Genetic Recommend Generating Method with Real-time Fitness Function Adaption*

Minchul Jung, Jehwan Oh, and Eunseok Lee

School of Electrical and Computer Engineering, Sunkyunkwan University,
300 Cheoncheon-dong, Jangan-gu, Suwon, Gyeonggi-do 440-746, South Korea
{mcjung, hide7674, eslee}@ece.skku.ac.kr

Abstract. Existing recommender systems generate recommendation usually using user's information previously collected. The information reflects the user`s tastes, but it doesn`t include user`s intend at that time. So existing recommender systems sometimes generate not suitable recommendation because of difference between user`s current purpose and the information of past time. In this paper, we propose genetic recommend generating method for overcome this problem. Our method analyzes user`s real-time click-stream for grabbing user`s current intention, then uses genetic algorithm for generating appropriate recommendation. To reflect user`s real-time intention, the proposed method adapts fitness function of genetic algorithm continuously. To evaluate the proposed approach, we compare the proposed method with existing CF methods using the web-server log data collected from Internet jewelry shop. And we confirm that the proposed approach can generate more accurate recommendation then compared methods.

Keywords: Genetic algorithm, Recommender system, User preference, User profile, Real-time recommendation

1 Introduction

Due to recent advances in information and communication technology and growth of The Internet users, the amount of information on the Internet has growing exponentially. It makes difficult to find information on the Internet for user. To resolve this problem, the recommender system that offers the appropriate information for the user has been actively studying.

General recommender systems create a recommendation based on information about the user. Information is collected using explicit and implicit method [2]. The existing recommender systems generate recommendation based on the above-mentioned information-gathering techniques to determine the user's preference. However, this recommendation does not consider the intent of the user. So existing

* This work was supported by the MKE 21st Century Frontier R&D Program in Korea and a result of subproject UCN 08B3-B1-10M, ITRC IITA-2008-(C1090-080-0046), Grant No. R01-2006-000-10954-0, Basic Research Program of the Korea Science & Engineering Foundation.

recommender systems sometimes generates inappropriate recommendation because of difference between user`s current purpose and information of past time. Indeed, the inappropriate recommendation makes a user angry because the recommendation was generated based on information about the gift to someone else before [3]. Furthermore, the existing recommender systems are not considering the intent of the current user. So the generated recommendation is not appropriate for the current situation [3].

In this paper, we propose genetic recommendation generating method offers real-time recommend in order overcoming drawbacks of the existing systems. The proposed method generates lists of recommendation using genetic algorithm, and using a fitness function of genetic algorithm that estimate the suitability of recommended lists (that is the best recommendation for the current user). In addition, our method collects user`s real-time click-stream to analyze the user's intent continually, and then adapts the fitness function of genetic algorithm to fit the user's current intention.

In this paper, we compared proposed method with existing recommender systems to evaluate the proposal. The experiments have been conducted using the server log files from the Internet jewelry store. We used precision and coverage to evaluate the proposed method. As a result, we prove that the accuracy of the proposed method is better than existing methods.

In addition, we measure the change of the accuracy as the amount of real-time user behavior increasing to evaluate whether the proposed technique can reflect the intent of the user in real-time. As a result, we verified the accuracy of recommendation is increasing as real-time user behavior information is increasing.

2 User Model

In this paper, we use two user models: one is based on the previously collected information as the model of the existing recommender system, and the other one is the real-time user's model. The first user model represents the user's preference for a particular product. But it cannot determine the user intention at the time of purchase, and does not fit with the user's current intent. So, in this paper, the model is only used assistant information for the fitness function of genetic algorithm.

In this paper, to collect information of the user, we gathered the user click stream implicitly. The benefits of implicit method is to collect much information than the explicit way, and do not bother users.

The click stream is represented as log format, and the user's behavior in a specific product page can be analyzed based on this information. In addition, user`s item preference can be inferred by using user`s behavior score. For example, if the user purchases the products, the preference score is 10 points, 2 points to see the large image of product, and interest in the detailed item information is adding 5 points. This score is stored in the DB and is used for analyzing the user's preference in the past [4].

The real-time user information is also analyzed based on real-time click stream. The visited categories and its frequency, the visited products, search keywords, and search options are included in the information. Based on this information, we can determine what the user wants. Here, we divide the intent of user as three states: the

user does not have a purpose (just wandering between items), the user is interested in a particular category, and the user is interested in a particular product. Depending on the user's actions the importance of each state gradually changes. It is reflected in the fitness function of the genetic algorithm to generate the appropriate recommend list for current situation.

3 Genetic Recommendation Generating Method

3.1 About the Genetic Algorithm

A genetic algorithm is a search technique used in computing to find exact or approximate solutions to optimization and search problems [1]. It is after the model of the nature ability to adapt environment. The process of genetic algorithm is shown in Figure 1.

Step. 1 Generate initial population.
Step. 2 Evaluate each individual using fitness function.
Step. 3 Iterate until stopping criteria
  Step. 3.1 Selection and Crossover
  Step. 3.2 Mutation
  Step. 3.3 Evaluate each individual using fitness function.
  Step. 3.4 Make next generation
  Step. 3.5 Evaluate stopping criteria

Fig. 1. The process of genetic algorithm

3.2 Chromosome Encoding

Fig. 2. An example of individual

In genetic algorithm, the individual in a group is represented by a set of chromosomes [1]. In this paper, we represent each individual as a list of recommendation provided to the user. Each individual is made up of 10 chromosomes and each chromosome is represented a product in the recommendation list. Recommended products are represented in the product ID on the product DB. Figure 2 shows an individual used in this paper.

3.3 Population

In genetic algorithm, to mimic natural selection, set the size of the population. Individuals have the low fitness score are abandoned to maintain the size of individuals after the operation of genetic algorithm [1]. As the size of the population is increases, the population has more diversity. But it makes the computing time of genetic algorithm increasing. In this paper, we set the number of individual as 10 to generate real-time recommendation and to maintain certain rate of diversity.

Genetic algorithm creates the initial individual as predefined population randomly [1]. However, the individual of this paper reflects the tastes of the user. So 5 individuals are generated randomly, and another 5 individuals are generated based on existing user's information.

3.4 A Fitness Function

In genetic algorithm, a fitness function is used to evaluate the superiority of the each individual and the higher score indicates that the individual is better [1].

One of our goal is to identify and adapt the user's real-time intent for provide a list of the appropriate recommendation. The information collected in real-time to reflect on the evaluation function for generate more appropriate recommendation. For more information about this will be discussed in Section 4.

3.5 Genetic Operations

During each successive generation, a proportion of the existing population is selected to breed a new generation. Individual solutions are selected through a fitness-based process, where fitter solutions as measured by a fitness function are typically more likely to be selected [1].

In this paper, we choose a roulette wheel selection. A form of fitness-proportionate selection in which the chance of an individual's being selected is proportional to the amount by which its fitness is greater or less than its competitors' fitness. Conceptually, this can be represented as a game of roulette - each individual gets a slice of the wheel, but more fit ones get larger slices than less fit ones. The wheel is then spun, and whichever individual owns the section on which it lands each time is chosen.

In genetic algorithm, crossover operation selects genes from parent chromosomes and creates a new offspring [1]. In this paper, we used uniform crossover because the genes of each individual can be replaced with an equal probability. Each gene is equal to the intersection to be replaced as independently (see Figure 3).

Mutation in the genetic algorithm is used for maintaining diversity as a chance to insert new genes in a group [1]. In this paper, we set the probability of mutation as 5%, and use random exchange method. This mutation method is changed one or more genes randomly at selected individuals.

Fig. 3. Uniform cross-over

3.6 Stopping Criteria

Genetic algorithm has the condition to terminate it. If the condition is reached, we believe that the genetic algorithm to achieve the purpose of the process and terminates the process. The terminate condition is usually certain amount of the generation or the fitness score [1]. In this paper, it is hard to set the goal fitness score because the user`s purpose is changing in real-time. Therefore, after the genetic process is passed 10 times the process is set to shut down conditions.

4 Real-time Fitness Function Adaption

Genetic algorithm used in this paper is applied to the evaluation criteria divided into 3: the diversity of recommendation, the suitability of recommendation, and the user`s product preference. The diversity of recommendation means how much the recommended list includes the product that the user has not seen before. And the suitability of recommendation indicates how much the recommended list is appropriate for user`s real-time intention. It is measured based on the frequency of visits of the category. Finally, the user preference of product means how similar the product in the recommendation list to the product favored previously.

The formula to calculate the diversity of the recommend is the same as expression (1): N_i means the number of new products included in the individual i, and α is a constant value for normalization.

$$V(i) = \alpha \times N_i \quad \textbf{(1)}$$

The formula to calculate the suitability of the recommendation is at expression (2): C_{ij} means the visit frequency of the category j belong to the individual i, where a category of products that recently visited 10 categories, and visited ratio. β is a constant value for normalization.

$$S(i) = \sum \beta \times C_{ij} \quad \textbf{(2)}$$

The formula to calculate the user preference of product is shown in expression (3): sim_{ij} indicates the similarity between item j in the individual i and the item previously favored. λ is a constant value for normalization.

$$P(i) = \sum \lambda \times sim_{ij} \quad (3)$$

The overall evaluation function is the expression (4). And ω_N, ω_S, ω_P represents the weight of each evaluation criteria to maintain the sum makes 3. In the early time with the value of 1:1:1, but as the intention of the current user changes, the goal of criteria also changes for making the adapted recommendation. In addition, the constant value of α, β, and λ should be set differently depending on the policy, in this paper the values were set to 1:1:1.

$$f(i) = \omega_N \times N(i) + \omega_S \times S(i) + \omega_P \times P(i) \quad (4)$$

As mentioned in Section 2, the goal of the user is divided into 3 kinds of state: the user does not have special purpose, the user is interested in a particular category, and the user is interested in a particular product. If the user does not have the purpose just to see a variety of products, the diversity of the recommend is important to generate a list of recommends. If the user is interested in a specific category, the suitability of a particular category and a variety of products within that category is important. If the user is interested in a specific product, the most similar product is included in a list of recommended items. It is hard to catch the purpose of the user at a time, depending on the user's behavior gradually changed for each state, bring the importance of the evaluation function is reflected in the genetic algorithm to generate the appropriate recommendation list for the current situation.

5 Evaluation Environments

In this paper, we used the data collected from the Internet jewellery shop by 2 months to evaluate it. The data includes the user's behaviour, purchasing information, and click stream records from the web server log file. The number of user is 137, and the number of products is 230. In addition, the number of collected click stream is nearly 35000.

Analysis of the user's preferences is based on the behaviour of the user, and each user's behaviour is analyzed by the score. In addition, we generate item similarity matrix using Item-to-Item CF technique [7] for referring the fitness function of genetic algorithm.

In this paper, we evaluated the proposed method in terms of the accuracy and the diversity of the recommendation. We use two measures: the accuracy was measured by the *Precision* and the diversity was measured by the *Item Coverage*. The measures were typically used in the evaluation of the recommender systems.

6 Evaluation Results

6.1 Comparison Results

In this paper, we compared the proposed method with two existing recommender systems: user-based CF and item-based CF [6]. We divided the test data in the ratio of 80 to 20 to evaluate: 80 used for learning data, and 20 used for verification data. The existing systems make recommendation for each user with the top 30 items in the list of recommended items. And the proposed method use the top 3 of recommendation (the number of products is 30) generated at the end of each transaction.

The comparison is executed in terms of the accuracy and the diversity of the recommendation. The results of comparison are presented in Table 2. According to the experiment result, our proposed method is better than existing method in precision 8~10% and in coverage 11-17%.

Table 1. The comparison results between proposed method and existing method

Method	Precision	Coverage
User-based CF	0.3889	0.7564
Item-based CF	0.4017	0.8156
Proposed method	0.4820	0.9242

6.2 Measuring the Accuracy of Real-time Recommendation

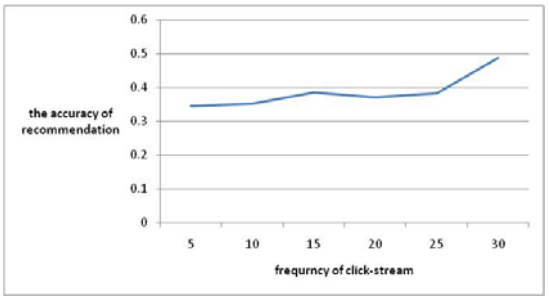

Fig. 4. An evaluation result for accuracy of real-time recommendation

To measure the accuracy of the system's real-time recommendation, we divide the data by 80:20. And we select the user data who`s transactions was occurred over 15. Then we generate recommendation using each user's recent 3 transaction data and measure the accuracy of recommendation based on the items in the top 2 individuals. The result is presented in Figure 4. The results show as user's real-time information was increase; the accuracy of the recommendation also seems to rise.

7 Conclusions and Future Work

In this paper, we proposed real-time genetic recommendation method in order to overcome the existing recommendation techniques are not reflect the current user`s intend. The proposed method collects the user's click stream, and analyzes the user's intent. Then the method generates recommendation fit the user's intention by using the genetic algorithm. The fitness function of genetic algorithm adapts user`s real-time intention continually. By comparing with existing method, the experimental results show the proposed method could deliver more diverse and more accurate recommendation than existing methods. In addition, we measured the change of the accuracy as the amount of real-time user behavior increases to evaluate whether the proposed technique can reflect the intent of the user in real-time. As a result, we verified that the accuracy of recommendation is increasing as real-time user behavior information increases.

Additional future research will be optimizing the elements of genetic algorithm for the recommendation research domain to improve the accuracy of recommendation. In addition, the context information can also be applied on real-time recommendation method to improve the performance of analyzing real-time intent of the user.

References

1. Mitsuo, G., Runwei, C.: Genetic algorithm & Engineering Optimization, John Wiley & Sons (2000)
2. Gediminas, A., Alexander T.: Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. IEEE Transaction on Knowledge and Data Engineering 17(6), pp.734—749 (2005)
3. Gorgoglione, M., Palmisano, C., Tuzhilin, A.: Personalization in Context: Does Context Matter When Building Personalized Customer Models? In: 6th IEEE International Conference on Data Mining, pp.222—231 (2006)
4. Cho, J., Kwon, K., Park, Y.: Collaborative Filtering Using Dual Information Sources. IEEE Intelligent Systems 22(3), pp.30—38 (2007)
5. Zanker M., Jessenitschnig M., Jannach D., Gordea S.: Comparing Recommendation Strategies in a Commercial Context. IEEE Intelligent Systems 22(3), pp.69—73 (2007)
6. Zan, H., Daniel, Z., Hsinchun, C.: A Comparison of Collaborative-Filtering Recommendation Algorithms for E-commerce. IEEE Intelligent Systems 22(5), pp.68—78 (2007)
7. Linden, G., Smith, B., York, J.: Amazon.com Recommendation Item-to-Item Collaborative Filtering. IEEE Internet Computing, pp.76—80 (2003)