Ephemeral Context to Support Robust and Diverse Music Recommendations

Pavel Kucherbaev [1] **Nava Tintarev** [1] **Carlos Rodriguez** [2]

Abstract

While prior work on context-based music recommendation focused on fixed set of contexts (e.g. *walking*, *driving*, *jogging*), we propose to use multiple sensors and external data sources to describe momentary (ephemeral) context in a rich way with a very large number of possible states (e.g. *jogging fast along in downtown of Sydney under a heavy rain at night being tired and angry*). With our approach, we address the problems which current approaches face: 1) a limited ability to infer context from missing or faulty sensor data; 2) an inability to use contextual information to support novel content discovery.

1. Introduction

Imagine the following persona: Anna is a university student in graphic design. She is active, easy going, and organized; keeping a structured well-planned calendar. She is 21 and high on the trait of Openness to Experience (Tkalcic & Chen, 2015), and enjoys traveling to new places and meeting new people, as well as discovering and listening to new music using online radio services.

To improve Anna's listening experience, we consider her context. To decide which aspects of context to look at, we ran an exploratory crowdsourced survey. Out of 103 respondents 97 listen to different music based on their mood (e.g., happy, calm, sad), 92 based on their activity (e.g., commuting, jogging); 38 based on the ambience (e.g. sunny, rainy, loud, quiet), and 32 based on the location (e.g., a city park, a beach). From free-text answers we know that preferences of respondents also change according to the weather, time of day, people around, headphones or speakers used, upcoming concerts, the difficulty of work they do, languages they learn, reminiscence, and music that they just heard somewhere. Several respondents suggested

---
[*]Equal contribution [1]Delft University of Technology, Delft, The Netherlands [2]University of New South Wales, Sydney, Australia. Correspondence to: Pavel Kucherbaev <p.kucherbaev@tudelft.nl>.

combined causes, such as friends that are around and the activity they performed together.

We propose that Anna's online radio would suggest unexpected pleasant surprises based on her momentary unique (and therefore ephemeral) context. The contribution of this paper is using multiple sensors and external data sources to make the inference of life-logging events richer (through combinations of inputs), and more reliable (using multiple sensors to improve fault tolerance). It describes how rich context information can be combined in a large number of ways to improve the diversity of recommendations, which will lead to more opportunities for music discovery.

2. Related work

The type of contextual recommendations that can be made is shaped by sensors and signal processing used. Nowadays it is possible to accurately detect activities such as biking, driving, running, or walking based on smartphone sensors (Liao et al., 2015), or based on environmental sound cues (Shaikh et al., 2008). It is also possible to detect personality traits based on phone call patterns and social network data of the user (de Oliveira et al., 2011). Similarly, interest in an object can be inferred based on ambient noise levels, and positions of people and objects in relation to each other (Dim & Kuflik, 2015). In the SenSay system, phone settings and preferences are set based on detected environmental and physiological states (Siewiorek et al., 2003).

With improvements in smartphone technology, there is a lot of potential for using rich contextual information to improve recommendations, in particular considering that people prefer to listen to different music in different contexts (Cunningham et al., 2008; Schedl et al., 2014; Su et al., 2010). Among the first to propose a context-aware music recommendation system are Park et al. (Park et al., 2006). They used weather data (from sensors and external data sources), and user information, to predict the appropriate music genre, tempo, and mood. Music can also be recommended based on user's heart beat to bring its rate to a normal level (Liu et al., 2009); activities detected automatically (e.g. running, walking, sleeping, working, studying, and shopping) (Wang et al., 2012); driving style, road type, landscape, sleepiness, traffic conditions, mood, weather, natural phenomena (Baltrunas et al., 2011); and
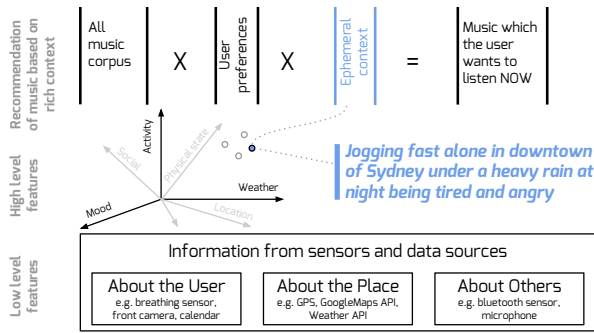
Figure 1. An example of music recommendation using as input ephemeral context constructed from high-level features inferred from smartphone sensors, wearables, and external data sources.

emotional state to help to transition to a desired state (Han et al., 2010). Soundtracks have also been recommended for smartphone videos based on location (using GPS and compass data for orientation), and extra information from 3rd party services such as Foursquare (Yu et al., 2012).

These examples use sensors and external data sources for music recommendation. Some of these context-aware music discovery systems recommend not just relevant, but new music to users (Wang et al., 2014). Our contribution is to combine rich context in a way that is a) fault tolerant, and b) aims to facilitate music discovery, by constructing a momentary ephemeral context.

3. Approach

Here we give an example of how our approach could work to recommend music to Anna by detecting her ephemeral context based on high-level features (e.g. activity, mood, or the weather), which are inferred from low-level sensor data (see Figure 1) and discuss the benefits of such approach.

From low-level to high-level features. We infer that Anna's *activity* is "jogging" from the pattern of her smartphone's accelerometer and GPS, and because this activity was also planned in her calendar. Her *speed* is classified as "fast" for jogging, because her speed is 15km/h, while usually she runs 13km/h. For the *social* component she is classified as "alone", since her Bluetooth sensor does not see Bluetooth sensors of her friends' smartphones and the microphone does not recognise voices around. The *location* is "downtown of Sydney" based on the coordinates given by her smartphone's GPS, and the point of interest identified by GoogleMaps API, as well as recent reviews about it from FourSquare. The *weather* is "heavy rain" according to the moisture sensor of her phone and the weather forecast for the location from Weather API. The *time of day* is "night", because her smartphone time is 23:56. Her *physical state* is "tired" based on the high heart rate measured

by her smart bracelet, and the respiratory pattern coming from her breathing sensor. Her *mood* is detected as "angry" by her smartphone front camera (Busso et al., 2004) and based on her public interactions on social media (e.g., angry emoticon). We combine these high-level features to construct a momentary ephemeral context, which becomes: "*Jogging fast alone in downtown of Sydney under a heavy rain at night being tired and angry*".

From individual recommenders to a hybrid one. We propose to use several individual recommenders focused on different sets of high-level features (e.g. a recommender looking only at location, weather, and time). A hybrid recommender later weights recommendendations of each individual one, based on explicit preferences of Anna, and the reliability of underlying high-level features, if detected at all. Anna can change weights to make an emphasis on a certain aspect, such as location or activity, depending on the way she wants to explore music. We provide an interactive web-based demonstration[1] of how such a hybrid recommender might work based on ephemeral context.

Benefits. Our approach allows us to effectively address fault tolerance and leverage music discovery:

Fault tolerance. Different factors are used as a measure of fault tolerance. For example, if GPS and calendar locations are different, the system will omit location-based recommendations from the hybrid recommender.

Music discovery. Since ephemeral context frequently changes, the recommendations supplied will vary from moment to moment, leading to more opportunities for music discovery (e.g. 8 high-level features taking 8 values each, give more than 16mln combinations).

4. Outlook

Our next steps will be dedicated to *identification of combinations of high-level features* influencing music preferences, possibly via a user study; to *evaluation of recommendations' relevance* with cultural preferences in mind, potentially using crowdsourcing (von Ahn, 2008), and to study how to make such rich user profiling compliant with *privacy* concerns (Coles-Kemp et al., 2011). We also plan to study how to improve the transparency of ephemermal recommendations using textual and visual explanations. As such, we aim to deliver a streaming music experience, driven by context, while giving the user a sense of transparency and control. We are confident that music discovery through rich context is a very promising research topic, allowing streaming services to provide better personalized experiences to their listeners.

---

[1]The demo, the code, and the results of the exploratory survey are available at `https://github.com/pavelk2/ICML2017Demo`.

References

Baltrunas, Linas, Kaminskas, Marius, Ludwig, Bernd, Moling, Omar, Ricci, Francesco, Aydin, Aykan, Lüke, Karl-Heinz, and Schwaiger, Roland. *InCarMusic: Context-Aware Music Recommendations in a Car*, pp. 89–100. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

Busso, Carlos, Deng, Zhigang, Yildirim, Serdar, Bulut, Murtaza, Lee, Chul Min, Kazemzadeh, Abe, Lee, Sungbok, Neumann, Ulrich, and Narayanan, Shrikanth. Analysis of emotion recognition using facial expressions, speech and multimodal information. In *Proceedings of the 6th International Conference on Multimodal Interfaces*, ICMI '04, pp. 205–211, New York, NY, USA, 2004. ACM.

Coles-Kemp, Lizzie, Reddington, Joseph, and Williams, Patricia AH. Looking at clouds from both sides: The advantages and disadvantages of placing personal narratives in the cloud. *Information Security Technical Report*, 16(3):115–122, 2011.

Cunningham, Stuart, Caulder, Stephen, and Grout, Vic. V.: Saturday night or fever? context-aware music playlists. In *In: Proc. Audio Mostly*, 2008.

de Oliveira, Rodrigo, Karatzoglou, Alexandros, Concejero Cerezo, Pedro, Armenta Lopez de Vicuña, Ana, and Oliver, Nuria. Towards a psychographic user model from mobile phone usage. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pp. 2191–2196, New York, NY, USA, 2011. ACM.

Dim, Eyal and Kuflik, Tsvi. Automatic detection of social behavior of museum visitor pairs. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 4(4):17, 2015.

Han, Byeong-Jun, Rho, Seungmin, Jun, Sanghoon, and Hwang, Eenjun. Music emotion classification and context-based music recommendation. *Multimedia Tools Appl.*, 47(3):433–460, May 2010.

Liao, J., Wang, Z., Wan, L., Cao, Q. C., and Qi, H. Smart diary: A smartphone-based framework for sensing, inferring, and logging users'; daily life. *IEEE Sensors Journal*, 15(5):2761–2773, May 2015.

Liu, Hao, Hu, Jun, and Rauterberg, Matthias. Music playlist recommendation based on user heartbeat and music preference. In *Proceedings of the 2009 International Conference on Computer Technology and Development - Volume 01*, ICCTD '09, pp. 545–549, Washington, DC, USA, 2009. IEEE Computer Society.

Park, Han-Saem, Yoo, Ji-Oh, and Cho, Sung-Bae. A context-aware music recommendation system using fuzzy bayesian networks with utility theory. In *Proceedings of the Third International Conference on Fuzzy Systems and Knowledge Discovery*, FSKD'06, pp. 970–979, Berlin, Heidelberg, 2006. Springer-Verlag.

Schedl, Markus, Breitschopf, Georg, and Ionescu, Bogdan. Mobile music genius: Reggae at the beach, metal on a friday night? In *Proceedings of International Conference on Multimedia Retrieval*, ICMR '14, pp. 507:507–507:510, New York, NY, USA, 2014. ACM.

Shaikh, Mostafa Al Masum, Molla, M. K. I., and Hirose, K. Automatic life-logging: A novel approach to sense real-world activities by environmental sound cues and common sense. In *2008 11th International Conference on Computer and Information Technology*, pp. 294–299, Dec 2008. doi: 10.1109/ICCITECHN.2008.4803018.

Siewiorek, Daniel, Smailagic, Asim, Furukawa, Junichi, Krause, Andreas, Moraveji, Neema, Reiger, Kathryn, Shaffer, Jeremy, and Wong, Fei Lung. Sensay: A context-aware mobile phone. In *Proceedings of the 7th IEEE International Symposium on Wearable Computers*, ISWC '03, pp. 248–, Washington, DC, USA, 2003. IEEE Computer Society.

Su, Ja-Hwung, Yeh, Hsin-Ho, Yu, Philip S., and Tseng, Vincent S. Music recommendation using content and context information mining. *IEEE Intelligent Systems*, 25(1):16–26, January 2010.

Tkalcic, Marko and Chen, Li. *Personality and Recommender Systems*, pp. 715–739. Springer, 2015.

von Ahn, L. Human computation. In *2008 IEEE 24th International Conference on Data Engineering*, pp. 1–2, April 2008.

Wang, Mian, Kawamura, Takahiro, Sei, Yuichi, Nakagawa, Hiroyuki, Tahara, Yasuyuki, and Ohsuga, Akihiko. Context-aware music recommendation with serendipity using semantic relations. In *Revised Selected Papers of the Third Joint International Conference on Semantic Technology - Volume 8388*, JIST 2013, pp. 17–32, New York, NY, USA, 2014. Springer-Verlag New York, Inc.

Wang, Xinxi, Rosenblum, David, and Wang, Ye. Context-aware mobile music recommendation for daily activities. In *Proceedings of the 20th ACM International Conference on Multimedia*, MM '12, pp. 99–108, New York, NY, USA, 2012. ACM.

Yu, Yi, Shen, Zhijie, and Zimmermann, Roger. Automatic music soundtrack generation for outdoor videos from contextual sensor information. In *Proceedings of the 20th ACM International Conference on Multimedia*, MM '12, pp. 1377–1378, New York, NY, USA, 2012. ACM.