Affective recommender systems: the role of emotions in recommender systems

Marko Tkalčič
University of Ljubljana Faculty of electrical engineering
Tržaška 25, Ljubljana, Slovenia
marko.tkalcic@fe.uni-lj.si

Andrej Košir
University of Ljubljana Faculty of electrical engineering
Tržaška 25, Ljubljana, Slovenia
andrej.kosir@fe.uni-lj.si

Jurij Tasič
University of Ljubljana Faculty of electrical engineering
Tržaška 25, Ljubljana, Slovenia
jurij.tasic@fe.uni-lj.si

ABSTRACT

Recommender systems have traditionally relied on data-centric descriptors for content and user modeling. In recent years we have witnessed an increasing number of attempts to use emotions in different ways to improve the quality of recommender systems. In this paper we introduce a unifying framework that positions the research work, that has been done so far in a scattered manner, in a three stage model. We provide examples of research that cover various aspects of the detection of emotions and the inclusion of emotions into recommender systems.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*

Keywords

recommender systems, emotions

1. INTRODUCTION

In the pursuit of increasing the accuracy of recommender systems, researchers started to turn to more user-centric content descriptors in recent years. The advances made in affective computing, especially in automatic emotion detection techniques, paved the way for the exploitation of emotions and personality as descriptors that account for a larger part of variance in user preferences than the generic descriptors (e.g. genre) used so far.

However, these research efforts have been conducted independently, stretched among the two major research areas, *recommender systems* and *affective computing*. In this paper we (i) survey the research work that helps improving recommender systems with affective information and (ii) we provide a unifying framework that will allow the members of the research community to identify the position of their activities and to benefit from each other's work.

Decisions@RecSys 2011, Chicago, USA

2. THE UNIFYING FRAMEWORK

When using applications with recommender systems the user is constantly receiving various stimuli (e.g. visual, auditory etc.) that induce emotive states. These emotions influence, at least partially (according to the bounded rationality model [16]) the user's decisions on which content to choose. Thus it is important for the recommender system application to detect and make good use of emotive information.

2.1 Describing emotions

There are two main approaches to describe the emotive state of a user: (i) the *universal emotions model* and (ii) the *dimensional model*. The universal emotions model assumes there is a limited set of distinct emotional categories. There is no unanimity as to which are the universal emotions, however, the categories proposed by Ekman [10] (i.e. happiness, anger, sadness, fear, disgust and surprise) appear to be very popular. The dimensional model, on the contrary, describes each emotion as a point in a continuous multidimensional space where each dimension represents a quality of the emotion. The dimensions that are used most frequently are valence, arousal and dominance (thus the VAD acronym) although some authors refer to these dimensions with different names (e.g. pleasure instead of valence in [20] or activation instead of arousal in [13]). The circumplex model, proposed by Posner et al. [24], maps the basic emotions into the VAD space (as depicted in Fig. 1)

2.2 The role of emotions in the consumption chain

During the user interaction with a recommender system and the content consumption that follows, emotions play different roles in different stages of the process. We divided the user interaction process in three stages, based on the role that emotions play (as shown in Fig. 2): (i) *the entry stage*, (ii) *the consumption stage* and (iii) *the exit stage*.

The work surveyed in this paper can be divided in two main categories: (i) generic emotion detection algorithms (that can be used in all three stages) and (ii) usage of emotion parameters in the various stages. This paper does not aim at providing an overall survey of related work but rather to point out good examples of how to address various aspects of recommender systems with the usage of techniques

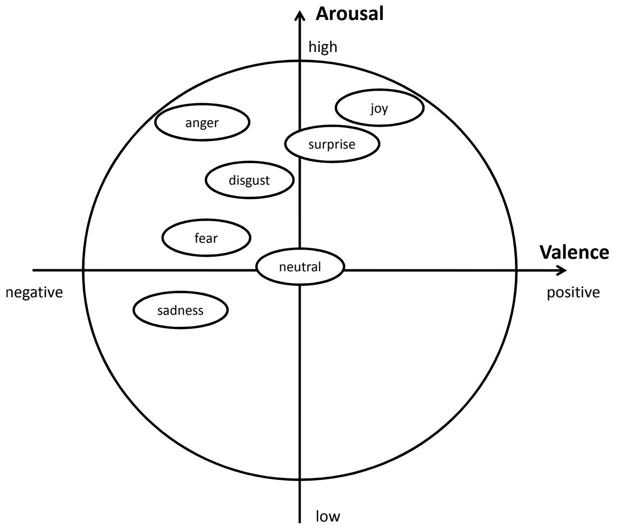

Figure 1: Basic emotions in the valence-arousal plane of the dimensional model

borrowed from affective computing.

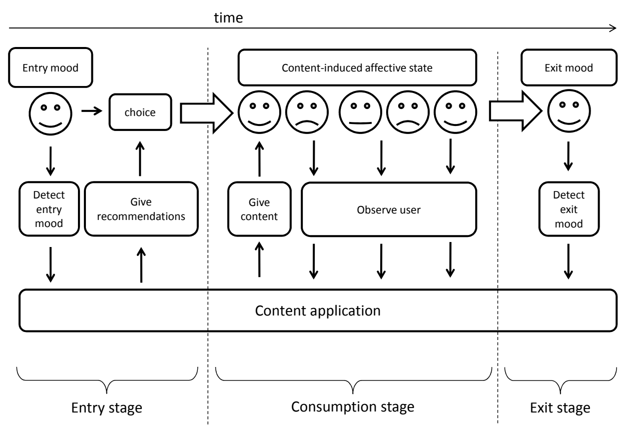

Figure 2: The unifying framework: the role of emotions in user interaction with a recommender system.

In the remainder of the paper we address each stage separately by surveying the existing research work and providing lists of open research areas. At the end we discuss the proposed framework and give the final conclusions.

3. DETECTING AFFECTIVE STATES

Affective states of end users (in any stage of the proposed interaction chain) can be detected in two ways: (i) explicitly or (ii) implicitly. The implicit detection of emotions is more accurate but it's an intrusive process that breaks the interaction. The implicit approach is less accurate but it's well suited for user interaction purposes since the user is not aware of it. Furthermore, Pantić et al. [22] argued that explicit acquisition of users' affect has further negative properties as users may have side-interests that drive their explicit affective labeling process (egoistic tagging, reputation-driven tagging or asocial tagging).

The most commonly used procedure for the explicit asessment of emotions is the Self Assessment Manikin (SAM) developed by [7]. It is a questionnaire where users assess their emotional state in the three dimensions: valence, arousal and dominance.

The implicit acquisition of emotions is usually done through a variety of modalities and sensors: video cameras, speech, EEG, ECG etc. These sensors measure various changes of the human body (e.g. facial changes, posture changes, changes in the skin conductance etc.) that are known to be related to specific emotions. For example, the Facial Action Coding System (FACS), proposed by Ekman [9], maps emotions to changes of facial characteristic poionts. There are excellent surveys on the topic of multimodal emotion detection: [31, 22, 14]. In general, raw data is acquired from one or more sensors during the user interaction. These signals are processed to extract some low level features (e.g. Gabor based features are popular in the processing of facial expression video signals). Then some kind of classification or regression technique is applied to yield distinct emotional classes or continuous values. The accuracy of emotion detection ranges from over 90% on posed datasets (like the Kanade-Cohn dataset [18]) to slightly better than coin tossing on spontaneous datasets (like the LDOS-PerAff-1 dataset [29]) [27, 6].

4. ENTRY STAGE

The first part of the proposed framework (see Fig. 2) is the entry stage. When a user starts to use a recommender system, she is in an affective state, the *entry mood*. The entry mood is caused by some previous user's activities, unknown to the system. When the recommender system suggests a limited amount of content items to the user, the entry mood influences the user's choice. In fact, the user's decision making process depends on two types of cognitive processes, the rational and the intuitive, the latter being strongly influenced by the emotive state of the user, as explained by the bounded rationality paradigm [16]. For example, a user might want to consume a different type of content when she is happy than when she is sad. In order to adapt the list of recommended items to the user's entry mood the system must be able to detect the mood and to use it in the content filtering algorithm as contextual information.

In the entry part of user-RS interaction one of the aspects where emotions can be exploited is to influence the user's choice. Creed [8] explored how the way we represent information influences the user's choices.

It has been observed by Porayska-Pomsta et al. [23] that in tutoring systems there is a strong relation between the entry mood and learning. They analysed the actions that a human tutor took when the student showed signs of specific affective states to improve the effectiveness of an interactive learning environment.

A user modeling approach that maps a touristic attraction with a piece of music that induces a related emotion has been developed by Kaminskas and Ricci [17]. Their goal was to find an appropriate musical score that would reinforce the affective state induced by the touristic attraction.

Using the entry mood as a contextual parameter (as described by Adomavicius and Tuzhilin in [2]) could improve the recommender's performance. Both Koren et al. [19] and Baltrunas et al. [5] suggest using the matrix factorization approach and enrich it with contextual parameters. At the context-aware recommender systems contest in 2010 [1] the goal was to select a number of movies that would fit the user's entry mood. The contest winners' contribution, Shi et al. [25] used several approaches amog which the best was the joint matrix factorization model with a mood-specific regularization.

As an extension to the usage of emotions as contextual information an interesting research area is to diversify the recommendations. For example, if a user is sad, would it be better to recommend happy content to cheer her up or to recommend sad content to be in line with the current mood? Research on information retrieval results diversification is getting increased attention, especially after the criticism of the recommendation bubble has started[2]. Although we are not aware of any work done on results diversification connected with emotions, a fair amount of work has been done on political news aggregators in order to stimulate political pluralism [21].

5. CONSUMPTION STAGE

The second part of the proposed framework is the consumption stage (see Fig. 2). After the user starts with the consumption of the content she experiences affective responses that are induced by the content. Depending on the type of content, these responses can be (i) single values (e.g. the emotive response to watching an image) or (ii) a vector of emotions that change over time (e.g. while watching a movie or a sequence of images). Figure 3 shows how emotions change over time in the consumption stage. The automatic detection of emotions can help building emotive profiles of users and content items that can be exploited for content-based recommender algorithms.

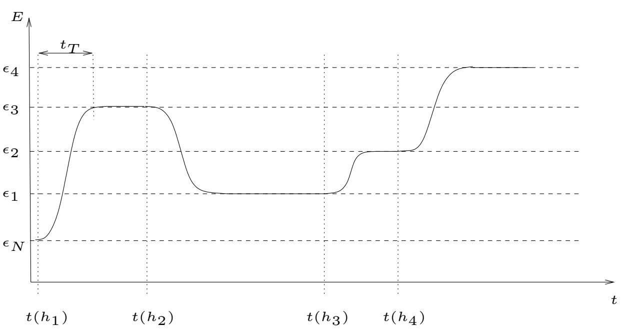

Figure 3: The user's emotional state ϵ is continuously changing as the time sequence of the visual stimuli $h_i \in H$ induce different emotions.

Using emotional responses for generating implicit affective tags for content is the main research area in the consumption section. Pantić et al. [22] argued why the usage of automatic emotion detection methods improves content tagging: the minimization of the drawbacks caused by egoistic tagging, reputation-driven tagging and asocial tagging. They also anticipate that implicit tagging can be used for user profiling in recommender systems.

Joho et al. [15] used emotion detection from facial expressions to provide an affective profile of video clips. They used an item profile structure that labels changes of users emotions through time relative to the video clip start. The authors used their approach for summarizing highlights of video clips.

Hanjalić et al. [11] approached the summarization of video highlights from the other side: they used the source's low level features (audio and video) to detect higlihjts without taking into account the responses of end users.

The research work described so far in this section is interesting because allows us to model the content items (images, movies, music etc.) with affective labels. These affective labels describe the emotions experienced by the users who consume the items. In our previous work [26] we have shown that the usage of such affective labels over generic labels (e.g. genre) significantly improves the performance of a content-based recommender system for images. We used explicitly acquired affective metadata to model the items and the users' preferences. However, in another experiment [28], where we used implicitly acquired affective metadata, the accuracy of the recommender system was significantly lower but still better than with generic metadata only.

In a similar experiment, Arapakis et al. [4] built a recommender system that uses real time emotion detection informaion.

6. EXIT STAGE

After the user has finished with the content consumption she is in what we call the exit mood. The main difference between the consumption stage and the exit stage is that the exit mood will influence the user's next actions, thus having an active part, while in the consumption stage the induced emotions did not influence any actions but were a passive response to the stimuli. In case that the user continues to use the recommender system the exit mood for the content just consumed is the entry mood for the next content to be consumed.

The automatic detection of the exit mood can be useful as an indicator of the user's satisfaction with the content. Thus the detection of the exit mood can be seen as an unobtrusive feedback collection technique.

Arapakis et al. [3] used the exit mood, detected through videos of users' facial expressions, as an implicit feedback in their recommender system for video sequences.

In an experiment with games, Yannakakis et al. [30], used heart rate activity to infer the "fun" that the subjects experience in physical interactive playgrounds.

7. OPEN RESEARCH AREAS

We identified four main areas where further research should be conducted in order to build true affective recommender systems: (i) using emotions as context in the entry stage, (ii) modeling affective content profiles, (iii) using affective profiles for recommending content and (iv) building a set of datasets.

Although some work has been carried out on exploiting the entry mood we believe that there is still the need to answer tha basic question of the entry stage: *which items to recommend when the user is in the emotive state A?*. We

[1] http://www.dai-labor.de/camra2010/

[2] http://www.thefilterbubble.com/

further believe that there are firm differences between what the user wants now and what is good for a user on a long run. Thus bringing the research on results diversification (see the work done in [21, 1]) into affective recommender systems is a highly important topic.

Affective content profiling is still an open question, especially profiling content items that last longer than a single emotive response. The time dependancy of content profiles has also o strong impact on the algorithms that exploit the profiles for recommending items.

With the except of the LDOS-PerAff-1 dataset [29] (which is limited in the amount of content items and users), the research community does not have a suitable dataset upon which to work. It is thus required that a large-scale dataset, compareable to the MovieLens or Netflix datasets, is built.

8. CONCLUSION

In this paper we have provided a framework that describes three ways in which emotions can be used to improve the quality of recommender systems. We also surveyed some work that deals with parts of the isuues that arise in te pursuit of affective recommender systems.

An important issue in recommender systems, especially when it comes to user-centric systems, is to move from data-centric assessment criteria to user-centred assessment criteria. We have not addressed this issue in this paper as it appears to larger dimensions. The recsys community has so far relied on metrics borrowed from information retrieval: confusion matrices, precision, recall etc. (see [12] for an overview). However recommender systems are used by end users and thus the assessment of the end users should be taken more into account. We suggest to move towards metrics that take into account the user experience as pointed out in `http://www.usabart.nl/portfolio/KnijnenburgWillemsen-UMUAI2011_UIRecSy.pdf`.

9. REFERENCES

[1] G. Adomavicius and Y. Kwon. Improving Aggregate Recommendation Diversity Using Ranking-Based Techniques. *IEEE Transactions on Knowledge and Data Engineering*, (99):1–1, 2011.

[2] G. Adomavicius, R. Sankaranarayanan, S. Sen, and A. Tuzhilin. Incorporating contextual information in recommender systems using a multidimensional approach. *ACM Transactions on Information Systems (TOIS)*, 23(1):103–145, 2005.

[3] I. Arapakis, J. Jose, and P. Gray. Affective feedback: an investigation into the role of emotions in the information seeking process. *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, (January):395–402, 2008.

[4] I. Arapakis, Y. Moshfeghi, H. Joho, R. Ren, D. Hannah, and J. M. Jose. Enriching user profiling with affective features for the improvement of a multimodal recommender system. *Proceeding of the ACM International Conference on Image and Video Retrieval - CIVR '09*, (i):1, 2009.

[5] L. Baltrunas. Exploiting contextual information in recommender systems. *Proceedings of the 2008 ACM conference on Recommender systems - RecSys '08*, page 295, 2008.

[6] M. S. Bartlett, G. C. Littlewort, M. G. Frank, C. Lainscsek, I. R. Fasel, and J. R. Movellan. Automatic Recognition of Facial Actions in Spontaneous Expressions. *Journal of Multimedia*, 1(6):22–35, Sept. 2006.

[7] M. Bradley and P. Lang. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59, 1994.

[8] C. Creed and R. Beale. Using emotion simulation to influence user attitudes and behavior. *HCI workshop at BCS 2005*, pages 1–3, 2005.

[9] P. Ekman. Facial expression and emotion. *American Psychologist*, 48(4):384, 1993.

[10] P. Ekman. Basic Emotions. In *Handbook of Cognition and Emotion*, pages 45—-60. 1999.

[11] A. Hanjalic. Adaptive extraction of highlights from a sport video based on excitement modeling. *IEEE Transactions on Multimedia*, 7(6):1114–1122, Dec. 2005.

[12] J. L. Herlocker, J. A. Konstan, L. Terveen, and J. T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst*, 22(1):5–53, 2004.

[13] S. V. Ioannou, A. T. Raouzaiou, V. a. Tzouvaras, T. P. Mailis, K. C. Karpouzis, and S. D. Kollias. Emotion recognition through facial expression analysis based on a neurofuzzy network. *Neural networks : the official journal of the International Neural Network Society*, 18(4):423–35, May 2005.

[14] A. Jaimes and N. Sebe. Multimodal human–computer interaction: A survey. *Computer Vision and Image Understanding*, 108(1-2):116–134, 2007.

[15] H. Joho, J. M. Jose, R. Valenti, and N. Sebe. Exploiting facial expressions for affective video summarisation. *Proceeding of the ACM International Conference on Image and Video Retrieval - CIVR '09*, page 1, 2009.

[16] D. Kahneman. A perspective on judgment and choice: mapping bounded rationality. *The American psychologist*, 58(9):697–720, Sept. 2003.

[17] M. Kaminskas and F. Ricci. Location-Adapted Music Recommendation Using Tags. *User Modeling, Adaption and Personalization*, pages 183–194, 2011.

[18] T. Kanade, J. Cohn, and Y. Tian. Comprehensive database for facial expression analysis. In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, pages 46–53. IEEE, 2000.

[19] Y. Koren. Collaborative filtering with temporal dynamics. *Communications of the ACM*, 53(4):89, Apr. 2010.

[20] A. Mehrabian. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in Temperament. *Current Psychology*, 14(4):261–292, Dec. 1996.

[21] S. Munson, D. X. Zhou, and P. Resnick. Sidelines: An algorithm for increasing diversity in news and opinion aggregators. *Proceedings of ICWSM09 Conference on Weblogs and Social Media. San Jose, CA.*, 2009.

[22] M. Pantic and A. Vinciarelli. Implicit human-centered tagging [Social Sciences. *IEEE Signal Processing*

Magazine, 26(6):173–180, Nov. 2009.

[23] K. Porayska-Pomsta, M. Mavrikis, and H. Pain. Diagnosing and acting on student affect: the tutor's perspective. *User Modeling and User-Adapted Interaction: The Journal of Personalization Research*, 18(1-2):125–173, 2007.

[24] J. Posner, J. a. Russell, and B. S. Peterson. The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, 17(3):715–34, Jan. 2005.

[25] Y. Shi, M. Larson, and A. Hanjalic. Mining mood-specific movie similarity with matrix factorization for context-aware recommendation. *Proceedings of the Workshop on Context-Aware Movie Recommendation*, pages 34–40, 2010.

[26] M. Tkalčič, U. Burnik, and A. Košir. Using affective parameters in a content-based recommender system for images. *User Modeling and User-Adapted Interaction: The Journal of Personalization Research*, pages 1–33–33, Sept. 2010.

[27] M. Tkalčič, A. Odić, A. Košir, and J. Tasič. Comparison of an Emotion Detection Technique on Posed and Spontaneous Datasets. *Proceedings of the 19th ERK conference, Portorož, 2010*, 2010.

[28] M. Tkalčič, A. Odić, A. Košir, and J. Tasič. Impact of Implicit and Explicit Affective Labeling on a Recommender System's Performance. *Joint Proceedings of the Workshop on Decision Making and Recommendation Acceptance Issues in Recommender Systems (DEMRA 2011) and the 2nd Workshop on User Models for Motivational Systems: The affective and the rational routes to persuasion (UMMS 2011)*, page 112, 2011.

[29] M. Tkalčič, J. Tasič, and A. Košir. The LDOS-PerAff-1 Corpus of Face Video Clips with Affective and Personality Metadata. *Proceedings of Multimodal Corpora: Advances in Capturing, Coding and Analyzing Multimodality (Malta, 2010), LREC*, page 111, 2009.

[30] G. N. Yannakakis, J. Hallam, and H. H. Lund. Entertainment capture through heart rate activity in physical interactive playgrounds. *User Modeling and User-Adapted Interaction*, 18(1-2):207–243, Sept. 2008.

[31] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang. A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. *IEEE Trans. Pattern Analysis & Machine Intelligence, Vol. 31, No.*, 1pp:39–58, 2009.