



REVUE ALGERIENNE DES SCIENCES

Section -A-

Sciences de la Nature et de la Vie

Sciences techniques

<http://univ-eltarf.dz/fr/>



Contextual Collaborative Filtering Recommendation Using Bio-inspired Algorithm

IbtissemGasmi^a, HassinaSeridi-Bouchelaghem^b, Fouzia Anguel^{a, c}

^aDepartment of Computer Science, Chadli BENDJEDIDUniversity, El Tarf, Algeria

^bLISCOLaboratory, BadjiMokhtar University, Annaba, Algeria

^cLabGEDLaboratory, BadjiMokhtar University, Annaba, Algeria

Informations	Abstract
Keywords : Recommender system Collaborative filtering Contextual information Genetic algorithm Movielens *Correspondance : gasmibtissem@gmail.com	Collaborative filtering algorithm has been successfully employed in e-commercial and academic application. However, traditional collaborative filtering methods use only users' evaluations and do not consider the change of the users' preferences over context. They give a same treatment on user's score in different contextual situations. In order to solve above problem, this paper proposes a new item-based collaborative filtering algorithm in order to incorporate contextual information of users and items. The proposed algorithm endows the rating of a user on an item with a weight function which takes into account the importance of each contextual information for users. Experimental results from Movielens data set show that the proposed algorithm outperforms the traditional item-based collaborative filtering algorithm.

1. Introduction

Nowadays, it is very difficult for users to find the most desirable information in a short time due to the increasing information in the world [1]. A recommender system offers a feasible solution to this problem. It aims to provide information or items (such as music, movie, books, etc.) that are likely of interest to a user, it has been successfully employed in e-commerce and academic applications [2]. Recommender systems can be largely divided into two categories: content-based recommendation and collaborative filtering recommendation. Content-based approach tries to predict the utility of items for a particular user based on the significant features describing the content of an item [3, 4]. Whereas, Collaborative Filtering approach ignores the contents of the item, and makes recommendations only based on the items previously rated by other users [5]. Despite its success and popularity, this technique suffers from several limitations including sparsity, system scalability, and synonymy. Moreover, the traditional collaborative filtering method does

not incorporate contextual information in the recommendation process. They ignore the changes of user's interests over context because they suppose that user's preferences are stationary [6]. However, in the real world, item's evaluation can be influenced by additional and varying contextual information. In addition, every user has a different importance or priority on each contextual feature. For instance, in many situations, evaluation time, user's gender or item's genre plays an important role in defining a user's preference for an item. In this paper, we propose a Context Aware Collaborative Filtering algorithm to incorporate the contextual information in the recommendation process. Firstly, the proposed algorithm uses the Genetic Algorithm to calculate the influence of each contextual parameter on users' preferences. Then it computes the weight function for each rating when calculating user predictions. The proposed weight function is a simple linear combination that reflects the similarity between two user's contexts.

The rest of the paper is organized as follows. Section 2 presents related works. Section 3 describes in details the proposed

algorithm. Section 4 presents experimental works. It provides data set details, evaluation metric and results of different experiments. The final section provides a conclusion.

2. Related works

More recently, a number of researches attempted to improve the quality of recommendation systems by taking contextual information into account. Such systems are called Context aware system [7-10]. Context is a multifaceted concept; it includes any information that can be used to characterize the situation of an entity [11].

Si et al. [12] present a system that incorporates user's dynamic contexts into the traditional Collaborative Filtering algorithm. Instead of using ratings, the authors use a dynamic profile vector. Park and Kahng study the temporal dynamics in users' music listening behaviors by considering periodicity of the time dimension and popularity change [6]. Sun et al. [1] divide the user's rating history into several periods, then analyze user's interests distributing in these periods by a phrased forecast method. To integrate context into Collaborative Filtering, Chen determines relevant ratings for a given context using the similarity of context proposed [13].

At the same time, several researchers use machine learning methods to ameliorate recommender systems. Genetic algorithms are randomized search and optimization technique based on the evolutionary ideas of natural selection and genetics. It has mainly been used in two aspects of a Recommender System: clustering [14, 15] and hybrid methods [16, 17].

Abdul Hameed et al. propose a method combining Genetic Algorithms based clustering algorithm with Collaborative Filtering [18]. Salehi et al. consider weights of implicit or latent attributes of materials for learner as chromosomes in Genetic Algorithm and according to historical rating they optimize the weights [19]. Then, they generate recommendation by Nearest Neighborhood Algorithm using the optimized weight vectors. Karatzoglou et al. present a Genetic Collaborative Filtering model that is based on a generalization of matrix factorization to address contextual recommendation problems [20].

3. Proposed algorithm

The challenge of collaborative filtering system is to improve the quality of recommendations for the users. Traditional collaborative filtering methods do not exploit any contextual information about the users or

items. They ignore the changes of user's interests over context because they suppose that user's preferences are stationary. Thus, ratings produced at different contexts are weighted equally. Nevertheless, in real life, to say that two people are similar is not based solely on their opinions on a specific topic, but also on other factors such as their age and profession. Under the framework of item-based collaborative filtering, this paper proposes a new hybrid algorithm which considers the influence of each contextual parameter on users' preferences. The weight function is introduced to endow each score with a weight value according to context. In order to assign the optimal value for each contextual information, the proposed algorithm uses a genetic algorithm.

3.1 String Representation

The chromosome is a string of length 5. It is encoded with real numbers between 0 and 1; the number of genes in each chromosome is equal to the number of features (user's age, user's gender, user's occupation, movie's genre, time evaluation). A simple chromosome may look like as follows:

W1	W2	W3	W4	W5
----	----	----	----	----

Where w_1 is the weight associated to the user's age.

w_2 is the weight associated to the user's gender.

w_3 is the weight associated to the user's occupation.

w_4 is the weight associated to the movie genre.

w_5 is the weight associated to the time evaluation.

the sum of the weights in a chromosome should be equal to 1.

$$\sum_i w_i = 1 \quad (1)$$

The initial population is formed by generating N random strings represented in figure 1.

3.2 Fitness function

In this paper the fitness function will be the Mean Absolute Error (MAE) of the Recommender System. The fitness function used for evaluating the fitness of our Genetic Algorithm is determined by equation (2):

$$\text{fitness} = \frac{\sum_i |p_i - q_i|}{n} \quad (2)$$

For each prediction-rating pair $\langle p_i, q_i \rangle$, the fitness function treats the absolute error between p_i and q_i ,

The lower the fitness value is, the more accurately the recommendation engine predicts user ratings.

In order to calculate the fitness value, it is necessary to follow the next steps for each user x :

- Calculating the set of K neighborhood of user u :

The most important phase is to measure the similarities between users [4]. Firstly, the proposed algorithm collects those users who shared common items with the active user. Then, the similarity weight between the active user and his neighborhood is computed by Pearson correlation coefficient:

$$corr_{u,v} = \frac{\sum_I (R_{u,i} - \bar{R}_u)(R_{v,i} - \bar{R}_v)}{\sqrt{\sum_I (R_{u,i} - \bar{R}_u)^2 \sum_I (R_{v,i} - \bar{R}_v)^2}} \quad (3)$$

$R_{u,i}$ denotes the rating of user u on item i .

I denotes the set of common items for users u and v .

$\bar{R}_u$ is the average ratings of user u .

- Calculating the prediction :

The predicted value of the user u to unrated item i is computed using the following equation:

$$p_{u,i} = \bar{R}_u + \frac{\sum_{j=1}^K (R_{j,i} \times corr_{u,j} \times f(u,j))}{\sum_{j=1}^K (corr_{u,j} \times f(u,j))} \quad (4)$$

Where $corr_{u,j}$ is the similarity weight between users u and j .

$R_{j,i}$ is the rating of user j for item i .

K is the number of neighborhood of user u .

$f(u,j)$ represents the degree of similarity between the contexts of user u and j .

$0 \leq f(u,j) \leq 1$ is the weight function defined as:

$$f(u,j) = \sum_l w_l f_l(x_u, x_j) \quad (5)$$

$$f_l(x_u, x_j) = \begin{cases} 1 & \text{if } x_u = x_j \\ 0 & \text{if } x_u \neq x_j \end{cases}$$

Where w_l is the weight for contextual feature $l \in [\text{user's age, user's gender, user's occupation, movie's genre, time evaluation}]$.

x_u and x_j are the feature's value of users u and j , respectively.

3.3 Selection

Selection is a genetic operator that gives preference to better chromosome. The goodness of each individual depends on its fitness which is determined by an objective function. The probability of an individual to be selected is computed using the following equation:

$$P_i = F_i / \sum_{j=1}^N F_j \quad (6)$$

N is the size of the population.

F_i represents the fitness degree of an individual i .

3.4 Crossover

Crossover is a genetic operator that exchanges genes between two chromosomes for generating two new chromosomes. In this work, single point on both parents' chromosome is selected (figure 1). All data beyond that point in either string are exchanged to produce two children for the next generation.

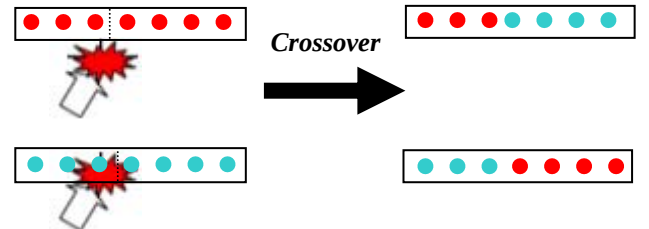


Fig. 1 :Single-point crossover

3.5 Mutation

Mutation is a genetic operator that alters one or more gene values in a chromosome (figure 2). A random position is chosen in the chromosome and all data beyond that point are replaced by a random number between 0-1. Each chromosome applied mutation with a probability p_m .



Fig. 2 :Single-point mutation

3.6 Termination criterion

In this paper the processes of fitness computation, selection, crossover, and mutation

were evolved for M number of iterations or stopped when no change had occurred for over 10 generations. The best chromosome seen up to the last generation provides the solution to our problem.

4. Experiments

In order to evaluate the effectiveness of the proposed algorithm, we compare the results of the proposed system to those traditional Collaborative Filtering algorithms. We repeated the experiments by using different parameters to show how the feature weights affect the fitness accuracy and performance.

4.1. Dataset

In order to evaluate the recommendation quality, the algorithm was tested using Movielens1, a data set widely used to evaluate collaborative filtering algorithm. This data set is provided by the University of American Minnesota GroupLens project group. It contains 10000054 ratings on a scale from 1 to 5 and 95580 tags applied to 10681 movies by 71567 users. Users were selected at random for inclusion and each one had rated at least 20 movies.

MovieLens dataset contains much contextual information (such as age, gender, occupation and postal code) which can be exploited to improve the prediction accuracy of Recommender Systems. The information of a user includes: UserID, Gender, Age and Occupation and those of the item are MovieID, Title, Genres.

4.2. Prediction and recommendation results

Experiments were performed to show the performance of the proposed recommendation method. The Pearson correlation coefficient is used for finding the neighbor users with different values for the number of neighbors k. Table 1 informs about the MAE error obtained for Movielens when applying Pearson correlation. The results show that each recommendation method has obtained its best result for a different number of neighbor users. Moreover, for any value of K (the number of neighbors for each user) the results obtained with our genetic algorithm are better than the ones obtained with the traditional collaborative filtering algorithm.

So, experimental result from Movielens data set validates that the weight function proposed in this paper improves the quality of recommendations. In fact, this function incorporates the contextual factors that influence the users' decisions. It integrates also a degree of importance for each contextual parameter.

Tab. 1: Mean absolute error comparison of different number of neighbors

K	Traditional Collaborative filtering	Proposed algorithm
8	2,32	1.21
9	2,03	1.15
10	1,75	1.29
11	1,39	1.20
12	1,14	1.08
13	0,91	0.83
14	0,83	0.75

Moreover, parameters setting for the Genetic Algorithm could be found that gives better results. The values of Genetic Algorithm parameters are listed in table 2.

Tab. 2: The parameters of the Genetic Algorithm

Population Size	80
Crossover Rate	0.8
Mutation Rate (pm)	0.2
Crossover method	Single point
Selection method	Tournament

Table 3 shows the best approximate weighting values generated by the Genetic Algorithm. Obviously, we assume that time evaluation, gender and occupation still play an important role in preference prediction.

Tab. 3: The weight of each contextual parameter

Weight of age	0,22
Weight of gender	0,20
Weight of occupation,	0,10
Weight of genre	0,15
Weight of time evaluation	0,33

5. Conclusion

Collaborative filtering based item has been known to be the most successful recommendation technique [4]. However, Traditional collaborative filtering methods are

¹ <http://www.grouplens.org/>

only rely on the user's preferences, which is insufficient in real world problem. In recent years, contextual information is more and more important in collaborative filtering. Thus, Several researchers try to exploit contextual features to improve the prediction accuracy of Recommender Systems.

This paper proposes a new hybrid collaborative filtering algorithm that takes contextual information into account. Genetic Algorithm is used to find an optimal combination of weight features. Then, The weight function is introduced to endow each score with a weight value according to context. Experimental results from Movielens data set show that the new algorithm outperforms the traditional item-based collaborative filtering algorithm.

References

- [1] SunJ., Zhao J.and YuX., (2009). A Collaborative Filtering Algorithm with Phased Forecast. *Rough Sets and Knowledge Technology*. vol. 5589, pp. 698-705.
- [2] Serrano W., (2019). Intelligent Recommender System for Big Data Applications Based on the Random Neural Network. *Big Data and Cognitive Computing*.
- [3] PazzaniM. J.(1999). A Framework for Collaborative, Content-Based and Demographic Filtering. *Artificial Intelligence review*, vol.13, pp.393-408.
- [4] BurkeR.(2002). Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction*., vol.12, pp. 331-370.
- [5] Zhou J. and LuoT. (2009). Towards an Introduction to Collaborative Filtering", *Proceedings of the International Conference on Computational Science and Engineering*, pp. 576-581.
- [6] Park C. H. and KahngM.(2010). Temporal Dynamics in Music Listening Behavior: A Case Study of Online Music Service. *Proceedings of IEEE/ACIS 9th International Conference on Computer and Information Science*, pp. 573-578.
- [7] Gasmi I., Seridi-Bouchlaghem H., Labar H. (2012) Collaborative filtering recommendation based on research habit of users. In: *Proceedings of the third IEEE International Conference on Multimedia Computing and Systems*, Tangier, Morocco.
- [8] Gasmi I., Seridi-Bouchlaghem H., Labar H., Baareh A. (2015). Collaborative filtering recommendation based on dynamic changes of user interest. *Intelligent Decision Technologies*, vol. 9, no. 3, pp. 271-281.
- [9] HusseinT., LinderT., GaulkeW., and ZieglerJ.(2009). Context-aware recommendations on rails. In *Proceedings of the workshop on context aware recommender systems*, pp. 22-25, New York, USA.
- [10] Bobadilla J., Ortega F., Hernando A., Alcalá J.,(2011). Improving collaborative filtering recommender system results and performance using genetic algorithms. *Knowledge-Based Systems*, Vol. 24, N°. 8, pp. 1310 - 1316.
- [11] DeyA. K.(2001). Understanding and using context. *Personal and Ubiquitous Computing*. vol. 5 (1), pp.4-7.
- [12] SiH., KawaharaY., KurasawaH., MorikawaH., and AoyamaT.(2005). A context aware collaborative filtering algorithm for real world oriented content delivery service", *Proceedings of the 7th international conference on ubiquitous computing, Metapolis and urban life workshop*, Tokyo, Japan, pp. 65-69.
- [13] ChenA., (2005).Context-aware collaborative filtering system. *Proceeding of The Springer International Workshop on Location and Context-Awareness*, Berlin, Heidelberg, vol. 3479, pp. 244-253.
- [14] Kim K.andAhnH.(2008). A recommender system using GA K-means clustering in an online Shopping market.Expert system with application, , vol. 34 (2), pp. 1200-1209.
- [15] Zhang F. and ChangH.Y. (2006).A collaborative filtering algorithm employing genetic clustering to ameliorate the scalability issue. *Proceedings of the IEEE International Conference on e-Business Engineering*, pp. 331-338.
- [16] Y. Ho, S. Fong and Z. Yan.(2007). A hybrid GA-based collaborative filtering model for online recommenders. *Proceedings of the International Conference on e-Business*, Barcelona, Spain, 200-203.
- [17] Al-ShamriM.Y. and BharadwajK.K.(2008) Fuzzy-genetic approach to recommender systems based on a novel hybrid user mode.Expert system with application, pp. 1386-1399.
- [18] M. Abdul Hameed, O. Al Jadaan and S. Ramchandram. (2010). Information Theoretic approach to Cold Start; Problem using Genetic Algorithms. *Proceedings of the International Conference on Computational Intelligence and Communication Networks*.
- [19] SalehiM., PourzaferaniM. and RazaviS.(2013)"Hybrid attribute-based recommender system for learning material using genetic algorithm and a multidimensional information model", *Egyptian Informatics Journal*, vol.14(1), pp. 67-78.
- [20] aratzoglouA., AmatriainX., BaltrunasL. and OliverN. (2010)MultiverseRecommendation: N-dimensional Tensor Factorization for Context-aware Collaborative Filtering. *Proceedings of the fourth ACM conference on Recommender Systems*, New York, USA, pp. 79-86,.