

## Recommendation Systems Based on Association Rule Mining for a Target Object by Evolutionary Algorithms

Hossein Hatami Varzaneh <sup>a</sup>, Behzad Soleimani Neysiani <sup>b\*</sup>, Hassan Ziafat <sup>c</sup>, Nasim Soltani <sup>d</sup>

<sup>a</sup> Department of Computer Engineering, Faculty of Computer & Electrical Engineering, Kashan Branch, Islamic Azad University, Kashan, Isfahan, Iran

<sup>b</sup> Department of Computer Engineering, Faculty of Computer & Electrical Engineering, University of Kashan, Kashan, Isfahan, Iran

<sup>c</sup> Young Researcher and Elite Club, Natanz Branch, Islamic Azad University, Natanz, Isfahan, Iran

<sup>d</sup> Department of Software Engineering, Allame Naeini Higher Education Institute, Naein, Isfahan, Iran

### Abstract

Recommender systems are designed for offering products to the potential customers. Collaborative Filtering is known as a common way in Recommender systems which offers recommendations made by similar users in the case of entering time and previous transactions. Low accuracy of suggestions due to a database is one of the main concerns about collaborative filtering recommender systems. In this field, numerous researches have been done using associative rules for recommendation systems to improve accuracy but runtime of rule-based recommendation systems is high and cannot be used in the real world. So, many researchers suggest using evolutionary algorithms for finding relative best rules at runtime very fast. The present study investigated the works done for producing associative rules with higher speed and quality. In the first step Apriori-based algorithm will be introduced which is used for recommendation systems and then the Particle Swarm Optimization algorithm will be described and the issues of these 2 work will be discussed. Studying this research could help to know the issues in this research field and produce suggestions which have higher speed and quality.

### Keywords:

Recommender Systems;  
Collaborative Filtering;  
Association Rule Mining;  
Multi-Objective Evolutionary Algorithms;  
Particle Swarm Optimization;  
Genetic.

### Article History:

**Received:** 29 November 2017  
**Accepted:** 08 March 2018

## 1- Introduction

Online business success highly relies on the ability to present personal goods, services, and information items to the potential customers. This result in willingness toward recommender systems. Through statistical methods and knowledge discovery, these systems present services to the customers[1, 2]. Collaborating filtering system is one recommender system presents recommendation through detecting similar users based on enter date and previous transactions[3]. Collaborating filtering based recommender systems have many challenges such as recommend generation speed, database sparsity, scalability, recommends utility and so on. There have been great attempts for overcoming the collaborating filtering problems and these Attempts resulted in the high quality recommender generation. The present study reviews the previous studies in this area and examines the steps and resulted findings.

Recommender systems define the information systems able to analyze the previous behaviors and present the suggestions for the current issues. In other word, recommender systems try to guess the user`s thinking through his similar behavior or other similar users in order to get the best case most appropriately to the user`s taste[1, 2].

There are many types of recommender systems such as:

- Content-Based [4]: the working method of content filtering, based on item content analysis and trying to understand the discipline among them for generating the recommendation.

\* **CONTACT:** B.soleimani@grad.kashanu.ac.ir

**DOI:** <http://dx.doi.org/10.28991/esj-2018-01133>

© This is an open access article under the CC-BY license (<https://creativecommons.org/licenses/by/4.0/>).

- Knowledge-Based[5]: the recommender systems based on knowledge record knowledge about users, items, one specific item, and needs of the specific user.
- Memory-Based[6]: one database from users` known preferences is kept for all items and for each prediction, all computation is done all over this database.
- Collaborating Filtering (CF)[2, 7]: the collaboration filtering method uses the similar users` information to give a recommendation. Therefore, collaborating recommender systems collect the information regarding the users` preferences.

## 2- Challenges of Collaborating Filtering Based Recommender Systems

The collaborating filtering method faces with challenges such as dispersion in the database to improve the scalability, increasing recommendations utility, synonyms and so on[8]. Attempts regarding overcoming the problems of collaborating filtering are classified as the following:

- Traditional Method based on neighborhood[9]: CF recommender systems based on item chooses the nearest neighbor through computing similarity among items.
- The Clustering Method[10]: Through clustering items, the clustering algorithms reduce the search space and increase the scalability; through computing the similarities in members of one cluster the predictions would be generated.
- Case-based Reasoning[11]: Case-based reasoning is a decision-making method based on solutions to similar solutions . This method has a strong learning ability and can use the previous experience of encountering new problems. In the dynamic and variant situations when the conditions are unknown or unclear, this reasoning method is appropriate.
- Compound Algorithms[6]: Combination of CF with knowledge-based techniques and combination of GF with content based techniques.
- Association Rule Mining (ARM)[12, 13]: This method is used to detect the dependence on existing items in one transactional database so that presence of some items in transactions implies the presence of some other items in the same transactions.

## 3- Related Work of CF Recommender Systems Based on ARM

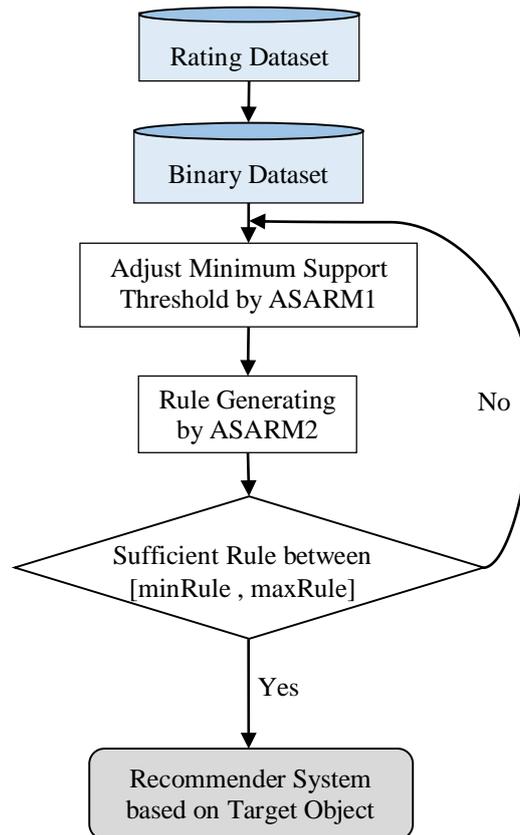
Generally speaking the association rule mining (ARM) could be considered as a two-step Process:

- Step One: finding all frequent items set
- Step Two: generating strong association rules for frequent item set found in step one

Because the computation in step one is quite expensive, commonly the algorithms focus on the optimization of the operation since the efficiency of the algorithm is measured considering the complexity of this step. Rules are as  $A \rightarrow B$  where  $A$  is rule`s tail (or consequent) and  $B$  is rule`s head (or antecedent). The criterion to call association rules strong is based on two parameters as support and confidence. The rules in which the values of support confidence are higher than minsupport and minconfidence are called strong rules. Apriori Algorithm is the first algorithms for finding frequent item set. According to Tyagi and Bharadwaj [13] generally speaking, support and confidence threshold values are effective on the quality of output rule. That is why Adaptive-Support Association Rule Mining (ASARM) was presented to effect rules` quality by the least support adjustment. But this not an appropriate algorithm because it performs the Apriori algorithms several times during the operation steps and output with low efficiency. Multi-Objective Particle Swarm Optimization Association Rule Mining (MOPSO-ARM) was introduced for finding quality rules and uses evolutionary algorithms. They emphasized on this point that this algorithm uses Particle Swarm Optimization (PSO) algorithm inherently appropriates for the continuous environment but the rules used in this evolutionary algorithm as particles are discrete[3, 14].

### 3-1- Adaptive-Support Association Rule Mining (ASARM)

Considering the fact that support and the confidence threshold values effect on the quality of Apriori algorithm output rules, the Adaptive-Support Association Rule Mining (ASARM) has been proposed [12, 15]. In this algorithm, rule generation has been done by a CBA-RG algorithm which is the evolutionary version of the Apriori algorithm. In 0 the steps of this algorithm are shown. This algorithm changes dataset rates to binary value based on average value of rates which results in a kind of like/dislike rate instead of showing how much like. Then it will change minimum supported threshold and generate rules by Apriori algorithm until enough rules have been made. Then will send the rules to recommendation system.



**Figure 1. ASARM Algorithm Performance.**

Shortcomings of ASARM Algorithm:

- Its efficiency decreases drastically because rules are generated several times.
- Indirect rules cannot be generated, because the system uses a database with transactions including object-target transaction. By indirect transaction, it means the non-existing rules in the database that could be made by indirect rules.

### 3-2- Evolutionary Algorithms

Today, one important research area is developing search methods based on natural evolution principles. In evolutionary computations, the basic concepts of natural evolution are inspired in an abstract way to look for finding an optimized solution for different problems. There are methods such as quick search, binary search, first deep, and first surface search that return the responses in a certain way. These algorithms, however, are not efficient when facing with a vast amount of data. Even more advanced algorithms such as iterative deepening are helpless at in finding a solution or desired area when facing with hyperspace. There is a line of the algorithm by which the problems are gained randomly and in an indefinite base. These types of problems are known as evolutionary algorithms trying to computerize the existing method in nature. The genetic algorithm is one evolutionary algorithm. The main idea of this algorithm is transferring inherent features by genes. Each gene in the chromosomes is the representative of one feature. The chromosomes are regenerated through crossover and mutation operation and create a new generation.

### 3-3- The Optimization of Multi-Objective Functions

Another concept in evolutionary algorithms is the optimization of functions. Optimization of different or sometimes opposite functions is done simultaneously. Parameters` inputs are different and multi-objectives and the output turn out single-objective.

1. There are two main principles for optimization of some objective functions.
  - Directing the search route toward getting the optimized responses.
  - Retaining and generating optimized response during response population.
2. Single-objective methods encounter some difficulties when facing with multi-objective functions:
  - Not finding several responses during one-time algorithm running.

- No guarantee for finding the differences and variant optimized responses.
  - For problems with discrete variables having several optimized responses the traditional methods could not be applied.
3. There are three solutions for multi-objective problems as the following:
- Putting weight for different objective functions and adding them up. In this method, functions gain one weight rate  $W$  considering the importance degree so that the sum of functions' weight equals one.

$$Optimize = w_1 f_1 + w_2 f_2 + w_3 f_3 \quad w_1 + w_2 + w_3 = 1 \tag{1}$$

- Using lexicographical method. In this method, the functions are ordered and ranked alphabetically. The functions with higher rank have more effects on the optimized responses.
- Applying Pareto Method. Selection based on Pareto method is One repetitive ranking procedure is performed to rank the population members and The ranking is based on the whole Population.

These types of algorithms are looking for different and optimized responses none of which is beaten/overcome by the other.

**3-4- Multi-Objective Particle Swarm Optimization Association Rule Mining (MOPSO-ARM)**

Another algorithm proposed based on ASARM is the multi-objective particle swarm optimization association rule mining (MOPSO-ARM) [13]. This algorithm tries to remove the shortcomings of ASARM. In Particle swarm optimization each rule is taken as one particle. Since all rules' head is fixed and the tails are different from each other, then only the tails are taken. This algorithm is applied for finding optimized particles floating in the space.

Particles are shown by support and confidence value in Figure 2 seven particle in Figure 2 are representative of rules

Which the horizontal axis parameter is supporting represented by  $f_1$  and the vertical axis parameter is confidence represented by  $f_2$ [16]. How to search for two locals guide for the  $P_3$  particle show in Figure 2. The search steps lbest1 using  $f_1$  as the first function and  $f_2$  as the second function by using the FindLocalBest algorithm As follows:

- First step: The set of all particles for a function  $f_1$  greater than or equal to is:

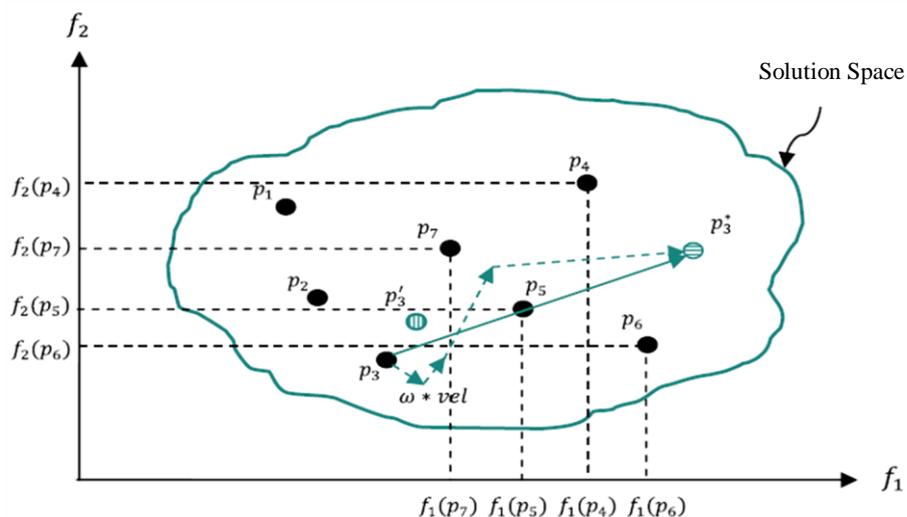
$$Max f_1 (P_3) = \{P_3, P_4, P_5, P_6, P_7\}$$

- Second step: The neighborhood of the particle  $P_3$  for  $M = 3$  is calculated as follows:

$$LocalNbd_M (P_3) = \{P_3, P_5, P_7\}$$

- Third step: The lbest1 particle  $P_3$  will be as follows:

$$lbest^1 (P_3) = \operatorname{argmax}_{p_j \in LocalNbd_M (P_3)} \{f_2 (p_3), f_2 (p_5), f_2 (p_7)\} = p_7$$



**Figure 2. Showing the Particles by Particles in a Two-Dimensional Space [13].**

The multiple-objective particle swarm optimization (MOPSO) finds quality rules. The output of this algorithm is two parameters as  $lbest^1$  and  $lbest^2$ . The speed direction and the best position in which the particle has been in are shown

by  $vel(p_i)$  and  $pbest$  respectively. The particles move in the space through the following relations and this repeats for 75 generations.

$$vel(p_i) = \omega \times vel(p_i) + c_1 r_1 (pbest(p_i) - p_i) + \quad (2)$$

$$c_2 r_2 (lbest^1(p_i) - p_i) + c_3 r_3 (lbest^2(p_i) - p_i)$$

$$p_i = p_i + vel(p_i) \quad (3)$$

To solve this multi-objective optimization problem weighting for multi-objective functions was applied [13] and the following relation is presented as association rule quality.

$$quality(A \rightarrow B) = W_1 \times support(A \rightarrow B) + W_2 \times confidence(A \rightarrow B) \quad (4)$$

Where  $W_1$  and  $W_2$  are allocated weights to the support and confidence parameters so that  $W_1 + W_2 = 1$  and  $W_1, W_2 \in [0,1]$ .  $W_1$  shows higher value and  $W_2$  shows lower value; highest quality of rule depends on the value of support rather than confidence and vice versa. These weights are selected experimentally. The sum of the quality of rules that are true for the target object is computed [13].

$$Score(target) = \sum quality(A \rightarrow B) \quad (5)$$

If computation is done to the target item, the target item is suggested to the user if the figure is higher than the defined threshold value. If the computation is done by the target user, that user is considered similar to the target user if the figure is higher than the defined threshold value and the similarity user's interests are recommended to the target user.

Once for data set with spare data and once for a data set with dense data, the algorithm is repeated and is compared to the 6 to 9 parameters.

$$precision = \frac{\text{The correct number of items proposed}}{\text{All item proposed}} \quad (6)$$

$$recall = \frac{\text{The correct number of items proposed}}{\text{All item user ratings}} \quad (7)$$

$$accuracy = \frac{\text{The number of items Categories correct}}{\text{All items categories}} \quad (8)$$

$$F1-measure = \frac{2 \times precision \times Recall}{precision + Recall} \quad (9)$$

Figure 3 show the steps of this algorithm. The shortcomings regarding this idea are as the following:

- Generally, the environment for this problem is discrete while MOPSO algorithm is for continuous environments naturally and it demands to discretize the operators, however, there are special algorithms for the discrete environment.
- The required operation for generating a new generation has moving nature and through repetitive and unnecessary mathematical operations of MOPSO algorithm, a lot of computations have been done resulting in more computation overloading.

### 3-5- Association Rules Mining Algorithm through Genetic Algorithm (GA-ARM)

Another algorithm similar to MOPSO-ARM has been proposed which is the association rules mining algorithm through genetic algorithm [17]. It has been stated that by using a genetic algorithm, the speed and exactness of the MOPSO-ARM were improved. The quality of generated rules is measured through Relation 5.

Shortcomings regarding the GA-ARM are:

- In the evolutionary genetic algorithm, several times referring to the algorithm to the database and repetitive computation bring about a slowdown in the recommender system [17, 18].
- Some objects or items may have few votes and this result in less support and confidence as well as be wasting of the time and decreasing the exactness of the recommender system. Moreover, more works on the ARM recommender systems have been done emphasizing to apply this system in other applications or showing other aspects of the systems and they are:

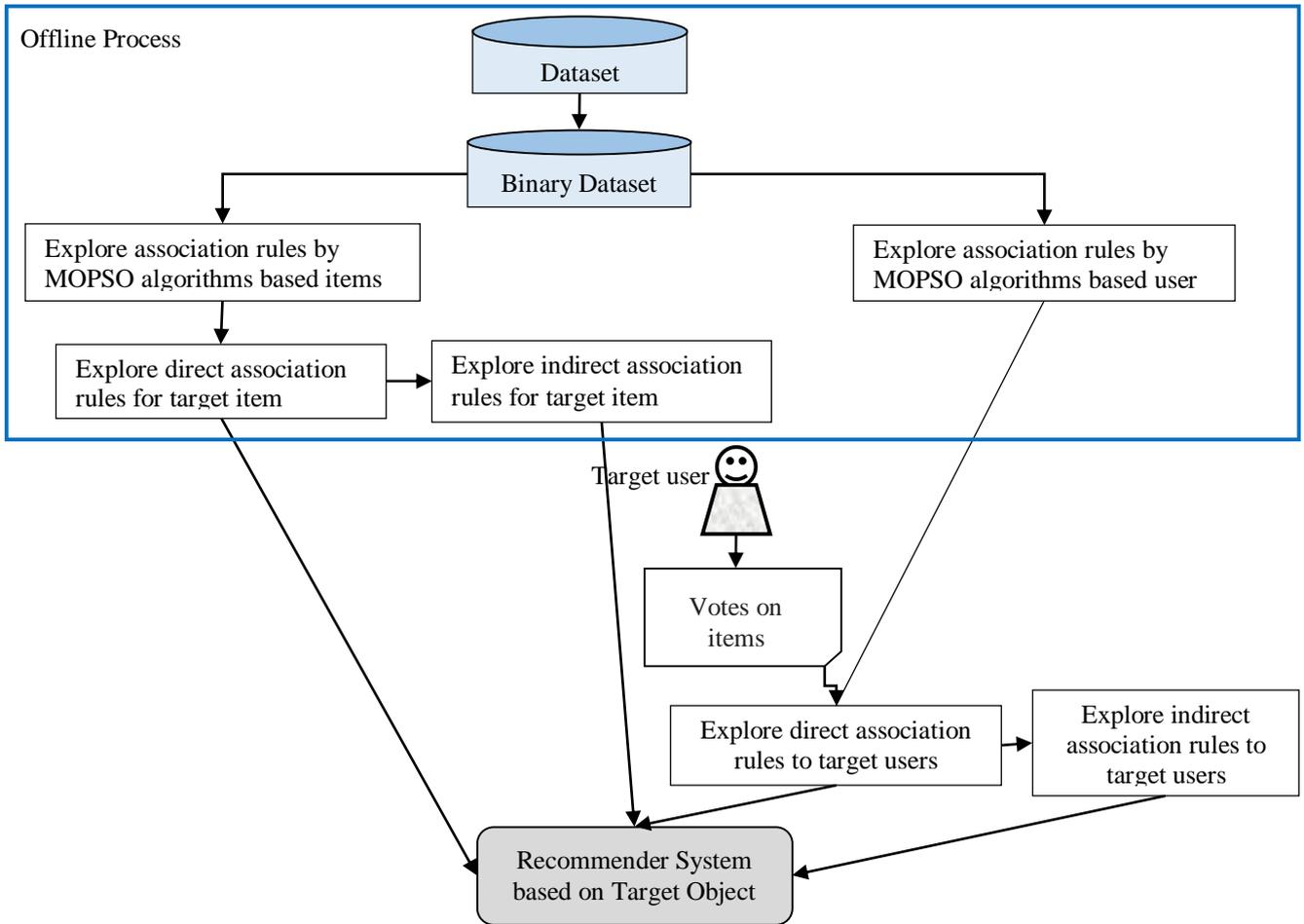


Figure 3. MOPSO Algorithm Performance [13].

- Particle swarm with collaborating filtering based on recommender system through fuzzy features (2015)[3]: through PSO algorithm and a fuzzy based system they concluded that in this method collaborating filtering is based on a more precise model memory but it has lower scalability.
- Development, position, and the procedure of the recommender systems: a comprehensive review of critical literature (2014) [19] : it is a review of recommender systems since 2001 to 2013 referring to PSO proving that there had not been a new study in this area and based on the reviews a study for improving the efficiency or the exactness has not been proposed yet.
- Personalized recommender systems based on evolutionary multi-objective optimization (2015)[20]: in this study, the task of a personalized recommender system is modeled as a multi-objective optimized problem and criteria such as exactness and variety are used for evaluation.
- Improving variety of recommendations by classification of users based on individual's information (2015)[21]: in this study, one classification system is introduced to improve the variety of recommendations, without losing recommender system's exactness.

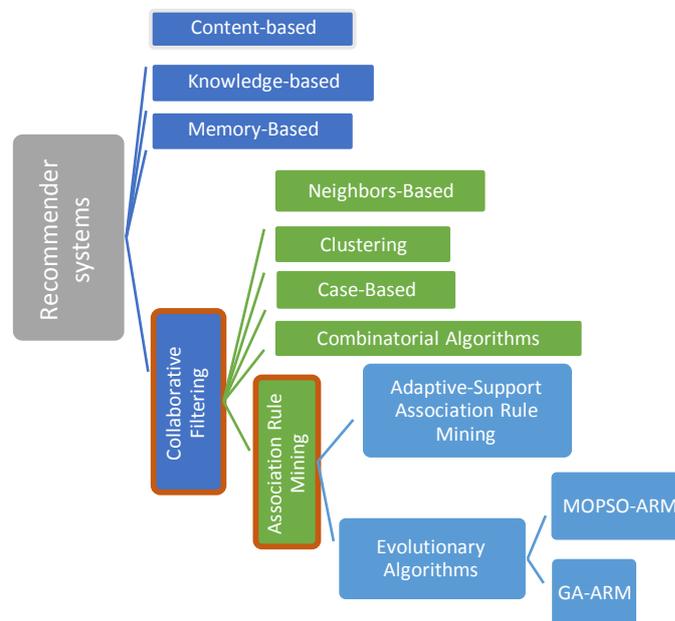
The details of the resources provided in this area are shown in Table 1.

Table 1. Review of the proposed methods.

Resource	Algorithm	Aim	Implemented by	Pros & Cons	Data set
ASARM	CBA-RG	Association Rules for recommender system based on observation	C++	Local Search, Slow	EachMovie
MOPSO-ARM	MOPSO	explore only high quality indirect and direct associations	NA	Local Search, Medium	MovieLens
GA-ARM	Genetic	recommender system based on Association Rules for a target object	Matlab	Local Search, Fast	MovieLens

## 4- Conclusion

The present study examined the recommender systems and among existing methods, the introduced algorithms were compared focusing on collaborating filtering based on association rules mining. ASARM has low efficiency because it produces the rules several times. Then the evolutionary algorithms and the concept of multi-objective optimization functions are put in this algorithm and MOPSO-ARM and ARM were analyzed by genetic algorithm. MOPSO is designed for continuous environments and the operators need to be discrete, the computation operations repeat redundantly in generating the generations and have extra overloaded and it could reduce the efficiency of the algorithm. Although making operators discrete have been overcome in the genetic algorithm and the efficiency has been improved, it still makes the recommender systems slow because the algorithm refers to the database and does repetitive computations. Based on the findings, the system is more efficient and faster if the repetitive computation and unnecessary referring to the database being removed and they are optimized. Moreover, by removing spare data the exactness and efficiency could be improved. Introduced recommender systems in the present study are based on association rules. Trough association rules among voted items, the recommender systems detect the similarities among the users and recommend generating. Another recommender system could be designed so that it detects similarities between the users through an anthology of users` voted items and generates more efficient recommendations. The criterion for choosing ARM and improving the recommender systems are appropriate for further studies. Figure 4 show the scope and related works in this research area. Clearly, there can be more methods to improve accuracy and performance of this systems because the results of prior works show that the accuracy of this system was less than 70% and runtime of them takes a long time yet.



**Figure 4. Scope of Evolutionary Association Rule Based Collaborative Recommendation Systems.**

## 12- References

- [1] Adomavicius, G., and 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, no. 6 (June 2005): 734–749. doi:10.1109/tkde.2005.99.
- [2] Resnick, Paul, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, and John Riedl. "GroupLens." *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work - CSCW '94* (1994). doi:10.1145/192844.192905.
- [3] Wasid, Mohammed, and Vibhor Kant. "A Particle Swarm Approach to Collaborative Filtering Based Recommender Systems through Fuzzy Features." *Procedia Computer Science* 54 (2015): 440–448. doi:10.1016/j.procs.2015.06.051.
- [4] Ansari, Asim, Skander Essegaier, and Rajeev Kohli. "Internet Recommendation Systems." *Journal of Marketing Research* 37, no. 3 (August 2000): 363–375. doi:10.1509/jmkr.37.3.363.18779.
- [5] Burke, Robin. *User Modeling and User-Adapted Interaction* 12, no. 4 (2002): 331–370. doi:10.1023/a:1021240730564.
- [6] Pennock, David M., Eric Horvitz, Steve Lawrence, and C. Lee Giles. "Collaborative filtering by personality diagnosis: A hybrid memory-and model-based approach." In *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence*, pp. 473-480. Morgan Kaufmann Publishers Inc., 2000.

- [7] Linden, G., B. Smith, and J. York. "Amazon.com Recommendations: Item-to-Item Collaborative Filtering." *IEEE Internet Computing* 7, no. 1 (January 2003): 76–80. doi:10.1109/mic.2003.1167344.
- [8] Wang, Qi, Jian Ma, Xiuwu Liao, and Wei Du. "A Context-Aware Researcher Recommendation System for University-Industry Collaboration on R&D Projects." *Decision Support Systems* 103 (November 2017): 46–57. doi:10.1016/j.dss.2017.09.001.
- [9] Ma, Hao, Irwin King, and Michael R. Lyu. "Effective Missing Data Prediction for Collaborative Filtering." *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '07* (2007). doi:10.1145/1277741.1277751.
- [10] Xue, Gui-Rong, Chenxi Lin, Qiang Yang, WenSi Xi, Hua-Jun Zeng, Yong Yu, and Zheng Chen. "Scalable Collaborative Filtering Using Cluster-Based Smoothing." *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '05* (2005). doi:10.1145/1076034.1076056.
- [11] Tyagi, Shweta, and Kamal K. Bharadwaj. "A Collaborative Filtering Framework Based on Fuzzy Case-Based Reasoning." *Proceedings of the International Conference on Soft Computing for Problem Solving (SocProS 2011) December 20-22, 2011* (2012): 279–288. doi:10.1007/978-81-322-0487-9\_27.
- [12] Lin, Weiyang, Sergio A. Alvarez, and Carolina Ruiz. *Data Mining and Knowledge Discovery* 6, no. 1 (2002): 83–105. doi:10.1023/a:1013284820704.
- [13] Tyagi, Shweta, and Kamal K. Bharadwaj. "Enhancing Collaborative Filtering Recommendations by Utilizing Multi-Objective Particle Swarm Optimization Embedded Association Rule Mining." *Swarm and Evolutionary Computation* 13 (December 2013): 1–12. doi:10.1016/j.swevo.2013.07.001.
- [14] Q. Wang, C. Zeng, W. Zhou, T. Li, L. Shwartz, and G. Y. Grabarnik, "Online Interactive Collaborative Filtering Using Multi-Armed Bandit with Dependent Arms," arXiv preprint arXiv:1708.03058, 2017.
- [15] Shridhar, M., and M. Parmar. "Survey on association rule mining and its approaches." (2017): 129-135.
- [16] Ting, Chuan-Kang, Rung-Tzuo Liaw, Ting-Chen Wang, and Tzung-Pei Hong. "Mining Fuzzy Association Rules Using a Memetic Algorithm Based on Structure Representation." *Memetic Computing* 10, no. 1 (January 5, 2017): 15–28. doi:10.1007/s12293-016-0220-3.
- [17] R. Mofidi, "Improve the performance of Rule based recommender systems by Genetic algorithms," Master, computer, University of Alame Naeeni, Naeen, Iran, 2015.
- [18] Sarkar, Sobhan, Ankit Lohani, and Jhareswar Maiti. "Genetic Algorithm-Based Association Rule Mining Approach Towards Rule Generation of Occupational Accidents." *Computational Intelligence, Communications, and Business Analytics* (2017): 517–530. doi:10.1007/978-981-10-6430-2\_40.
- [19] Xu, Jin, Johnson-Wahrmannand Karaleise, and Shuliang Li. "The development, status and trends of recommender systems: a comprehensive and critical literature review." (2014): 117-122.
- [20] Zuo, Yi, Maoguo Gong, Jiulin Zeng, Lijia Ma, and Licheng Jiao. "Personalized Recommendation Based on Evolutionary Multi-Objective Optimization [Research Frontier]." *IEEE Computational Intelligence Magazine* 10, no. 1 (February 2015): 52–62. doi:10.1109/mci.2014.2369894.
- [21] Z. E. Koupaei and M. Khayyambashi, "Improving Recommendation Diversity by Classifying Users Based on Demographic Information," *IJCCER*, vol. 3, pp. 68-74, 2015.