Expert Systems With Applications 89 (2017) 179–187

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Item recommendation using tag emotion in social cataloging services

Hyewon Lim*, Hyoung-Joo Kim

School of Computer Science and Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea

ARTICLE INFO

Article history:
Received 9 February 2017
Revised 26 July 2017
Accepted 27 July 2017
Available online 27 July 2017

Keywords:
Recommendation
Tag
Emotion
Social cataloging service
High-Order Singular Value Decomposition (HOSVD)

ABSTRACT

Due to the overload of contents, the user suffers from difficulty in selecting items. The social cataloging services allow users to consume items and share their opinions, which influences in not only oneself but other users to choose new items. The recommendation system reduces the problem of the choice by recommending the items considering the behavior of the people and the characteristics of the items.

In this study, we propose a tag-based recommendation method considering the emotions reflected in the user's tags. Since the user's estimation of the item is made after consuming the item, the feelings of the user obtained during consuming are directly reflected in ratings and tags. The rating has overall valence on the item, and the tag represents the detailed feelings. Therefore, we assume that the user's rating for an item is the basic emotion of the tag attached to the item, and the emotion of tag is adjusted by the unique emotion value of the tag. We represent the relationships between users, items, and tags as a three-order tensor and apply tensor factorization. The experimental results show that the proposed method achieves better recommendation performance than baselines.

1. Introduction

Numerous contents appear every day. Thousands of movies are made, and more than one million books are published worldwide in a year. While consuming various contents, people can link the content their own experiences and feelings, and they can interact with others about their interests through various social media. The social cataloging services, such as Goodreads,[1] LibraryThing,[2] and Movielens,[3] allow users to catalog items and share their opinions on them with others through ratings, tags, and reviews. These services usually deal with time consuming content such as books and movies. They provide only meta-data or a fraction of the actual item such as sample teasers or chapters rather than providing the item itself. Thus, users are more likely to choose items that they want to consume carefully based on their personal taste by referring to the estimation of other users.

The inundation of content causes users in social cataloging services to have difficulty in selecting items among plenty of information. Recommendation systems have been proposed to solve the problem, and various recommendation techniques have been studied (Kefalas, Symeonidis, & Manolopoulos, 2016; Qingbiao, Jie, & Xu, 2011). Collaborative filtering is the most widely used recommendation method based on user's past behavior. Since the purpose of the recommendation system is to provide the appropriate information to users and improve their gratification, it is necessary to pay attention to the subjective feedback of the user in addition to the information about the item. Conventional recommendation systems have utilized rating data as user's explicit feedback on items. Unlike rating, tagging data does not explicitly indicate the user's preference for the item, but it contains additional information about the user's experience since the user directly inputs the tag. Especially, a tag that reflects an individual's subjective opinion contains positive or negative valence or certain feeling; it become a cue for understanding how a user considers an item. Therefore, the utilization of tagging data for recommendation can support the user experience and complement the existing rating information, thereby providing the possibility of improving the recommendation performance. In this paper, a tag that reflects user's emotion will be called an *emotion tag*.

The user's emotions play an important role in selecting and consuming items. According to Tkalčič, Kosir, and Tasic (2011), the emotions obtained from the action just before consuming the item affect the user's selection of a new item, and during consumption, the emotion changes with the passage of time. After the consumption, the emotion affects the user's next action; it can be very useful to measure the user's satisfaction with the item. In the social cataloging system, a consideration of the emotion factors can increase the accuracy of the recommendation system, since rating and tagging items can be viewed as a behavior reflecting this post-consumption feeling.

* Corresponding author.
E-mail addresses: hwlim@idb.snu.ac.kr (H. Lim), hjk@snu.ac.kr (H.-J. Kim).

[1] http://www.goodreads.com.
[2] http://www.librarything.com.
[3] http://movielens.org.

http://dx.doi.org/10.1016/j.eswa.2017.07.046
0957-4174/

A user's rating means an overall estimation, i.e., an item is positive or negative, and the tags are a detailed and additional reason of the rating. Therefore, emotion tags can be interpreted differently depending on which valence is used. If the same tag is assigned to the different items, it can be understood as positive, negative, or sarcastic meaning depending on whether the user uses the tag to the item with the high rating or with the low rating. For example, "funny" means "peculiar" as well as "humorous". Therefore, it is necessary to consider the intention of the user in the tag for a better understanding of the user's preference.

In this paper, we propose a tag-based item recommendation approach considering the emotions contained in tags. To calculate the tag weight, we first normalize the rating data and assign the value to each tag to consider the user's overall assessment of the item. Then, we obtain the emotion value of the emotion tags based on SenticNet (Cambria, Speer, Havasi, & Hussain, 2010), which is the emotion lexical resource, and arrange the tag weight using the emotion value. In this process, the weight of the same tag can be changed according to the positive or negative valence of the item.

In general, the ternary relationships of users, items, and tags are described by the tripartite graph; however, it cannot reflect the ternary association, but only three pairs of relationship, i.e., user-item, user-tag, and item-tag. Therefore, we model the relationship of users, items, and tags as a three-order tensor, which is a multi-dimensional matrix, and use a High-Order Singular Vector Decomposition (HOSVD) (De Lathauwer, De Moor, & Vandewalle, 2000) as a tensor factorization approach to recommend the appropriate items for each user. The previous research has mainly used the existence of tags as the initial element of a tensor, but we utilized the tag weight based on the emotion as the initial value to provide enriched information of the ternary relationship. We evaluate the performance of the proposed method using Movielens data, which is a social movie cataloging service, and showed that considering the emotions of tags improve the recommendation quality.

The contribution of this paper is in proposing a tag-based recommendation method considering user's emotions in tags to improve recommendation performance. We propose a method to calculate the weight of the tag emotions to take into account the user's emotion using the user's ratings and an emotion dictionary. Our experiments show that user's emotion plays an important role in item recommendation. This paper is organized as follows: Section 2 describes related research. In Section 3, we explain the tag weighting scheme based on emotions and the HOSVD algorithm for item recommendation. Section 4 describes the performance evaluation of the proposed recommendation method. Section 5 discusses the conclusion and future research.

2. Related work

2.1. Tag-based recommendation

Users in social cataloging services use tags for the purpose of facilitating retrieval of items and for sharing their opinions and communicating with other users (Ames & Naaman, 2007). Xu, Fu, Mao, and Su (2006b) classify tags into five categories: content-based tags which describe the content or categories of an object (e.g., Lucene, Germany Embassy), context-based tags which represent time or location that object was created (e.g., San Francisco, 2005-10-19), attribute tags which show the properties of an object (e.g., Jeremy's Blog, Clay Shirky), subjective tags which explain user's opinion or emotion (e.g., funny, Cool), and organizational tags for personal usage (e.g., to-read, to-review, my paper). The former three are informative tags that describe the item itself, and the latter two are tags that contain the user's individual opinion; both can be used together (e.g., good performance).

Much of the research on tagging has focused on why users are tagging, how tagging differs depending on the system, and whether the community affects user's tagging behaviors (Ames & Naaman, 2007; Meo, Ferrara, Abel, Aroyo, & Houben, 2013; Nov & Ye, 2010; Sen et al., 2006). They have reported that most of the social media services understand the importance of tagging. Tags are being payed attention in many studies of recommendation system because it is not the fixed keyword but the user's own subject. Guy, Zwerdling, Ronen, Carmel, and Uziel (2010) integrated tags used in social networks of business systems and proposed an item recommendation method which combines user and tag information. The authors generate the user profile for recommendation based on the various user-tag relations such as used tags, incoming tags, and indirect tags. Zhang and Liu (2012) suggested a diffusion-based hybrid recommendation algorithm considering the two roles of the tags that organizes items and connects between user and item. They shows that the latter role of tags is more helpful to recommend items, and the hybrid approach shows the best result. Kim, Alkhaldi, El Saddik, and Jo (2011) modeled users based on their tags. They classified items into two sets, positive and negative, and calculate the tag weights of the items in both sets. After that, they found the relevant topics based on the tags for the recommendation. Research of Gedikli and Jannach (2010, 2013) has conducted to predict the rating of the item by making rating on the tag itself in order to improve the quality of the tag-based item recommendation. Kim and Kim (2012) suggested an item recommendation method based on implicit trust relationships derived from user's tagging information.

2.2. Emotions in recommendation

The relationship between the emotions and user's consumption have been studied in various fields, and many recommendation studies have focused on how the emotions before and after consumption affect the choice of the next item (Chang, 2009; Gardner, 1985; Oliver, 2008; Tkalčič et al., 2011; Winoto & Tang, 2010). Tkalčič et al. (2011) classified emotion into three stages when the user uses the recommendation system and introduced emotion detection methods and emotion usage at each stage. According to this study, user's emotion before consuming items affects user's item selection. On the consumption stage, one emotion or various emotions appear over time depending on the type of content. Finally, emotion after consumption affects the user's next behavior, which is an indicator of whether the user is satisfied with the item. Zheng, Mobasher, and Burke (2013) studied the role of emotion in recommendation algorithms. They studied the recommendation considering emotion feature in context-aware splitting algorithm and differential context modeling algorithm. The evaluation revealed that emotion features improve the recommendation performance. Winoto and Tang (2010) showed that the rating can be biased according to the user's pre-mood, and they proposed a recommendation method considering the rating bias. SenticRank (Xie et al., 2016) is the framework which maps the tag-based user profile to the sentiment space and ranks the resources suitable for the user's query. The research was conducted for personal search, but it can be applied to the recommendation system. Qingbiao et al. (2011) studied a sentiment enhanced tag-based recommendation method which utilizes the positive and negative polarities of tag synsets for calculating similarities between resources. Dong, OMahony, Schaal, McCarthy, and Smyth (2016) applied the sentiment to product recommendation. They proposed an approach to combine the product similarity and the product sentiment; the product sentiment is obtained by extracting features from the user's review and calculate the sentiment of each feature. Kim, Kim, and Jo (2014) proposed a recommendation method for tackling cold start and data sparsity issues in music recommendation systems. The authors developed UniTag, a tagging ontology, to assign meaning and scores to user tags. The UniTag ontology consists of UniMusic ontology for solving the semantic ambiguity of the tags and UniEmotion ontology for weighting the tags. The users' profiles are generated by combining the emotion weight of the tags given through the ontology and the number of plays of items. A collaborative filtering technique is applied to the user profiles to recommend music to users. This study is similar to our research in terms of using the emotions of tags, but the target content and the strategy to measure preference values is different. Our targeted content, such as movies or books, has a limited number of items to be repeated, and the users leave the feedback in the social cataloging service after consuming items outside the service. Therefore, it is difficult to use the amount of item consumption as the value of the preference in our research. We used the user's ratings as the preference value of the items. The preference based on the number of plays of music can relatively change with time, but the preference using the ratings has always the same value, and thus can be more consistently reflected.

To extract emotion, various emotion lexical resources were introduced such as SenticNet (Cambria et al., 2010), Emolex (Mohammad & Turney, 2010), ANEW (Affective Norms for English Words) (Bradley & Lang, 1999), and SentiWordNet (Baccianella, Esuli, & Sebastiani, 2010). Among them, SenticNet (Cambria et al., 2010) is an emotional vocabulary dictionary for concept-level sentiment analysis. Emotion is represented by affective dimensions consisting of pleasantness, attention, sensitivity, and aptitude and by polarity based on the dimensions. In this paper, we use SenticNet 4.0 which includes 50,000 concepts to extract the emotions contained in the tags.

2.3. Tensor factorization approaches

In recommendation systems using tags, the relationships between users, items, and tags are represented by a tripartite graph. It can capture the three pairs of relationship, i.e., user-item, user-tag, and item-tag, but loses co-existence information about users, items, and tags (Peng, Zeng, Zhao, & Wang, 2010). Expressing the ternary relationship with a multidimensional matrix instead of a tripartite graph can improve the quality of the recommendation because the ternary associations can be considered. Researchers have used a three-order tensor to represent the ternary relationship and apply the tensor factorization method to capture the latent semantic associations among them. HOSVD (De Lathauwer et al., 2000), which is one of the tensor factorization methods, has been applied in various recommendation studies. Symeonidis, Nanopoulos, and Manolopoulos (2008, 2010) have proposed a tensor-based recommendation approach. They used a three-order tensor for user, item, and tag relationships and applied the HOSVD technique. A Kernel-SVD combination algorithm was adopted to improve the accuracy of the recommendation. Peng et al. (2010) introduced the concept of hidden tag and hidden item to efficiently grasp the similarity between users and suggested a recommendation technique using Tucker decomposition. Xu, Zhang, and Liu (2006a) adopted CubeSVD (Sun, Zeng, Liu, Lu, & Chen, 2005), which has been investigated to improve personalized web search using HOSVD technique. They split an original tensor into several sub-tensors in order to reduce the sparsity of the tensor. Ifada and Nayak (2014) studied the scalability in the tensor reconstruction process and the ranking of the recommended items. The authors ranked the result of the tensor reconstruction in order to consider the user's previous activities.

In our study, we used the tensor model and applied HOSVD as the tensor factorization method. Most of the recommendation method using the tensor model have used a binary value as an elements of the tensor. However, we used the emotion value as an initial value of the tensor for considering user's preference and impression of items.

3. Proposed approach

In this section, we describe a tag-based recommender system based on the emotion tags. The social cataloging system allows users to rate and tagging items. The rating indicates the user's overall preference or interest in the item, and the tag provides additional rich information about the preference because they depict the features or personal impression of the item. Especially, the emotion tags that reveal user's opinions or feelings can be important cues to the improvement of recommendation.

3.1. Preliminaries

To describe our approach, we define the concepts and the entities used in this paper. Let $U = \{u_1, u_2, u_3, \ldots, u_{|U|}\}$ be the set of the users, $I = \{i_1, i_2, i_3, \ldots, i_{|I|}\}$ be the set of items, $R = \{r_1, r_2, r_3, \ldots, r_{|R|}\}$ be the set of the ratings, and $T = \{t_1, t_2, t_3, \ldots, t_{|T|}\}$ be the set of the tags, where $|U|$, $|I|$, $|R|$, and $|T|$ are the number of users, items, ratings, and tags respectively. We denote the relationship that users assign ratings and tags to items as Y and define it as:

$$Y = < U, I, Y_{rating}, Y_{tag} > \quad (1)$$

where the Y_{rating} and Y_{tag} are represented as a set of triples, such as:

$$Y_{rating} \subseteq \{< u, i, r >: u \in U, i \in I, r \in R\} \quad (2)$$

and

$$Y_{tag} \subseteq \{< u, i, t >: u \in U, i \in I, t \in T\} \quad (3)$$

If rating does not exist, we describe rating as $r = \emptyset$. If tag does not exist, $t = \emptyset$; however, in this paper, we only consider the users who tagged items.

3.2. Weighting of tags

3.2.1. Rating-based tag weight

User's decisions always reflect emotion (González, De La Rosa, Montaner, & Delfin, 2007). When a rating and tags are assigned to an item, tags play a role in supporting rating except when used for personal classification and retrieval as an organizational tag. We assume that the rating is the result of condensing the user's feelings about the item so that the tag has a positive or negative valence based on the rating. The rating-based tag weight based on the rating is calculated as follows:

$$weight_{base}(t_{u,i}) = \begin{cases} r_{u,i}(t) & \text{if } r_{u,i} \neq \emptyset \\ 0 & otherwise \end{cases} \quad (4)$$

where $r_{u,i}$ is the rating which a user u assigned to an item i, and $r_{u,i}(t)$ is the normalized rating value for a tag t used by the user u in the item i. If the original rating is used as the tag weight, a bias may occur because the range of the rating given to the item varies depending on the user. Thus, we vectorize each user's ratings and normalize them into a unit vector. $r_{u,i}(t)$ is described as follows:

$$r_{u,i}(t) = \frac{r_{u,i}}{\sqrt{\sum_{i=1}^{|I|} r_{u,i}^2}} \quad (5)$$

3.2.2. Emotion-based tag weight

Emotion tags contain the more detailed intensity of polarity than the rating; the local weight of each tag use the emotion value of each tag. To obtain the value, the following steps are executed for each tag t:

(1) A special character removal. Remove the special characters contained in the tag. (e.g., awesome!)
(2) A proper noun removal. Proper nouns can be used as tags, and they may contain the emotion words (e.g., Jennifer Love Hewitt, Barnes & Noble). In this case, the words do not affect the emotion of the tag; thus, those tags are removed.
(3) Calculating local weight of tag. We use the emotion dictionary to calculate $weight_{emotion}(t_{u,i})$.
   i. If the tag exists in the emotion dictionary, the emotion value of the tag is used as the tag weight.

$$weight_{emotion}(t_{u,i}) = EmotionScore(t_{u,i}) \tag{6}$$

   where $EmotionScore(t)$ gives the emotion value of tag $t_{u,i}$ in the range from -1 to 1 if the tag is included in the emotion dictionary, otherwise the value is 0.
   ii. If the tag does not exist in the emotion dictionary and is composed of more than two words, the weight is calculated in units of words.
      a. Tokenizing. If the tag is composed of two or more words (e.g., great performance, this book is good), tokenize it in units of words.
      b. Lemmatizing and stemming. Each word appears in the form of a root word.
      c. Calculating local weight of tag. The weight of the tag is calculated based on the emotion value of each word which comprise the tag:

$$weight_{emotion}(t_{u,i}) = \frac{1}{|term_{emotion}|}\sum_{j=1}^{|term|} EmotionScore(term_j) \tag{7}$$

      where a *term* is the word which is consist of a tag t, $|term|$ is the number of words in the tag, $|term_{emotion}|$ is the number of the emotion words.

If no emotion value is finally obtained through step 3, the local weight of tag becomes 0.

3.2.3. Calculating overall tag weight

The total weight of each tag $weight(t_{u,i})$ is calculated using a unified model as proposed in the previous research (Kefalas & Manolopoulos, 2017; Symeonidis, Tiakas, & Manolopoulos, 2011; Yuan, Cong, Ma, Sun, & Thalmann, 2013) by combining the rating-based weight and the emotion tag-based weight as follows:

$$weight(t_{u,i}) = (1-\alpha)\times weight_{base}(t_{u,i}) + \alpha\times(weight_{emotion}(t_{u,i})\times 0.5) \tag{8}$$

where α is the parameter to control the influence of the emotion of the tag. The range of $weight(t_{u,i})$ adjusted from $(-1, 1)$ to $(-0.5, 0.5)$ so that the width of the range became similar to $weight_{base}(t_{u,i})$ while maintaining the polarity. The appropriate value of α is selected empirically. If the tag has no emotion value, only the rating-based weight is used to calculate the total weight ($\alpha = 0$).

If the user tagged an item with several tags, the average weight of the tags is used as the weight of each tag of the item.

$$weight(t_{u,i}) = \frac{1}{|t_{u,i}|}\sum_{k=1}^{|t_{u,i}|} weight(t^k_{u,i}) \tag{9}$$

where $t^k_{u,i}$ is the kth tag that user u assigned to item i.

3.3. Tensor factorization

Tensor factorization is the recommendation technique which deals with the multidimensional data. In this paper, we applied HOSVD (De Lathauwer et al., 2000) to exploit the latent relationships among objects. HOSVD is one of the tensor factorization methods that applied SVD to a tensor which is a n-dimensional matrix. We model the ternary relationship among users, items, and tags with three-order tensor and apply HOSVD obtain the reconstructed tensor. The list of the recommended items is generated according to the latent associations in the reconstructed tensor. Fig. 1 illustrates HOSVD and Fig. 2 describes the process of HOSVD. We briefly introduce the technique.

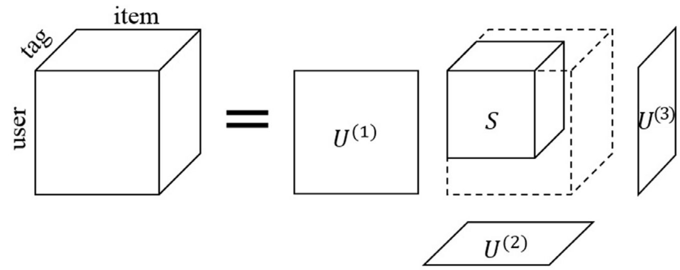

Fig. 1. The illustration of HOSVD.

First, an initial three-order tensor $\mathcal{A} \in \mathcal{R}^{|U|\times|I|\times|T|}$ is constructed, where $|U|$, $|I|$, $and|T|$ are the number of users, items, and tags, respectively. Then, tensor $\mathcal{A}$ is unfolded for all n modes. Through the unfolding process, the tensor is transformed to 2D matrices. Three new matrices A1, A2, and A3 are created as follows:

$$\begin{aligned} A1 &\in \mathcal{R}^{I_u\times I_i I_t}\\ A2 &\in \mathcal{R}^{I_i\times I_t I_u}\\ A3 &\in \mathcal{R}^{I_u I_i\times I_t}\end{aligned} \tag{10}$$

where I_u, I_i and I_t are the tensor dimensions. Next, SVD is applied to each of the three unfolded matrices.

$$\begin{aligned} A1 &= U^{(1)}\cdot S_1\cdot V_1^T\\ A2 &= U^{(2)}\cdot S_2\cdot V_2^T\\ A3 &= U^{(3)}\cdot S_3\cdot V_3^T\end{aligned} \tag{11}$$

Using the initial tensor $\mathcal{A}$ and the left singular vectors of the unfolded matrices, a core tensor $\mathcal{S}$ is constructed, which contains the ternary association between user, item, and tag.

$$\mathcal{S} = \mathcal{A}\times_1 U_1^{(1)^T}\times_2 U_2^{(2)^T}\times_3 U_3^{(3)^T} \tag{12}$$

Finally, a reconstructed tensor $\hat{\mathcal{A}}$ is computed, which is an approximation tensor of $\mathcal{A}$, with a core tensor $\mathcal{S}$. The reconstructed tensor has new entries as well as original entries.

$$\hat{\mathcal{A}} = \mathcal{S}\times_1 U_1^{(1)}\times_2 U_2^{(2)}\times_3 U_3^{(3)} \tag{13}$$

3.4. A running example

To facilitate the understanding of our approach, let us consider the following example. Suppose users assign ratings and tags to movies as shown in Table 1. In this running example, we assume that t_1 =“brave”, t_2 = “disgusting”, t_3 = “humanity”, and t_4 = “funny”, and each emotion value on the emotion dictionary is $t_1 = 0.306$, $t_2 = -0.41$, $t_3 = 0.105$, and $t_4 = 0.619$. The usage data is modeled as a three-order tensor $\mathcal{A} \in \mathcal{R}^{3\times5\times4}$ and each activity has a weight. The weight is calculated by user’s rating and the emotion value of each tag. Firstly, the rating-based weights for user u’s tags $weight_{base}(t_{u,i})$ are calculated by the Eq. (5). In the case of u_1, the weights of tags are 0.436 for tag t_1 and t_2, and 0.654 for tag t_3. Secondly, the emotion tag based weight $weight_{emotion}(t_{u,i})$ is computed by the Eq. (6) or (7) if the emotion dictionary includes the tag or the terms comprising the tag. Finally, the overall tag weights are calculated by combining both weights based on

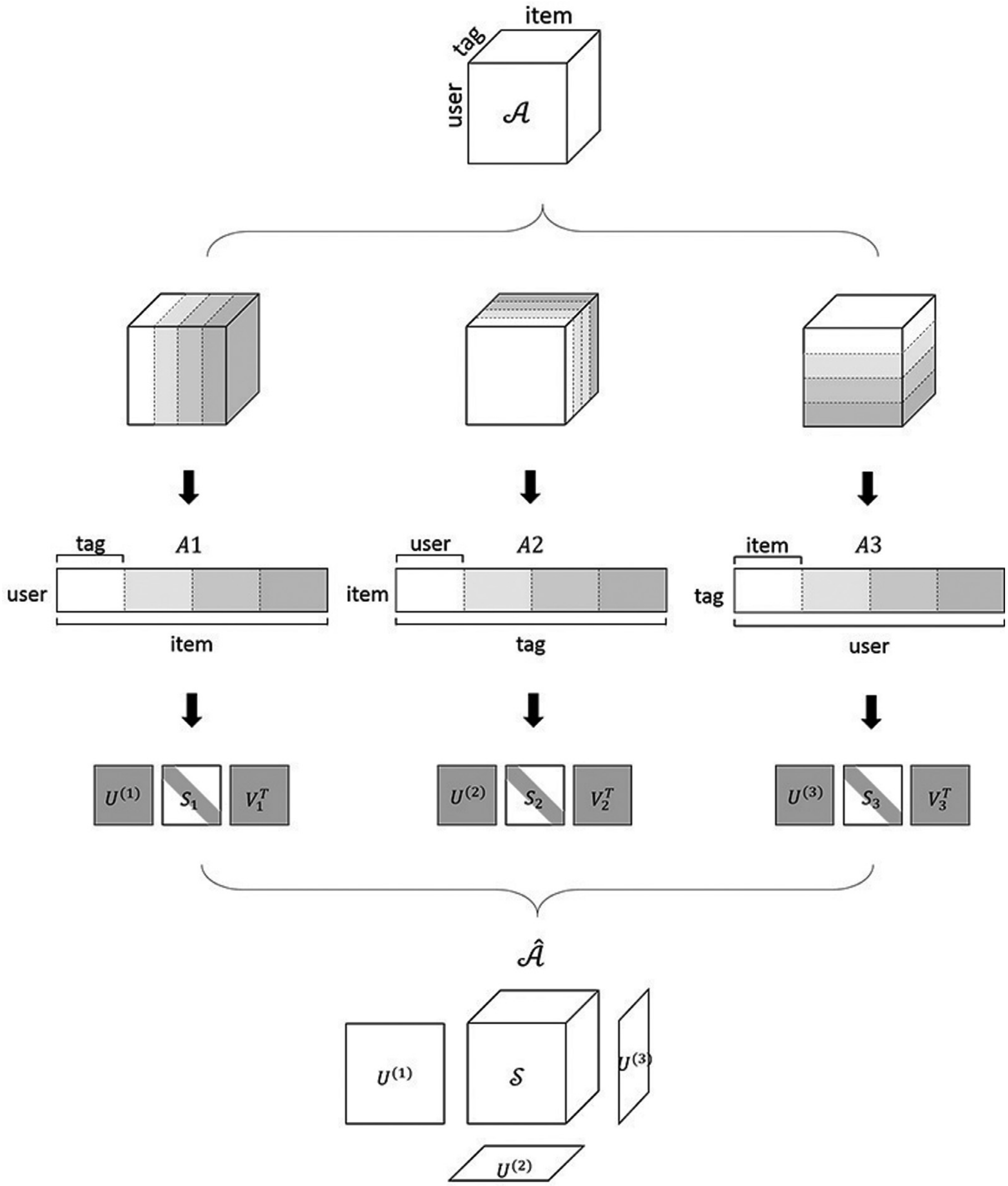

Fig. 2. The process of HOSVD.

Table 1
A usage data of the running example.

User	Movie	Tag	Rating
1	2	1	2
1	2	2	2
1	3	1	2
1	4	3	3
2	1	3	3
2	2	2	2
2	4	3	5
3	1	3	4
3	5	1	4
3	5	4	4

Table 2
An initial tensor $\mathcal{A}$. The parameter α for calculating the overall weight is set to 0.2.

User	Movie	Tag	$weight(t_{u,i})$ for $\mathcal{A}$
1	2	1	0.342
1	2	2	0.342
1	3	1	0.378
1	4	3	0.533
2	1	3	0.398
2	2	2	0.218
2	4	3	0.658
3	1	3	0.471
3	5	1	0.506
3	5	4	0.506

Eqs. (8) and (9). For instance, if α is set to 0.2, the tag weight for $< u_1, m_2, t_1 >$ is calculated as $(0.8 \times 0.436) + 0.2 \times (0.306 \times 0.5) = 0.378$. Since u_1 tags m_2 with two tags, the final tag weight is 0.342, which is the average weight of t_1 (0.378) and t_2 (0.307). The final weight $weight(t_{u,i})$ is reported in Table 2.

After the tensor factorization process, we have the reconstructed tensor $\hat{\mathcal{A}}$ and new entries are generated as described in

Table 3
A reconstructed tensor $\hat{\mathcal{A}}$ from the usage data. New entries are generated as highlighted.

User	Movie	Tag	$weight(t_{u,\,i})$ for $\hat{\mathcal{A}}$
1	**1**	**3**	**0.11**
1	2	1	0.24
1	2	2	0.22
1	3	1	0.14
1	4	3	0.62
2	1	3	0.28
2	2	2	0.19
2	**3**	**1**	**0.12**
2	**3**	**2**	**0.11**
2	4	3	0.57
3	1	3	0.51
3	5	1	0.51
3	5	4	0.43

Table 3. These entries become the candidate for the recommendation.

4. Experimental evaluation

4.1. Dataset

We use the dataset of Movielens,[4] which is a social movie cataloging service, to evaluate the performance of the proposed approach. The dataset has 71,567 users, 10,681 movies, 10,000,054 rating history, and 95,580 tagging histories. There are 15,230 distinct tags, and 4009 users use tags at least once. On average, each user rates 143 movies and has 10 distinct tags. Among the users who have tagging history, 40% of the users use only one tag. There are very few active users who rate many movies, especially those who use tags. The users have 10 distinct tags and 140 movies on average. 75% of the users have less than equal to 5 tags, and most of the users have more than 10 movies. The movies has 9 tags on average, with a maximum of 139. 95% of the users assign less than equal to 2 tags on a movie.

We limited the data for the experiment to users and movies with tagging history. In order to obtain the dense data, we applied p-core (Batagelj & Zaveršnik, 2002) at level k to the dataset, which means that each user, item, and tag occurred at least k times; This process removes unfamiliar items and less frequently used tags. We applied $k = 5$, and finally 210 users, 544 movies, and 365 tags are used.

For implementing our approach, we used a list of actors, directors, writers, and producers provided by IMDB[5] for eliminating proper nouns and Natural Language Toolkit (NLTK) (Bird, 2006) for tag processing. As the emotion dictionary, SenticNet 4.0 (Cambria et al., 2010) was utilized. For the data reduction process of HOSVD, we preserve 80% of the information in the original diagonal matrix S_i $(1 \le i \le 3)$.

4.2. Experimental results

The dataset is divided into five subsets for 5-fold cross validation. For each fold, we selected 80% of each user's history as the training set and the remaining 20% as the test set. The ratio between the emotion tag and the un-emotional tag, which is called the *ordinary tag*, in each training set is reported in Table 4. In all training sets, the emotion tag accounted for more than 50%; namely, it seems meaningful to reflect the emotion tags that each user expressed the feelings of each movie with various intensities to the recommendation.

[4] https://grouplens.org/datasets/movielens/.
[5] http://www.imdb.com.

Table 4
The ratios between the emotion tag and the ordinary tag in each training set.

Set	Emotion Tag	Ordinary Tag
1	52.25%	47.75%
2	50.14%	49.86%
3	50.84%	49.16%
4	51.16%	48.84%
5	50.57%	49.43%

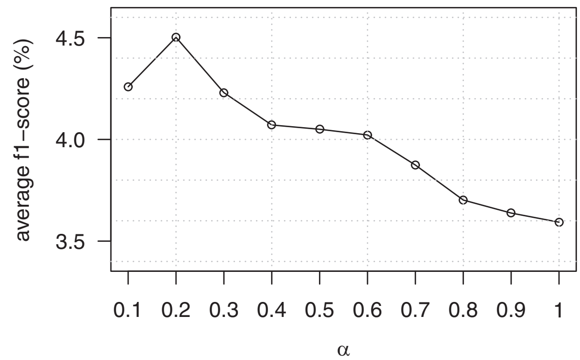

Fig. 3. The average f1-score according to the change of α when n movies (n = 1, 2, 3, 4, 5, 10, 15, 20, 25, 30) are recommended.

We conducted the experimental evaluation to find an appropriate value of the parameter α. Fig. 3 shows the average f1-score according to the change of α value between 0.1 and 1.0 when n movies (n = 1, 2, 3, 4, 5, 10, 15, 20, 25, 30) are recommended. When α is 0.2, we had the best result on the dataset. The parameter for controlling the emotion tag-based weight is assigned the small value so that the rating-based weight of the tag does not change significantly when both weights are combined. The significant change in the rating-based weight means a change in the rating given by the user, and also a change in users the overall impression of the item. The control parameter α is set to 0.2 for the rest of our experiments.

Next, we conducted the experiment for the performance analysis. To compare the performance, we consider the following methods:

- Baseline. The previous research (Symeonidis, Nanopoulos, & Manolopoulos, 2010) set the weight based on the existence of the tag. If user u tagged item i with tag t, the weight is 1.

$$weight\,(t_{u,i}) = \begin{cases} 1 & \text{if} < u, m, t > \in Y \\ 0 & otherwise \end{cases} \tag{14}$$

- Rating-based only (Rate). Regardless of the type of the tags, user's rating based weight $weight_{base}(t_{u,\,i})$ is used as the tag weight.

$$weight\,(t_{u,i}) = weight_{base}\,(t_{u,i}) \tag{15}$$

- Emotion Tag-based only (ET). The tag weight is calculated based only on the emotion tags, excluding the user's rating. In this case, if the weight of a tag which is not included in the emotion dictionary is assigned 0, the ternary relationship is regarded as non-existence; 0 means that the user did not attach the tag to the movie. In order to solve this problem, the range of $weight_{emotion}(t_{u,\,i})$ was adjusted from $(-1, 1)$ to $(0, 2)$.
  The emotion value of each tag cannot be regarded as a general preference for the movie. Therefore, if a movie is tagged with multiple tags, the average of the tag weights is given to each

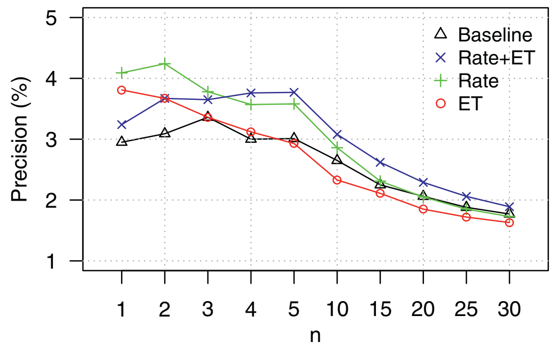

Fig. 4. The comparisons of precision as the number of recommended item increases.

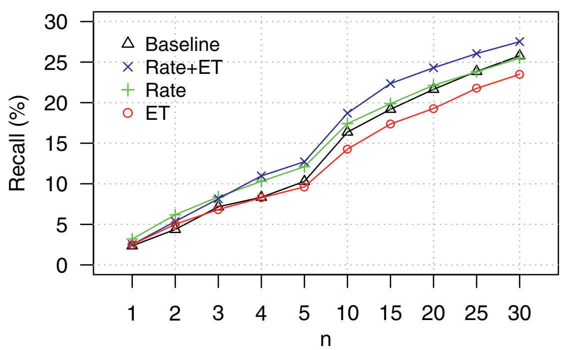

Fig. 5. The comparisons of recall as the number of recommended item increases.

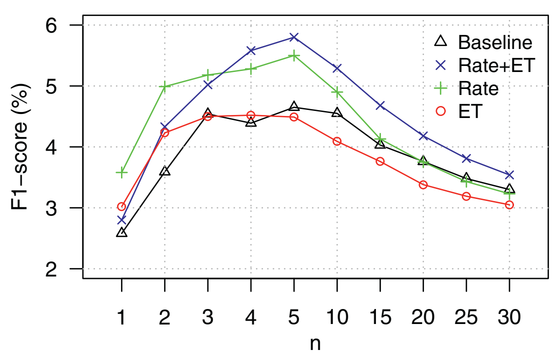

Fig. 6. The comparisons of f1-score as the number of recommended item increases.

tag as the weight. The $weight(t_{u,\,i})$ is calculated as follows:

$$weight\,(t_{u,i}) = \frac{1}{|t_{u,m}|}\sum_{j=1}^{|t_{u,m}|}(weight_{emotion}(t_j) + 1) \tag{16}$$

where $|t_{u,\,m}|$ is the number of tags attached at a movie m by a user u.

Our approach (Rate+ET) was compared with the methods in top-n movie recommendations. As the measures for evaluating the results, the precision, recall, and f1-score were used.

$$\begin{aligned} Precision &= \frac{the\ number\ of\ correct\ positive\ predictions}{the\ number\ of\ positive\ predictions} \\ Recall &= \frac{the\ number\ of\ correct\ positive\ predictions}{the\ number\ of\ positive\ examples} \\ f1-score &= \frac{2\times Precision\times Recall}{Precision + Recall} \end{aligned} \tag{17}$$

The results are shown in Figs. 4, 5, and 6: the precision, recall, and f1-score of four approaches respectively. The x-axis of each

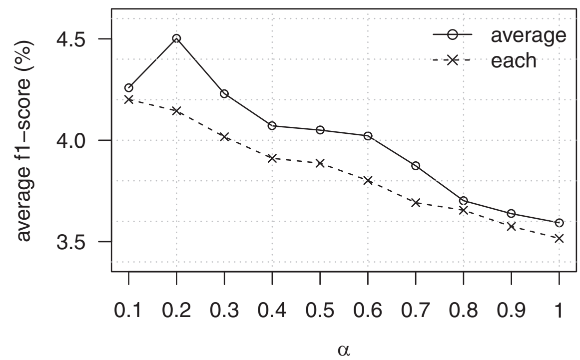

Fig. 7. When a movie has the multiple tags, the difference of f1-score between using an average value of the tags and using each value of the tags as the weight of each tag.

graph represents the number of recommended movies, and the y-axis represents the values of precision, recall, and f1-score, respectively. The results indicate that the recommendation method considering user's emotions shows better performance. Among them, we find that the approach considering the detailed emotions with overall valence (Rate+ET) is generally better than the other methods. In comparison with the rating-based method, the approach using both the rating-based and the emotion tag-based weight shows the better results as n is increased. It implies that reflecting the users' subjective emotion for the items enables a better understanding of the users' preferences. In the case of the method utilizing only the emotion tag, the performance is degraded as compared with the other methods. This is because not all the user's tags are in the emotion tag category, and tags express what the user felt with only a few keywords; thus, even if the average of the emotional values of all the tags attached to a movie are utilized as the tag weight, it may not represent the user's overall satisfaction and preference. All the differences in the results are statistically significant with $p <\ 0.05$.

We also investigated how the multiple tags on the same item are handled during the calculation of the tag weight. In the proposed approach, if there is more than one tag on the item, the average value of the tags is used as the weight of each tag (Eq. (9)). However, if an item has multiple tags, there are two ways to compute the weight of the tags. One is to use the weight of each tag as it is, and the other is to give the average weight of the tags attached to the item as explained in the proposed method.

Fig. 7 indicates the average f1-score according to the change of α value when n movies are recommended by calculating tag weights in the two ways. The solid line is for the case of using the average value, and the dotted line is for the case of using each weight of the tags. When the average value is given as the weight of the tags on the same movie, the performance is better than using each weight of the tag. Fig. 8 describes the difference in the results depending on how multiple tags in the movie are handled when the tag weights are computed based only on the emotion tag based method. What the solid and dotted lines mean is the same as in Fig. 7. In this case, using the average value as the weight of tags shows better performance because using respective weights can reflect the various emotions that user expresses about the movie but cannot adopt the general preference deriving from the tags.

When a user tags a movie, positive and negative tags can be used together to describe a detailed emotion for the movie; it can also affects the way to handling of multiple tags. It indicates that using the average value as the weight of multiple tags can help to reflect the users overall preference for the movie.

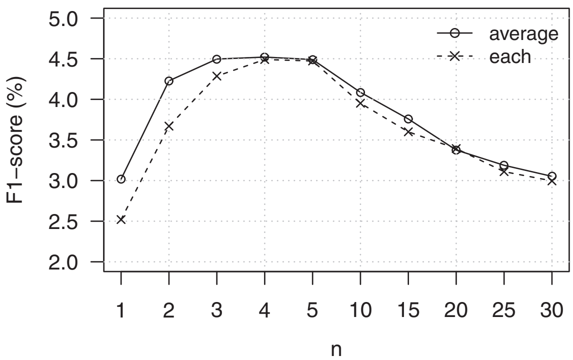

Fig. 8. When a movie has the multiple tags and tag weights are computed based on the emotion tag based method, the difference of f1-score between using an average value of the tags and using each value of the tags as the weight of each tag.

5. Conclusion

Users in social cataloging services catalog items and share their experiences with others. Overloading of various contents causes users to have difficulty in selecting items. The recommendation system reduces the problem of the selection by recommending the item considering the behavior of the user and the characteristics of the contents.

In this study, we propose a tag-based recommendation method considering the emotions reflected in the user's tags. The user's estimation of the item is made after consuming the item; thus, the user's emotions are reflected directly, and they can play an important role in the recommendation system. The rating has an overall positive or negative valence for the item, and the tag is the detailed reason for the estimation. Therefore, when user rated and tagged an item, we utilize the rating of the item as the basic feeling of the tag and adjust the tag weight with the unique emotion value of the tag based on SenticNet, which is the emotion dictionary.

To solve the problem that ternary relationships of users, items, and tags are mapped to two-dimensional relations and cannot reflect the association of three entities, we express those relationships as a three-order tensor and apply HOSVD, which is one of the tensor factorization methods, to the tensor. The proposed recommendation method is compared with the cases where the weight of the tag is calculated only by rating, only with the emotion value of the tag, and by the tag's existence. The result indicates that our approach improves the recommendation performance.

As the future research, if the emotion dictionary is extended using synonyms and antonyms, the coverage of the emotion dictionary is increased, and it can affect the recommendation quality. Also, since users tend to use few tags to the items, the tag expansion can improve the quality of the recommendation.

Acknowledgement

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (No. R0113-15-0005, Development of a Unified Data Engineering Technology for Largescale Transaction Processing and Real-time Complex Analytics).

References

Ames, M., & Naaman, M. (2007). Why we tag: Motivations for annotation in mobile and online media. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 971–980). ACM.
Baccianella, S., Esuli, A., & Sebastiani, F. (2010). Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining.. In *LREC: 10* (pp. 2200–2204).
Batagelj, V., & Zaveršnik, M. (2002). Generalized cores. arXiv preprint cs/0202039.
Bird, S. (2006). Nltk: The natural language toolkit. In *Proceedings of the COLING/ACL on interactive presentation sessions* (pp. 69–72). Association for Computational Linguistics.
Bradley, M. M., & Lang, P. J. (1999). Affective norms for English words (ANEW): Instruction manual and affective ratings. *Technical Report*. The center for research in psychophysiology, University of Florida. Technical report C-1.
Cambria, E., Speer, R., Havasi, C., & Hussain, A. (2010). Senticnet: A publicly available semantic resource for opinion mining.. *AAAI fall symposium: Commonsense knowledge*.
Chang, H.-C. (2009). Emotion barometer of reading: User interface design of a social cataloging website. In *CHI'09 extended abstracts on human factors in computing systems* (pp. 3371–3376). ACM.
De Lathauwer, L., De Moor, B., & Vandewalle, J. (2000). A multilinear singular value decomposition. *SIAM Journal on Matrix Analysis and Applications, 21*(4), 1253–1278.
Dong, R., OMahony, M. P., Schaal, M., McCarthy, K., & Smyth, B. (2016). Combining similarity and sentiment in opinion mining for product recommendation. *Journal of Intelligent Information Systems, 46*(2), 285–312.
Gardner, M. P. (1985). Mood states and consumer behavior: A critical review. *Journal of Consumer Research, 12*(3), 281–300.
Gedikli, F., & Jannach, D. (2010). Rating items by rating tags. In *Proceedings of the 2010 workshop on recommender systems and the social web at ACM RecSys* (pp. 25–32).
Gedikli, F., & Jannach, D. (2013). Improving recommendation accuracy based on item-specific tag preferences. *ACM Transactions on Intelligent Systems and Technology (TIST), 4*(1), 11.
González, G., De La Rosa, J. L., Montaner, M., & Delfin, S. (2007). Embedding emotional context in recommender systems. In *Data engineering workshop, 2007 IEEE 23rd international conference on* (pp. 845–852). IEEE.
Guy, I., Zwerdling, N., Ronen, I., Carmel, D., & Uziel, E. (2010). Social media recommendation based on people and tags. In *Proceedings of the 33rd international ACM SIGIR conference on research and development in information retrieval* (pp. 194–201). ACM.
Ifada, N., & Nayak, R. (2014). Tensor-based item recommendation using probabilistic ranking in social tagging systems. In *Proceedings of the 23rd international conference on world wide web* (pp. 805–810). ACM.
Kefalas, P., & Manolopoulos, Y. (2017). A time-aware spatio-textual recommender system. *Expert Systems with Applications, 78*, 396–406.
Kefalas, P., Symeonidis, P., & Manolopoulos, Y. (2016). A graph-based taxonomy of recommendation algorithms and systems in lbsns. *IEEE Transactions on Knowledge and Data Engineering, 28*(3), 604–622.
Kim, H., & Kim, H.-J. (2012). Improving recommendation based on implicit trust relationships from tags. In *Proceedings of the 2nd international conference on computers, networks, systems, and industrial applications* (pp. 25–30).
Kim, H. H., Kim, D., & Jo, J. (2014). A unified music recommender system using listening habits and semantics of tags. *International Journal of Intelligent Information and Database Systems 4, 8*(1), 14–30.
Kim, H.-N., Alkhaldi, A., El Saddik, A., & Jo, G.-S. (2011). Collaborative user modeling with user-generated tags for social recommender systems. *Expert Systems with Applications, 38*(7), 8488–8496.
Meo, P. d., Ferrara, E., Abel, F., Aroyo, L., & Houben, G.-J. (2013). Analyzing user behavior across social sharing environments. *ACM Transactions on Intelligent Systems and Technology (TIST), 5*(1), 14.
Mohammad, S. M., & Turney, P. D. (2010). Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text* (pp. 26–34). Association for Computational Linguistics.
Nov, O., & Ye, C. (2010). Why do people tag?: Motivations for photo tagging. *Communications of the ACM, 53*(7), 128–131.
Oliver, M. B. (2008). Tender affective states as predictors of entertainment preference. *Journal of Communication, 58*(1), 40–61.
Peng, J., Zeng, D. D., Zhao, H., & Wang, F.-y. (2010). Collaborative filtering in social tagging systems based on joint item-tag recommendations. In *Proceedings of the 19th ACM international conference on information and knowledge management* (pp. 809–818). ACM.
Qingbiao, Z., Jie, F., & Xu, G. (2011). Incorporating sentiment analysis for improved tag-based recommendation. In *Proceedings of the 2011 IEEE ninth international conference on dependable, autonomic and secure computing* (pp. 1222–1227). IEEE.
Sen, S., Lam, S. K., Rashid, A. M., Cosley, D., Frankowski, D., Osterhouse, J., … Riedl, J. (2006). Tagging, communities, vocabulary, evolution. In *Proceedings of the 2006 20th anniversary conference on computer supported cooperative work* (pp. 181–190). ACM.
Sun, J.-T., Zeng, H.-J., Liu, H., Lu, Y., & Chen, Z. (2005). Cubesvd: A novel approach to personalized web search. In *Proceedings of the 14th international conference on world wide web* (pp. 382–390). ACM.
Symeonidis, P., Nanopoulos, A., & Manolopoulos, Y. (2008). Tag recommendations based on tensor dimensionality reduction. In *Proceedings of the 2008 ACM conference on recommender systems* (pp. 43–50). ACM.
Symeonidis, P., Nanopoulos, A., & Manolopoulos, Y. (2010). A unified framework for providing recommendations in social tagging systems based on ternary semantic analysis. *IEEE Transactions on Knowledge and Data Engineering, 22*(2), 179–192.
Symeonidis, P., Tiakas, E., & Manolopoulos, Y. (2011). Product recommendation and rating prediction based on multi-modal social networks. In *Proceedings of the fifth ACM conference on recommender systems* (pp. 61–68). ACM.

Tkalčič, M., Kosir, A., & Tasic, J. (2011). Affective recommender systems: The role of emotions in recommender systems. In *Proceedings of the RecSys 2011 workshop on human decision making in recommender systems* (pp. 9–13). Citeseer.

Winoto, P., & Tang, T. Y. (2010). The role of user mood in movie recommendations. *Expert Systems with Applications, 37*(8), 6086–6092.

Xie, H., Li, X., Wang, T., Lau, R. Y., Wong, T.-L., Chen, L., … Li, Q. (2016). Incorporating sentiment into tag-based user profiles and resource profiles for personalized search in folksonomy. *Information Processing & Management, 52*(1), 61–72.

Xu, Y., Zhang, L., & Liu, W. (2006a). Cubic analysis of social bookmarking for personalized recommendation. In *Asia-Pacific web conference* (pp. 733–738). Springer.

Xu, Z., Fu, Y., Mao, J., & Su, D. (2006b). Towards the semantic web: Collaborative tag suggestions. *Collaborative web tagging workshop at WWW2006, Edinburgh, Scotland*.

Yuan, Q., Cong, G., Ma, Z., Sun, A., & Thalmann, N. M. (2013). Time-aware point-of-interest recommendation. In *Proceedings of the 36th international ACM SIGIR conference on research and development in information retrieval* (pp. 363–372). ACM.

Zhang, Z.-K., & Liu, C. (2012). Hybrid recommendation algorithm based on two roles of social tags. *International Journal of Bifurcation and Chaos, 22*(07), 1250166.

Zheng, Y., Mobasher, B., & Burke, R. D. (2013). The role of emotions in context-aware recommendation. *Decisions@ RecSys, 2013*, 21–28.