

# Movies recommendation system using collaborative filtering and k-means

Phongsavanh Phorasim\* and Lasheng Yu

School of Information Science and Engineering, Changsha, Hunan, China

Received: 16-November-2016; Revised: 04- February-2017; Accepted: 10-February-2017

©2017 ACCENTS

## Abstract

*The purpose of this research is to develop a movie recommender system using collaborative filtering technique and K-means. Collaborative filtering is the most successful algorithm in the recommender system's field. A recommender system is an intelligent system that can help a user to come across interesting items. This paper considers the users  $m$  ( $m$  is the number of users), points in  $n$  dimensional space ( $n$  is the number of items) and we present an approach based on user clustering to produce a recommendation for the active user by a new approach. We used k-means clustering algorithm to categorize users based on their interests. We evaluate the traditional collaborative filtering and our approach to compare them. Our results show the proposed algorithm is more accurate than the traditional existing one, besides it is less time consuming than the previous existing methods.*

## Keywords

*Recommendation system, Collaborative filtering, K-means, Clustering, Data mining.*

## 1.Introduction

Currently as a new technic of advertisement, we use sets of information such as movies, music and other form of technique to convince customers. Technology is developing so fast and dissemination of knowledge has increased, as well as the needs of consumers that are more complex. Manufacturers and suppliers had difficulty in offering products and services that meet customer needs for the convenience of buying the service because of business which makes the competition even more active. In this era of competition, complex information causes overload problems which in turn are time consuming.

Recommendation systems are information filtering system that aids users in predicting rating or preference of an item under users' consideration. The systems offer users alternate selections without having to work out all the details by themselves. As overwhelming information explosion renders searching, extraction, analysis, and processing hideous and formidably time-consuming operations, recommender systems became a favorable decision tool or assistant to offload such undesirable tasks. Worse yet, activities involving human are inevitably subject to human errors that can lead to poor or wrong decisions.

Recommender systems usually use collaborative filtering algorithms or a combination of the collaborative filtering algorithms and the other filtering algorithms to find users who have similar tastes and suggest items [1-3]. We conducted the movies recommender system by using collaborative filtering and k-means which uses the techniques of data mining to divide user in the groups before entering into collaborative filtering. After entering into the system, it will process the information and give the required prediction thus saving the user time.

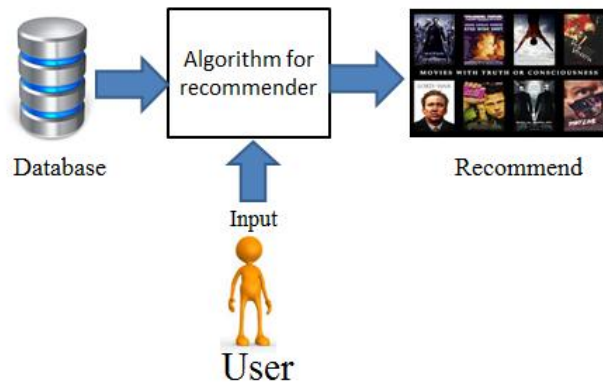
## 2.Literature review

### 2.1Recommendation system

Recommender systems are beneficial to both service providers and users [4]. They reduce transaction costs of finding and selecting items in an online shopping environment [5]. A recommendation system is a system used to present or talk about personal experienced to the customers by presenting information filtered to them, specifically for each customer through several channels of communication. This increases the chances of the information reaching the customers faster and reduces the overloaded information. Recommendation systems have also proved to improve the decision making process and quality [6]. General recommendation systems contain four parts: database, human-computer interface, algorithm,

\*Author for correspondence

recommendation components which has been shown in Figure 1.



**Figure 1** General recommendation systems

## 2.2 Data mining

Data mining is the process of discovering interesting knowledge, such as associations, pattern, changes, significant structures and anomalies, from large amounts of data stored in databases or data warehouse or other information repositories [7]. The main techniques for data mining include classification and prediction, clustering, outlier detection, association rules, sequence analysis, time series analysis and text mining, and also some new techniques such as social network analysis and sentiment analysis. Detailed introduction of data mining techniques can be found in other volumes and journals on data mining [8-10].

## 2.3 Clustering

Clustering techniques have been applied in different domains, such as pattern recognition, image processing, statistical data analysis and knowledge discovery [11]. Clustering provides division of data into groups of similar objects (observations, events) based on the information found in the data describing the objects or their relationships. Clustering algorithm tries to partition a set of data into a set of sub-clusters in order to discover meaningful groups that exist within them [12]. The goal is to compare the objects in a group to be similar (or related) to one another and different from (or unrelated to) the objects in other groups. The algorithm used in the residential similarity or proximity by calculating when measuring distance between vectors of information such as Euclidean distance. Clustering information is useful in analyzing the data which do not know the exact details of the target class or group number.

## 2.4 K-means

K-means clustering is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining. K-means clustering aims to partition an observation into  $k$  clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. In practice, in the management of the independent variables with more than two variables that use the same to replace two element vectors ( $x_1, x_2$ ) with the same positions to  $n$  element vectors ( $x_1, x_2, \dots, x_n$ ). To determine the group in advance, a random selection process starting from the center of each group, then measured the distance during each of the data center group of the conditions of the nearest or minimum distance in the group data [8].

### The typical k-means clustering method

- 1) Clusters the data into  $k$  groups where  $k$  is predefined.
- 2) Select  $k$  points at random as cluster centers.
- 3) Assign objects to their closest cluster center, according to the Euclidean distance function.
- 4) Calculate the centroid or mean of all objects in each cluster.
- 5) Repeat steps 2, 3 and 4 until the same points are assigned to each cluster in consecutive rounds. The formula use for cluster center (centroid) [13].

Clustering begin from  $K=2, 3, 4, \dots, K_n$  until suitability, this case use Euclidian distance is the basis of the group results. This is called the sum of squared errors (SSE)

$$E_k^2 = \sum_{i=1}^k e_k^2 \quad (1)$$

Where:  $e_k^2 = \sum_{i=1}^{n_k} (x_{ik} - M_k)^2$ ; Or

$$\sqrt{\sum_{i=1}^{n_k} (x_{ik} - M_k)^2}$$

Where:  $M$  is center of group  $k$ ,  $X$  is data of point in  $ik$ .

## 2.5 Collaborative filtering

Collaborative filtering (CF) is a popular recommendation algorithm that bases its predictions and recommendations on the ratings or behavior of other users in the system [14]. The fundamental assumption behind this method is that other users' opinions can be selected and aggregated in such a way that provide a reasonable prediction of the active user's preference. Intuitively, they assume that, if users agree about the quality or relevance of any items, then they will likely agree about other items. For example, if a group of users likes the same things as Mary, then Mary is likely to like the things they

like which she hasn't yet seen. There are other methods for performing recommendation, such as finding items similar to the items liked by a user using textual similarity in metadata content-based filtering (CBF). Now a popular technique for finding similarity has three parts: adjusted cosine-based similarity, cosine-based similarity and correlation-based similarity.

CF is dependent on the correlation-rate (Co-Rate) to find user's feature nearest to all users from the database. So a major cause significant problems such as scalability problem. This is a challenge for the designers, because the system will have increased amount of active users, which will result into user similarity hence processing takes long time [15]. So in case of a larger system, this solution can be achieved by designing the system to segment users into groups before entering into collaborative filtering to reduce the time to process user similarity and the results will be faster.

$$R_{sim}(t,c) = \frac{\sum_{i \in I} (R_{i,t} - \bar{R}_t)(R_{i,c} - \bar{R}_c)}{\sqrt{\sum_{i \in I} (R_{i,t} - \bar{R}_t)^2} \sqrt{\sum_{i \in I} (R_{i,c} - \bar{R}_c)^2}} \quad (2)$$

Where:  $t$  is the user of data target and  $c$  is the user of data comparison.

$R_{sim}(t,c)$  is the correlation similarity between user of data  $t$  and  $c$ .

$R_{i,t}$  and  $R_{i,c}$  is the user rating for user of data  $t$  and  $c$ .

$\bar{R}_t$  and  $\bar{R}_c$  is the k-mean rating for user of data  $t$  and  $c$

Collaborative filtering is very important for the introduction of the result of the value forecast satisfaction to predict the preferences of users on one piece of information [14]. These values were calculated from the similarity between pairs of items. The next step is to target users by using two types of weight sum and regression herein speak only as part of how the weight sum prediction is only due to the method used in this research. Weight sum prediction prophecy is the satisfaction target number of information pieces, each piece  $k$  Rating is a weighted value based on the resemblance the formula to calculate.

$$P_{t,i} = \frac{\sum_{k \in K} (R_{sim}(t,i) * R_{i,k})}{\sum_{k \in K} (|R_{sim}(t,i)|)} \quad (3)$$

$P_{t,i}$  is the satisfaction that the target user  $t$  will have a piece of data  $i$ .

### 3. Implementation

The implementation of recommendation system, k-means and collaborative filtering techniques steps are as follows:

#### 3.1 Study the problems and needs of the system

Studied the algorithm used for k-means and collaborative filtering techniques in using web application PHP language to also study the need for the data to recommend screening of the movies which are close to the needs of most applications, including legacy systems that provide educational information using historical data, that allows users to set the preferences movies on a scale of one to five (5) level as shown in *Table 1*.

**Table 1** The movies liking form rating scale

Liking	Rating scale
Very like	5
Like	4
Normal	3
Not like	2
Do not like	1

#### 3.2 Data collection

Collected information is to relate and to meet the objectives that have been set in the system. This part shows the basic data used to develop the system for the group of users with k-means and set the data used to create a database of the system. The data has been considered from website movie lens project <http://grouplens.org/datasets/>. This data is clustered in ten groups as shown in *Table 2*.

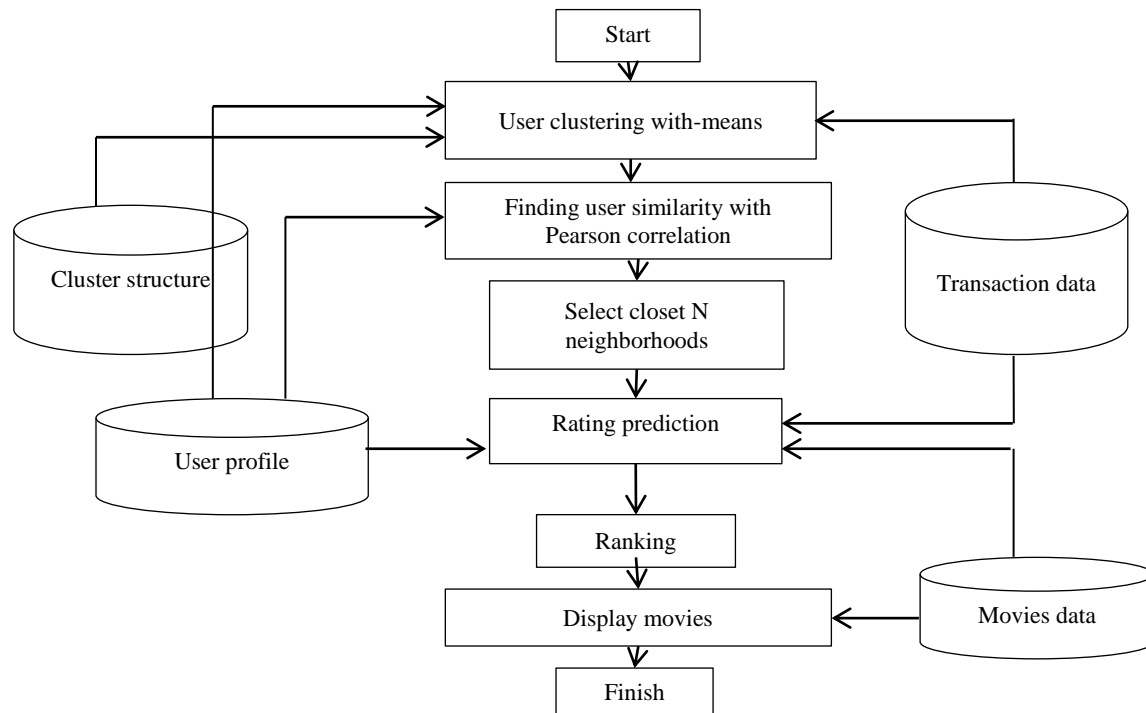
**Table 2** The nature of the group is to divide with k-means

Group	Member of group
1	36
2	70
3	56
4	52
5	61
6	39
7	41
8	52
9	72
10	20

### 3.3 Systems analysis

The processing model is shown in *Figure 2*.

The system can be divided into three parts including the user, the movie manager and system administrator.

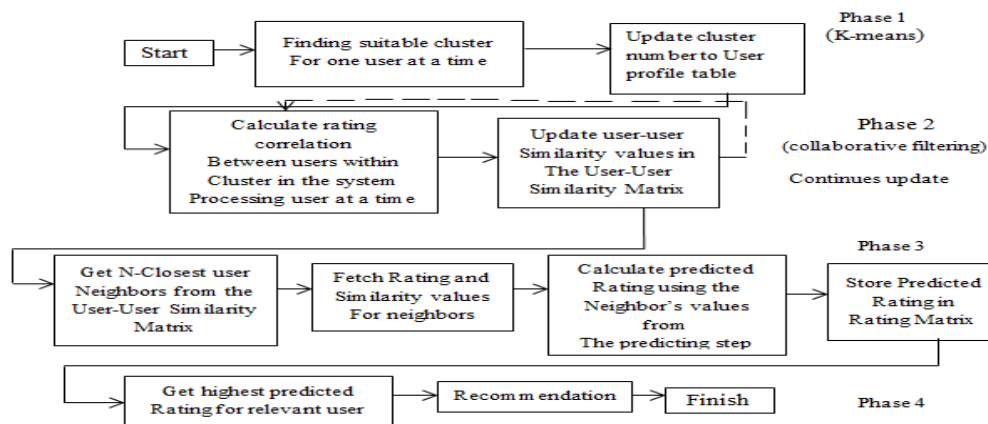


**Figure 2** Processing model

*Figure 2* shows the work after active user gives a rating to the movie, when you watched and requests a recommendation system of other movies. Then, the system will cluster active user to the group which you are recommended by using k-means. After that, the system will take that profile active user to compare with other users which are under same cluster with active user to find out what you are most similar to the active user, using Pearson correlation coefficient

as the comparison tool. Then selects the most similar to the current user of 30 people that bring profile to all of those to predict rating movies of the current user that has never visited before with weighted-sum prediction. After that, they were ranked and rated the movies with up to 10 of the current users.

### 3.4 Development system



**Figure 3** The movie guide

Figure 3 explains the four different phases in the system. The first phase is a group that provides users with rating information of the movie that provides with a variable of the grouping. Phase 2 to 4 is a processing collaborative filtering.

Collaborative filtering will calculate the rating between users within the cluster, after that go and compare the similarity values in the user-user similarity matrix. It will select closest user similarity and pull data similarity values for neighbors that come out and calculate the predicted rating and put in store to search highest predicted rating for relevant user and recommendation to the user.

### 3.5 Processing k-means and collaborative filtering

The system will search group for users by using k-means to find the distance between users, the group of users and clustering of users. The system clustering with the k-means algorithms by measuring the distance of each data point from the center of the 10 groups by using Euclidean distance and calculated information will be stored in the database.

**Table 3** User gives movies rating

User_Id	Movie_Id	Rating
501	124	5
501	133	2
501	140	4

Table 3 shows the movie rating given by the users. Each group were compared as shown in Table 4.

**Table 5** Table user given rating

	Father of the bride	Golden eye	Casino	Four rooms	Money train	Get shorty	Assassins
User_1	1	4	?	3	?	3	2
User_2	?	3	?	2	1	3	?
User_3	3	?	1	?	2	1	2
User_4	3	?	4	?	1	4	2
User_5	?	1	2	1	4	?	?

The system will be searched by users who are similar, and compare user 1 to all the others who have pieces of information rating 3 (Casino) include the user with 3, 4 and 5, so at this stage to make a comparison between the User 1 to 3, 4 and 5 respectively.

Table 6 shows the calculation of the similarities by Person correlation:

**Table 4** The data of centroid for three movies

Movie -ID	K 1	K 2	K 3	K 4	K 5	K 6	K 7	K 8	K 9	K10
124	3	3	3	2	1	4	5	3	3	4
133	2	3	3	4	3	3	3	2	2	3
140	3	3	3	4	2	3	4	3	2	3

The system performs clustering algorithm with k-means of measuring the distance from the point of a centroid cluster of the 10 groups. The equation for calculating the Euclidean distance as a group of users with one or K1 below:

$$D = \sqrt{(3 - 5)^2 + (2 - 2)^2 + (3 - 4)^2} = 2.23$$

Similarly, calculate the distance of the user groups with 2 to 10, which will have the full results below.

The distance of a user or a group 1 k 1 equals 2.23

The distance of a user or a group 2 k 2 equals 2.44

The distance of a user or a group 3 k 3 equals 2.44

The distance of a user or a group 4 k 4 equals 3.60

The distance of a user or a group 5 k 5 equals 4.24

The distance of a user or a group 6 k 6 equals 1.73

The distance of a user or a group 7 k 7 equals 1

The distance of a user or a group 8 k 8 equals 2.23

The distance of a user or a group 9 k 9 equals 2.82

The distance of a user or a group 10 k 10 equals 1.73

So the above outcome determines the evidence that such users are spaced from the least common group to the greater, so the system will provide the user 501 in seven groups.

After that, the system will search for a user similarity based on the definition and will create a matrix of data between users on the movie as shown in Table 5.

$$\bar{R}_t = (1+3)/2 = 2; \quad \bar{R}_c = (3+1)/2 = 2;$$

$$\bar{R}_t = 2 \text{ and } \bar{R}_c = 2$$

$$\begin{aligned} \text{RSim}(t_1, c_3) &= \frac{(1-2)(3-2) + (3-2)(1-2)}{\sqrt{(1-2)^2 + (3-2)^2} \sqrt{(3-2)^2 + (1-2)^2}} \\ &= \frac{(-1)(1) + (1)(-1)}{\sqrt{(-1)^2 + (1)^2} \sqrt{(1)^2 + (-1)^2}} \\ &= \frac{-1}{\sqrt{2} \sqrt{2}} \\ &= -0.5 \end{aligned}$$

The calculation is similar to Person correlation to the resemblance of making comparisons will get result e like this:

The similarities between User\_1 and User\_3 = -0.5

The similarities between User\_1 and User\_4 = 0.5

The similarities between User\_1 and User\_5 = 0

After that, the process of bringing the user that looks for rating similar to the target the number of K to predict satisfaction as possible by weight sum equation, Here, the size of the user is 2(K=2).

Therefore, only User 4 and 5 which is similar have considered for calculation.

$$K = \{c_4, c_5\}$$

$$P_{U1,I3} = \frac{(Rsim(t_1, c_4) * R_{U4,I3}) + (Rsim(t_1, c_5) * R_{U5,I3})}{|Rsim(t_1, c_4) + Rsim(t_1, c_5)|} = \frac{(0.5 * 4) + (0 * 2)}{|(0.5) + (0)|} = \frac{2}{0.5}$$

$$P_{U1,I3} = 4$$

The above calculation of the predicted rating User1 is equal to User4 for movie casino.

**Table 6** The calculation of the similarities between the User 1 and User 3

	Father of the bride	Golden eye	Casino	Four rooms	Money train	Get shorty	Assassins
User_1	1	4	?	3	?	3	2
User_3	3	?	1	?	2	1	2

### 3.6 Performance

The purpose of clustering is to know how many people in the groups and the centroid of the group. Then bring centroid to a cluster group for new user to the group by k-means algorithm. In this paper, we use WEKA software. It is used to cluster a group of users, the data downloaded from the website movie lens project data. It is so big to choose 500 users, 80

movies records and rating. First, we should transform data from excel to CSV because this file is supported by WEKA. We can delete something if unnecessary or not relate to your data. In our case we only use user, movies name and rating shown in the figure below (Figure 4).

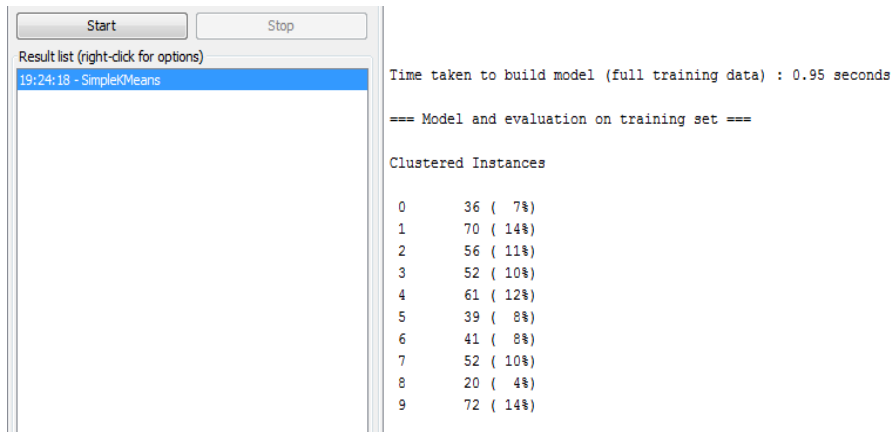
A1		User_ID																									
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U						
1	User_ID	Toy Story	Jumanji	(1997)	Grumpier	Waiting to	Father of	Heat	(1995)	Sabrina	(1995)	Tom and	Sudden D	GoldenEy	American	Dracula: D	Balto	(1999)	Nixon	(1994)	Cutthroat	Casino	(1995)	Sense and	Four Room	Ace Ventu	Moner
2	Us_2		1	4	3	2	5	2	2	3	4	3	1	4	3	2	5	2	2	3	4						
3	Us_3		4	1	2	2	2	3	1	4	1	3	4	1	2	2	2	3	1	4	1						
4	Us_4		4	1	2	3	3	1	5	4	2	2	4	1	2	3	3	1	5	4	2						
5	Us_5		5	5	4	1	2	2	3	4	4	5	5	5	4	1	2	2	3	4	4						
6	Us_6		4	4	4	3	4	3	1	1	2	5	4	4	4	3	4	3	1	1	2						
7	Us_7		2	1	3	3	4	1	5	2	3	3	2	1	3	3	4	1	5	2	3						
8	Us_8		2	4	1	3	2	1	5	4	4	2	2	4	1	3	2	1	5	4	4						
9	Us_9		3	5	5	4	3	3	2	2	3	4	3	5	5	4	3	3	2	2	3						
10	Us_10		5	2	2	1	4	3	2	3	4	3	5	2	2	1	4	3	2	3	4						
11	Us_11		4	3	3	4	4	1	3	5	5	5	4	3	3	4	4	1	3	5	5						
12	Us_12		1	5	1	2	5	4	4	4	2	3	1	5	1	2	5	4	4	4	2						
13	Us_13		2	4	2	2	4	5	3	3	3	4	2	4	2	2	4	5	3	3	3						
14	Us_14		4	3	4	1	1	2	4	2	4	2	4	3	4	1	1	2	4	2	4						
15	Us_15		5	2	3	1	2	2	5	1	3	3	5	2	3	1	2	2	5	1	3						
16	Us_16		2	4	1	5	4	4	4	3	1	5	2	4	1	5	4	4	4	3	1						
17	Us_17		2	5	2	3	2	1	2	4	2	1	2	5	2	3	2	1	2	4	2						
18	Us_18		3	5	4	4	3	2	1	5	4	2	3	5	4	4	3	2	1	5	4						
19	Us_19		4	4	2	2	4	5	3	2	5	3	4	4	2	2	4	5	3	2	5						
20	Us_20		5	3	3	2	4	3	4	1	3	4	5	3	3	2	4	3	4	1	3						
21	Us_21		2	5	5	4	2	5	5	3	5	2	2	5	5	4	2	5	5	3	5						
22	Us_22		1	4	4	3	2	1	2	2	4	4	1	4	4	3	2	1	2	2	4						

**Figure 4** The data user given movies rating



The data is selected and converted into a comma separated values (CSV) and formatted using the CSV converter and then cluster imported to WEKA program. It is the first open WEKA program, explorer, pre-process, and then open your file and click the cluster, after that you will see so many options and then click on the choose button to choose

simple k-means. Next left click on the smile k-means –N 2-A “weka.core. Euclidean Distance –R first-last” then it will show function. On these pages we can set the number of clusters depending on how many groups do you want. In our paper, we use 10 clusters to get the results as shown in *Figure 5 and 6*.



**Figure 5** Number of group and member of group

Cluster centroids:				
Attribute	Cluster#			
	Full Data (499)	0 (36)	1 (70)	2 (56)
User_ID	US_2	Us_9	Us_8	Us_11
Toy Story (1995)	3.1383	2.1111	2.2714	4.1607
Jumanji (1995)	3.2565	3.8889	3.8571	2.6964
Grumpier Old Men (1995)	3.2745	2.7222	2.3429	3.4107
Waiting to Exhale (1995)	3.2084	3.3611	4.0286	4.1786
Father of the Bride Part II (1995)	3.3327	3.0278	2.9571	3.1071
Heat (1995)	3.1042	3.1111	2.4857	3.4464
Sabrina (1995)	3.1815	2.2778	2.6714	4.2679
Tom and Huck (1995)	2.9679	2.1667	2.4429	4.2143
Sudden Death (1995)	3.2345	3.4722	2.9571	3.7857
GoldenEye (1995)	3.2545	3.8611	2.8286	3.4821
American President, The (1995)	3.1383	2.1111	2.2714	4.1607
Dracula: Dead and Loving It (1995)	3.2565	3.8889	3.8571	2.6964
Balto (1995)	3.2745	2.7222	2.3429	3.4107
Nixon (1995)	3.2084	3.3611	4.0286	4.1786
Cutthroat Island (1995)	3.3327	3.0278	2.9571	3.1071
Casino (1995)	3.1042	3.1111	2.4857	3.4464
Sense and Sensibility (1995)	3.1815	2.2778	2.6714	4.2679
Four Rooms (1995)	2.9679	2.1667	2.4429	4.2143
Ace Ventura: When Nature Calls (1995)	3.2345	3.4722	2.9571	3.7857
Money Train (1995)	3.3032	3.7778	3.0286	3.375
Get Shorty (1995)	3.1844	3.6111	3.1	2.8214
Copycat (1995)	3.2906	3.0833	3.6429	3.0893
Assassins (1995)	3.2986	3.1667	3.4857	3.0714
Powder (1995)	3.2193	2.7222	3.3571	3.1468

**Figure 6** Centroid of group

*Figure 5* shows that group 1 includes 36 people, group 2 includes 70 people, group 3 includes 56 people, group 4 includes 52 people, group 5 includes 61 people, group 6 includes 39 people, group 7 includes 41 people, group 8 includes 52 people, group 9 includes 20 people and the last group 10

includes 72 people, on the same time we will see *Figure 6*. That is centroids all of group for movies. Furthermore, by comparing this set of different groups, using the centroid method we figured out that while the group is increasing, members are getting improved in number too.

#### 4. Conclusion and future work

Collaborative filtering is the most successful and popular algorithm in the recommender system's field. It helps customers to make a better decision by recommending interesting items. Even though this algorithm is the best, it suffers from poor accuracy and high running time. To solve these problems, this paper proposed a recommendation approach based on user clustering by using the Euclidian distance to calculate two users to cluster dataset. This method combines clustering and neighbors' vote to generate predictions. In the future there may be techniques, fuzzy c-means in the group stages of the first system to provide a more effective segmentation.

#### Acknowledgment

None.

#### Conflicts of interest

The authors have no conflicts of interest to declare.

#### References

- [1] Zhan J, Hsieh CL, Wang IC, Hsu TS, Liao CJ, Wang DW. Privacy-preserving collaborative recommender systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*. 2010; 40(4):472-6.
- [2] Gong S. A collaborative filtering recommendation algorithm based on user clustering and item clustering. *Journal of Software*. 2010; 5(7):745-52.
- [3] Manvi SS, Nalini N, Bhajantri LB. Recommender system in ubiquitous commerce. In *international conference on electronics computer technology 2011* (pp. 434-8). IEEE.
- [4] Pu P, Chen L, Hu R. A user-centric evaluation framework for recommender systems. In *proceedings of the fifth ACM conference on recommender systems 2011* (pp. 157-64). ACM.
- [5] Hu R, Pu P. Acceptance issues of personality-based recommender systems. In *proceedings of the third ACM conference on recommender systems 2009* (pp. 221-4). ACM.
- [6] Pathak B, Garfinkel R, Gopal RD, Venkatesan R, Yin F. Empirical analysis of the impact of recommender systems on sales. *Journal of Management Information Systems*. 2010; 27(2):159-88.
- [7] Witten IH, Frank E, Hall MA. *Data mining: practical machine learning tools and techniques*. Morgan Kaufmann Publishers, Elsevier; 2011.
- [8] Han J, Kamber M. *Data mining: concepts and techniques*. Elsevier; 2011.
- [9] Hand DJ, Mannila H, Smyth P. *Principles of data mining*. MIT press; 2001.
- [10] Zhao Y. *R and data mining: examples and case studies*. Academic Press; 2012.
- [11] Kuźelewska U. Advantages of information granulation in clustering algorithms. In *international conference on agents and artificial intelligence 2011* (pp. 131-45). Springer Berlin Heidelberg.
- [12] McSherry D. Explaining the pros and cons of conclusions in CBR. In *European conference on case-based reasoning 2004* (pp. 317-30). Springer Berlin Heidelberg.
- [13] Felfernig A, Friedrich G, Schmidt-Thieme L. Guest editors' introduction: recommender systems. *IEEE Intelligent Systems*. 2007; 22(3):18-21.
- [14] Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*. 2005; 17(6):734-49.
- [15] Adomavicius G, Tuzhilin A. *Recommendation technologies: survey of current methods and possible extensions*. Information Systems Working Papers Series, 2004.



**Yu Lasheng**, Vice professor in Central South University of China. He is the member of ACM and CCF. He is ACM/ICPC golden medal coach. He received the B.Sc. degree in Computer Science, the Master degree and a Ph.D. degree in Control Theory and Control Engineering from Central South University. He is the editor of the *Journal of Convergence Information Technology* and *Advances in Information Sciences and Service Sciences*, etc. He is also the reviewer for the journals such as *Future Generation Computer Systems*, *Journal of Parallel and Distributed Computing* and *Artificial Intelligence Review*, etc. He has published at least 70 papers on Agent technologies or Algorithms, and has published 3 books. He has organized and implemented many projects which have created great achievements in the society. His main research interests include agent technologies and applications, structure and algorithm, smart computing etc.  
Email: yulasheng@csu.edu.cn