

Performance Improvement of a Movie Recommendation System based on Personal Propensity and Secure Collaborative Filtering

Woon-hae Jeong*, Se-jun Kim*, Doo-soon Park** and Jin Kwak***

Abstract—There are many recommendation systems available to provide users with personalized services. Among them, the most frequently used in electronic commerce is ‘collaborative filtering’, which is a technique that provides a process of filtering customer information for the preparation of profiles and making recommendations of products that are expected to be preferred by other users, based on such information profiles. Collaborative filtering systems, however, have in their nature both technical issues such as sparsity, scalability, and transparency, as well as security issues in the collection of the information that becomes the basis for preparation of the profiles. In this paper, we suggest a movie recommendation system, based on the selection of optimal personal propensity variables and the utilization of a secure collaborating filtering system, in order to provide a solution to such sparsity and scalability issues. At the same time, we adopt ‘push attack’ principles to deal with the security vulnerability of collaborative filtering systems. Furthermore, we assess the system’s applicability by using the open database MovieLens, and present a personal propensity framework for improvement in the performance of recommender systems. We successfully come up with a movie recommendation system through the selection of optimal personalization factors and the embodiment of a safe collaborative filtering system

Keywords—Collaborative Filtering, Movie Recommendation System, Personal Propensity, Security, Push Stack

1. INTRODUCTION

With the expansion of the Internet and the advent of smartphones, people now are able to get easy access to and extend their activities on the Web. The domestic Internet usage rate reached 78%, and the number of smartphone users totaled 30 million in 2012 [1]. Under these environmental changes, today’s users have an enormous amount of information within their reach and can see it increase exponentially. In the flood of such information, they want to have a means to search for desired information easily and quickly. In attempts to achieve this, portals have created and posted a list of search words that have been most sought, while shopping malls have provided a lineup of hot-selling items. Users want to search for their required information in the flood of information easily and quickly. To cope with these demands, portals have created and

Manuscript received October 9, 2012; accepted January 7, 2013.

Corresponding Author: Jin Kwak

* Dept. of Computer Science, Soonchunhyang Univ., Chungnam, Korea (onguri@gmail.com)

** Dept. of Computer Software Engineering, Soonchunhyang Univ., Chungnam, Korea (parkds@sch.ac.kr)

*** Dept. of Information Security Engineering, Soonchunhyang Univ., Chungnam, Korea (jkwak@sch.ac.kr)

posted a list of search words that have been most sought, and shops have provided information on most sought-after products in their catalogues. However, these methods do not seem to offer help with the user's decision-making due to their inability to reflect the characteristics of individual users. Personalized services mean providing consumers with products and services most suitable to individual tastes based on their personal information, but without explicit questions about their desired products and services [2]. Among personalized services, recommender systems make recommendations of services or products that target customers might like. A wide range of recommender systems have been developed and also widely adopted by such Internet shopping malls as Amazon and CD Now [3]. Among these, collaborative filtering systems have been known to be the most successful method, and have found a variety of applications in Web pages, movies, thesis, and newspaper articles [4-6].

Collaborative filtering, which is the most widely used technique in movie recommendation systems, is a method of making automatic recommendations of certain movies by creating profiles based on diverse kinds of information collected from multiple users, and subsequently making predictions based on these profiles, about the interests of a user who has preferences similar to other like-minded users. In a collaborative filtering system, users give preference ratings to items based on their taste. After that, the system calculates preference similarities among users from such ratings, and makes predictions about a user's rating for a product which the user has not rated yet. A collaborative filtering system is designed to analyze a user's personal propensities based on many users' ratings given on the same product and predicts the user's rating for a product which the user has not rated yet. Accordingly, many users are required to make ratings on many items to come up with better recommendations. However, users are normally unable to assess all the items in the system, which always presents a fundamental problem named 'rating sparsity' to the collaborative filtering systems [7].

In addition to the sparsity issue, there also exist scalability and transparency issues which have been pointed out as restrictions to the system [8, 9]. Furthermore, the system also displays vulnerability on security. Collaborative filtering systems make recommendation of new products using the rating data given to such items as purchase history. Accordingly, it is impossible to fundamentally block out malicious attacks; it is necessary to detect such ill-intentioned users and attack targets in order to prevent the attacks.

Malicious attacks against collaborative filtering systems can be divided into two kinds - push attack and nuke attack. Push attacks are intended to get the highest ratings, whereas nuke attacks attempt to get the lowest ratings [10].

This paper uses personal propensity factors in order to cope with sparsity and scalability issues inherent in collaborative filtering systems use in movie recommendation systems, and to thus improve the performance of these systems. With a lack of rating data, no collaborative filtering system can make optimal recommendations to target customers, though good recommendations are the original objective of such a system or of any personalized service. This is due to a difficulty in making judgments as to whether a target user has similar propensities to those that users have shown in their ratings used in the system. There are a couple of methods of dealing with this sparsity issue.

The first method is that, in a collaborative filtering system, a nearest neighbor algorithm is applied and a group of personal propensities are used to its inverse, and the other is that personal propensities assessed and selected by demographic methods are applied to the system. This paper adopts the second method. Personal propensities can be defined as components or factors or

elements typical of an individual by which the individual can be characterized. The more such personal propensities are available, the better a user can be identified and personalized; however, we here suggest a method of optimizing such personal propensity. Further, we make assessment of the method and performance of the system Huh [11], Lee [12], and Jang [13] have established among other previously available methods.

This paper consists of 6 sections. Section 2 describes related works. Section 3 evaluates a method of selecting optimal personal propensities suggested by the paper. Section 4 explains a movie recommendation procedure proposed in this paper. Section 5 deals with the application results and performance evaluation of the movie recommendation system and Section 6 wraps up this paper with its conclusion.

2. RELATED WORK

This section describes, through the use of personal propensity factors and a safe collaborative filtering system, such details as recommender systems, personalization techniques, methods of selecting collaborative filtering systems, and detection of attacks to collaborative filtering systems that are required for the recommendation of movies.

2.1 Recommender system

A recommender system makes recommendation of products that are suitable for a customer's demands, based on the analysis of such information as products that many customers are interested in, demographic data, and past purchasing activity [8]. Personalized services tailored to individual tastes have been emphasized in e-commerce transactions. Personalization means the process of quickly responding on the Internet to a customer's needs that are unique and specific. Web personalization is defined as activity made on the Internet by an individual in response to his/her interests or tastes [14]. The reasons why personalized service is important are that customers can reduce their attempts to search for products, and companies not only increase customers' loyalty to their e-commerce sites through the recommendation of proper products but also build attachment between them and their customers [15].

2.2 Personalization techniques

The personalization techniques for recommender systems include:

(1) Content-based recommender system

This system analyzes item information and recommends certain items to users. It is suitable for recommendation of such items as texts, documents, news, and web pages whose contents are abundant and easy to analyze [16].

(2) Rule-based filtering

This technique specializes in the acquisition of users' information profiles by means of questions to users about their interests and preferences. Users' profiles can be obtained by asking questions about the users' tastes and preferences on particulars and collecting and analyzing their answers. This filtering system recommends to users or provides them with

information about products that are considered to be suitable given a user's psychological and preference information based on such profiles [15].

(3) Demographic filtering

This system makes recommendations using users' information such as age, sex, and education level [17]. Demographic attributes have an advantage of making an easy analysis of users' preferences regarding various kinds of items and item categories.

(4) Collaborative filtering

This system makes recommendations by utilizing each user's assessment information [18]. As a collaborative filtering system makes use of rating information, it has an advantage of performing recommendations without the information on a user or on a specific item.

(5) Learning agent-based filtering system

This personalization technique utilizes learning agents that are designed to trace users' attributes, habits, and personal preferences through the analysis of log files including records of visit to websites and their frequency, access location, and time [16].

A recommender system is a program which makes predictions about relations among customers or among items, and searches for items that a user may be expectedly to desire. The purpose of many studies on recommender systems have mainly focused on their capability of how likely they are able to recommend products that a customer is satisfied with. Collaborative filtering is the method most frequently used to categorize similarities among items.

Fig. 1 shows a neighborhood-based algorithm that has generally been used in collaborative filtering systems [7]. The active user calculates distances to other users and selects as its neighbors the number of users who are located at nearest distances. The distance between users can be calculated using the Pearson correlation coefficient, the mean-square difference, or vector similarity. In [18], the Pearson correlation coefficient produced a better result than the vector similarity, and [19] showed that the Pearson correlation coefficient brought about a better outcome though a selection of either a too small or too large number of neighbors might lead to a reduction in its prediction capability. When distances to other users have been calculated, a predicted score for an item can be computed by summing other users' rated scores in proportion to their



Fig. 1. Collaborative filtering algorithm

distance weights, using the following equation [7].

$$P_{a,i} = r_a + \frac{\sum_{u=1}^n (r_{u,i} - \bar{r}_i) * W_{a,u}}{\sum_{u=1}^n W_{a,u}} \quad (1)$$

Equation (1) is introduced to calculate a distance through the Pearson correlation coefficient. ‘ $P_{a,i}$ ’ indicates an active user ‘a’’s prediction about an item ‘i’. ‘n’ is the number of the neighboring users, ‘ $r_{u,i}$ ’ means the rating of a user ‘u’ on an item ‘i’, and ‘ $w_{a,u}$ ’ is defined as the weighted similarity between the active user ‘a’ and its neighbor ‘u’ [5].

The Pearson correlation coefficient has a number close to ‘1’ when a user A rates a movie high that a user B has also rated high and user A also gives a low rating to a movie that the user B has given a low rating; and, it is close to ‘0’ when the vice versa holds true.

2.3 Approaches to collaborative filtering

There are two kinds of collaborative filtering: user-based collaborative filtering and item-based collaborative filtering.

(1) User-based collaborative filtering

This approach is to calculate distances to quantify how closely two users match each other in respect with a certain common item. For example, if user No 1 and No 2 put in same ratings in the same movie, the distance will be 0. On the other hand, assuming they give different ratings, the distance will be farther depending on the difference.

(2) Item-based collaborative filtering

Most recommender systems utilize an item-based collaborative filtering technique rather than a user-based one. For instance, when users who like movie No 1 also like movie No 2, the distance between two movies is regarded as being close.

User-based and item-based algorithms face problems when an item is recommended. That is because any similar item, if not the exactly same item, can't be applied as a base, for the criterion for application is very strict. For instance, in the event someone who has a great interest in action movies is asked to give ratings in a recommender system, the basic algorithm is not able to discern whether s/he is actually interested in action movies. Here, the concept of dimensional reduction is introduced.

Dimensional reduction, for example, can be explained in an immense matrix drawn up to investigate preference relations among many people and many movies. An abstraction stage should be introduced at the time of acquisition of the data from movie ratings. Many people are classified by similar criteria; then, new items are substituted using the similarity criteria to be classified into categories of same taste. The more actively dimensional a reduction is applied, the more effective a recommendation will be.

Collaborative filtering has been researched and applied in various fields. GroupLens [5], a news article recommendation system, Video Recommender [4], a video recommendation system, Ringo [6], a music recommendation system, and PHOAKS [20], a user-related search engine on the World Wide Web have used a collaborative filtering system, and Amazon, CDNow, Drug-

store, and MovieFinder have recommended their products using collaborative filtering systems.

2.4 Detection of attacks on collaborative filtering systems

Dellarocas [21] summarized several types of attack that have made on well-known systems used by online commercial communities including eBay. We introduce examples of such attacks on currently available collaborative filtering algorithms and similar systems and suggest a prediction algorithm to minimize the aftereffects of such attacks.

Lam and Riedl [22] analyzed manipulated ratings by users and categorized recommendation attacks by type. A method of influencing a recommender system involves attackers who get access to the system and directly give items their ratings. In the end, the ratings manipulated by the attackers led the system to produce wrong judgments. Lam and Riedl made an analysis of such elements as manipulated ratings that influenced the system's recommendations.

Even though studies on attacks on recommender systems have been under way, none has yet come up with a solution to effectively predict random and arbitrary attacks on them. A possible method of effectively forecasting random attacks involves analysis of the trend of rating streams and prediction of possible random attacks on recommender systems [23].

3. OPTIMAL TECHNIQUES FOR RATING OF PERSONAL PROPENSITY

This paper used for our movie recommendation system the open data called MovieLens, which contains a total of 10,000 records from 943 raters on 1,682 movies [24].

Use of personal propensity factors for recommendation may vary. It may produce different weights depending on what items have been recommended. If recommendation of items for the regulation of physical composition should be made, personal propensity variables might include such items as height, weight, and MBI, as these factors may possess higher importance than other personalization factors.

The movie recommendation system described in this paper needs to deal with personal propensity variables with a view to identifying individual movie preferences; therefore, it makes analysis of personal propensities to be used for the recommendation of suitable movies to users by utilizing personalization factors used in existing personal propensity identification methods and collaborative filtering-based movie recommendation systems [11-13]. Main personalization factors in this paper include genre, character, age, blood type, and region, which are used to constitute nearest neighbors in the system's attempts to make recommendation of movies. Adoption

```
1|24|M|technician|85711
2|53|F|other|94043
3|23|M|writer|32067
4|24|M|technician|43537
5|33|F|other|15213
6|42|M|executive|98101
7|57|M|administrator|91344
8|36|M|administrator|05201
9|29|M|student|01002
10|53|M|lawyer|90703
11|39|F|other|30329
```

Fig. 2. MovieLens' user data (part)

196	242	3	movie	1	2	3	4	5	6	7	8	9	10	11	12	13	
186	302	3	user	1	5	3	4	3	3	5	4	1	5	3	2	5	5
22	377	1		2	4								2			4	
244	51	2		3										4			
166	346	1		4	3												
298	474	4		5						2	4	4				2	
115	265	2		6	4					5	5	5	4		3	5	
253	465	5		7			5			3				3			
				8						4							
				9					5	4							
				10	4		4			4		4		4	5	3	
				11						4	5			2	2		

Fig. 3. Rating data and Users' movie rating matrix (part)

of such diverse predictor variables ranging from blood type to age, sex, genre, character, and region were used here rests on an assumption that the use of a greater number of personal propensity variables will generally consider more personalization factors, thus contributing to the construction of a better recommender system [25].

Since the MovieLens data used in this paper does not provide personalization factors, it needs to extract such factors from the supplied information. Fig. 2 shows part of the personal information that was provided by MovieLens.

Fig. 3 shows parts of rating data and a matrix generated from the rating data given by users who viewed certain movies. The column of Fig.3 is movie number and the row of Fig. 3 is user number.

Personalization factors include region, and so we need to search for the region in the data. As the data had been collected in the States, the region was distinguished by zip codes.

The US zip codes can be identified at USPS.com. Fig. 4 shows part of a list of regions where all the 943 raters are located.

The United States comprises 51 states, which are given a number from 1 in the order of what is provided by the encyclopedia [17]. These numbers are uniquely granted to compute how many raters reside in each state. There were some raters who could not be referred to a specific region in the course of data processing and thus were excluded for that reason, leaving a total of 906 persons. Thus, the data were corrected to reflect the change. Moreover, some of the states in the modified data do not produce sufficient data, for examples, Alabama corresponding to the number 1 and showing only three raters and Arkansas having only one rater; therefore, the data has been changed to a larger category of four regions (Northeast, South, Midwest, and West)

As far as personality is concerned, criteria are required for judgment; for this purpose, the MBTI (Myers-Briggs Type Indicator) assessment is introduced to create 16 type combinations. The MBTI is a self-reported character type indicator that is rooted in the psychological type theory and developed into a questionnaire for easy use in daily life. The assessment has a total of 16 combinations of psychological types (ISTJ, ISFJ, INFJ, IMTJ, ISTP, ISFP, INFP, INTP,

No	Age	gender	regions	Male	Female	1 Alabama	AL	1
1	24	1	3			2 Alaska	AK	2
2	53	2	5			Arizona	AZ	3
3	23	1	9			Arkansas	AR	4
4	24	1	35			California	CA	5
5	33	2	38			Colorado	CO	6
6	42	1	47			Connecticut	CT	7
7	57	1	5					
8	36	1	46					

Fig. 4. A list of regions indicated by zip codes (part)

Input data

* ID :	<input type="text"/>	ID redundancy check
* PASSWORD :	<input type="text"/>	
* PASSWORD Confirm :	<input type="text"/>	
e-mail :	<input type="text"/>	@ hanmail.net

additional inputs

Date of birth :	Year : 1987	Month : 1	Day : 3
Preferred genre :	Drama		
Personality(I/E) :	Extraversion		
Personality(S/N) :	Sensing		
Personality(T/F) :	Feeling		
Personality(J/P) :	Judging		
血型 :	AB		

Fig. 5. A window for additional inputs at the time of member registration

ESTP, ESFP, ENFP, ENTP, ESTJ, ESFJ, ENFJ, and ENTJ), which can be used to identify an individual’s personality type [26].

Fig. 5 is window for additional inputs at the time of member registration.

Blood types are classified into four normal classification types (A, B, AB, and O), without consideration given to any specific details.

This paper analyzed factors from the six personal propensity variables that may generally influence movie recommendation systems and assessed their performances. The combinations that can be created by the six personal propensities are 2 raised to the power of 6, which is equal to 64. We obtained 63 combinations exclusive of one that indicates nothing, as shown in Fig. 6. Of these combinations, we tried to search for ones that may have great effects on movie recommendation systems. For this purpose, 63 combinations of factors were applied to make analysis of the movies found in the MovieLens data.

10 movies were chosen to be used in this movie recommendation and analyzed for 63 combinations, coming up with most influential factors among the 6 personal propensities in the rec-

1Element	2Element	3Element	4Element	5Element	6Element
1 Age	7 Age Sex	22 Age Sex	42 Age Sex	57 Sex Genre Blood type	
2 Sex	8 Age Genre	23 Age Sex	43 Age Sex	58 Sex Genre Blood type Region	
3 Genre	9 Age Blood type	24 Age Sex	44 Age Sex	59 Sex Genre Character Region	
4 Blood type	10 Age Character	25 Age Sex	45 Age Sex	60 Sex Blood type Character Region	
5 Character	11 Age Region	26 Age Genre	46 Age Sex	61 Genre Blood type Character Region	
6 Region	12 Sex Genre	27 Age Genre	47 Age Sex	62 Genre Blood type Character Region	
	13 Sex Blood type	28 Age Genre	48 Age Genre	63 Blood type Character Region	
	14 Sex Character	29 Age Blood type	49 Age Genre	64 Age Sex Genre Blood type Character Region	
	15 Sex Region	30 Age Blood type	50 Age Genre		
	16 Genre Blood type	31 Age Character	51 Age Blood type		
	17 Genre Character	32 Sex Genre	52 Sex Genre		
	18 Genre Region	33 Sex Genre	53 Sex Genre		
	19 Blood type Character	34 Sex Genre	54 Sex Genre		
	20 Blood type Region	35 Sex Blood type	55 Sex Blood type		
	21 Character Region	36 Sex Blood type	56 Genre Blood type		
		37 Sex Character			
		38 Genre Blood type			
		39 Genre Blood type			
		40 Genre Character			
		41 Blood type Character			

Fig. 6. Combinations for analysis from 6 personal propensities

Table 1. Part of the list of recommended movies

Age	Sex	Preferred genre
“Saint of Fort Washington, The” (1993)	“Great Day in Harlem, A” (1994)	“Wrong Trousers, The” (1993)
“Star Kid” (1997)	“Saint of Fort Washington, The” (1993)	“Close Shave, A” (1995)
“Santa with Muscles” (1996)	“Marlene Dietrich: Shadow and Light” (1996)	“Casablanca” (1942)
.	.	.
.	.	.
.	.	.
Blood type	Personality	Region
“Prefontaine” (1997)	“Casablanca” (1942)	“Someone Else’s America” (1995)
“Star Kid” (1997)	“Maya Lin: A Strong Clear Vision” (1994)	“Entertaining Angels: The Dorothy Day Story” (1996)
“Great Day in Harlem, A” (1994)	“Close Shave, A” (1995)	“Great Day in Harlem, A” (1994)
.	.	.
.	.	.
.	.	.
Age, genre, bloody type, personality, & region	Sex, genre, blood type, & character	Sex, blood type, & region
“Casablanca” (1942)	“Santa with Muscles “(1996)	“Great Day in Harlem, A” (1994)
“Close Shave, A” (1995)	“Close Shave, A” (1995)	“Casablanca” (1942)
“Wrong Trousers, The” (1993)	“Wrong Trousers, The” (1993)	“Close Shave, A” (1995)
.	.	.
.	.	.
.	.	.
Age, blood type, region	Sex, personality, region	Age, sex, region
“Star Kid” (1997)	“Someone Else’s America” (1995)	“Star Kid” (1997)
“Pather Panchali” (1955)	“Pather Panchali” (1955)	“Pather Panchali” (1955)
“Casablanca” (1942)	“Wrong Trousers, The” (1993)	“Casablanca” (1942)

ommendation of movies. Table 1 shows part of the list of recommended movies prepared after 63 combinations of personalization factors had been applied to the MovieLens data. Table 1 shows part of the list of recommended movies prepared after 63 combinations of personalization factors had been applied to the MovieLens data.

The number of commonly recommended movies was calculated from each combination to find recommended movies from the recommended movie list having significance. The number of those movies that were most frequently found in each of 63 combinations is considered to be meaningful. The numbers were counted as shown in Fig. 7.

The graph showed through the analysis of 63 combinations that such factors as genre, personality and age are most important. The use of these three factors could yield the most common significance no matter how the six personal propensities were applied.

In summary, at first it is assumed that the more personal propensities are used, the more effectively a system can recommend most suitable movies; however, the three personal propensities of age, genre, and personality are sufficient to make recommendation of good movies to an indi-

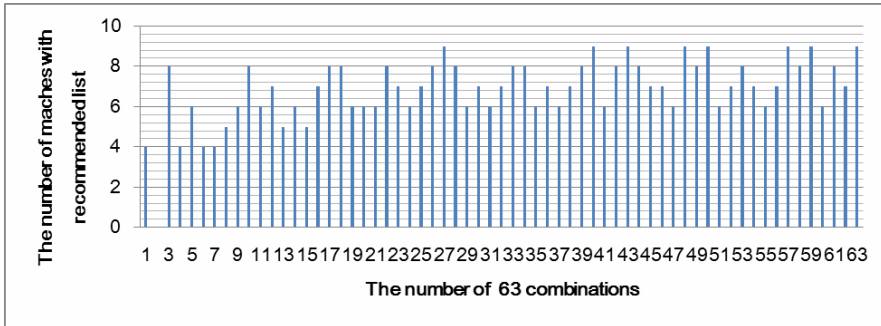


Fig. 7. Analysis graph for 63 combinations

vidual’s taste. That is to say, we are suggesting that these three propensity variables of age, genre, and character are the most important factors in movie recommendation systems.

4. RECOMMENDING PROCEDURE OF OUR MOVIE RECOMMENDATION SYSTEM

This paper attempts to combine existing collaborative filtering techniques with a personal propensity identification method and suggests a movie recommendation system which may produce best predictions about a list of movies that target users may like most. The was done by extracting users who may have purchase preferences or tastes similar to those target users, with much consideration given to such personalization factors as genre, personality, and age, which have been shown to play a critical role in determining an individual’s personal characteristics. The method suggested in this paper has the algorithm 1.

The method suggested in this paper establishes nearest neighbors via an existing collaborative filtering technique so as to generate a recommended movie list, in the event there are an ade-

Algorithm 1. Movie recommendation system

```

1 : INPUT:  $U_k$  : user, R : rating table, N : recommended recipient user
2 : OUTPUT: RL : recommended movie list
3 : Procedure: if count rating(N,R) < sufficient amount then
4 :  $\dot{N} \leftarrow$  Identify Profile
5 : for each Identify Profile  $U_i$  in  $U_k$  do
6 : clustering( $\dot{N}, U_k$ )
7 : find nearest cluster N
8 : RL  $\leftarrow$  recommended with nearest cluster with R
9 : else if Collaborative Filtering(N,R)
10 : push attack
11 : end if
    
```

quate amount of users' rating data available. However, if there are not enough data available, an optimal recommendation cannot be made due to the sparsity issue, in which case the use of such personalization factors as genre, personality, and age that can best reflect individual characteristics will compensate for such sparsity to eventually enable the system to make optimal recommendations. For acquisition of such personal propensity information, the system can collect it explicitly by receiving it when a target user logs in. For this purpose, the system requires users to put in three personalization factors of genre, personality, and age when they log in. And, users are clustered along with other users who have given their ratings by means of the K-means clustering technique. After the clustering, the system carries out the push attack principle for a safe recommender system. Then, a recommendation list is established by using the movie rating data given by those clustered viewers including the target viewers.

5. APPLICATION RESULTS AND PERFORMANCE EVALUATION OF THE MOVIE RECOMMENDATION SYSTEM

Fig. 8 shows a search screen generated from the movie data of the movie recommendation system proposed in this paper.

Fig. 8 shows average ratings that have been computed with the ratings given by all users who had viewed the relevant movie, along with movie title and genre, which are stored in the database

Fig. 9: This page shows a list of movies that are recommended by the system, using a target customer's personal propensities.

In addition, for assessment of the performance, our system was compared with the movie recommendation system suggested by Huh [11], Lee [12], and Jang [13], among other previously available systems, where a group of personal propensity variables had been used. For performance assessment, we used the MovieLens data to make comparison between a list of movies recommended by the system suggested in this paper and that created by the Huh, Lee, and Jang.

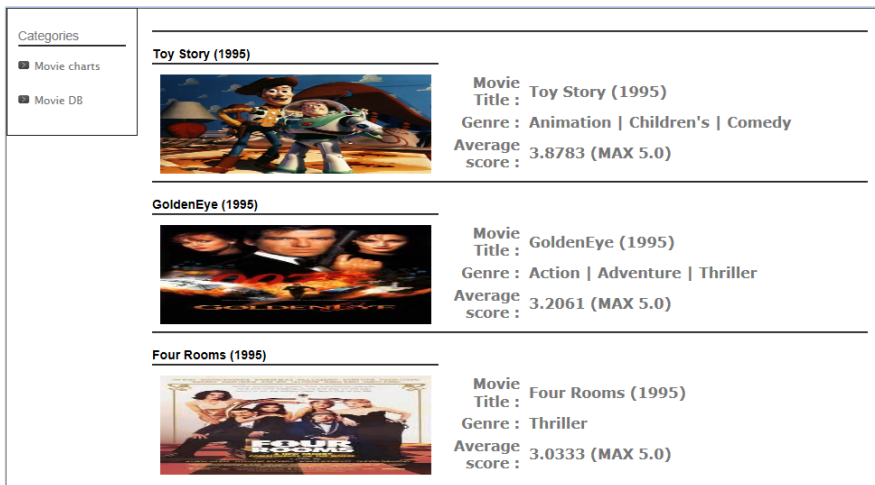


Fig. 8. Movie DB Search Screen



Fig. 9. Results of movie recommendation using personal propensities

Table 2 shows the results of recommended movies using Huh, Lee, Jang, and our method.

Table 2 shows that such personal propensities as blood type, sex, and region do not have considerable effects on personalization. To sum up, the table indicates that the variables of genre, personality, and age are enough personal propensity factors to product good results.

Table 2. Lists of recommended movies using method, Huh, Lee, Jang and our method

Blood type[Lee]	Age, sex, genre, personality[Huh]	Age, sex, genre, blood type, personality, region[Jang]	Age, genre, personality [our method]
“Prefontaine” (1997)	“Casablanca” (1942)	“Casablanca” (1942)	“Casablanca” (1942)
“Star Kid” (1997)	“Close Shave, A” (1995)	“Close Shave, A” (1995)	“Close Shave, A” (1995)
“Great Day in Harlem, A” (1994)	“Wrong Trousers, The” (1993)	“Wrong Trousers, The” (1993)	“Wrong Trousers, The” (1993)
“Pather Panchali” (1955)	“Rear Window” (1954)	“Rear Window” (1954)	“Rear Window” (1954)
“Some Mother’s Son” (1996)	“Wallace & Gromit: The Best of Aardman Animation” (1996)	“Wallace & Gromit: The Best of Aardman Animation” (1996)	“Wallace & Gromit: The Best of Aardman Animation” (1996)
“Close Shave, A” (1995)	“Third Man, The” (1949)	“Third Man, The” (1949)	“Third Man, The” (1949)
“Everest” (1998)	“Manchurian Candidate, The” (1962)	“Manchurian Candidate, The” (1962)	“Manchurian Candidate, The” (1962)
“Wrong Trousers, The” (1993)	“Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb” (1963)	“Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb” (1963)	“Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb” (1963)
“Casablanca” (1942)	“Citizen Kane” (1941)	“Citizen Kane” (1941)	“Citizen Kane” (1941)
“Wallace & Gromit: The Best of Aardman Animation” (1996)	“One Flew Over the Cuckoo’s Nest” (1975)	“One Flew Over the Cuckoo’s Nest” (1975)	“One Flew Over the Cuckoo’s Nest” (1975)

5. CONCLUSION

An ever-increasing quantity of data has made the search for desired information still more difficult for users. Personalized services have attracted attention as a means to cope with this problem. A personalized service can be referred to as a recommender system in which a target user is provided with her/his preferred services or products. Recommender systems that are frequently used in e-commerce include collaborative filtering systems. However, they have faced such issues as sparsity, scalability, and transparency. To provide a solution to these problems, this paper presents a recommender system using personal propensity variables. Unlike many previous papers that have adopted a number of predictor variables for personalized recommendations, this paper has attempted to develop a system designed to make optimal recommendations without input of too many of unnecessary factors. The system could make recommendations tailored or personalized to an individual user through application of three independent factors of genre, personality, and age rather than use of the all available personal propensity variables. This system uses as personalization factors only the information that a user puts in when logging into the system, with which it effectively provides a solution to such issues as sparsity and scalability and is thus able to supply an optimal list of recommended movies to users even without an adequate number of rating data.

REFERENCES

- [1] 2011 Internet Use Survey Summary Report, KISA, Nov. 2011 pp.2-27.
- [2] M. D. Mulvenna, S. S. Anand, A. G. Büchner, "Personalization on the Net using Web mining," *Communications of the ACM*, Vol.43 Issue 8, Aug. 2000, pp.122-125.
- [3] J. Lee, S. Park, "Performance Improvement of a Movie Recommendation System using Genre-wise Collaborative Filtering," *Journal of Intelligence Information Systems*, Vol.13, No.4, 2007, pp.65-78.
- [4] W. Hill, L. Stead, M. Rosenstein and G. Furnas, "Recommending and evaluating choices in a virtual community of use," *Proceeding of the SIGCHI conference on Human factors in computing systems*, 1995, pp.194-201.
- [5] P. Resnick, N. Iacovou, M. Sushak, P. Bergstrom, and J. Riedl, "GroupLens: an open architecture for collaborative filtering of Netnews," *Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work*, Chapel Hill, NC, 2001, pp.175-186.
- [6] U. Shardanand and P. Maes, "Social information filtering: algorithms for automating 'word of mouth'," *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1995, pp.210-217.
- [7] H. Lee, H. Kim, "Improving Collaborative Filtering with Rating Prediction Based on Taste Space," *Journal of Korean Institute of Information Scientists and Engineers*, Vol.34, No.5, 2007, pp.389-395.
- [8] B. Sarwar, G. Karypis, J. Konstan, J. Riedl, "Item-based collaborative filtering recommendation algorithms," *Proceeding of the 10th international conference on World Wide Web*, 2001, pp.285-295.
- [9] P. Li, and S. Yamada, "A Movie Recommender System Based on Inductive Learning," *IEEE Conf. on Cybernetics and Intelligent System*, 2004, pp.318-323.
- [10] B. Mobasher, R. Burke, R. Bhaumik, and J.J. Sandvig, "Attacks and Remedies in Collaborative Recommendation," *Intelligent Systems*, IEEE, 2007, pp.56-63.
- [11] J. Heo, D. Park, Y. Jeong, "A Personalized Movie Recommendation System Based on the Personal Propensity and Collaborative Filtering," *Proceedings of Korea Multi Media Society Spring Annual Conference*, Vol.12, No.2, 2009, pp.325-326.

- [12] S. Lee, D. Park, M. Hong, H. Lee, Y. Jeong, "A Personalized Movie Recommendation System Based on the Blood Types and Collaborative Filtering," Proceedings of Korea information Processing Society Fall Annual Conference, 2009, Vol.16, No.2.
- [13] S. Jang, D. Park, Y. Jeong, "Performance Improvement of a Movie Recommendation System based on the Personal Propensity and collaborative Filtering," Proceedings of Korea Information Processing Society Spring Annual Conference, 2010, Vol.17, No.1.
- [14] B. Mobasher, H. Dai, T. Luo, and M. Nakagawa, "Discovery and Evaluation of Aggregate Usage Profiles for Web Personalization," Data Mining and Knowledge Discovery, Vol.6, No.1, 2002, pp.61-82.
- [15] T. Hong, H. Lee, B. Su, "A Web Personalized Recommender Systems Using Clustering - based CBR", Journal of Intelligence Information Systems, Vol.11, No.1, 2005, pp.107-121.
- [16] R. J. Mooney and L. Roy, "Content-based Book Recommending Using Learning for Text Categorization," In Proceedings of ACM SIGIR '99 Workshop Recommender Systems: Algorithms and Evaluation, 1999.
- [17] J. Le and O.R. Zaiane, "Combining Usage, Content, and Structure Data to Improve Web Site Recommendation," Fifth International Conference on Electronic Commerce and Web Technologies (EC-Web '04), 2004, pp.305-315.
- [18] J. S. Breese, D. Heckeman and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence, 1998, pp.43-52.
- [19] J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl, "An algorithmic framework for performing collaborative filtering," Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, 1999, pp.230-237.
- [20] L. Terveen, W. Hill, B. Amento, D. McDonald, J. Creter, "PHOAKS: a system for sharing recommendations," Communications of the ACM Vol.40, No.3, 1997, pp.59-62.
- [21] C. Dellarocas, "Immunizing Online Reputation Reporting Systems against Unfair Rating and Discriminatory Behavior," Proceedings of the ACM Conference on Electronic Commerce, 2000.
- [22] S. Lam, and J. Riedl, "Shilling Recommender Systems for Fun and Profit," Proceedings of the 13th International WWW Conference, 2004.
- [23] Y. Kim, J. Kim, "Attack Detection in Recommender Systems Using a Rating Stream Trend Analysis," Proceedings of Korean Society for Internet Information Journal, Vol.12, No.2, 2011, pp.85-101.
- [24] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom and J. Riedl, "GroupLens : An Open Architecture for Collaborative Filtering of Netnews," Proceedings of the ACM Conf. on Computer Supported Cooperative Work 1994, pp.175-186.
- [25] S. Kim, W. Jeong, D. Park, "A Study on the Performance Evaluation and Improvement of Personalized Movie Recommendation System," Proceedings of Korea information Processing Society Autumn Annual Conference, Vol.19 No.2, 2012.
- [26] B. Myers, M. H. McCaulley, N. L. Quenk, A. L. Hammer, *MBTI Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator®*, Third Edition, Consulting Psychologists Press, Palo Alto, CA, 1998.



Se-jun Kim

He received a BS degree in Computer Software Engineering from Soonchunhyang University in 2012. Currently, he is a master degree student in the Department of Computer Science and Engineering at Soonchunhyang University. His research interests include data mining and parallel processing.



Woon-Hae Jeong

He received the BS degree in Computer Software Engineering from Soonchunhyang University in 2012. Currently, he is a master degree student in the Department of Computer Science and Engineering at Soonchunhyang University. His research interests include data mining and security.



Doo-Soon Park

He received his PhD in Computer Science from Korea University in 1988. Currently, he is a professor in the Department of Computer Software Engineering at Soonchunhyang University, South Korea. He is President of Korea Information Technology Association (KITA) and Director of Culture Technology at Soonchunhyang University. He was editor in chief of JIPS (Journal of Information Processing Systems) at KIPS (Korea Information Processing Society) from 2009 to 2012. He was Dean of the Engineering College at Soonchunhyang University from 2002 to 2003, and was the Director of the u-Healthcare Research Center at Soonchunhyang University from 2006 to 2007. He was a visiting professor at the University of Colorado from 2004 to 2005. He has served as an organizing committee member or chair of many international conferences including CUTE 2012, CSA 2012, 3DOC-12, ICCCT 2012, AIM summer 2012, FutureTech 2012, AIM 2012, CUTE 2011, ISPA 2011, CUTE 2010, ICA3PP 2010, ICUT 2009, CSA 2009, HPCC-09, and ISA2009. His research interests include parallel processing, data mining, and multimedia information processing. He is a member of IEEE, ACM, KMS, and KIPS. Contact him at parkds@sch.ac.kr.



Jin Kwak

He received his BS (2000), MS (2003), and PhD (2006) from Sungkyunkwan University (SKKU) in Korea. Prior to joining the faculty at Soonchunhyang University (SCH) in 2007, He attended Kyushu University in Japan as a visiting scholar. After that, he served in MIC (Ministry of Information and Communication, Korea) as a Deputy Director. Also, he has served as a Dean of DISE (2009-2010) and Vice-Dean of the College of Engineering (2009) in SCH. Now he is a professor in the Department of Information Security Engineering (DISE) at SCH.

Also, now he is a Director of SCH BIT Business Incubation Center and a Director of Industry-University and Institute Partnership Division center at SCH. His main research areas are cryptology, information security applications and information assurance.