User Modeling, Adaptation, and Personalization: 21th International Conference, UMAP 2013, Rome, Italy, June 10-14, 2013 Proceedings. Lecture Notes in Computer Science, Volumen 7899. Springer, 2013. 331-334

DOI: http://dx.doi.org/10.1007/978-3-642-38844-6_30

Modeling Emotions with Social Tags

Ignacio Fernández-Tobías, Iván Cantador, Laura Plaza

Universidad Autónoma de Madrid, 28049 Madrid, Spain
{i.fernandez, ivan.cantador, laura.plaza}@uam.es

Abstract. We present an emotion model based on social tags, which is built upon an automatically generated lexicon that describes emotions by means of synonym and antonym terms. Using this model we develop a number of methods that transform social tag-based item profiles into emotion-oriented item profiles. We show that the model's representation of a number of basic emotions is in accordance with the well known psychological circumplex model of affect, and we report results from a user study that show a high precision of our methods to infer the emotions evoked by items in the movie and music domains.

Keywords: emotions, social tagging, folksonomies.

1 Introduction

Emotions are intense feelings that are directed at someone or something. In adaptive and personalized systems, emotions are usually considered as contextual signals that can lead to enhanced approaches in a wide array of applications, such as constructing user behavior models [2], tailoring search results [3], and recommending items [5], to name a few. Hence, modeling, capturing and exploiting emotions present challenging problems.

In this paper we focus on the emotion modeling task, and restrict our attention to situations where emotions are expressed in (and can be extracted from) text contents – such as reviews in blogs, and annotations in social tagging systems –, differently to e.g. situations where emotions are recognized in either the visual or auditory modalities.

2 An Emotion Lexicon

Among the existing dimensional models of emotion, the circumplex model [4] is a dominant one. It suggests that emotions are distributed in a two-dimensional circular space formed by two independent dimensions: *arousal* and *pleasure*. Figure 1a shows such distribution. Arousal represents the vertical axis and reflects the intensity of an emotion; and pleasure represents the horizontal axis and reflects whether an emotion is positive or negative. The center of the circle represents medium levels of arousal and pleasure. Any emotion can be represented at any level of arousal and pleasure, including a neutral level of one or both of such factors. The figure shows the distribution of 16 basic emotions. We also consider this set of emotions.

Our dimensional model is built upon an automatically generated lexicon $\mathcal{L} = \{\mathrm{t}_1, \ldots, \mathrm{t}_K\}$ composed of synonym and antonym terms t_k of the emotions' names – which are adjectives (e.g. *happy*, *sad*), as shown in Figure 1a. The synonym and antonym terms of each emotion's name are obtained from the online thesaurus provided by Dictionary.com (http://thesaurus.com).

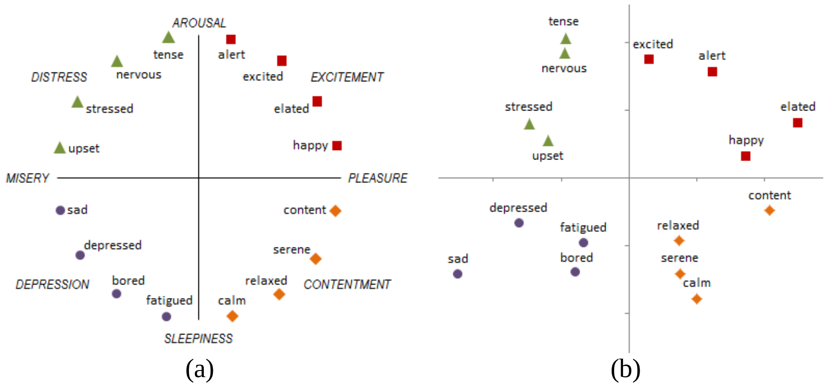

Figure 1. Two-dimensional distributions of basic emotions established in the circumplex model (1a) and automatically obtained in our tag-based model (1b).

Once the lexicon $\mathcal{L}$ is generated, an emotion $\mathrm{e}_i \in \mathcal{E}$ is represented as a vector $\mathbf{e}_i = (e_{i,1}, \dots, e_{i,K}) \in \mathbb{R}^K$, in which the component $e_{i,k}$ corresponds to the term $\mathrm{t}_k \in \mathcal{L}$, and is a numeric value defined as:

$$e_{i,k} = \begin{cases} tfidf(\mathrm{t_k}, \mathrm{e_i}) & \text{if } t_k \in synonyms(\mathrm{e_i}) \\ -tfidf(\mathrm{t_k}, \mathrm{e_i}) & \text{if } t_k \in antonyms(\mathrm{e_i}) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The component $e_{i,k}$ is greater than 0 if the term t_k is a synonym of the emotion e_i, lower than 0 if t_k is an antonym of e_i, and 0 otherwise. Its absolute value corresponds to the TF-IDF weight of t_k computed by considering the lexicon $\mathcal{L}$ as the *collection vocabulary*, and the set $\mathcal{E}$ of emotions (described as sets of synonym and antonym terms) as the *collection documents*. With the proposed vector representation, we can measure (dis)similarities between emotions. Specifically, we can use the well known cosine similarity $sim(\mathrm{e_i}, \mathrm{e_j}) = \cos(\mathbf{e}_i, \mathbf{e}_j)$.[1]

To validate the correspondences between our computational model and the theoretic circumplex model, Figure 1b shows the projections of the emotion vectors into a two-dimensional space by applying Principal Component Analysis. We can see that our model locates all the 16 basic emotions in their corresponding quadrants. More interestingly, in our model the axes defined by the two most informative components are related to the *arousal* and *pleasure* factors of the circumplex model.

3 Emotion-oriented Tag-based Profiles

The proposed representation of emotions lets transform social tag-based item profiles (i.e., the items' annotation sets) into emotion-oriented profiles. Formally, let an emotion $e_i \in \mathcal{E}$ be defined as in formula (1). For an item (object) o_n, let $\mathbf{o}_n = (o_{n,1}, \dots, o_{n,|\mathcal{T}|}) \in \mathbb{R}^{|\mathcal{T}|}$ be the item's *tag-based profile*, where $o_{n,i}$ corresponds to the tag $\mathrm{t}_i \in \mathcal{T}$ of the item's folksonomy. Then, from such profile, we define the item's *emotion-oriented profile* as $\mathbf{q}_n = (q_{n,1}, \dots, q_{n,|\mathcal{E}|}) \in [-1,1]^{|\mathcal{E}|}$, where the *i*-th component corresponds to the emotion $\mathrm{e}_i^C \in \mathcal{E}$, and its weight is computed as $q_{n,i} = cos(\mathbf{o}_n, \mathbf{e}_i)$.

Moreover, for each emotion-oriented profile, we consider two alternatives for defining the emotion vectors $\mathbf{e}_i$: *basic vectors*, whose components correspond to terms of the lexicon, as defined in formula (1), and *extended_N* vectors, whose components correspond to the *N* folksonomy tags that cooccur most frequently (in the tag-based item profiles) with the terms of the basic vectors. These tags are not necessarily synonyms/ antonyms of the lexicon terms, and it is not clear whether they can be valuable.

4 Experiments

We conducted a user study in which participants, recruited via social networking sites, were presented with sets of movies or musicians (no combinations of both) from the HetRec'11 social tagging datasets [1]. They were requested to freely state which emotions they considered as relevant for each item (movie or musician), thus manually (and collectively) creating emotion-oriented item profiles, which we considered as ground truth. A total of 72 users participated, evaluating 178 movies and 132 musicians. They generated 713 evaluation cases, assigning an average of 3.30 and 4.18 emotions to items in the movies and music domains, respectively.[1]

To evaluate the quality of the emotion-oriented profiles generated by our methods with respect to the ground truth profiles, we compared them by means of the Precision at position *k*, P@*k*, which, for a particular item, is defined as the percentage of the top *k* emotions returned by a method that are relevant for the item, as stated by the users of our study. Table 6 shows average precision values of the different methods and a random emotion ranking method. The extended method was the best performing approach in both domains with P@1 values close to 70%, and was outperformed by the basic method in the music domain for P@2 and P@3. In general, the methods performed in the music domain better than in the movies domain.

Table 1. Average P@*k* values of the considered emotion-oriented profiles.

Emotion	**movies**				**music**			
vector model	**#evals**	**P@1**	**P@2**	**P@3**	**#evals**	**P@1**	**P@2**	**P@3**
random	165	0.297	0.305	0.302	129	0.327	0.339	0.345
basic	107	0.598	0.528	0.514	109	0.606	**0.670**	**0.636**
extended_10	77	**0.675**	**0.643**	**0.589**	11	**0.636**	0.636	0.546

Acknowledgments. This work was supported by the Spanish Government (TIN2011-28538-C02) and the Regional Government of Madrid (S2009TIC-1542).

References

1. Cantador, I., Brusilovsky, P., Kuflik, T.: Second Workshop on Information Heterogeneity and Fusion in Recommender Systems. In: *RecSys'11*, pp. 387-388 (2011)
2. Hastings, J., Ceusters, W., Smith, B., Mulligan, K.: The Emotion Ontology: Enabling Interdisciplinary Research in the Affective Sciences. In: *CONTEXT'11*, pp. 119-123 (2011)
3. Meyers, O. C.: A Mood-based Music Classification and Exploration System. MSc Thesis. School of Architecture and Planning, MIT (2007)
4. Russell, J. A.: A Circumplex Model of Affect. *Journal of Personality and Social Psychology* 39(6), pp. 1161–1178 (1980)
5. Winoto, P., Ya Tang, T.: The Role of User Mood in Movie Recommendations. *Expert Systems with Applications* 37(8), pp. 6086-6092 (2010)

[1] The emotion lexicon, profiles, and evaluation tool are available at http://ir.ii.uam.es/emotions