Analysis of Mood Tags for Multimedia Content Recommendation in Social Networks

Chang Bae Moon
ICT-Convergence Research Center
Kumoh National Institute of Technology
Gumi, South Korea
cb.moon@kumoh.ac.kr

Jong Yeol Lee
Department of Computer Software Engineering
Kumoh National Institute of Technology
Gumi, South Korea
soyeum@kumoh.ac.kr

Dong-Seong Kim
Department of IT Convergence Engineering
Kumoh National Institute of Technology
Gumi, South Korea
dskim@kumoh.ac.kr

Byeong Man Kim
Department of Computer Software Engineering
Kumoh National Institute of Technology
Gumi, South Korea
bmkim@kumoh.ac.kr

***Abstract*—** The propensity of Web information purchasers is changing from cost effectiveness that emphasizes price with performance to cost satisfaction that emphasizes purchaser's psychological satisfaction. One of the methods to improve user's cost satisfaction in recommending multimedia contents is to use mood inherent in multimedia contents. SNS services based on mood folksonomy are applications using such a method. However, in the applications based on folksonomy, the problem caused by synonyms exists. This paper suggests a cost-satisfactive multimedia contents recommendation method to solve the problem on synonym. To this end, we utilizes the value of Arousal & Valence which express mood of multimedia content as its internal tag. A method that defines the relation between AV values of multimedia contents and AV values of mood tags is suggested by considering synonym and the correlation between them is analyzed. From the analysis, it is shown that the AV values of multimedia contents with a mood tag and its synonyms reside in the area of the mood tag in the Thayor model.

Keywords—multimedia content; cost satisfaction; multimedia content mood; multimedia content recommendation; mood tag; social network

I. INTRODUCTION

Checking the propensity of information purchasers in Web, they emphasize cost effectiveness which mainly emphasizes performance compared to price. In modern society, not only cost effectiveness, but also consumption that emphasizes cost effectiveness with psychological satisfaction is emerging [1]. Since multimedia contents recommendation methods using mood recommend multimedia contents of the mood desired by user, such methods increase user's satisfaction and so it is closely related to cost satisfaction. As the studies related to mood-based multimedia contents recommendation, the studies of [2, 3, 4, 5, 6, 7] exist; the studies of [2, 3, 4] define tag mood with mood vector; the study of [5] recommends multimedia contents by using mood; the study of [6, 7] provide mood color of multimedia contents.

Folksonomy and Taxonomy are existed for classification. In the Folksonomy, unlike the hierarchical taxonomic structure of a traditional classification system, folksonomy has a flat structure and is essentially a classification system with user participation, instead of specialized management by a librarian or curator. Typical cases employing folksonomy include Instagram, Facebook, Last.fm, IPTV, and YouTube, which are SNSs. The advantage of folksonomy is that it circumvents taxonomy's problems of expandability and exclusivity. However, folksonomy has limitations regarding the tags used for content; e.g., analogues and new coinages.

To counter the problems faced by SNSs, this paper proposes a method that utilizes values for arousal and valence (AV) to express the mood of multimedia content in internal tags. Here, the relation between the AV values of multimedia content and the AV values of mood tags is defined by considering synonyms, and the correlation between them is analyzed.

II. RELATED STUDIES

As the studies defining mood, there are a method using Russell model [2], a method using Hevner model [3] and a method using Thayer model [4]. As Russell model & Hevner model are based on adjective, they have demerit that meaning is redundant and adjective expression is equivocal in expression. What supplements the equivocal demerit of Russell model & Hevner model is the expanded 2nd model of Thayer, which uses 12 words. In the 2nd dimension mood model of Thayer, the mood of multimedia contents is expressed as vector value composed of Arousal & Valence. While arousal indicates the strength of stimulation felt by hearer in multimedia contents, valence indicates the stability of sound and there are [5~7] studies in multimedia contents recommendation method that uses this model. In the case of actual environment like social network or folksonomy, however, synonym & new coinage can be used besides 12 mood words. Considering this point in multimedia contents recommendation may be said to be significant.

The study to solve the problem of synonym based on folksonomy, tagging level and new coinage is study of [5]. This study suggests method that utilizes mood vector (strength of Arousal & Valence) as internal tag of music. Namely, as both music and tag are expressed as AV value, music that has synonym tag can be also searched. In the study, however, as

testee assigns a 2-dimensional mood vector of a music piece after listening to it, it is hard to apply the method to large amount of data and so it can be said to be a method that uses data of small quantity in specific area. However, in this paper, social based data, namely the global based mood data that has been built for several years is used, where data are tagged with mood tags not with AV values.

The studies [8-10] use Folksonomy tags of music contents, where mood of music is defined by using mood tags of music contents in folksonomy site (last.fm) and mood of music is trained by SVM (Support Vector Machine). In the case of existing studies, acoustic and musical characteristics contained in music are used and thus, there is demerit in applying to other multimedia contents with different characteristics like image and video. In the proposed method, however, only multimedia contents tags are used and thus, there is merit that expansive application to all Web service having tag like Instagram, Facebook, Last.fm, IPTV, YouTube is possible.

III. MULTIMEDIA CONTENT RECOMMENDATION CONSIDERING SYNONYMS BASED ON MOOD TAGS

As described in Figure 1, our multimedia contents recommendation method considering synonym based on mood tags consists of 4 phases, the first of which is construction of multimedia contents information, the second of which is generation of AV table of multimedia contents and the third of which is generation of AV table of tags and the last of which is multimedia contents recommendation based on these two tables.

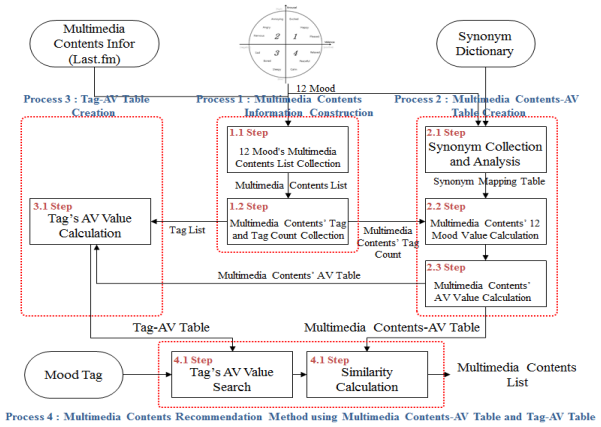

Fig. 1. Multimedia Contents Recommendation Structure

A. Constructing multimedia contents information

The process to construct information of multimedia contents is divided into 2 steps, the first of which is collecting list of multimedia contents of 12 moods and the second of which is collecting tags of multimedia contents and tag counts. In this paper, the study is performed only for music contents instead of all type of multimedia contents though the proposed method can be applied to all type of multimedia contents.

B. Generating AV table of multimedia contents

The process to construct AV table of multimedia contents consists of three stages; the first is to collect & analyze synonym, the second is calculating mood vector of multimedia contents and the third is AV value calculation of multimedia contents. In order to classify multimedia contents tags into 12 mood, namely to construct synonym mapping table, outcome of synonym analysis in study [1] is utilized.

The 12 mood count value, called by mood vector, of multimedia contents can be calculated as formula (1) by using synonym mapping table and tag information of multimedia contents.

$$W_k = (M_{1,k}, M_{2,k}, ..., M_{11,k}, M_{12,k}),$$
$$M_{m,k} = \sum_{i=1}^{n} t_i^{m,k}, \qquad (1)$$
$$t_i^{m,k} = \begin{cases} c_i^k & if \quad Tag_i^k \in S^m \\ 0 & otherwise \end{cases},$$

where, k is the index of multimedia content, m the mood index ($1 \le m \le 12$), i the tag index, W_k the mood vector of multimedia content k, $M_{m,k}$ the count value of mood m for multimedia content k. Tag_i^k the *i'th* tag of multimedia contents k, S^m the synonym list corresponding to mood m, c_i^k the tag count of tag Tag_i^k.

The mood vector of multimedia content k can be indicated as 12 points on AV coordinate of Thayer 2 dimensional model. The counter value of each mood indicates the distance from the origin on relevant mood axis.

$$AV_k = (A_k, V_k),$$
$$A_k = \frac{1}{12} \sum_{m=1}^{12} M_{m,k} \times cos(\theta^m), \qquad (2)$$
$$V_k = \frac{1}{12} \sum_{m=1}^{12} M_{m,k} \times sin(\theta^m),$$

where, AV_k is the AV value of multimedia content k, A_k the Arousal value of multimedia content k, V_k the Valance value of multimedia content k, m the mood index ($1 \le m \le 12$), θ^m the angle pertaining to mood m, $\theta^m = \theta^{m-1} + 30$ ($2 \le m \le 12$) and $\theta^1 = 15°$. $M_{m,k}$ is the value which pertains to mood m for multimedia contents k, same with the one in formula (1).

C. Generating AV table of tags

To search multimedia contents by using AV values of multimedia contents, the AV value of a tag is also necessary. The AV value of a tag is calculated by average of the AV values of multimedia contents including the mood tag.

D. Recommendation of multimedia contents by mood tags

After getting the AV value of a tag in user query from the AV table of tags, the similarity between it and the AV value of

each multimedia content is calculated. Multimedia content with high similarity is preferentially recommended. In this paper, before analysis of recommendation performance, AV values distribution of multimedia contents and tags are analyzed on the basis of 12 moods of Thayer 2-dimensional model by using the ANOVA Test.

IV. EXPERIMENTS

To analyze performance of the proposed method, first, some 50,000 multimedia contents and their tags are constructed by using API of Web site, Last.fm. In details, some 3,200 of tag 'angry', some 1,900 of 'annoying', some 40 of 'bored', some 10,000 of 'calm', some 70 of 'excited', some 10,300 of 'happy', some 170 of 'nervous', some 4,800 of 'peaceful', some 10 of 'pleased', some 1,900 of 'relaxed', some 10,200 of 'sad' and some 7,500 of 'sleepy' are included in the dataset. By using the dataset, the AV table of multimedia contents and the AV table of tags are constructed. Then, the distribution of AV values of contents and tags are analyzed by using these tables.

To confirm whether 12 mood groups of multimedia contents have mutually independent distribution, in this paper, ANOVA test and the homogeneity test of dispersion are performed for their AV values. The 12 moods (happy, sad, annoying, pleased, excited, nervous, bored, sleep, calm, peaceful, relaxed, angry) are selected as the independent variables, and A value and V value of each multimedia content are selected as the dependent variables. The experimental results of these tests are seen in Table 1 and 2. Since all p-value is 0.000, the null hypothesis H0 can be dismissed and the alternative hypothesis H1 can be adopted. Thus, the AV values of multimedia contents are mainly distributed around the AV value of tags and the distributions of 12 moods are distinguishable. Also, this implies the proposed method can show better performance when multimedia contents whose similarity is high with AV value of a query tag are recommended.

TABLE I. ANOVA TEST (SOS : SUM OF SQUARES)

		SoS	DF	Mean Square	F	P-Value
Arousal	Mood	715046.500	11	65004.22	2134.111	0.000
	Error	2426323	79657	30.460		
	Total	3141369	79668			
Valence	Mood	442061.200	11	40187.38	1617.705	0.000
	Error	1978856	79657	24.842		
	Total	2420918	79668			

TABLE II. LEVENE TEST

	Levene	DF1	DF2	P-Value
Arousal	346.4137	11	79657	0.000
Valence	228.9353	11	79657	0.000

V. CONCLUSION

In this paper, to partially solve the problem of synonym existing in Folksonomy in SNS service, a multimedia contents recommendation method that uses AV values of mood tags instead of tags themselves was suggested. In this paper, before analysis of recommendation performance, AV values distribution of multimedia contents and tags are analyzed on the basis of 12 moods by using the ANOVA Test. Through the ANOVA Test and the homogeneity test of dispersion by using the AV table of multimedia contents, it was shown that distribution differences of AV values exist between 12 moods. In the future works, the performance of recommendation will be analyzed by using the proposed method.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2017R1D1A1B03033733, 2018R1C1B6001042, 2018R1A6A1A03024003); This research was supported by the MSIT(Ministry of Science, ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2019-2014-1-00639) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation).

REFERENCES

[1] C. B. Moon, J. Y. Yi, D.-S. Kim, B. M. Kim, Analysis of Overlapping Mood Tags Based on Synonyms, Korea Computer Congress 2018 (KCC 2018), KIISE (2018), June 20-22; ICC JEJU, Korea, pp.667-669 (2018).

[2] Russel, J. A., A Circumplex Model of Affect, Journal of Personality and Social Psychology, Vol.39, No.6, pp.1161-1178 (1980).

[3] Hevner, K., Experimental studies of the elements of expression in music, The American Journal of Psychology, Vol.48, No.2, pp.246-268 (1936).

[4] Thayer, R. E., The Biopsychology of Mood and Arousal, Oxford University Press (1990).

[5] C. B. Moon, H. S. Kim, B. M. Kim, Music Retrieval Method using Mood Tag and Music AV Tag based on Folksonomy, Journal of KIISE, Vol.40, No.9, pp.526-543 (2013).

[6] C. B. Moon, H.S. Kim, H. A. Lee, B. M. Kim, Analysis of Relationships Between Mood and Color for Different Musical Preferences, Color Research & Application, Vol.39, No.4, pp.413-423 (2014).

[7] C. B. Moon, H.S. Kim, D. W. Lee, B. M. Kim, Mood Lighting System Reflecting Music Mood, COLOR research and application, Vol.40, No.2, pp.201-212 (2015).

[8] Ness, S. R., Theocharis, A., Tzanetakis, G. and Martins, L. G., Improving automatic music tag annotation using stacked generalization of probabilistic svm outputs, Proc. of the 17th ACM international conference on Multimedia, pp.705-708 (2009).

[9] Laurier, C., Sordo, M., Serra, J. and Herrera, P., Music mood representations from social tags, Proc. of the 10th International Society for Music Information Conference, Kobe, Japan, pp.381-386 (2009).

[10] Kim, J., Lee, S., Kim, S. and Yoo, W. Y., Music mood classification model based on arousal-valence values, Proc. of 13th International Conference on Advanced Communication Technology (ICACT), 2011, pp.292-295 (2011).