griffiths, S. Cunningham, and J. Weinel. A discussion of musical features for automatic music playlist generation using affective technologies. In *Audio Mostly Conference*, AM '13, 2013.

[3] A. Hannak, E. Anderson, L. F. Barrett, S. Lehmann, A. Mislove, and M. Riedewald. Tweetin' in the rain: Exploring societal-scale effects of weather on mood. In *International AAAI Conference on Weblogs and Social Media*, ICWSM '12, 2012.

[4] E. Howarth and M. S. Hoffman. A multidimensional approach to the relationship between mood and weather. *British Journal of Psychology*, 75(1):15–23, 1984.

[5] P. Knees and M. Schedl. A survey of music similarity and recommendation from music context data. *ACM Trans. Multimedia Comput. Commun. Appl.*, 10(1):2:1–2:21, 2013.

[6] C. Laurier. *Automatic Classification of musical mood by content-based analysis.* PhD thesis, Universitat Pompeu Fabra, 2011.

[7] S. Nirjon, R. F. Dickerson, Q. Li, P. Asare, J. A. Stankovic, D. Hong, B. Zhang, X. Jiang, G. Shen, and F. Zhao. Musicalheart: A hearty way of listening to music. ACM SynSys '12, 2012.

[8] K. Okada, B. F. Karlsson, L. Sardinha, and T. Noleto. Contextplayer: Learning contextual music preferences for situational recommendations. In *ACM SIGGRAPH '13*, 2013.

[9] J. Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161–1178, 1980.

[10] X. Wang, D. Rosenblum, and Y. Wang. Context-aware mobile music recommendation for daily activities. In *ACM Multimedia '12*, 2012.

[11] P. Zhou, Y. Zheng, Z. Li, M. Li, and G. Shen. Iodetector: A generic service for indoor outdoor detection. In *ACM SenSys '12*, 2012.