A Survey on the Acceptance of Listening Context Logging for MIR Applications

Sebastian Stober, Matthias Steinbrecher, Andreas Nürnberger

Faculty of Computer Science
Otto-von-Guericke-University Magdeburg, D-39106 Magdeburg, Germany
{sebastian.stober, matthias.steinbrecher, andreas.nuernberger}@ovgu.de

Abstract. In order to enrich music information retrieval applications with information about a user's listening habits, it is possible to automatically record a large variety of information about the listening context. However, recording such information may violate the user's privacy. This paper presents and discusses the results of a survey that has been conducted to assess the acceptance of listening context logging.

1 Introduction

Information about a user's listening habits may be very valuable in order to build personalized applications in music information retrieval (MIR). E.g., several studies [7,1,9,4] indicate that there might be meaningful user-specific genres emerging from usage patterns that people consciously or unconsciously use when they access music collections or describe music. Such "idiosyncratic genres" [7] might be a valuable extension or even a replacement for – long discussed but not yet agreed upon – objective genre classification schemes. Personalizing not only the process of genre classification but also the classes (i.e. the genres) used could lead to a higher degree of personalization and user satisfaction. Alternatively, information about the usage context can be used directly to browse a music collection or to enrich a similarity-based structuring as orientation aid or as separate content-describing facet.

In previous work [8], we motivated and demonstrated how widely available environmental data can be exploited to allow organization, structuring and exploration of music collections by personal listening contexts. We described a logging plug-in for music players that automatically records data about the listening context and discuss possible extensions for more sophisticated context logging. Based on data collected in a small user experiment, we showed how data mining techniques can be applied to reveal common usage patterns. This data could finally be used by a prototype user interface based on elastic lists for browsing by listening context. A variety of possibilities for automatic context logging was proposed. However, for most of them privacy appears to be the most important issue as more sophisticated methods come closer to surveillance. So the question is not, what is technically possible but how much information about his activities a user is willing to share.

The paper is structured as follows: In Section 2 we present the different options for automatic context logging addressed in this survey. Section 3 briefly describes the survey design and the context. Results of the survey are presented in Section 4. In Section 5 we investigate possible correlations between the background of the participants and their acceptance of context logging. Finally, Section 6 concludes our work.

2 Automatic Context Logging

According to the definition by Dey [2], any information that can be used to characterize the situation of a person, place or object of consideration makes up its context. He differentiates four types of primary context: location, identity, time, and activity [3]. The activity is especially important because it is closely related to the idiosyncratic genres identified in [7]. Moreover, knowledge about the user's current activity might enable a more sophisticated modellng of the listening context with regard to the listening modes defined in [5].

The following options for automatic context logging were addressed:

- **Music metadata**
  Together with a timestamp, this is the minimal information required to build a user's listening profile. It usually comprises the title, artist, album of a song. In the music information retrieval domain, there exists already a variety of systems that capture time, (user) identity and location, e.g. the *Audioscrobbler*[1] plug-in of the music community website *last.fm*[2]. Similarly, *iTunes Genius* sends information about a user's music collection and playlists to a central server. This information is then used to generate Genius playlists through collaborative filtering.
- **Ambient noise**
  Most devices that are capable of playing music have also a built-in microphone. This could be used to record short snippets (1 or 2 seconds) of the environmental soundscape in the gap between two consecutive songs. (Alternatively, signal processing methods could be applied to remove the known music signal from the recorded one. This way, information could even be collected when music is playing.) From the resulting sound snippets, noise profiles could be generated that have enough information to classify the soundscape into general categories like "*silence*", "*people talking*", "*nature sounds*", "*traffic sounds*" or "*party*" thus giving valuable information about location and activity. Using noise profiles instead of the actual recording would further not allow sensible information to be extracted.
- **GPS position**
  Many mobile devices nowadays have a GPS receiver. Periodically recording the position would provide valuable information about the location. Together with the inferred speed of travel this could be linked with specific activities or travel.

[1] http://www.audioscrobbler.net/
[2] http://last.fm/

- **Keyboard and mouse events per minute**
  Assuming the listener is using a computer, further possibilities arise. Detecting whether and how much the mouse and keyboard are used in a sliding time window yields evidence about the user's activity. For instance, low keyboard and mouse activity may indicate reading or browsing whereas high keyboard activity may refer to writing a text or programming. It might not even be necessary to derive such higher level activity description. The low level information might be already sufficient to distinguish activity contexts. Note that no information about the actual keys would be recorded such that it would be impossible to reconstruct the data that was entered.
- **Currently running applications**
  Together with the previous this is probably the best source of information about the listener's current activity if a computer is used but also comes close to surveillance. Through a client-side tracking software, it could be periodically logged which applications are currently running and which application has the focus. The social networking website *wakoopa*[3] already uses such a technique to monitor which programs and web application are used by its members and build profiles.
- **Facial expression**
  Many notebooks and mobile phones have a built-in webcam. With such a camera, the facial expression of the user could be classified periodically using e.g. image processing methods applied in the context of human- computer interaction [6]. Note that only the classifications and no actual images need to be logged.
- **Bio-information**
  Listening context information can also comprise direct information about the user's current condition. For instance, the adaptive system for playlist generation called *PAPA* (Physiology and Purpose-Aware Automatic Playlist Generation) [10] as well as the already commercially available *BODiBEAT* music player[4] uses sensors that measure certain bio-signals (such as the pulse) of the user as immediate feedback for the music currently played. This information is then used to learn which characteristics of music have a certain effect on the user. Based on this continuously adapting model playlists for certain purposes can be created. Alternatively, the sensor information could be used to derive listening contexts.
- **Ambient light**
  Recently, some notebooks are equipped with illuminance sensors to adapt the display brightness. Data from such sensors could be exploited, too. Whilst this might provide only little information about the listening context, logging this would result in almost no interference with the listener's privacy.
- **Status**
  There exists a variety of applications that allow a user to set his current status: instant messenger applications come with predefined states such as

[3] http://wakoopa.com/
[4] http://www.yamaha.com/bodibeat/

> "*online*", "*away*" or "*occupied*" and the option to specify an additional custom status message. Further, many social networking websites or micro-blogs such as *twitter*[5] allow their members to specify what they are currently doing. Depending on how much effort a user puts into updating the status as a measure of communication, this information may be very valuable to describe the listening context.

Except for the bio-information, all data could be gathered at low costs by using only built-in hardware. Further, it can be measured without distracting the user from his current activity. Clearly this is an advantage over simply asking the user, what he is currently doing. (The latter would require a user action without a directly recognizable benefit so that there is hardly any motivation for the user to cooperate.)

3 Survey Design and Context

The design of the survey emerged from pre-surveys with a small number of participants. The original key question targeted MIR applications in general but was considered as too abstract. It had to be rephrased so that it described a scenario that participants not familiar with the MIR domain could easily grasp. Therefore, the chosen motivation was to learn personalized genres for sorting a personal music collection:

> Current music players allow to sort music according to genres. Unfortunately, genres are often either too general (e.g. rock/pop) or far too specific (e.g. "scottish lo-fi post-rock" for the band "Mogwai") such that they are not very helpful for sorting. An alternative is currently investigated within the AUCOMA[6] project of the DKE research group: It might be possible to learn individual "genres" that reflect a user's listening habits (e.g. "breakfast music", "car driving music", "party music"). These could be used to structure the music collection according to individual listening habits. For the identification of different listening situations, the player could record a variety of information.
> BUT: Recording such information may violate your privacy! Therefore, please tell us what information your music player may record about you!

The survey contained 8 questions that can be categorized into 4 topics, each presented in detail with the results in a respective subsection of Section 4:

1. Demographic information comprising, gender, age and country of residence.
2. General relation to music.
3. Use of (web-) applications that collect, access and expose to some extend private data of their users.
4. Acceptance of logging information about the listening context.

[5] http://twitter.com/
[6] http://www.dke-research.de/aucoma

Table 1. Countries with more than 5 participants in the survey.

Country	Number of Participants	Percentage
Germany	323	70.07%
USA	24	5.21%
Austria	14	3.04%
France	10	2.17%
Turkey	7	1.52%
Switzerland	6	1.30%
Spain	5	1.08%

The last topic represents the main question of the survey whereas the others were added to be able to estimate a possible bias of the participants. Further, correlations between the background of the participants (especially in relation to music) were expected that could be identified.

The survey was conducted in two parts: Between March 3rd and March 8th, 2009 a paper questionnaire in German was filled out by 156 fare visitors of the German CeBIT 2009 fare. Based on this questionnaire, an online questionnaire was designed afterwards to extend the scope of the survey. It was open to the public from March 16th until June 15th, 2009. The questions of the online questionnaire were identical to those of the paper version. However, the questions were split across multiple (screen-) pages and an English translation was added for international participants. 305 persons filled out the online questionnaire resulting in 461 participants in total.

4 Survey Results

4.1 Demographic background of the participants

From the 461 participants of the survey, 101 (i.e. 21.9%) were female and 354 (i.e. 76.8%) were male. 6 persons did not answer this question. The average age was 29.25 with a standard deviation of 10. Table 1 shows the countries with more than 5 participants in the survey. In total, persons from 34 countries participated where 4 persons did not state their country of residence.

The majority of the participants was from Germany. This is primarily due to the fact that part of the survey was conducted amongst visitors of the German CeBIT fare. Further, the survey was advertised at the Otto-von-Guericke-University. Thus, it can be assumed that there are many German students amongst the participants. Most of the international participants were probably be informed about the survey by the announcements posted on popular MIR mailing list such as the *music-ir* list maintained by IRCAM.

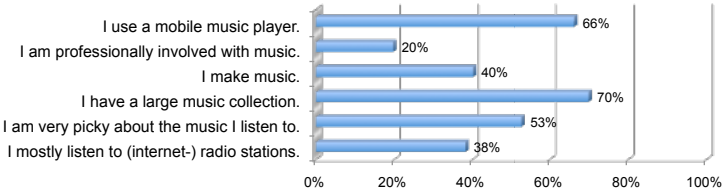

Fig. 1. Statements selected that describe the person's general relation to music.

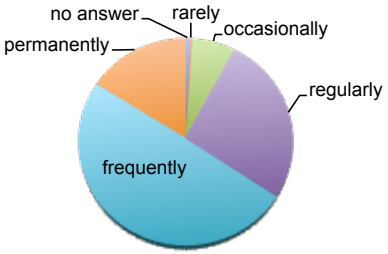

Fig. 2. Answers for the question: *How frequently do you listen to music?*

4.2 General relation to music

6 statements were given that described a person's general relation to music. The participants were ask to check all of these that were applicable to themselves. Figure 1 shows the statements and how often each one was checked.

Further, the participants were asked, how frequently they listen to music. The possible answers for this question were deliberately formulated in a fuzzy way as participants of a pre-survey found this less complicated. The distribution of the answers is shown in Figure 2.

4.3 Use of (web-) applications that collect, access and expose to some extend private data of their users

For 10 categories, one or two popular representatives were chosen exemplarily and depicted by their logos:

1. *last.fm* - a music community website, that builds a detailed profile of each user's musical taste by recording details of all the songs the user listens to via a plugin installed into the user's music player.
2. *flickr* – an image and video community website that allows users to share personal photographs.

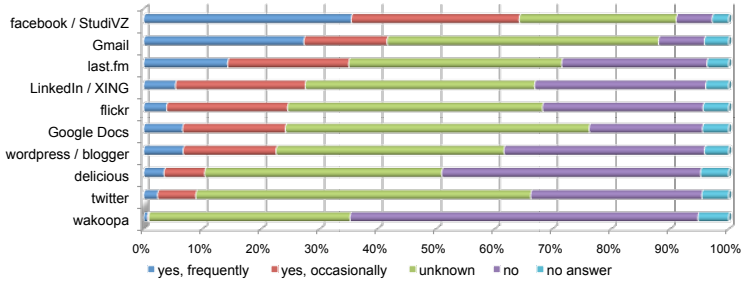

Fig. 3. Answers to the question: *Do you use the following (or comparable) applications?*

3. *delicious* – a social bookmarking website for storing, sharing, and discovering web bookmarks.
4. *wakoopa* – a social networking website that monitors which programs and web application are used by its members through a client-side tracking software.
5. *wordpress / blogger* – two popular blog publishing services.
6. *twitter* – a micro-blogging service that enables its users to send and read short text-based messages of up to 140 characters.
7. *LinkedIn / XING* – two business-oriented social networking websites mainly used for professional networking.
8. *facebook / studiVZ* – two networking websites, the latter being rather popular amongst german-speaking students.
9. *Gmail / Google Mail* – a webmail service.
10. *Google Docs* - a web-based word processor, spreadsheet, presentation, and form application.

Participants were asked whether and how often they use each of these 10 applications. Figure 3 shows the distribution of the answers for the different applications. For better readability, the applications were sorted by decreasing number of users (i.e. participants that answered either *"yes, frequently"* or *"yes, occasionally"*).

4.4 Acceptance of logging information about the listening context

Here, answers were mandatory in contrast to the preceding questions as this was the key point of the survey. Figure 4 shows the distribution of the answers for this question sorted by the decreasing number of persons that answered *"yes (unconditionally)"*. Furthermore, there was the possibility to leave a comment, if for some case the answer was *"maybe"* or *"no"*.

Clearly, logging the music metadata is the variant with the highest acceptance. Given that about 34% of the participants use last.fm, an unconditional

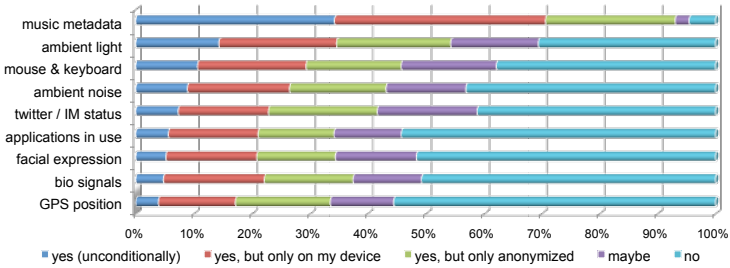

Fig. 4. Answers for the question: *Would you allow your music player (as software or as a self-contained device) to log the following information in order to enable it to learn personalized genres for sorting your music collection? (It is assumed, that you can pause the logging anytime you find it inappropriate.)*

acceptance a little less than 34% is not surprising. It is a surprise that less than 15% would allow logging of the ambient light – supposedly the least sensible information of all. Logging mouse and keyboard events per minute, ambient noise or the status via twitter or instant messaging have medium acceptance. For the latter option, a much higher acceptance was anticipated given that the twitter and instant messaging status is by default visible to anyone and thus not an information that people tend to keep private. On the other end of the spectrum, logging the GPS position is the least popular option. Surprisingly, it is even slightly less accepted than logging bio signal which is definitely most obtrusive. Taking additionally into account the conditional acceptance (i.e. either local or anonymized logging), the participants would rather agree with logging of bio signals than facial expression or applications in use.

In the comments, many participants state that they are concerned about their privacy. Some are against any kind of data collection whereas others do not intend to log data unless they are convinced of the benefit, i.e. an actual improvement of the MIR system. Some even expressed their doubts that such context information could actually be relevant to learn personalized genres. On the other hand, there is fear that users could be patronized by their "intelligent music-player" and would no longer be in control of the music selection. Another fear is that the collected information could be used for marketing purposes. A few participants remark that the additional logging functions would require more storage and processing power or that the development in the end would increase the costs for the hard- or software. Furthermore, some wrote that they would be worried that data once recorded could leak out of the system, e.g. if someone hacked the server or even their computer.

5 Analysis of factors that influence the decision

While the findings presented so far allow us to create a coarse picture of the idiosyncrasies of the survey participants, we are now going to address possible dependencies in greater detail. Broadly speaking, we will present how the acceptance of the amount of context logging varies under certain conditions such as gender, age and other attributes specified in the survey. For this the possible answers to the 9 different logging options discussed in Section 2 were assigned weights from 0 to 2 in the following way: Unconditionally accepting the logging was weighted by a factor of 2. Acceptance of local context information storage on the user's device got assigned the weight 1 while both the claim for anonymized storage and the indecisiveness (indicated by a *"maybe"* answer) amounted 0.5 to the score. Finally, a *"no"* led to zero weight. A new "logging" attribute was introduced by summing up the respective weights of all 9 possible logging contexts (yielding scores between 0 and 18) and subsequently binning it into 6 equidistant groups which served as indicators for the general acceptance to logging context information.

In the remainder of this section we will investigate some selected conditional distributions of this new logging attribute. We explain the applied visualization along with the answer to the question how the logging acceptance varies among the country of residence of the respective survey participant. Figure 5 depicts a sequence of stacked bar charts, one for each distinct condition instance, here: country of residence. Every bar chart represents the relative frequency of the values of the logging attribute, starting with the lowest acceptance at the bottom and increasing acceptance vertically. The widths of the bar charts correspond to the probability of the condition instance itself. Since the majority of participants were Germans, the majority of 70% from Table 1 is reflected by the broad column in the left of Figure 5. It is interesting to see that the next smaller population (USA) has considerable fewer reservations against collecting context information as can be seen by the respective column 195 which shows a large portion of participants that belong the highest acceptance group.

Figure 6 shows the logging acceptance distribution conditioned on both gender and age group. The majority of survey participants depicted by the third bar chart were male persons between 20 and 30 years old. It was unexpected that this subgroup showed a lesser affinity to context logging in general than the next older subgroup represented by the first bar chart in the row. Another noteworthy subgroup comprises the persons under 20 years of age. First, they show the most drastic difference in the distribution when further conditioned on gender. And second, this subgroup seems to be quite reluctant to context logging as there is no participant that belongs to the highest acceptance group. This is insofar interesting as we anticipated a rather airy handling of private information.

The distribution shown in Figure 7 shows the result of investigating the personal relation to music (refer to Section 4.2 and Figure 1) with respect to the logging acceptance. For this we aggregated the six different attributes into one quantifying the overall affinity to music. This new attribute has six values with

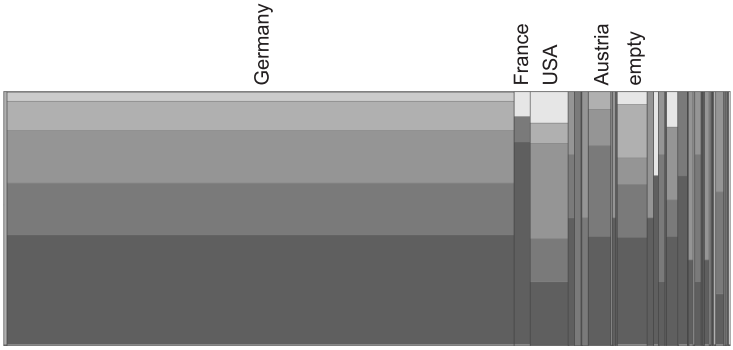

Fig. 5. Distributions of logging acceptance given the country of residence. Brighter areas belong to higher acceptance.

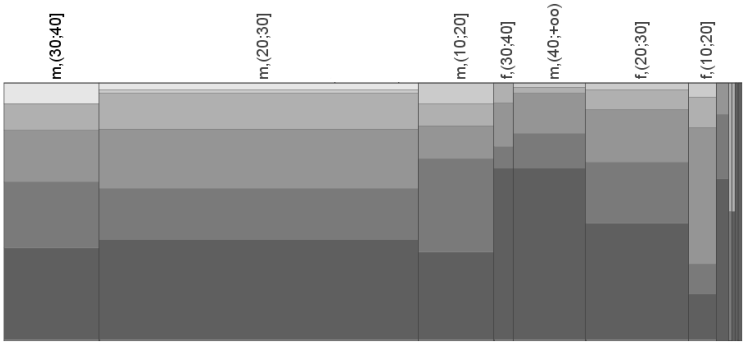

Fig. 6. Distributions of logging acceptance given the age group and gender of the survey participant. Brighter areas belong to higher acceptance.

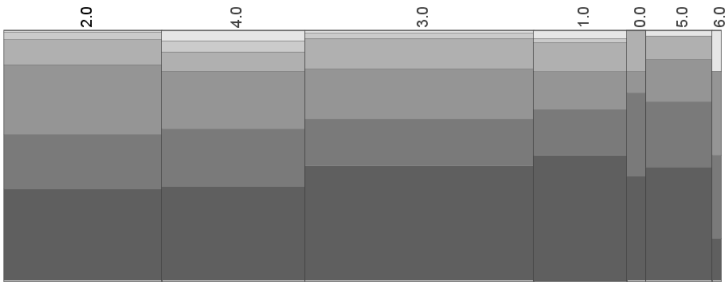

Fig. 7. Distributions of logging acceptance given the strength of affinity towards music. Brighter areas belong to higher acceptance.

0 representing no relation of music at all and 6 denoting strongest relation (by having ticked all possible relations in the survey). Here, no significantly differing subgroups could be identified, neither for the aggregated affinity to music nor for the individual relations.

The remaining question to be answered was how the involvement and participation in the several web applications and online communities (refer to Section 4.3) influences the logging acceptance. To assess the overall involvement we again aggregated the respective attributes into a new one with a domain of four possible values: "low" if the participant used up to 2 web applications, "medium" if he used 3 or 4, "high" if he used 5 or 6 and "very high" of he used more than 6. Brighter segments in the chart of Figure 8 correspond to a higher usage. One can clearly observe a correlation between the level of logging acceptance and the intensity of using online communities and web applications, respectively: The more a participant was involved into online communities (and thus: the more he was used to give away personal information), the higher the acceptance of logging.

6 Conclusions

As expected, privacy is an important issue for most of the participants. For many participants of the survey the sophisticated logging methods addressed come close to surveillance. Thus users must be fully informed about the extend of the logged data and in full control of whether they want this data to be logged or not. Many participants are skeptical and first need to be convinced of the benefits from providing their context information. Furthermore, even if the logged context information proves to be indeed useful for MIR applications, the

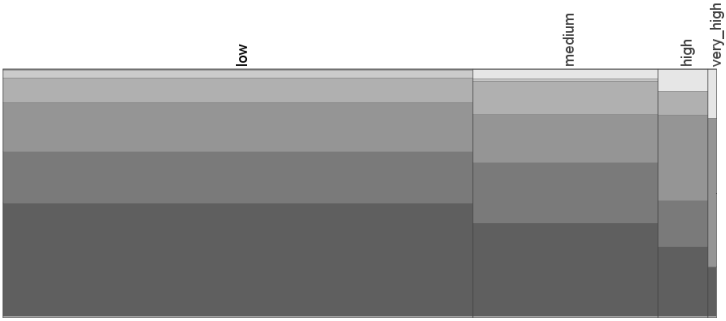

Fig. 8. Distributions of logging acceptance given the usage intensity of online communities and web applications. Brighter areas belong to higher acceptance.

users still want to be in full control of the music selection. This is an important guideline for developing future personalized MIR applications.

Acknowledgements

This work is supported by the German Research Foundation (DFG) and the German National Merit Foundation. The authors would like to thank all participants of the survey and everyone who helped in preparation, especially Korinna Bade, Janet Feigenspan, Tilo Hähnel, Christiane Karthaus, Stephan Lau, Miriam Mezger, and Marcus Nitsche. Henner Graubitz, Christian Moewes and the TTZ Magdeburg helped to conduct the survey at the CeBIT 2009 fare. For the online version of the questionnaire, the open source online survey application LimeSurvey[7] was used.

References

1. D. Bainbridge, S. J. Cunningham, and J. S. Downie. How people describe their music information needs: A grounded theory analysis of music queries. In *Proc. of ISMIR'03*, 2003.
2. A. K. Dey. Understanding and using context. *Personal Ubiquitous Comput.*, 5(1):4–7, 2001.
3. A. K. Dey and G. D. Abowd. Towards a better understanding of context and context-awareness. In *Computer Human Interaction (CHI) 2000 Workshop on the What, Who, Where, When, Why and How of Context-Awareness*, April 2000.

[7] http://www.limesurvey.org/

4. X. Hu, J. S. Downie, and A. Ehmann. Exploiting recommended usage metadata: Exploratory analyses. In *Proc. of ISMIR'06*, 2006.
5. D. Huron. Listening styles and listening strategies. Society for Music Theory 2002 Conference, November 2002.
6. S. Ioannou, G. Caridakis, K. Karpouzis, and S. Kollias. Robust feature detection for facial expression recognition. *Journal on Image and Video Processing*, 2007(2):5–5, 2007.
7. S. Jones, S. J. Cunningham, and M. Jones. Organizing digital music for use: an examination of personal music collections. In *Proc. of ISMIR'04*, 2004.
8. V. Laube, C. Moewes, and S. Stober. Browsing music by usage context. In J. J. Burred, A. Nürnberger, G. Peeters, and S. Stober, editors, *Proceedings of the 2nd Workshop on Learning the Semantics of Audio Signals (LSAS)*, pages 19–29, Paris, June 2008. IRCAM.
9. J. H. Lee and J. S. Downie. Survey of music information needs, uses, and seeking behaviours: Preliminary findings. In *Proc. of ISMIR'04*, 2004.
10. N. Oliver and L. Kreger-Stickles. PAPA: Physiology and purpose-aware automatic playlist generation. In *Proc. of ISMIR'06*, 2006.