Profile Generation from TV Watching Behavior Using Sentiment Analysis

Yasufumi Takama
Tokyo Metropolitan University
ytakama@sd.tmu.ac.jp

Yuki Muto
Hitachi Information Systems

Abstract

This paper proposes a method for generating user profile from user's TV watching behavior using sentiment analysis. Personalized technologies such as information recommendation are currently hot topic of Web intelligence. Among them, TV program recommendation is expected to be one of the practical applications in near future, as digital TV service and partner robots providing personalized support are getting into our living environment. The proposed method does not estimate user's interest in a TV program only from its watching time as most of existing methods do, but also from user's utterances by applying sentiment analysis. The method employs fuzzy inference for estimating the user's rating of a TV program. A user profiles with bookmark format is generated based on the estimated rating of TV programs. Experiments are performed by collecting TV watching logs with diary-based approach, and the results show the proposed method can generate a user profile that can reflect user's interests.

1. Introduction

Personalized technologies such as information recommendation and relevant feedback are currently hot topic of Web intelligence. Among them, TV program recommendation is expected to be one of the practical applications in near future, as digital TV service and partner robots are getting into our living environment. Digital TV service provides us with vast number of TV programs, and we could not fully enjoy that without intelligent support, i.e., personalized TV program recommendation. On the other hand, it is said that TV is bound up with the ordinary, natural rhythms of our daily life [7]. Therefore, TV could be a base for establishing friendly relationship between humans and partner robot / home robots [5, 6], which expect to co-exist in our living environment.

This paper proposes a method for generating user profile from user's TV watching behavior. The method generates positive and negative profile with bookmark format for each user, based on iEPG and estimated rating of watched TV programs. The method does not estimate user's interest in a TV program only from its watching time as most of existing methods do, but also from user's utterances during watching the program by applying sentiment analysis. The method employs fuzzy inference for estimating the user's rating of a TV program.

Experiments are performed and the results show the proposed method can generate a user profile that can reflect user's interests.

2. Related works

TV program recommendation systems have been studied for realizing personalized TV guides [2, 4, 8, 9, 10]. In order to recommend appropriate TV programs, a user profile that reflects his / her preference on selecting TV programs should be estimated. The typical features used for generating user profile are watching time and attributes of TV programs [2, 8, 9, 10]. It is assumed that a user watches a TV program for a long time if s/he found it interesting, while a user switches to another channel when s/he found it uninteresting. Attributes of TV programs, such as genres and performers (actors), are used to estimate common features of TV programs that a user is interested in. Such attributes can be obtained from metadata, such as TV-anytime [2 9, 10], and iEPG (Internet Electronic Program Guide).

Users' behaviors of watching TV programs (viewing behaviors) have also been investigated [1, 4, 7]. In the studies, various types of viewing behaviors, such as concerned viewing and diversion viewing, are observed according to period of time in a day [4, 7]. It is also observed that user's watching behavior in weekday is more stable than that in weekend [1, 7].

3. Profile generation algorithm

3.1. Profile format

This paper employs a user profile that has the form of bookmark, which is employed by various applications such as ordinary web browsers. Figure 1 shows an example of the user profile. The user profile has a 3-layered structure, and stores the URIs of watched TV programs

0-7695-3028-1/07 $25.00 © 2007 IEEE
DOI 10.1109/WI-IATW.2007.17

as leafs. Each user has two profiles with the same format; positive and negative profile. Positive profile stores TV programs a user is interested in, and negative profile stores TV programs s/he is not interested in. The first layer has 15 predetermined categories, each of which corresponds to a genre of TV programs used in iEPG description, such as drama, news, sports, and music. Each category in the first layer can have multiple subcategories as the second layer. Each subcategory represents an arbitrary topic, which characterizes TV programs in it. The example of subcategories is performers, baseball teams, etc. A TV program must belong to only one category, while it can belong to multiple subcategories. The detail of generating subcategories is described in Section 3.4.

Each URI in a profile has several attributes such as estimated rating and flag of recommendation. The estimated rating of a TV program is determined based on the algorithm that is proposed in Section 3.2 and 3.3. A flag of recommendation of a TV program is 'true' if it was recommended by the system, otherwise 0.

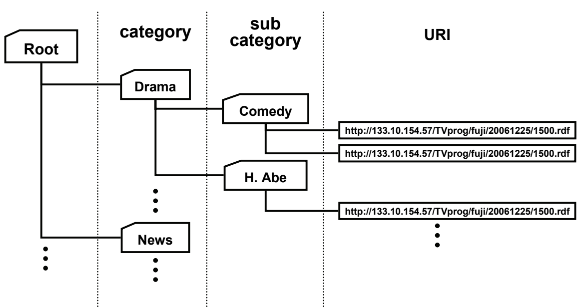

Figure 1. Profile format

3.2. Evaluation of watched TV program

In our study, user's selection of TV channels and his/her utterances during watching TV are stored as the log of the user's TV watching behavior. From the log, the rating of a TV program by the user is estimated based on the watching time and utterances, which is calculated by dividing into the following 3 scores.

- $Score_w(p)$: a score based on watching time
- $Score_{SO}(p)$: a score based on sentiment expressions
- $Score_f(p)$: a score based on frequency of utterances

The score of a TV program p based on watching time is defined by Eq. (1), where $wtime(p)$ and $airtime(p)$ indicate the amount of time (minutes) a user watched and its airtime, respectively.

$$Score_w(p) = wtime(p) / airtime(p). \quad (1)$$

It is often assumed that user's utterances during watching a TV program reflect his/her impression and evaluation on the program. For example, a user would say "I do not like [a performer's name]." if the performer s/he does not like appeared in the program. A user would also say "Great! Daisuke won!" when his favorite pitcher won a baseball game. Such utterances contain sentiment expressions [3], which can be used for estimating program's rating as well as for generating subcategories as proposed in Section 3.4. This paper tries to extract the triple of <*program, object, SO-score*> from sentiment expressions. A SO-score (semantic orientation score) represents the degree of positive / negative sentiment of a user against an object in the program. A dictionary of sentiment expression is prepared, in which a word (sentiment expression) is stored with its SO-score (ranging from -1.0 to 1.0). A SO-score of an object o, $SO(o)$, is calculated based on the SO-scores of the sentiment expressions that are co-occurred with o in the user's utterances. Equation (2) defines $SO(o)$, in which $S_e(o)$ indicates a set of sentiment expressions e_i that co-occur with o in utterances, and $SO(e)$ is SO-score of e.

$$SO(o) = \frac{2}{\pi}\arctan\left(\sum_{e_i \in S_e(o)} SO(e_i)\right). \quad (2)$$

The score of a program p based on sentiment expressions, $Score_{SO}(p)$, is calculated by Eq. (3), where $S_o(p)$ is a set of objects appeared in the utterances during watching p, N_p is the sum of frequencies of all objects ($S_o(p)$) in the utterances, and $N(o)$ is the frequency of o in the utterances.

$$Score_{SO}(p) = \sum_{o_i \in S_o(p)} \frac{N(o_i)}{N_p} SO(o_i). \quad (3)$$

Finally, $Score_f(p)$ is defined by Eq. (4), which is based on the assumption that a user says something frequently while watching p, if s/he is interested in it.

$$Score_f(p) = \frac{F(p)}{\text{ceiling}(wtime(p)/10)}, \quad (4)$$

where $F(p)$ is the number of utterances during watching p, and ceiling(x) is a ceiling function. The score indicates the frequency of utterances per 10 minutes.

3.3. Fuzzy inference for rating estimation

We suppose that watching time and the frequency of utterances do not reflect directly user's interest in a TV program, but his/her attention to it. For example, a user would find a TV program uninteresting after he watched it. It would also be possible that a user frequently complain the program or performers in the program. Therefore, this paper proposes to estimate the user's rating of a TV program by 2-step fuzzy inference. We employ fuzzy inference because of its easiness for describing inference rules. At the first step, the degree of attention by a user to a program, $Score_a(p)$, is calculated from $Score_w(p)$ and $Score_f(p)$. Figure 2 shows the membership functions and rule matrix used for calculating $Score_a(p)$. As for fuzzy inference, Min-max composition and center of gravity method for defuzzification are used.

At the second step, the estimated rating of a program, *Score*(*p*) is calculated from $Score_a(p)$ and $Score_{SO}(p)$. Figure 3 shows the membership functions and rule matrix used for the calculation. As $Score_{SO}(p)$ can have negative value, the value of *Score*(*p*) ranges from -1.0 to 1.0.

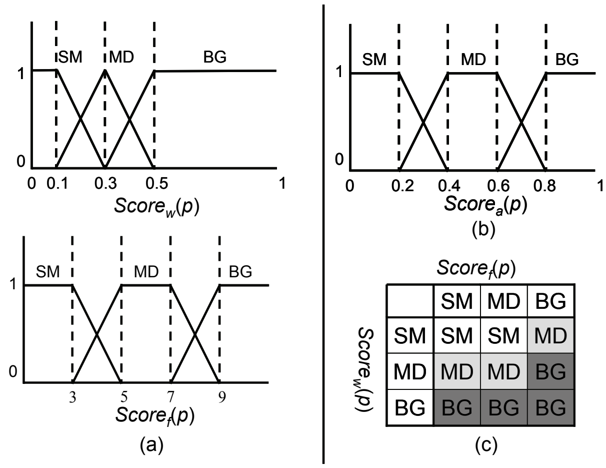

Figure 2. Fuzzy inference for degree of attention: Membership functions for antecedent (a) and consequent (b), rule matrix (c)

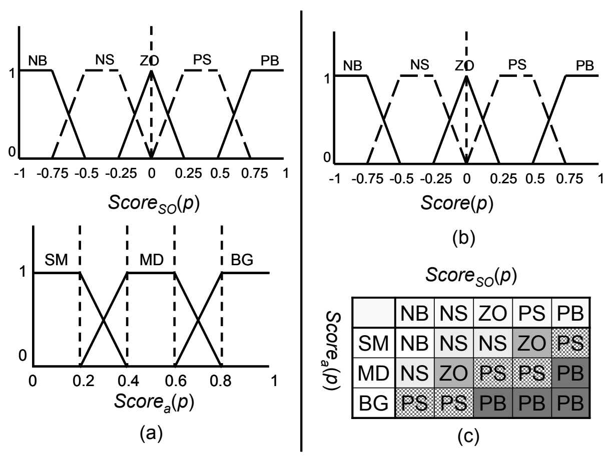

Figure 3. Fuzzy inference for rating: Membership functions for antecedent (a) and consequent (b), rule matrix (c)

3.4. Generation of category

A user profile is generated and updated periodically, based on the corresponding user's log of watching behavior. Generation of a user profile is performed according to the following steps.

1. Classification of watched TV programs
2. Categorization of TV programs into categories
3. Categorization of TV programs into subcategories

In step 1, watched TV programs are classified into 2 sets, S_p and S_n. The S_p contains TV programs of which estimated rating *Score*(*p*) is 0 or positive, while S_n contains those with negative rating. Positive (negative) profile is generated from S_p (S_n) with step 2 and 3. In step 2, TV programs are categorized into 15 categories as noted in Sec. 3.1. The category of a TV program is determined based on the "genre" metadata in its iEPG description.

In step 3, TV programs in each category are further categorized into subcategories. As noted in Sec. 3.1, a TV program can belong to multiple subcategories in a category. The title of a TV program, performers in the program, and keywords that are appeared in the "summary" metadata in iEPG are used as attributes of the TV program. As our study focuses on Japanese TV programs, morphological analysis is applied to extract nouns as keywords. Subcategories are selected from attributes in the following 2 ways. It is noted that positive profile contains only TV programs of which the *Score*(*p*) exceeds a threshold, while negative profile stores all programs in S_n.

Subcategory generation based on speech recognition:

When a TV program to be stored in a profile has an object *o*, of which *SO*(*o*) exceeds a threshold, the subcategory that corresponds to *o* is generated and the program is categorized into it.

Subcategory generation from common attributes:

It is supposed that TV programs that a user is interested in have some common features such as TV genre and favorite performers. This paper extracts an attribute as a subcategory, if the number of TV programs that has it exceeds a threshold in recent *N*-day log of a user. It is also noted all programs in S_p are considered in this step.

User profile is single daily updated by merging a temporal profile into current user profile. A temporal profile is generated from today's log, except that "subcategory generation from common attributes" use recent *N*-day log as well. Newly-appeared TV programs and subcategories are merged into current profile, while some existing TV programs are removed. The rating of the stored TV program is updated according to Eq. (5), where τ is half period and Δd is elapsed day after being stored in the profile. The program is removed from the profile if its score gets lower than a threshold. The subcategory that contains no TV program in it is also removed.

$$Score(p) \leftarrow Score(p) \cdot e^{-\frac{\log 2}{\tau}\Delta d}. \quad (5)$$

4. Experimental Results

Due to the limitation on our current experimental environments, experiments are performed based on the data collected with diary-based approach. We asked 20 test subjects to note TV programs they watched from December 25, 2006 to January 7, 2007. Each subject noted watched TV programs along with the watching time, rating of the program (11 scale: -5 to +5), and comments on it. Comments are used to extract sentiment expressions, instead of applying speech recognition to utterances.

Table 1 shows summary of the generated profiles: the number of watched TV programs, the numbers of subcategories in positive (POS) and negative (NEG) profiles.

Table 1. Summary of generated profiles

Sbj	# of programs	# of subcategories (POS)	# of subcategories (NEG)
A	97	31	4
B	65	21	0
C	51	22	0
D	51	52	0
E	49	19	2
F	47	29	0
G	39	62	0
H	31	18	3
I	31	5	1
Total	461	259	10

In order to evaluate generated subcategories, 10 subcategories are selected from each positive profile (except subject I) in descending order of estimated rating, and evaluated by the corresponding subject. All subcategories in positive profile for subject I and in negative profiles for all subjects are evaluated in the same way. Table 2 summarizes the evaluation result, in which subcategories are classified into 4 types according to test subjects' evaluations, such as follows.

Type 1: A subcategory is appropriate for characteristic of watched programs and reflects a subject's interest.
Type 2: A subcategory is appropriate for characteristic of watched programs but does not reflect a subject's interest.
Type 3: A subcategory is not appropriate for characteristic of watched programs but reflects a subject's interest.
Type 4: A subcategory is neither appropriate for characteristic of watched programs nor reflect a subject's interest.

In Table 2, 72% of subcategories in positive profile are evaluated as appropriate (type 1 and 2), while negative profile contains only subcategories test subjects are not interested in. This result shows that the proposed method can generate user profiles that correctly reflect subjects' interests.

9. Conclusion

This paper proposed a method for generating user profile from user's TV watching behavior. In addition to watching time, the method makes use of user's utterances during watching a TV program by applying sentiment analysis. A user profiles with bookmark format is generated based on user's ratings of TV programs, which are estimated by fuzzy inference. Experiments are performed by collecting TV watching logs with diary-based approach, and the results show the proposed method can generate a user profile containing subcategories that reflect user's interests. Future plan includes the experiment by using speech recognition module.

Table 2. Evaluation of subcategories

Type	Positive profile	Negative profile
1	50(0.59)	0(0.0)
2	11(0.13)	0(0.0)
3	2(0.02)	0(0.0)
4	22(0.26)	10(1.0)

References

[1] N. Daita, J. M. Pires, M. Cardoso, and H. Pita, "Temporal Patterns of TV Watching for Portuguese Viewers," *2005 Portuguese Conference on Artificial Intelligence*, pp. 151-158, 2005.

[2] Y. B. Fernandez, J. J. P. Arias, M. L. Nores, A. G. Solla, and M. R. Cabrer, "AVATAR: An Improved Solution for Personalized TV based on Semantic Inference," *IEEE Transactions on Consumer Electronics*, Vol. 52, No. 1, pp. 223-231, 2006

[3] T. Inui and M. Okumura, "A Survey of Sentiment Analysis," *Journal of Natural Language Processing*, Vol. 13, No. 3, pp. 201-241, 2006, in Japanese.

[4] T. Isobe, M. Fujiwara, H. Kaneta, N. Uratani, and Toshiya Morita, "Development and Features of a TV Navigation System," *IEEE Trans. on Consumer Electronics*, Vol49, No.4, pp.1035-1042, 2003.

[5] N. Matsuhira and H. Ogawa, "Trends in Development of Home Robots Leading Advanced Technologies," Toshiba Review, Vol. 59, No. 9, 2004 (written in Japanese).

[6] Y. Muto, Y. Iwase, S. Hattori, K. Hirota, and Y. Takama, "Web Intelligence Approach for Human Robot Communication under TV Watching Environment," SCIS&ISIS2006, pp. 426-429, 2006.

[7] A. Taylor and R. Harper, "Switching On to Switch Off: a Analysis of Routine TV Watching Habits and Their Implications for Electronic Programme Guide Design," usableiTV, 1, pp. 7-13, 2002.

[8] H. Zhang, S. Zheng, and J. Yuan, "A Personalized TV Guide System Compliant with MHP," *IEEE Transactions on Consumer Electronics*, Vol. 51, No. 2, pp. 731-737, 2005.

[9] H. Zhang and S. Zheng, "Personalized TV Program Recommendation based on TV-Anytime Metadata," *Proc. of the Ninth International Symposium on Consumer Electronics (ISCE 2005)*, pp. 242-246, 2005

[10] Y. Zhiwen and X. Zhou, "TV3P: An Adaptive Assistant for Personalized TV," *IEEE Transactions on Consumer Electronics*, Vol. 50, No. 1, pp. 393-399, 2004.