International Journal of Multimedia and Ubiquitous Engineering
Vol.9, No.5 (2014), pp.231-240
http://dx.doi.org/10.14257/ijmue.2014.9.5.23

Personalized TV Contents Recommender System Using Collaborative Context tagging-based User's Preference Prediction Technique

Haesung Lee[1], Joonhee Kwon[2]

[1,2]Department of Computer Science, Kyonggi University
San 94-6, Yiui-dong, Yeongtong-ku, Suwon-si, Gyeonggi-do, Korea
{seastar0202, kwonjh}@kgu.ac.kr

Abstract

Increasing amounts of TV contents are being disseminated through advanced broadcast, satellite and smart TV technologies leading to an information overload. Despite the excessive increase of TV contents, there are few considerations in the development of novel methods for personalized smart TV content recommender services. Most existing personalized TV content services are mainly focus on using individual profiles and require user's immediate participation in rating their experienced contents. In this paper, we propose the context tagging-based user's preference prediction mechanism by extending the widely known recommender algorithm, collaborative filtering (CF) in order to increase the user's satisfaction about the recommender service. The development of the prototype system shows the usefulness of proposed mechanism. And the experiment confirms that the proposed mechanisms improve the recommender system.

Keywords: *Personalized recommender system, context awareness, social network analysis, clustering, collaborative filtering, smart TV*

1. Introduction

With the TV channel explosion, many people face an exhausting task of having to find something to watch on TV that fits their interest. This means that viewer receive much more information than they can actually manage, which many lead them to believe that they are missing programs that could likely interest them.

Recently, the TV contents recommender system has recently been issued due to the diffusion of Smart TVs and TV-related content explosion. The major motivation for the realization of the TV contents for each user from a large number of available selections. Recent many studies have proposed a collaborative filtering-based TV contents recommender system [1-3]. The major feature of these studies is based on a user history of past viewing data. Recommendations are provided by matching the prediction of user's preferences with TV contents. In other words, TV contents that are similar to those the user preferred in the past are recommended. Therefore, CF-based TV contents recommender systems completely rely on profiles of viewers and TV contents respectively.

However, the CF-based mechanism has widely known problems. The so-called *grey-sheep* problem caused difficulties when a particular user's ratings history does not help the system identify a set of similar users. And, the *cold-start* problem is caused by new users in the system which has not submitted any ratings. Thus, the system is not able to infer user user's preferences and generate recommendations until a few items have been rated. In addition, for a CF-based system to work well, several users must evaluate each item which causes user's inconvenience. Moreover, it is common for the user-item matrix

ISSN: 1975-0080 IJMUE

to be sparsely populated, making it difficult to identify similar users and contents, due to a lack of ratings overlap.

In this paper, we propose the context tagging-based user's preference prediction mechanism by extending the widely known recommender algorithm, collaborative filtering (CF). Our proposed mechanism solves common problems of CF-based systems. In addition, the prediction of user's preference on TV contents is more accurate than the existing recommender system.

The rest of this paper is organized as follows. In section 2, previous studies concerning personalized TV contents recommender systems and CF-based systems are described. Also, the use of social network and a clustering technique which are mainly concerned in this work are explained as background. Section 3 describes the detail of collaborative context tagging-based user's preference prediction algorithm. In Section 4, the architecture of the TV content recommender system which applied with proposed recommender mechanism is described. And we present the implementation of TV content recommender system prototype. In section 5, the evaluation results are described. Finally, Section 6 concludes and points out direction of future work.

2. Related Works and Background

Recommendation techniques come in two common methods, *collaborative filtering* (CF) and *content-based filtering*. Recently, many researches try to build recommender systems that help the user in filtering TV contents using these common basic recommendation methods [4-6].

2.1 TV contents recommender systems

Among previous TV contents recommender systems, the TV advisor developed by Das and Horst makes use of explicit techniques to generate recommendations for a TV viewer [7]. Milan Bjelica determined the most important factors of a TV contents recommender and proposed an analyzing model for the estimation of user interest based on a content-based approach [8]. SeungGwan Lee *et al.* suggested a personalized TV contents recommender system for the cloud computing environment [9]. These studies are based on a content-based approach involving contents and user attributes (actor, genre, running time, program title, time, and date) of contents viewed by each viewer in the past. Ana Belen *et al*. took a hybrid content-based and item-based collaborative filtering approach to recommend TV contents [10].

But these previous researches which commonly consider recommender techniques require the user to take the initiative and explicitly specify their interest, in order to get high quality recommendations causing TV viewer's inconvenience. Besides, they have few consideration about solving widely known recommender problems, *grey-sheep* and *cold-start* which are main cause of .reducing the performance of recommender systems. In this paper, we aim to solve these problems of recommender techniques by extending CF techniques.

2.2 Context awareness for personalization

Context was initially defined as the location of the user, the identity of people near the user, the objects around, and the changes in these elements [11]. The importance of including and using the contextual information in recommendation systems was demonstrated in [12]. Similarly, Oku *et al*. incorporate additional contextual dimensions such as time, companion, and weather into the recommendation process and use machine learning techniques to provide recommendations in a restaurant recommender system [13]. They empirically show that the context-aware approach significantly outperforms the corresponding non-contextual approach.

In this paper, we use TV viewer's contexts for improving personalized recommendation accuracy and user's satisfaction with recommendations.

3. Preference Prediction based on Collaborative Context Tagging

In the domain of recommender systems, collaborative filtering (CF) is common methods for predicting user's preference on specific information or item. Generally, CF involves matching the rating of the current users for objects (*e.g.* TV contents or products) with those of similar users in order to produce recommendations for objects not yet rated or seen by the active user. Actually, however viewer's preference on TV contents is dynamically changed based on situations or contexts with which the viewer is faced. For example, the viewer who is a single woman in her 20s would prefers to watch TV contents about the latest fashion trends than TV contents about the trip to Paris during the Valentine's Day season. Therefore, it is needed to not only use the viewer's past viewing data but also consider the viewer's current context in order to more exactly predict viewer's preference. In this paper, we propose the context aware viewer's preference prediction technique through extending existing CF.

3.1. Collaborative context tagging model

Fig.1 shows the model defined for collaborative context tagging-based prediction of viewer's preference. The model is composed viewers, TV contents, user's ratings for TV contents and the context with which the viewer faced when he or she watches the TV content. The context tags describe the situation when the viewer watches the specific TV content. If the type of context tag is 'Season' and value is 'Valentine's Day, this context shows that the viewer is in Valentine's Day season.

$$C = \{C_1, C_2, ..., C_k\} \quad (1)$$

In the context tagging model, the types of contexts are varied. (1) shows the set of contexts tags which are tagged with the specific TV content R_i .

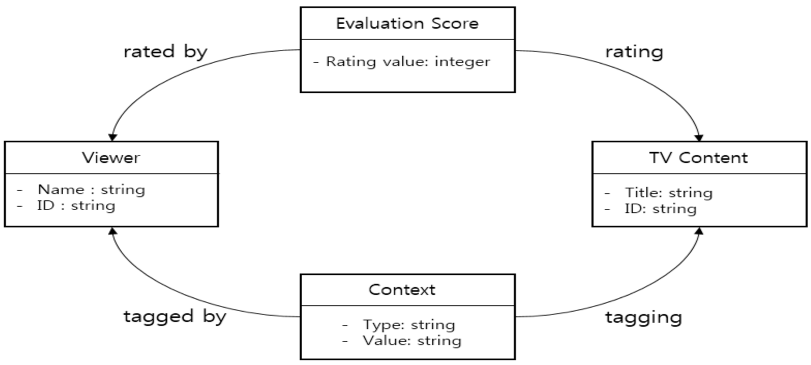

Figure 1. The model for collaborative context awareness-based viewer's preference prediction

That is, the number of context tags tagged with the TV content is more than one and the type of context tags are varied. C_t means the type of contexts, *e.g.* time, season, date or user's profile and $t \in 1 .... k$. The context tag is automatically tagged with the specific TV content through context tagging mechanism [14].

Context tagging mechanism captures viewer's current contexts such as user's ID, weather or date. And, the mechanism decides to whether the context is tagged with the specific TV content. For the decision of context tagging, we use arbitrary threshold.

Table 1 shows the automatic context tagging mechanism, Algorithm 1. The *TW* (Tagging Weight) defines the relative importance of the context tag of the specific TV content. If *TW* is higher than the arbitrary threshold, the context tag is tagged with the TV content. Through Algorithm 1, context tags automatically tagged with the TV content without viewer's intervention. And, it is possible defined context aware CF.

Table 1. Automatic Context Tagging Algorithm

Algorithm 1. *Automatic Context Tagging*

Input: R is the set of TV contents, C is the set of contexts, TH is the arbitrary threshold

i := 1;
while (i == the number of all contents in R)
  for each (C_j related with R_i)
  begin
    TW_{ij} = (the frequency of C_j related with R_i) /
        (the number of all contexts related with R_i);
    if ($TW_{ij} > TH$) **then** C_j is tagged with R_i
do i := i + 1
end while

3.2. Collaborative context awareness-based viewer's preference prediction

For predicting the specific user's preference, general CF composes similar users only based on past rating data. Therefore, existing CF can't consider dynamically changed preferences. For considering not only viewer's preference but also dynamically changed preferences, we additionally concern context similarity between viewers considering the context tagging model described in the Section 3.1. Proposed recommender technique is composed of three steps. Firstly, the preference similarity between viewers is computed. (2) is the equation for calculating the preference similarity between users. SIM_{ab} is the preference similarity between user a and b.

$$SIM_{ab} = \frac{\sum_{i=1}^{k}(r_{ai} - \overline{r_a})\cdot(r_{bi} - \overline{r_b})}{m} \quad (2)$$

$\overline{r_a}$ and $\overline{r_b}$ are the value of the average calculated from all scores rated by user a and b respectively. m is the average of total scores rated a and b. Secondly, the context importance of the specific TV content is computed. (3) is the equation for the user a's context importance, $CCW_{a,i,c}$ of the content, i, with the set of current contexts, C.

$$CCW_{a,i,c} = \sum_{k \in C} r_{aik} \cdot CW_{ik} \quad (3)$$

CW_{ik} is the importance of the content i of the context k. For example, if the content A is more frequently tagged with 'winter: season' than 'summer: season', the context tag 'winter: season' is more important with the content A. (4) is the equation for CW_{ik}.

$$CW_{ik} = \log \frac{N_i}{n_{ik}} \quad (4)$$

N_i is the total frequency of context tags tagged with the content, i. And n_{ik} is the frequency of the context tag, k tagged with the content, i. Finally, the viewer's preference prediction is calculated by converging *SIM* and *CCW*. (5) is the equation for the viewer's preference on the TV content concerning the collaborative context awareness-based prediction mechanism.

$$PF_{a,i,c} = \overline{r_a} + a\sum_{b=1}^{k}(CCW_{a,b,c} - \overline{r_b})\cdot SIM_{ab} \quad (5)$$

$PF_{a,i,c}$ means user a's preference on the TV content, i ,when current contexts is c. That is, with similar neighbors, the preference of the specific viewer in the situation can be predicted.

4. Prototype of Personalized TV Content Recommender Services

For showing the usability of our proposed recommender technique in the real smart home environment, we develop the prototype system which pervasively recommends TV contents to a TV viewer. Figure 2 presents overall architecture of developed recommender prototype.

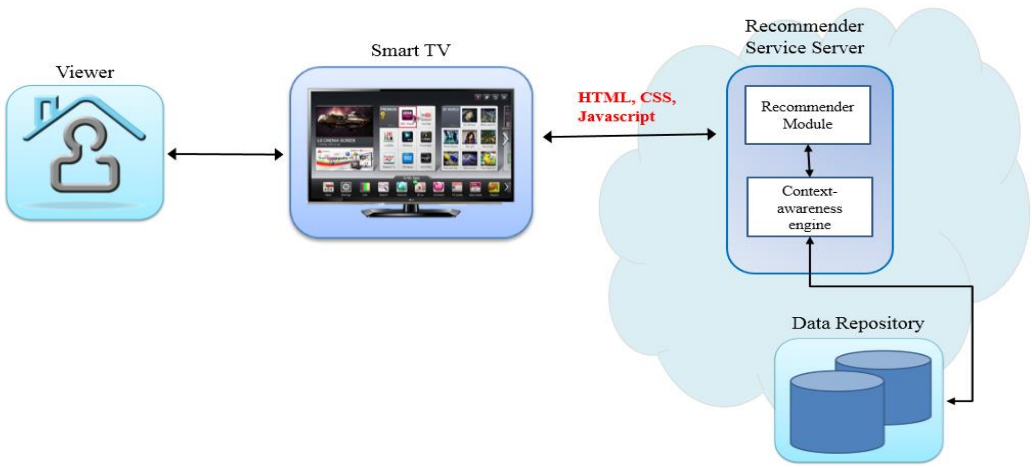

Figure 2. The Architecture of the Personalized TV Content Recommender System

The architecture of developed prototypes consists of a *Recommender Service Server* and Smart *TV Application*. *Smart TV Application* sends current viewer's ID to *Recommender Service Server*. Then the *Context-awareness engine* in the *Recommender Service Server* calculates the situation similarity based on data which are stored in *Data Repository* with context-awareness-based recommender model defined in Section 3.1.

Table 2. Implementation Environment

Hardware & Devices	Software & Implementing Tools
- Intel® Core™ i7-3770 CPU @ 3.40GHz	- IIS (Internet Information Service)
- 8GB RAM	- .Net Framework
- LG Smart TV 42 LA 6650	- C#, HTML, JavaScript, CSS, ASP
	- MS SQL Server 2008

The *Recommender Module* composes recommendations which are proper with a certain TV viewer's current contexts through Internet. A viewer can interact with the smart TV recommender application through a HTML5 enabled browser and the smart TV application is connected to Internet with back-end recommender system implemented with web technologies such as Ajax or XML programming. Table 2 describes the implementation environment with which the prototype system is developed. The smart TV application communicates with the recommender server through Internet.

For the demonstration of implemented prototype system, we arbitrarily define the context of the user 'Stephanie' as **[Context #1**: <u>Gender</u>: female; <u>Age</u>: thirties; <u>Marital Status</u>: single; <u>Season</u>: winter, <u>Weather</u>: raining, <u>Time</u>: morning**]**. Figure 3 is showing the developed prototype. As seen on dotted line-rectangles of Figure 3, the developed prototype recommends TV contents for the user 'Stephanie' with the context aware-based recommender technique proposed in this paper. In addition, the prototype provides TV contents which are viewed by people who have similar similarities with the user 'Stephanie'. Also, the prototype recommends TV contents with the consideration of time changes making the user's satisfaction improved. Therefore, the prototype developed with our proposed recommender technique could recommend more personalized TV contents efficiently.

Figure 3. The Prototype of the Personalized TV Content Recommender

5. Experiment

For the comparative evaluations between proposed recommendation technique and the general CF algorithm, we compose synthetically generated 1000 user's profile data, 1000 context tags, and 1000 movie contents. For these synthetic data, we randomly generated 1000 data. Each data consists of a user, a content, a rating value and contexts according to the context-tagging model defined in Section 3.1. In addition, we additionally develop a context generator module for the efficient experiment. The context generator randomly generates sets of contexts such like (1) defined in Section 3.1.

In this work, we many consider two general methods of the experiment. Many recommender systems use *Mean Absolute Error* (MAE) for evaluating the performance [15]. MAE between ratings and predications is a widely used in the domain of recommender systems. MAE is a measure of the deviation of recommendation from true user-specified rating values. For each rating-prediction pair $< p_i, q_i >$, the metrics treats the absolute error between this pair. The MAE is computed by first summing these absolute errors of the N corresponding rating-prediction pairs and then computing the average. Formally,

$$MAE = \frac{\sum_{i=1}^{N} | p_i - q_i |}{N} \tag{6}$$

Also, we use *Receiver Operating Characteristic* (ROC) sensitivity to measure decision-support accuracy. Decision-support accuracy measure how well predictions help users experience or accepts recommended contents. ROC curve is the plot sensitivity versus 1-specificity for a certain user. Sensitivity is defined as the probability that a good item is accepted by the user and specificity is defined as the probability that a bad item is rejected by the user. Therefore, we consider an item good if the user gave it a rating of 4 or above, otherwise we consider the item bad. For composing user rating for an item, we assume that if the item is tagged with a certain context tag, the user rating about the item is determined by tagging frequency of the situation tag. If the tagging frequency of the situation tag on a certain item is 4, then the user rating is 4. Additionally, in case of our proposed user's preference prediction mechanism, we consider the impact of applying context similarity in the accuracy of prediction about items which have high possibility to satisfy a certain user who needs for proper services in a specific situation. For this, we minus the fixed value, from MAE of our proposed CF only.

Table 3. Experiment Results

Algorithm	MAE	ROC
CF-based Recommendation System	1.41	0.6
Proposed recommendation technique	0.59	0.88

The results of our experiments are summarized in Table 3. With the results of the experiment, our recommendation method performs better than the general recommendation method in which pure CF is applied. All the difference in *MAE* is statistically significant. In case of the *ROC*, the result implies that our recommendation method, compared to the previous method, does a better of job of recommending personalized and authorized contents to the user who is faced with a specific contexts.

6. Conclusions and Future Works

Recently, the TV contents recommender system has recently been issued due to the diffusion of Smart TVs and TV-related content explosion. The major motivation for the realization of the TV contents for each user from a large number of available selections. Generally, many CF-based TV contents recommender systems completely rely on profiles of viewers and TV contents respectively. However, these TV contents recommender systems have many limitations caused from disadvantages of CF.

In this paper, we proposed the context tagging-based user's preference prediction mechanism by extending the widely known recommender algorithm, collaborative filtering (CF). With the result of the experiment, we know that our proposed mechanism solves common problems of CF-based systems. In addition, the prediction of user's preference on TV contents is more accurate than existing recommender systems calculated. However, we do not take the experiment with the large dataset without any consideration about the scalability of our proposed recommendation method. Therefore, we will improve our algorithm by considering the factor of scalability in the next stage of our work.

Acknowledgement

This work was supported by the Gyonggi Regional Research Center (GRRC) and Contents Convergence Software (CCS) research center.

References

[1] M. Bjelica, "Towards TV Recommender System: Experiments with User Modeling", IEEE Trans. Consum. Electron., vol. 56, Issue 3, **(2010),** pp. 707-712, pp. 1763-1769.

[2] SG Lee, D. Lee and S. Lee, "Personalized DTV program recommendation system under a cloud computing environment", *IEEE Trans. Consum. Electron.*, vol. 56, issue 2, **(2010),** pp. 1034-1042.

[3] B.-M. and A. Belén, "A hybrid content-based and item-based collaborative filtering approach to recommend TV programs enhanced with singular value decomposition", Information Sciences 180.22, **(2010),** pp. 4290-4311.

[4] B. Schilit and M. Theimer, "Disseminating active map information to mobile Hosts", IEEE Network, vol. 8, no. 5, **(1994),** pp. 22–32.

[5] G. Adomavicius, R. Sankaranarayanan, S. Sen and A. Tuzhilin, "Incorporating Contextual Information in Recommender Systems Using a Multidimensional Approach", ACM Transactions on Information Systems, vol. 23, no. 1, **(2005),** pp.103-145.

[6] K. Oku, S. Nakajima, J. Miyazaki, and S. Uemura, "Context-aware SVM for context dependent information recommendation", In Proceedings of the 7th international Conference on Mobile Data Management **(2007).**
[7] L. A. Adamic and E. Adar, "Friends and neighbors on the web", Social Networking, vol. 25, no. 3, **(2003),** pp. 211–230.
[8] D. M. Ehrlich, "Social network survey paper", Int. J. Learn. Intellectual Capital, vol. 3, no. 2, **(2006),** pp. 167–177.
[9] L. Garton, C. Haythorntwaite and B. Wellman, "Studying online social networks", J. Comput.-Mediated Commun. [Online], vol. 3, no. 1. http://jcmc.indiana.edu/vol3/issue1/garton.html **(1997)**.
[10] R. R. Yager, "Concept representation and database structures in fuzzy social relational networks", IEEE Trans. Syst., Man, Cybern. A, Syst., Humans, vol. 40, no. 2, **(2010),** pp. 413–419.
[11] P. Kazienko and K. Musiał, "On utilizing social networks to discover representatives of human communities", Int. J. Intell. Inf. Database Syst., vol. 1, no. 3/4, **(2007),** pp. 293–310.
[12] A. Bruns, "Methodologies for mapping the political blogosphere: An exploration using the IssueCrawler research tool", First Monday [Online], vol. 12, no. 5, http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1834/718 **(2007)**.
[13] K. Lewis, J. Kaufman, M. Gonzalez, A. Wimmer and N. Christakis, "Tastes, ties, and time: A new social network dataset using Facebook.com, Social Networking", vol. 30, no. 4, **(2008)**, pp. 330–342.
[14] R. Jin, W. Chen and T. W. Simpson, "Comparative studies of metamodelling techniques under multiple modelling criteria, Structural and Multidisciplinary Optimization, vol. 23, issue 1, **(2001)**, pp. 1-13.
[15] C. J. Willmott, and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance", Climate Research, vol. 30, **(2005)**, pp. 79−82.

Authors

Hae-Sung Lee, she is a Ph.D. candidate of Computer Science at Kyonggi University, Korea. Her research areas include Context-aware Computing, Social Network, Information Retrieval and Mobile Computing. She works for software development on areas of data search in ubiquitous environment. She received her B.S., M.S. in Computer Science from Kyonggi University,Korea. Contact her at seastar0202@ kyonggi.ac.kr.

Joon-Hee Kwon, she is an associate professor of Computer Science at Kyonggi University, Korea. She was a visiting research professor at the Computer Science Department at New Jersey Institute of Technology. Her research areas include Context-aware Computing, Information Retrieval, Social Network, Web 2.0 and Mobile Database. Her research projects focus on areas of d1ata search using social network and Web 2.0 in ubiquitous environment. She received her B.S., M.S. and Ph.D. in Computer Science from Sookmyung Women's University, Korea. Contact her at kwonjh@kyonggi.ac.kr.

