Employing User-Generated Tags to Provide Personalized as well as Collaborative TV Recommendations

Andreas Thalhammer
Semantic Technology Institute
University of Innsbruck
Technikerstraße 21a
6020 Innsbruck, Austria
andreas.thalhammer@sti2.at

Günther Hölbling
Chair of Distributed Information Systems
University of Passau
Innstraße 43
94032 Passau, Germany
hoelblin@fim.uni-passau.de

Dieter Fensel
Semantic Technology Institute
University of Innsbruck
Technikerstraße 21a
6020 Innsbruck, Austria
dieter.fensel@sti2.at

ABSTRACT

Within the Web, the annotation of content has become a common way to provide efficient navigation and recommendation of resources. In the future, TV sets with integrated Web capabilities will offer tagging as a tool for content organization in the realm of home entertainment. The recommendation of TV content is a challenging task as a system has to consider each user's individual preferences without getting too specific. We present a strategy which employs user-generated tags in a flexible way to address this issue. Our approach provides two different ways of semantic ranking for TV program lists: The first allows a higher ranking of programs that fit well to the user's personal likings. The second introduces collaborative aspects and therefore promotes a community-driven approach rather than an individual way of recommendation.

Categories and Subject Descriptors

H.3.3 [**INFORMATION STORAGE AND RETRIEVAL**]: Information Search and Retrieval—*Information filtering*; H.5.1 [**INFORMATION INTERFACES AND PRESENTATION**]: Multimedia Information Systems

General Terms

HUMAN FACTORS, MEASUREMENT

1. INTRODUCTION

In recent years, the fusion of television and the Web has already begun. In this context, the integration of content from the Web into television and vice versa are two important and not yet completed tasks. Considering the characteristics of both information sources, the following turns out: while television is consumed mostly passively, Web content usually offers a high degree of user interaction. However, in the next years, this distinction will become more and more blurred. In particular, television will offer common ways of interaction that are currently only well known from the Web, especially social annotation of content.

Our approach applies user-generated tags in order to provide recommendation of TV content. As a result of an information filtering process, we provide two rankings of a program list, each of which is based on the same data - but employs different ways of user modeling. The personalized ranking focuses on the semantic similarity between the user's preferences for certain tags and the annotations of upcoming programs. The collaborative ranking measures the similarity between tag clouds in the same way, but from a community point of view as it considers tags from other users as well.

Both approaches are meant to address the problem of overspecialization (that occurs in solely content-based systems [1]) through social discovery: tags are user-generated and describe the semantics of an item. As a matter of fact, the semantics of a TV program do not necessarily correlate with the descriptions from the metadata.

Note that, in this paper, the terms "personalized" and "collaborative" are used in the context of social annotation.

2. TAG-BASED TV RECOMMENDATION

The field of recommendation uses two common kinds of ratings, implicit (extracted from user transactions) and explicit (the user is explicitly asked) ones [1]. As for the latter, instead of using the common numerical ratings, Sen et al. [4] suggest using tags in order to provide a more individual and accurate way of expressing what users like about a specific item. Thus, by employing tags, we switch from the "degree of the user's preference" point of view to the level of "what actually are the user's preferences (in her own words)".

In [5], it is suggested that resource recommendation should be performed by applying traditional collaborative filtering methods on user-item, user-tag, and item-tag datasets. In our system we refine this idea for the television domain by focusing on the item-tag data in combination with different representations of a single user profile.

2.1 Input data

As input for the recommendation approach, we consider two entities: The first is an upcoming program, which has not been tagged yet. The second is a user profile that contains the history of previously watched programs along with the tags assigned by the users.

Finding tags for an upcoming program, which is a candidate for recommendation, is a non-trivial task as users commonly assign tags after and not before media consumption. There exist various options to tackle this problem: Keywords can be extracted from the program descriptions (as it is done by tvister[1]) and reused as tags. Furthermore, a

[1]tvister - `http://www.tvister.de/`

professional team could tag upcoming programs in advance. It is clear that the creation of tag clouds in both of these ways differs from the dynamic process of community tagging. In [2], we found a feasible way to address this issue by applying a machine learning approach in combination with a client-server architecture. We use this approach in order to provide an efficient prediction of tags that are very similar to the ones that real users assign. Table 1 exemplifies the result of our tag prediction step by showing generated tags and their weights for three TV programs.

For the personalized and the collaborative part of recommendation, we use two different representations of the same user profile (containing the tagging history). To accentuate this, we refer to table 2, which shows a small user profile that is present in our dataset [2]. The tag cloud of each TV program contained in the user history is presented in different ways. The personalized approach does not consider tags from other users, but only the ones the current user assigned. This results in a binary representation, as a user either assigns a particular tag for a program, or not. To address this issue, we weight each tag by the user's individual preference for it (total number of usages in her profile). An example is provided by the left side of table 2. The collaborative approach incorporates the tags of all users that annotated one specific program and weights each tag by its total number of usages for a specific program. This way of presenting tag clouds is the most common one within the Web. The right column of table 2 exhibits an example of this notation.

The comparison of upcoming programs with the personalized version of the user profile and also with the collaborative one, results in two different program rankings.

In the following, we refer to the user profile as either the personalized or the collaborative representation.

2.2 Similarity Measure

A similarity measure is used to compare the tag clouds of the programs in the user profile to the ones of the upcoming programs. We represent tag clouds as vectors in a similar way as items are represented as vectors of user ratings in the collaborative filtering domain [1]: each dimension represents a single tag and each entry denotes the respective weight. These vectors can be compared to each other by measuring their degree of similarity. With the use of generated tags [2] we discovered a discrepancy between the tag weights of the upcoming programs (predicted) and the ones of the programs in the user profiles (accumulated). This deviation is related to the different perception of the same rating scale in user-item scenarios that is explained in [3]. Therefore, we decided to employ Pearson correlation as a similarity measure to mitigate this effect. For two programs $p, q \in P$, having attached the tags T_{pq} with the weight w, this results in the following formula:

$$sim(p,q) = \frac{1}{2}\left(\frac{\sum_{t \in T_{pq}} (w_{p,t} - \bar{w}_p)(w_{q,t} - \bar{w}_q)}{\sqrt{\sum_{t \in T_{pq}} (w_{p,t} - \bar{w}_p)^2 \sum_{t \in T_{pq}} (w_{q,t} - \bar{w}_q)^2}} + 1\right)$$

The value $\bar{w}_k$ stands for the average tag weight of the program k:

$$\frac{1}{|\vec{k}|} \sum_{t \in T_k} w_{k,t}$$

Note that the resulting similarity score lies between 0 and 1 with a neutral point at 0.5 and equality at 1.

2.3 Score Aggregation

After having measured the similarity of the new program to all programs in the user profile, we need to aggregate these scores to a final one for each upcoming program. It does not make sense to aggregate the similarity scores of all programs in the profile as users do often like more than one program genre. Hence, we only aggregate the similarity scores of the k-nearest neighbors (k-NN) in order to aim at specific genres the user prefers. This helps to obtain more accurate results as inter-genre measurements often result in a neutral similarity score that would be incorporated in every aggregation. For the aggregation of the scores of the k-NN, we use a weighted average approach with the scores

Table 1: Predicted tags and their weights for upcoming programs.

Alarm für Cobra 11 - Die Autobahnpolizei	action (3.937), polizei (1.999), krimi (1.995), spannend (1.990), autos (0.989), aufregend (0.898), aktion (0.896), serie (0.879)
Asterix - Sieg über Cäsar	film (1.617), comic (1.613), geschichte (1.597), spielfilm (0.859), spass (0.859), comik (0.859), lustig (0.859), zeichentrick (0.858)
Die Simpsons	zeichentrick (6.357), homer (3.653), comedy (3.552), kult (3.535), lustig (3.434), humor (2.529), serie (1.957), cartoon (1.773), simpsons (1.711), entspannung (1.549), amerika (1.498), fun (0.761), marge (0.899), james brooks (0.824), chillen (0.750), neue folgen (0.736), bart (0.726)

Table 2: Two different representations of a single user profile.

Program	Personalized	Collaborative
Die Simpsons	comic (1), satire (2), spass (1)	zeichentrick (1), satire (1), lustig (1), comic (1), spass (1)
Broken comedy	fun (1), lustig (3), satire (2)	fun (1), lustig (1), satire (1)
Navy CIS	spannend (2)	gerichtsmedizin (1), spannend (1), ncis (1), navy (1)
Die Simpsons	cartoon (1), lustig (3)	cartoon (1), lachen (1), chillen (1), lisa (1), homer (1), lustig (1), kult (2), bart (1)
Stargate	sci-fi (1)	sci-fi (1)
Verführung einer Fremden	spannend (2), thriller (1)	spannend (1), thriller (1)
switch reloaded	lustig (3) , verarsche (1)	parodie (1), verarsche (1), satire (1), lustig (1), tv (1)

Table 3: Personalized vs. Collaborative: 3-NN and the aggregated scores (gray) of the upcoming programs.

	Personalized	Collaborative
Die Simpsons	Die Simpsons (0.677) switch reloaded (0.651) Broken Comedy (0.636)	Die Simpsons (0.742) Die Simpsons (0.702) Broken Comedy (0.611)
	0.655	0.690
Asterix - Sieg über Cäsar	Die Simpsons (0.648) Die Simpsons (0.619) switch reloaded (0.619)	Die Simpsons (0.776) Broken Comedy (0.572) switch reloaded (0.554)
	0.629	0.650
Alarm für Cobra 11 - Die Autobahnpolizei	Navy CIS (0.679) Verführung einer Fremden (0.659) Stargate (0.499)	Navy CIS (0.588) Verführung einer Fremden (0.626) Stargate (0.499)
	0.623	0.576

being the weight (what results in squaring the similarity scores). For the programs within the k-nearest neighbors $kNN \subseteq Profile$ and the upcoming program p_{new} this leads to the following formula:

$$agg(p_{new}) = \frac{1}{\sum\limits_{p \in kNN} sim(p_{new}, p)} \sum_{p \in kNN} sim(p_{new}, p)^2$$

The aggregated score of an upcoming program can be interpreted as the user's degree of preference for it. In our approach, this value is used to provide a ranking within the list of upcoming programs.

3. PROOF OF CONCEPT

By using the upcoming programs of table 1 and the profile of table 2 as input data, we do now exemplify how the aforementioned profile representations can provide different scores and rankings. It needs to be pointed out that, for the reasons of clarity and brevity, the chosen user profile is very small (only seven tagged programs) and also the short list of upcoming programs does not relate to a real case scenario (usually more than 200 concurrent programs).

The personalized as well as the collaborative rankings, shown in table 3, demonstrate that a top-N recommendation is possible with only few ratings. By considering tags, similarities between TV programs can be determined although they are not strongly correlated through content or metadata. In our case, the Asterix movie nearly gets the same score (on both sides) as the upcoming Simpsons episode although it has no direct correlation (through TV metadata or content) to one of the user's previously watched programs. In contrast, the upcoming Simpson episode does have this link: the user has already watched two episodes before. Therefore the reasonably high score of the Asterix movie indicates that, even with a small profile, the use of tags as semantic descriptors might help to overcome the common problem of overspecialization. This also underlines our efforts to provide collaborative semantic tag prediction [2].

It is apparent that the ranking of the three programs in table 3 is the same for the personalized and for the collaborative representation of the user profile. However, as the differences between the scores in both lists indicate, the ranking would strongly differ taken a larger and more realistic number of upcoming programs (> 200) into account. The similarity scores of the collaborative ranking highlight the community factor of the ranking. The personalized part of the recommender system highly relies on the user's taste and therefore implements her individual preferences.

For a single top-N listing, it is possible to linearly combine both types of scores for each program.

4. CONCLUSION AND OUTLOOK

This paper presents two feasible and promising approaches to provide top-N recommendations through collaborative tagging. Moreover, it is demonstrated that the utilization of user-generated tags might help to overcome the problem of overspecialization in the emergent domain of TV recommendation.

For the future work, we plan to conduct a thorough evaluation of the proposed approach that also includes a user survey. Furthermore, the similarity measurements can be enhanced through lemmatization of tags in combination with ontology matchings between tag clouds.

5. REFERENCES

[1] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Trans. on Knowl. and Data Eng.*, 17:734–749, June 2005.

[2] G. Hölbling, A. Thalhammer, and H. Kosch. Content-based tag generation to enable a tag-based collaborative TV-Recommendation System. In *8th Int'l Conf. on Interactive TV&Video*, pages 273–282, 2010.

[3] T. Segaran. *Programming collective intelligence.* O'Reilly, 2007.

[4] S. Sen, J. Vig, and J. Riedl. Tagommenders: connecting users to items through tags. In *18th Int'l Conf. on World Wide Web*, pages 671–680, 2009.

[5] K. H. L. Tso-Sutter, L. B. Marinho, and S.-T. Lars. Tag-aware recommender systems by fusion of collaborative filtering algorithms. In *Proc. of the 2008 ACM symposium on Applied computing*, SAC '08, pages 1995–1999, 2008.