A Genetic Based Recommender System

BOUACHA Ismail
Department of Second Cycle
ESTI Annaba
Algeria
bouacha.ismail@gmail.com

***Abstract*— In this paper we propose a genetic based recommender system. Our goal is to predict relevant items (films) from a huge space of data based on users ratings. We use MovieLens data set (which contains data on ratings made by users about films) to validate our approach. Our idea consists on finding for an active user, the most similar group of users using genetic algorithms. After that, we recommend items (films) based on prediction of the appreciation that the active user can give for each one. Several metrics have been used to prove the efficiency of our approach such Mean Absolute Error, precision and recall.**

Keywords—Recommender System; MovieLens; Genetic Algorithm;Similarity

I. INTRODUCTION

Recently, recommender systems have led to a revolution in all areas: e-commerce [5], e-learning [6], entertainment [8], etc. Indeed, the large amount of data available on the web makes the choice of items (product, film, book, etc.) difficult for the user. On the other hand, companies are trying to increase their profits by proposing items deemed relevant to users. Recommendation systems seek to satisfy users by offering them items that are supposed to be interesting for them, using techniques such as content-based filtering [7], demographic filtering [4] and collaborative filtering [9]. The first type recommends items to an active user based on his history. It uses items preffered in the past to find new ones. The second filtering, uses information like sexe, age , etc in recommendation process. The last filterting uses similarity between users to recommend items. It is also possible to hybrid these techniques to exploit the advantage of each one.

Several researchers have proposed to use genetic algorithms in recommender systems [3]. However, most of them tried to find the optimal similarity metric or to propose a new one [2]. In both cases, proposed works require training and testing data to compute and test similarity between users.

We propose in this paper to use genetic algorithm and collaborative filtering to recommend items. Our idea consists on finding for an active user, the most similar group of users. After that, we apply a collaborative filtering to find relevant items. We believe that the first step can be seen as an NP-complete global optimization problem. We used genetic algorithm, to provide a near-optimal solution in reasonable time.

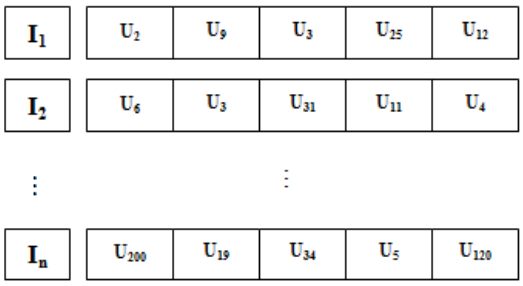
The rest of this paper is organized as follow: the next section shows our details of our proposition; next experimental results will be presented in section 3. Then we conclude in the last section.

II. PROPOSITION

A. Genetic Algorithm and similar users

First, A recommender system should be able to provide users with useful information on items that may be of interest to them. In this section, we present an innovative recommendation system that combines collaborative filtering technique and genetic algorithms. Genetic algorithms are based on the principles of evolution. An initial population of individuals could be generated randomly. Each individual represents a possible solution for a given problem. To evaluate the individual quality a fitness function should be defined. Genetic operators (crossover and mutation) could be used to generate new individuals from current ones. A new population is generated using selection operator until achieving the best fitness or exceeding a maximum iteration number.

Fig. 1. Population, individual and gene



As shown in the figure above, A population could be composed of n individuals I_i, each one represents a group of possible similar user U_i. To measure the similarity between users, several similarity metrics have been proposed such: Pearson Correlation [1], Cosine, etc.

To measure the fitness of a given solution (individual), we used the following function:

$$fitness(I_i) = SUM(PC(U_*, U_i)) / G \quad (1)$$

where: I_i : Individual i, PC: Pearson Correlation, U_*: the active user, U_i: User i in the individual, G : the number of users in the individual.

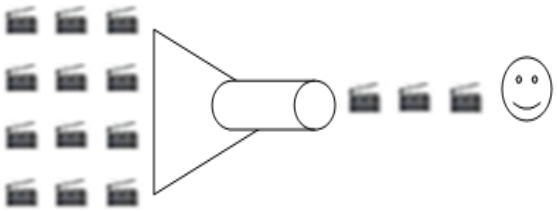
We applied one-point crossover and single point mutation techniques to generate new individuals. The crossover operator generates a new individual from one or two parent depending on a rate of crossover which is the probability of performing a crossover. We choose one part from each parent. The mutation operator updates on user from the users of the individual depending on mutation rate.

The selection step aims to maintain the most adapted individuals. For this reason, we used biased roulette to perform it.

B. *Collaborative filtering and recommendation*

The recommendation consists on estimating (predicting) the rating that the active user can attribute to each of the films to be recommended. This prediction is done by averaging the ratings assigned by similar users. Then we recommend the films considered relevant by setting a threshold.

Fig. 2. Recommendation

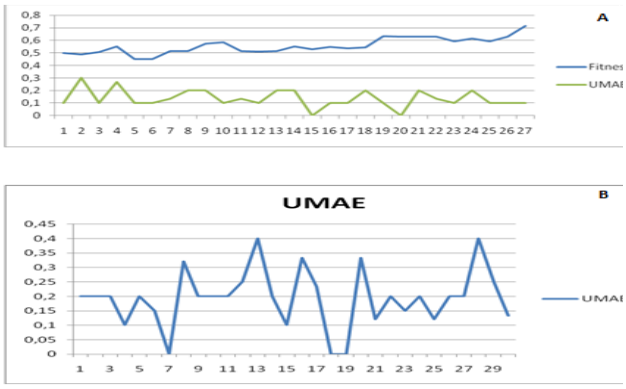


To recommend relevant items, we have two steps. The first one is the composition of items based on the group of similar users we have obtained by the genetic algorithm. The second step is the filtering of these items, based on the prediction rating for each one.

III. EXPERIMENTAL RESULTS

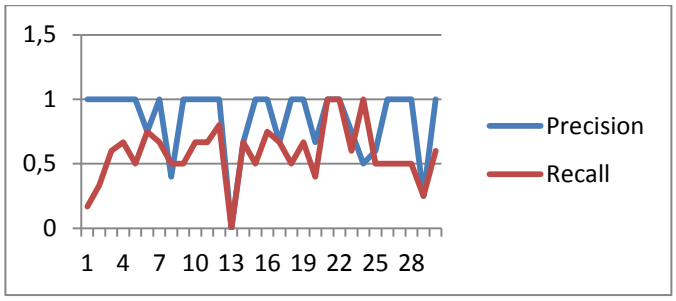
In this section we present results we have obtained after the implementation of the proposed approach. We used the most popular dataset in the domain of recommendation which is MovieLens dataset. It contains 100.000 ratings (each rating is between 1 and 5) made by 943 users about 1682 films. We have used UMAE metric to measure the efficiency of our approach. We present below results obtained by running our genetic recommender system 30 times.

Fig. 3. Results



The figure 3.A shows the evolution of fitness and UMAE error for the same user. We can see clearly that the best fitness value is 0.7 and the best UMAE is 0.1. The figure 3.B shows the evolution of UMAE for about 30 different users. We can see that the minimum UMAE is 0, the maximum UMAE is 0.4 and the average UMAE is less than 0.2. The figure 4 shows the evolution of precision and recall for about 30 different users. As shown in the figure the average precision value is about 0.8, and the average recall value is about 0.7.

Fig. 4. Precision & recall



The results we have obtained are encouraging and reflect the effectiveness of the proposed approach. Indeed the choice of the near optimal group of the most similar users ensures an acceptable error. On the other hand, we note that accuracy and recall are also considerable, despite the fact that the user group chosen contains less similar users, since we used an average of similarities. This last point gives the system a chance to recommend items that are appreciated by less similar users. This gives a diversification of recommended items.

IV. CONCLUSION

In this paper we presented a genetic algorithm approach for recommendation. It consists on finding for an active user, most similar users using Pearson Correlation. The group of similar users represents the source of recommendation which uses a prediction processing that help the system to find relevant items. MovieLens dataset have used as case study, and we also used metrics such as (Mean Absolute Error) MAE, (User Mean Absolute Error) UMAE, recall and precision to prove the efficiency of the proposed approach. First results seem to be interesting, and encourage us to improve this work by studying more datasets and more similarity metrics.

REFERENCES

[1] G. Adomavicius, A. Tuzhilin, 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 17 (6) (2005) 734–749.

[2] B. Alhijawi, Y. Kilani, Using Genetic Algorithms for Measuring the Similarity Values between Users in Collaborative Filtering Recommender Systems. Proceeding of the 15th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2016), 2016, 1-6.

[3] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutierrez, "Recommender ´ Systems Survey," Journal of Knowledge-Based Systems, vol. 46, no. 8, 2013, pp. 109–132.

[4] R. Burke, "Hybrid Recommender Systems: Survey and Experiments,"Journal of User Modeling and User-Adapted Interaction, 2002, pp. 331–370.

[5] J.J. Castro-Sanchez, R. Miguel, D. Vallejo, L.M. López-López, A highly adaptive recommender system based on fuzzy logic for B2C e-commerce portals, Expert Systems with Applications 38 (3) (2011) 2441–2454.

[6] H. Denis. Managing collaborative learning processes, e-learning applications. 29th Int. Conf. Inf. Tech. Interfaces, (2007) 345–350.

[7] K. Lang. NewsWeeder: Learning to filter netnews, 12th Int. Conf. Machine Learning, (1995) 331–339.

[8] S.K. Lee, Y.H. Cho, S.H. Kim, Collaborative filtering with ordinal scale-based implicit ratings for mobile music recommendations, Information Sciences 180 (11) (2010) 2142–2155.

[9] Y. Manolopoulus, A. Nanopoulus, A.N. Papadopoulus, P. Symeonidis, Collaborative recommender systems: combining effectiveness and efficiency, Exp. Syst. Appl. 34 (4) (2008) 2995–3013.