Moodplay: Interactive Mood-based Music Discovery and Recommendation

Ivana Andjelkovic
Media Arts and Technology
University of California, Santa Barbara
ivana@mat.ucsb.edu

Denis Parra
CS Department
Pontificia Universidad Catolica de Chile
dparra@ing.puc.cl

John O'Donovan
Dept. of Computer Science
University of California, Santa Barbara
jod@cs.ucsb.edu

ABSTRACT

A large body of research in recommender systems focuses on optimizing prediction and ranking. However, recent work has highlighted the importance of other aspects of the recommendations, including transparency, control and user experience in general. Building on these aspects, we introduce *MoodPlay*, a hybrid recommender system music which integrates content and mood-based filtering in an interactive interface. We show how *MoodPlay* allows the user to explore a music collection by latent affective dimensions, and we explain how to integrate user input at recommendation time with predictions based on a pre-existing user profile. Results of a user study (N=240) are discussed, with four conditions being evaluated with varying degrees of visualization, interaction and control. Results show that visualization and interaction in a latent space improve acceptance and understanding of both metadata and item recommendations. However, too much of either can result in cognitive overload and a negative impact on user experience.

1. INTRODUCTION

Recommender systems have become invaluable tools for helping users find useful information online. There are well established algorithms, such as Collaborative and Content-Based Filters and Matrix Factorization, used across a variety of domains to recommend digital content or merchandise. Due to its unique consumption characteristics, music falls into a domain where alternative approaches to the traditional recommendation problem can help. For instance, we can listen to the same track several times without decreasing satisfaction. Compared to other domains (e.g. movies), the consumption of music is fast and more context dependent. In this paper, we focus on building an interactive recommender system that suggests music artists based on contextual information. We present interaction mechanisms that allow the user to guide the system based on affective state, which in turn adapts to changes of listening context. There are several music recommender systems that employ different types of context (daily activity [48], time of the day [3], music genre [25], etc.). However, no previous work has integrated affective context for music discovery into a visual and interactive recommendation system. Throughout the paper, we use a broad term *affect* to refer to both mood and emotion. Moods, being more permanent and less intense than emotions, are commonly used in recommendation research as tags to describe music. On the other hand, most psychology models, including the one used in our system, focus on emotions. Experimental evidence shows a strong relation between emotion and music [22] and previous research in affect-based recommender systems produced improvements over their non-contextual alternatives [13, 44]. Furthermore, the importance of building interactive recommender interfaces that go beyond the static ranked list paradigm to improve user satisfaction with a system has been studied in the past [12, 16, 5, 21, 47, 34, 29] and it is supported by results showing that small improvements in accuracy do not always correlate with better user satisfaction [28, 23]. Our goal is to build a recommender system with an interactive interface that supports discovery of unknown, interesting items via interaction in an affective space. We frame our work around the following research questions: How can metadata such as affective information be visually represented for a recommender system? How can interaction, explanation and control be supported over such a visualization? What are the effects of such interactive visualizations on the user experience with a recommender system, and what is the right amount of interaction? In our effort to answer these questions, we have produced the following key contributions:

- **A novel visual interface for recommendation.** A visualization that maps moods and music artists in the same latent space, supporting item exploration and user control.
- **Affect-based recommendation method.** A novel, hybrid recommendation algorithm for affect-based and audio content-based music recommendation.
- **Enhanced interaction techniques.** We introduce several new interaction mechanisms for hybrid recommendation in a latent space. For instance, trail-based and radius-based techniques.
- **Structural model for interaction tasks.** We present an evaluation of the system through an online experiment (N=240). Empirical results show interesting relations between user interaction, trust, and user perception, summarized in a structural model. We propose further research on interface design for exploratory tasks in recommender systems.

UMAP '16, July 13 - 17, 2016, Halifax, NS, Canada

ISBN 978-1-4503-4370-1/16/07. . . $15.00

DOI: http://dx.doi.org/10.1145/2930238.2930280

2. RELATED WORK

Visual Approaches to Recommendations. MacNee *et al.* [28] and Konstan *et al.* [23], highlight the need for more user-centric research in recommendations, since small improvements in recommender accuracy do not necessarily improve users' satisfaction. However, research on visual and interactive interfaces has started to grow only in recent years. Examples include visualizations of music and work-related online communities – SFViz [15], [52], collaborative filtering recommenders with rich user interactions such

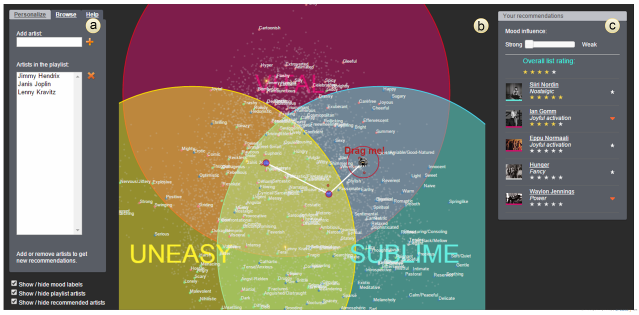

Figure 1: Screenshot of the MoodPlay interface, divided into three sections: (a) pane for entering artist names, (b) latent mood space visualization, (c) recommendation list, along with slider for adjusting mood influence

as PeerChooser [31] and SmallWorlds [16], and interactive visualizations for recommending conference talks – TalkExplorer [47] and SetFusion [34]. There is also a range of systems that support dynamic critiquing of an algorithm, such as Pu *et al.* [37] and Chen *et al.* [8]. For a detailed review of visual and interactive recommender systems, read Chen *et al.* survey [19]. To the best of our knowledge, *Moodplay* is the first interactive music recommender system that maps the artists in a latent, navigable, affective space.

Affect-based Recommendations. The important role of emotions in human decision-making [36, 30] has made affect an actively studied variable in context-aware recommender systems. For instance, Masthoff *et al.* [26] integrated affective state in a group recommender, while González *et al.* [14] incorporated the emotional context in a recommender for a large e-commerce learning guide. More relevant to our work, Park *et al.* [32] developed a music recommender that uses mood inferred from context information. More recently, Tkalcic *et al.* [45, 17] introduced a framework to identify the stages where emotion can be used for recommendation. *MoodPlay* models the user profile based on a set of artists, represents it in an affective latent space derived from GEMS model [51] and uses it to recommend new artists.

Recommendation of music artists. Recommendations in the music domain include approaches to recommend tracks [7, 24], albums [33], playlists [25, 2, 18] and artists [5, 20]. Particularly relevant to our aim at recommonding artists, Hijikata *et al.* [20] used a Naive Bayes recommender to recommend artists, while Bostandjev *et al.* [5] proposed a hybrid recommender system with a visual interactive interface – TasteWeights. Compared to the previous research in the field, we innovate by using artists' affective representation to compute similarity within a user-controllable recommendation interface.

Affect-based Visualizations of Music Collections. Russell's circumplex model of affect [39], which represents emotions and moods as a mixture of valence and arousal, is the most popular model used in affect-based visualizations. Yang *et al.* [49] incorporated it into their music retrieval method, and commercial applications such as Habu [43] and Musicovery [6] use it as a platform for music selection. However, many emotions cannot be uniquely characterized by valence and arousal values [9] and models derived from general research in psychology may not be suitable for musical emotions [50]. To address this issue, we propose a visual representation of music-specific affective dimensions, employing the hierarchical classification of emotions in the GEMS model by Zentner *et al.* [51].

Category	Sub-category	No. of moods	Example moods
Sublimity	Tenderness	24	Delicate, romantic, sweet
	Peacefulness	22	Pastoral, relaxed, soothing
	Wonder	24	Happy, light, springlike
	Nostalgic	9	Dreamy, rustic, yearning
	Transcendence	10	Atmospheric, spiritual, uplifting
Vitality	Power	29	Ambitious, fierce, pulsing, intense
	Joyful activation	32	Animated, fun, playful, exciting
Unease	Tension	32	Nervous, harsh, rowdy, rebellious
	Sadness	18	Austere, bittersweet, gloomy, tragic
	Fear *	10	Spooky, nihilistic, ominous
	Lethargy *	8	Languid, druggy, hypnotic
	Repulsiveness *	10	Greasy, sleazy, trashy, irreverent
Other *	Stylistic *	19	Graceful, slick, elegant, elaborate
	Cerebral *	12	Detached, street-smart, ironic
	Mechanical *	7	Crunchy, complex, knotty

Table 1: Structure and description of *MoodPlay* mood hierarchy. Categories and sub-categories marked with * are the expansions from the original GEMS model.

3. THE MOODPLAY SYSTEM

MoodPlay is accessible via web browser and its user interface consists of three sections: input, visualization and a recommendation panel. Users construct profiles by entering names of artists via an interactive drop-down list (Figure 1a). Based on the mood information associated with profile artists, the system positions a user avatar in a precomputed latent mood space (Figure 1b) and recommends new artists (Figure 1c).

Our interface design follows Shneiderman's visual information seeking mantra [41] by providing an overview, allowing zooming and panning the visualization, allowing filtering based on mood categories and toggling the visibility of details[1]. Recommendations are displayed as a ranked list in the right panel and the corresponding artist nodes are highlighted within the visualization. Upon clicking on names of the recommended artists, users are redirected to Last.fm where they can listen to the recommended music. They can also provide feedback by clicking on the stars below artist names.

Adaptivity of music recommenders is particularly important due to the dynamic nature of the listening context [42]. Thus, we model the change of a user's preference by enabling the movement of the avatar within visualization and maintaining an editable array of trail marks, weighted by distance from the current position (Figure 1.b). Finally, our recommendation approach accounts for the fact that mood-based similarity between artists does not necessarily match audio based similarity. Therefore, we allow users to adjust the mood influence via a slider control (Figure 1c) which dynamically re-sizes a catchment area around the current avatar position. The weaker the mood influence, the more we rely on audio similarity to calculate recommendations, and vice-versa.

3.1 A Visual Model of Affect

In order to show the relation between artists and moods in a two-dimensional space, we collected and analyzed Rovi mood metadata [10] for 4,927 artists. Each artist in our dataset is characterized by between 5-20 weighted moods out of 289, and represented with a vector $X \in \mathbb{R}^{289}$. Correspondence analysis [40] was used to reduce dimensionality to the 2D layout shown in Figure 1.

For the purpose of identifying potential clusters in our mood space, we explored whether our visual map fits into the hierarchical, music-specific emotion model proposed by Zentner *et al.* [51] - GEMS. This model consists of 3 main categories (vitality, uneasiness, sublimity), 9 sub-categories and 45 music relevant emotion

[1] Details in public video https://youtu.be/vH9q5ku8ocM

words distributed across different sub-categories. To perform our hierarchical classification of moods, we employed a WordNet [46] similarity tool [35] to calculate similarity scores between 289 Rovi and 45 GEMS mood words. The following steps were taken to reduce the observed classification error rate: (1) we created new mood categories to accommodate moods that do not belong to any of the GEMS categories, (2) 23 of the least frequently used mood tags in Rovi were discarded. Once the moods were classified, three clusters emerged in the 2D mood space. For visual explanation, we color each mood node according to the category it belongs to - vital moods are red, uneasy are green and sublime are blue, and we overlay Venn diagrams over the clusters.

Dataset. MoodPlay relies on a static music dataset of 4,927 artists, partially obtained randomly from Million Songs Dataset [4] and expanded by popular artists from the public EchoNest database [11] using proprietary metrics *familiarity* and *hotttness*. Mood data for each artist was obtained via Rovi API and the top ten most popular songs for each artist along with corresponding audio analysis data were obtained from EchoNest. Finally, artists in the recommendation list are linked to their external profile on Last.fm, where users can listen to artist songs.

3.2 Generating Recommendations

Recommendations are generated by the following three steps:

Offline computation of artist similarity. Artists' pairwise similarity, based on mood and audio content, is calculated offline and stored in two separate data structures. Mood-based similarity between any two artists is a function of their Euclidean distance in the affective space produced by correspondence analysis. To calculate audio-based similarity, we first identify the 10 most popular songs for each artist in our database via the EchoNest API and obtain audio analysis data for the the total of 49,270 songs from the same source. Following the approach by McFee et al. [27], obtained audio analysis data contains timbre, tempo, loudness and key confidence attributes, which are used to represent each song with a vector $v_i \in \mathbb{R}^{515}$. Finally, an accelerated approach for nearest-neighbor retrieval that uses maximum-variance KD-tree data structure was used to compute similarity between songs, since it is has a good balance of accuracy, scale and efficiency.

Online recommendation. During a user session, MoodPlay recommends new artists similar to the artists the user enters into her profile. First, we determine the overall mood by calculating the centroid of profile artist positions as a mean along x and y axes, where we then place the user avatar. Artists found within the adjustable radius around the centroid are all potential candidates for recommendation because they are considered to reflect the latent moods derived from the user's input. Among the candidate artists, we select the ten most similar to the user profile based on pre-computed audio similarity data, rank them by distance from user position and display first five as recommended artists.

Trail-based recommendation. In this novel, adaptive recommendation approach, users are allowed to move in the affective space while we keep track of each new position and apply a decay function to the preference trail when recommending new artists. Recommendations from the last position in the trail are assigned the greatest weight, because we presume that the most recent mood area of interest is the most relevant to user. The weights further decrease as a function of hop distance from the end of the trail. At each trail mark, we apply the recommendation algorithm described in the previous sub-sections, which produces an initial set of recommendation candidates. We then calculate adjusted distances d_a between each trail mark and corresponding recommendation candidates in the following way. First, we normalize distances between the trail mark and artists because radius can vary among trail marks. If the distances were not normalized, many relevant artists would be falsely considered irrelevant and would not appear in the final recommendation list. Next, we adjust the normalized distances for each trail mark based on the corresponding weights using the formula $d_a = d_n + \Delta \times (|T| - 1i)$, where d_n is a normalized distance, Δ is a decay constant, $|T|$ is a total number of trail marks and i is an iterator over the trail marks. After several tests, we found that weight constant Δ performs the best when calculated as: $\Delta = r_{min}/4$, where r_{min} is the minimal recommendation radius. The larger the value of Δ, the steeper the decay function is. Finally, the recommendation candidates are sorted based on adjusted distances, and top five are recommended to user.

Feature	(1)	(2)	(3)	(4)
Profile generation	x	x	x	x
Ordered list of recommendations	x	x	x	x
Display of latent mood space		x	x	x
Navigation in latent mood space			x	x
Hybridization control			x	x
Trail based recommendations				x
Number of subjects	68	60	51	61

Table 2: Availability of different features per experimental condition. Last row in the table shows the number of valid subjects in each condition.

4. EVALUATION

Evaluating recommender systems that contain interactive components is particularly challenging because of complex and potentially diverse interplay between the human participant and the automated algorithm. A crowdsourced study of 397 users was performed to allow for analysis of different interaction patterns with the MoodPlay recommender. After filtering out users we did not deem as valid, i.e., those who incorrectly answered attention check questions, 240 valid sessions remained.

Experimental Conditions. To understand the effects of mood-based interactions with a recommendation algorithm and to independently evaluate the influence of the MoodPlay visualization from an explanatory perspective, four conditions were tested, as shown in Table 2. The conditions have increasing visual and interaction complexity. Conditions (1) and (2) are based on a preexisting user profile while conditions (3) and (4) also allow for user input to the algorithm at recommendation time through interaction with the latent affective visualization. Figure 1 shows the full system, as tested in condition 4.

Participants. MTurk participants were paid a fixed amount of \$1.30 per study. Ages ranged from 18 to 65 with an average range of 25-30. 57% were male. When asked about music tastes, 80% said they listen to music frequently. Reported use of streaming services such as Pandora was normally distributed.

Procedure and Rating Collection. Participants accepted the study on MTurk and were redirected to a Qualtrics [38] pre-study with demographic and propensity related questions. Following this, they were randomly assigned a condition and performed the main task. Finally, participants gave qualitative feedback in a post study, also administered through the Qualtrics platform. During the main task, participant were asked to enter at least three profile items from a drop-down list, shown on the left in Figure 1. In all conditions, this profile was used to generate a list of 5 recommendations, that were shown on the right side of the screen. Ratings were collected for 5 items in an initial recommendation list, based on the user profile. Participants were then allowed to interact freely with the system to generate as many intermediate recommendation lists as they wished. For each of these lists, they were required to rate at least the top 2 items. Once satisfied, they again rated the full list of items for the final list. In addition to rating individual items, participants were required to provide an overall list rating for each list that was generated.

5. RESULTS AND DISCUSSION

Now that we have described the study design and setup, we present results in three areas. First, we describe a user interaction analysis, followed by a more holistic system evaluation using a structural equation model.

Limitations. During the study setup, a computational error was made during the indexing of artists and their positioning in the mood space. This resulted in a number of the artists being assigned to incorrect mood meta-data. In particular the error affected 37% of the artists significantly. The consequence of this error was that the first step in the hybrid recommendation phase –prediction of artists with similar mood, contained some noise. However, the second step, which is based on audio content features, was unaffected by the error. Accordingly, we focus our evaluation on user characteristics, interaction with the interface and experience, and place less attention on ratings-based analyses. A follow-up experiment is underway with a corrected model to assess these aspects in detail.

User Interaction. Conditions (1) to (4) in this experiment have increasing visual and interactive complexity. In order to understand the cost of observed differences in rating accuracy or user experience, an analysis of the time spent in the recommendation session was performed for each condition. In conditions (1) and (2), sessions lasted about 6 minutes on average, while in conditions (3) and (4), sessions averaged about 8 minutes. As expected, more time was spent in the interactive conditions (3 and 4). However, an interesting result was that from these two, people spent less time in the trail-based condition. While we do not have a significant result on rating accuracy, we did observe a trend towards higher ratings in condition 3. This will be examined more closely in the followup study.

Cognitive Load. Albers [1] states that learning new system interactions requires additional work and remembering, and users prefer to optimize their cognitive resources. In our study, we observed an effect relating the novel interactive features introduced and user's perception of understanding them. Condition (1) was perceived, as expected, significantly less confusing than the other three conditions ($p < 0.05$ in all 3 cases). Now, with respect to the agreement with the question *The system helped me understand and compare moods of different artists*, we conducted non-parametric Wilcoxon tests and we found that the agreement with the statement is close to significantly larger in condition (2) than in condition (1), W = 2389, $p = .019$, α-level= 0.0167 (we tested three hypotheses, so the original $\alpha = .05$ becomes $\frac{\alpha}{3} = .0167$). These results, although not conclusive, are indicative that the *Moodplay* visualization increases user understanding, but additional interactions (avatar, trails, etc.) might promote too much cognitive strain and they should be adjusted.

Structural Model. Since the MoodPlay system combines a recommendation algorithm, an interactive interface and subjective experiences of participants in the experiment, there are many variables that interact with each other. To study these interactions, several structural equation models [21] were tested over the personal characteristics of users from the pre-study; objective system aspects that were controlled in each condition; subjective aspects from the post study questionnaires and observed dependent variables from analysis of the system log data. Figure 2 shows the result of one such model with a reasonable fit to the data ($X^2(240) = 190, p < 0.05$). In this representation, edge thickness highlights the stronger effect sizes and values can be positive or negative, indicating effect direction. Notably, trust (both propensity and perceptive trust) plays an important role in how users perceive and understand recommendations. Visualization of the latent space causes an improvement in perceived accuracy. Gender influences degree of interaction, while participant age was more likely to influence the total time spent in the system, with older people spending more time on their interactions.

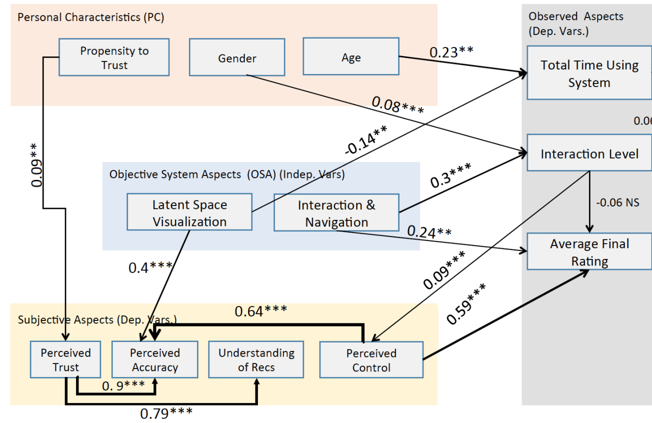

Figure 2: Structural equation model for variables in the experimental data, computed using Onyx. Significance levels are '*' p<.001, '**' p<.01, 'ns' p>.05. All factors in the model have been scaled to have a standard deviation of 1. Arrows are directed and edge values represent β co-efficients of the effect.**

6. CONCLUSION & FUTURE WORK

In this work we presented and evaluated *MoodPlay* –a hybrid recommender system for musical artists which introduces a novel latent space visualization based on mood tags. The system supports explanation and control of affective data through an interactive interface, and these data are applied to a user-controlled hybrid recommendation algorithm. Design and implementation of an online experiment (N=240) was presented to evaluate the *MoodPlay* system. Key findings include that participants generally liked exploring moods in the interactive latent space. Our study indicates some relation between level of interaction and cognitive strain. For example, introducing recommendation trails in the system might have produced a drop in several user experience metrics, this observed result opens an avenue for further research. In future work we will explore further the relation among personal characteristics, user perception and interaction with the ratings provided to the recommended suggestions. We will also investigate how different levels of interaction impact performance and cognitive load in information filtering tasks.

7. ACKNOWLEDGEMENTS

The author Denis Parra was supported by CONICYT, project FONDECYT 11150783. This work was partially supported by the U.S. Army Research Laboratory under Cooperative Agreement No. W911NF-09-2-0053; The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of ARL, NSF, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

8. REFERENCES

[1] M. J. Albers. Cognitive strain as a factor in effective document design. In *Proceedings of the 15th Annual International Conference on Computer Documentation*, SIGDOC '97, pages 1–6, New York, NY, USA, 1997. ACM.

[2] C. Baccigalupo and E. Plaza. Case-based sequential ordering of songs for playlist recommendation. In *Advances in Case-Based Reasoning*, pages 286–300. Springer, 2006.

[3] L. Baltrunas and X. Amatriain. Towards time-dependant recommendation based on implicit feedback. In *Workshop on context-aware recommender systems (CARSâĂŹ09)*, 2009.

[4] T. Bertin-Mahieux, D. P. Ellis, B. Whitman, and P. Lamere. The million song dataset. In *Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR 2011)*, 2011.

[5] S. Bostandjiev, J. O'Donovan, and T. Höllerer. Tasteweights: a visual interactive hybrid recommender system. In *Proceedings of the sixth ACM conference on Recommender systems*, pages 35–42. ACM, 2012.

[6] V. Castaignet and F. Vavrille. About musicovery. [Online; accessed 10-May-2015].

[7] O. Celma and P. Herrera. A new approach to evaluating novel recommendations. In *Proceedings of the 2008 ACM Conference on Recommender Systems*, RecSys '08, pages 179–186, New York, NY, USA, 2008. ACM.

[8] L. Chen and P. Pu. Interaction design guidelines on critiquing-based recommender systems. *User Modeling and User-Adapted Interaction*, 19(3):167–206, 2009.

[9] G. L. Collier. Beyond valence and activity in the emotional connotations of music. *Psychology of Music*, 35(1):110–131, 2007.

[10] R. Corp. *Rovi API*. http://developer.rovicorp.com/docs.

[11] EchoNest. *EchoNest API*. http://developer.echonest.com/docs/v4.

[12] B. Faltings, P. Pu, M. Torrens, and P. Viappiani. Designing example-critiquing interaction. In *Proceedings of the 9th international conference on Intelligent user interfaces*, pages 22–29. ACM, 2004.

[13] I. Fernández-Tobías, I. Cantador, and L. Plaza. An emotion dimensional model based on social tags: Crossing folksonomies and enhancing recommendations. In *E-Commerce and Web Technologies*, pages 88–100. Springer, 2013.

[14] G. Gonzalez, J. L. De La Rosa, M. Montaner, and S. Delfin. Embedding emotional context in recommender systems. In *Data Engineering Workshop, 2007 IEEE 23rd International Conference on*, pages 845–852. IEEE, 2007.

[15] L. Gou, F. You, J. Guo, L. Wu, and X. L. Zhang. Sfviz: interest-based friends exploration and recommendation in social networks. In *Proceedings of the 2011 Visual Information Communication-International Symposium*, page 15. ACM, 2011.

[16] B. Gretarsson, J. O'Donovan, S. Bostandjiev, C. Hall, and T. Höllerer. Smallworlds: Visualizing social recommendations. In *Computer Graphics Forum*, volume 29, pages 833–842. Wiley Online Library, 2010.

[17] B.-j. Han, S. Rho, S. Jun, and E. Hwang. Music emotion classification and context-based music recommendation. *Multimedia Tools and Applications*, 47(3):433–460, 2010.

[18] N. Hariri, B. Mobasher, and R. Burke. Context-aware music recommendation based on latenttopic sequential patterns. In *Proceedings of the Sixth ACM Conference on Recommender Systems*, RecSys '12, pages 131–138, New York, NY, USA, 2012. ACM.

[19] C. He, D. Parra, and K. Verbert. Interactive recommender systems: a survey of the state of the art and future research challenges and opportunities. *Expert Systems with Applications*, 2016. in press.

[20] Y. Hijikata, Y. Kai, and S. Nishida. The relation between user intervention and user satisfaction for information recommendation. In *Proceedings of the 27th Annual ACM Symposium on Applied Computing*, pages 2002–2007. ACM, 2012.

[21] B. P. Knijnenburg, S. Bostandjiev, J. O'Donovan, and A. Kobsa. Inspectability and control in social recommenders. In *Proceedings of the sixth ACM conference on Recommender systems*, pages 43–50. ACM, 2012.

[22] S. Koelsch. A neuroscientific perspective on music therapy. *Annals of the New York Academy of Sciences*, 1169(1):374–384, 2009.

[23] J. A. Konstan and J. Riedl. Recommender systems: from algorithms to user experience. *User Modeling and User-Adapted Interaction*, 22(1-2):101–123, 2012.

[24] B. Logan. Music recommendation from song sets. In *ISMIR*, 2004.

[25] F. Maillet, D. Eck, G. Desjardins, P. Lamere, et al. Steerable playlist generation by learning song similarity from radio station playlists. In *ISMIR*, pages 345–350, 2009.

[26] J. Masthoff. The pursuit of satisfaction: affective state in group recommender systems. In *User Modeling 2005*, pages 297–306. Springer, 2005.

[27] B. McFee and G. R. G. Lanckriet. Large-scale music similarity search with spatial trees. In A. Klapuri and C. Leider, editors, *ISMIR*, pages 55–60. University of Miami, 2011.

[28] S. M. McNee, J. Riedl, and J. A. Konstan. Being accurate is not enough: how accuracy metrics have hurt recommender systems. In *CHI'06 extended abstracts on Human factors in computing systems*, pages 1097–1101. ACM, 2006.

[29] S. Nagulendra and J. Vassileva. Understanding and controlling the filter bubble through interactive visualization: A user study. In *Proceedings of the 25th ACM Conference on Hypertext and Social Media*, HT '14, pages 107–115, New York, NY, USA, 2014. ACM.

[30] K. Oatley, D. Keltner, and J. M. Jenkins. *Understanding emotions* . Blackwell publishing, 2006.

[31] J. O'Donovan, B. Smyth, B. Gretarsson, S. Bostandjiev, and T. Höllerer. Peerchooser: visual interactive recommendation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1085–1088. ACM, 2008.

[32] H.-S. Park, J.-O. Yoo, and S.-B. Cho. A context-aware music recommendation system using fuzzy bayesian networks with utility theory. In *Fuzzy systems and knowledge discovery*, pages 970–979. Springer, 2006.

[33] D. Parra and X. Amatriain. Walk the talk: Analyzing the relation between implicit and explicit feedback for preference elicitation. In *Proceedings of the 19th International Conference on User Modeling, Adaption, and Personalization*, UMAP'11, pages 255–268, Berlin, Heidelberg, 2011. Springer-Verlag.

[34] D. Parra, P. Brusilovsky, and C. Trattner. See what you want to see: Visual user-driven approach for hybrid recommendation. In *Proceedings of the 19th International Conference on Intelligent User Interfaces*, IUI '14, pages 235–240, New York, NY, USA, 2014. ACM.

[35] T. Pedersen and M. Jason. *WordNet::Similarity*. http://maraca.d.umn.edu/cgi-bin/similarity/similarity.cgi.

[36] R. W. Picard. *Affective computing*. MIT press, 2000.

[37] P. Pu, B. Faltings, L. Chen, J. Zhang, and P. Viappiani. Usability guidelines for product recommenders based on example critiquing research. In F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, editors, *Recommender Systems Handbook*, pages 511–545. Springer US, 2011.

[38] Qualtrics. *Qualtrics*. https://www.qualtrics.com.

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

[40] N. J. Salkind, editor. *Encyclopedia of Research Design*. SAGE Publications, Inc., 0 edition, 2010.

[41] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium on*, pages 336–343. IEEE, 1996.

[42] S. Stober and A. Nürnberger. Adaptive music retrieval—a state of the art. *Multimedia Tools Appl.*, 65(3):467–494, Aug. 2013.

[43] G. Team. Habu music. [Online; accessed 10-May-2015].

[44] M. Tkalčič, U. Burnik, and A. Košir. Using affective parameters in a content-based recommender system for images. *User Modeling and User-Adapted Interaction*, 20(4):279–311, 2010.

[45] M. Tkalcic, A. Kosir, and J. Tasic. Affective recommender systems: the role of emotions in recommender systems. In *Proc. The RecSys 2011 Workshop on Human Decision Making in Recommender Systems*, pages 9–13. Citeseer, 2011.

[46] P. University. *WordNet*. https://wordnet.princeton.edu.

[47] K. Verbert, D. Parra, P. Brusilovsky, and E. Duval. Visualizing recommendations to support exploration, transparency and controllability. In *Proceedings of the 2013 international conference on Intelligent user interfaces*, IUI '13, pages 351–362, New York, NY, USA, 2013. ACM.

[48] X. Wang, D. Rosenblum, and Y. Wang. Context-aware mobile music recommendation for daily activities. In *Proceedings of the 20th ACM international conference on Multimedia*, pages 99–108. ACM, 2012.

[49] Y.-H. Yang, Y.-C. Lin, H. T. Cheng, and H. H. Chen. Mr. emo: music retrieval in the emotion plane. In A. El-Saddik, S. Vuong, C. Griwodz, A. D. Bimbo, K. S. Candan, and A. Jaimes, editors, *ACM Multimedia*, pages 1003–1004. ACM, 2008.

[50] M. Zentner and T. EEROLA. Self-report measures and models. *Handbook of Music and Emotion: Theory, Research, Applications*, 2011.

[51] M. Zentner, D. Grandjean, and K. R. Scherer. Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, 8(4):494–521, 2008.

[52] S. Zhao, M. X. Zhou, X. Zhang, Q. Yuan, W. Zheng, and R. Fu. Who is doing what and when: Social map-based recommendation for content-centric social web sites. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 3(1):5, 2011.