Recommender Systems Based on Evolutionary Computing: A Survey

Maryam Sadeghi, Seyyed Amir Asghari

Electrical and Computer Engineering Department, Kharazmi University, Tehran, Iran
Email: std_m.sadeghi@khu.ac.ir, asghari@khu.ac.ir

Abstract

Data mining techniques and information personalization have made significant growth in the past decade. Enormous volume of data is generated every day. Recommender systems can help users to find their specific information in the extensive volume of information. Several techniques have been presented for development of Recommender System (RS). One of these techniques is the Evolutionary Computing (EC), which can optimize and improve RS in the various applications. This study investigates the number of publications, focusing on some aspects such as the recommendation techniques, the evaluation methods and the datasets which are used.

Keywords

Evolutionary Computing, Genetic Algorithm, Recommender System

1. Introduction

Recommender systems have been generated and developed to help users for finding relevant information in the wide fields of information. Recommender system techniques have been proposed in various applications since the mid-1990s. The recommender systems try to recommend the most suitable items to the target users by investigating a user's interest in an item and the interactions between users and users or users and items. The RS have been developed for variety applications such as e-business, e-shopping, e-learning, e-tourism, e-library and etc. [1] [2]. Primary recommender systems were based on information retrieval. A few later recommender techniques were proposed in a way of filtering. Commonly used recommender techniques are divided into two groups: 1) Classical and traditional methods, and 2) Modern and hybrid approach. One of the several techniques having been investigated for the development of RS in the...
modern and hybrid approach is Evolutionary Computing (EC) techniques. Each recommendation approach has advantages and disadvantages. In this survey, first, the classical and traditional filtering ways are introduced and then the paper is focused on introducing and comparing the recent publications of EC approach in the recommender systems [2] [3] [4].

2. Basic Concepts

2.1. Item

Every object that can be recommended to a user is called item. If the item is useful for the user, it has positive value else its value is negative [2].

2.2. User

All operations in the recommender system offer to the target user. Target user has some public component such as age, income, marital status, education, profession and nationality, and also has some specific components such as favorite food, sports, music, movies, etc. Specific components extracting are very sensitive [1] [2].

2.3. Feedback

The most important concept in recommender system is user’s feedback on items. There are two types of explicit and implicit feedback [5] [6].

In explicit feedback, users explicitly rank items. Typically, users were asked to show their opinion with a number.

In implicit feedback, the recommendation is based on previous user’s interactions with the system [5].

3. Classical and Traditional Methods

Generally, recommender systems depend on how to do information filtering divide into three groups: 1) Collaborative Filtering, 2) Content-based Filtering and 3) Context-Aware.

3.1. Collaborative Filtering (CF)

Collaborative filtering recommends users by finding users with similar preferences. The CF filtering is classified into Memory based and Model based CF. Figure 1 shows the functional structure of the CF method [2].

Input should be in the form of rating matrix in the memory based CF. For example a music recommendation application system can be represented as user-music rating matrix. The set of users are represented as columns and items (music) as rows. The second type is model-based algorithms. This algorithm used rating matrix to create a model. Then by using this model do the recommendation. Memory-based algorithms in comparison with model-based algorithms have better and more accurate results [7].

There are different algorithms for CF filtering. Three common algorithms of this method are: 1) Random Algorithm, 2) Mean Algorithm and, 3) Neighbor-
CF has some limitations such as cold start, scarcity of user opinion and data validity.

3.2. Content-Based Filtering

Content is the base concept for building recommender applications in this method. Content-based filtering is based on item description rather than similarity of other users [2]. It uses machine learning algorithms to make a model of user preferences. Many approaches were proposed in the content-based recommendation, but if the content does not contain enough information recommendations do not guarantee accurate predictions [3]. This method has some limitations such as limited content analysis and cold start.

3.3. Context-Aware

There are various definitions for the concept of context. In this paper, each entity that can be used to personalize the information is called context. An entity can be an object or place or situation that is relevant with interaction of target user with application or a user [8] [9] [10]. The context information of a user depends on several cases, such as location, emotional state, personal characteristic etc. [11].

Different categories were chosen in the various papers. In this paper context is divided into two types: 1) Static 2) Dynamic [12] [13]. The context which changes rarely called static context such as contact list, user profile etc. The context which is extremely variable called dynamic context such as location of the user, temperature, time and etc. [8].

Context-aware recommender system approaches are classified into three approaches as shown in Figure 2: 1) Pre-filtering, 2) Post-filtering, and 3) Contextual modelling [8] [13].

In pre-filtering approaches as shown in Figure 2(a), the contextual information is used before all recommendations computing. The reduction-based approach is an example of pre-filtering approaches [14]. The ranked data which is not relevant to the context is filtered out before applying the recommendation.
algorithm. The important benefit of this approach is that it allowed using any recommender system technique. The pre-filtering approach used two-dimensional (users × items) recommendation to estimate the rating function [15] [16].

In post-filtering approaches, the result list of the recommendation is prepared and then filtered with contextual information as shown in Figure 2(b). The ratings are predicted using two-dimensional (2D) recommender system.

Contextual modeling approaches use explicitly the context information as a predictor for user’s rating an item. Hence, the approach formulates a multidimensional or MD (user × items × context) recommendation as shown in Figure 2(c) [17] [18] [19] [20].

4. Evolutionary Computing

As shown in Figure 3 Evolutionary Computing (EC) with Fuzzy sets, Artificial Neural Networks (ANNs), Swarm Intelligence (SI) and Artificial Immune Systems are subsets of computational intelligence [9].

The root of EC is back from Darwin theory of natural selection. Darwin’s theory explains that nature has limited resources [21] [22] [23]. Creatures that live are competing together because of limited resources. And try to extend their next generation. This generation is called survivor of fittest [24] [25] [26].

Table 1 shows the mapping between this theory and problem solving in the evolutionary computing [23].

In general, every problem has three parts: Input, Model and Output.
EC can solve the problem in three ways.

4.1. Optimization

As shown in Figure 4(a), in these methods input is unknown. Input should found in a way that it became optimize. Most of the problems are in this category.
Figure 3. Type of computational intelligence [9].

Figure 4. EC solving problem methods [22].

Table 1. Mapping between Darwin’s theory and evolutionary computing.

<table>
<thead>
<tr>
<th>Evolution</th>
<th>Problem Solving</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environment</td>
<td>Problem</td>
</tr>
<tr>
<td>Individual</td>
<td>Candidate Solution</td>
</tr>
<tr>
<td>Fitness</td>
<td>Quality</td>
</tr>
</tbody>
</table>

4.2. Modeling

As shown in Figure 4(b), in these methods model should found because Input and output relation is unknown. Input and output are known in these methods.

4.3. Simulation

As shown in Figure 4(c), the output is unknown. Inputs with regularity of the
model can simulate outputs [21] [22] [23].

EC were designed to solve complex problems. There are different types of evolutionary computing algorithms, as shown in Figure 3 Genetic Algorithm (GA), Genetic Programming (GP), Evolutionary Programming (EP) and Evolution Strategies (ES). EC is based on computational models such as natural selection, survival of the fittest and reproduction. Different EC algorithms have similar algorithmic characteristics and also they have many similarities in framework implementation [21].

Figure 5 shows a universal framework which includes three fundamental operations and two optional operations for most ECs [24] [25]. The first step in an EC algorithm is the “population initialization” step. Next step is entering evolutionary iterations with two operational steps, namely, “fitness evaluation and selection” and “population reproduction and variation” [26] [27]. The iteration continues until a termination, threshold is achieved. Sometimes EC algorithms need additionally perform such as an “algorithm adaptation” procedure or a “local search” (LS) procedure besides the above three necessary steps [27] [28] [29].

5. Categories Approaches

EC techniques can help recommender systems to have more powerful and more
There are three main approaches to applying EC in recommender systems which are 1) approaches in which EC are used to optimize weights of recommendation techniques or different component, 2) approaches utilizing EC for clustering of items or users and 3) hybrid and other approaches [21].

5.1. Approaches Based on Weights Optimizing

According to [30] [31] [32] [33], feedback feature arrays including user characteristics, item characteristics and the corresponding ratings were created user profile. An n-dimensional array as user preferences was considered and it included weights corresponding to the given features and it supported the profile matching algorithm within the RS. In these papers, techniques based on evolutionary computing had better performance than Neighborhood-based collaborative filtering (nCo) except [31] where the proposed method is not compared to any baseline. The public dataset of Movie Lens is used to test these papers.

In the [34] authors proposed a recommender system for location-based advertising based on user’s preferences and interaction of context with a user. It used the genetic algorithm to achieve optimal value, in calculating the similarity between two contexts from the multidimensional perspective. As shown in Figure 6 the proposed architecture which named Context-Aware CF using Genetic Algorithm (CACF-GA), consists of pre-arranging process and a cycle of three steps for 1) Filtering, 2) Context-Aware CF, and 3) Recommendation and Feedback. The dataset which used for experiments was gathered from students via a...
web-based system developed for this purpose. This data collection contains information of 150 places in five major commercial areas in Seoul, South Korea for shopping, eating, drinking, enjoying, and learning. The proposed model outperformed a standard Neighborhood-based collaborative filtering (nCo) and some simple context-aware recommendation techniques which provide the most accurate prediction results compared to comparative ones.

In [35], authors, first identified suitable attributes in a user profile and suggest an appropriate similarity. Second, EC are used to learn the weights of user's personal, behavioral and interaction attributes from social networks based on the comparison of individual features to increase recommendation effectiveness. At the end, it used trust propagation techniques towards handling sparsity problem of collaborative filtering. The proposed approach is tested only on synthetic data (simulated dataset of 20 users) and is not compared to any baseline algorithm. In its extended version [36], authors, compared the presented approach with a hybrid friend recommendation strategy based on interest similarity and social proximity [37]. It eventuated that approach [35] is superior on synthetic data than other approaches [37].

The [38] proposed an evolutionary approach for combining results of recommendation techniques in order to automate the choice of techniques and get fewer errors in recommendations. Experiments on Movie Lens data showed that the appropriate combination of the results of different recommendation techniques performed better than any one of collaborative filtering technique separately in the context addressed [39]. Depending on the technique and the experiment executed the improvement varies from 9.02% to 48.21%.

The [40] presented similarity function using the average rating for each user. Other approach used overall average rating for all users. This approach has compared to standard Neighborhood-based collaborative filtering (nCo) techniques on Movie Lens data. A series of comparison demonstrate the effectiveness of the proposed approach in terms of the quality of prediction.

The [41] has presented a metric to measure the similarity between users, which is applicable in collaborative filtering processes in recommender systems. Genetic algorithms are used to find an optimal similarity function between two sparse rating vectors relevant to users with elements showing their feedbacks on items. Experiments performed on unpublic and public real-world data from the movie domain, such as Movie Lens, Film Affinity and Netflix. This approach showed that The GA-metric runs 42% as fast as the correlation. This is an important advantage in the recommender system, especially when many different recommendation requests are made simultaneously (user to user).

The [42] used evolutionary computing to find the optimal aggregation of different recommendation algorithms by considering multi-objectives such as accuracy, diversity and novelty. In order to evaluate proposed hybrid, authors used Movie Lens dataset for movie recommendation and another dataset which consists of 19,150,868 user’s access to music tracks on the website Last.fm, for mu-

https://www.last.fm/.
5.2. Approaches Based on Clustering

In the [43], authors used clustering methods for segmentation of users into clusters of customers with similar needs. Utilizing from evolutionary computing to find initial seeds for k-means clustering. In its extended version [44], authors used genetic algorithms to find optimal cluster centers. In [43], experiments applied on public GVSu WWW dataset and in [44] dataset gathered from an Internet shopping mall. The proposed approach in compared with traditional clustering methods had better segmentation.

In [45] [46], authors proposed a memetic recommender system. This recommender system is based on the collaborative behavior of memes. It used the memetic algorithm [47] to propose a strategy to perform the local search for clustering users. This system compared with the traditional collaborative filtering technique, genetic algorithm and the Pearson algorithm based recommender system and leads to better accuracy. Experiments performed on Movie Lens dataset.

In the [48], a new recommendation method was proposed based on memetic algorithm. For testing this method used four highly sparse real-world datasets. These dataset are standard public [KDD-Cup 2000 (KDD), Restaurant and Consumer (RCM), and Entree Chicago Restaurant (ECR)] and the other is a private dataset. This approach evaluated recommendation performance and compared with frequency-based, user-based, item-based, k-means clustering-based, and genetic algorithm-based methods in terms of precision and recall. Authors also showed that proposed approach outperforms nCo approaches.

In the [49], an approach proposed which used evolutionary strategies to clustered users. In this approach, clusters are allowed to be overlapping which implicates that the average sparsity remains constant and independent of the number of created cluster. Experiments applied on the Jester dataset. The proposed approach compared to nCo utilizing k-means clustering and showed effectiveness.

5.3. Hybrid and Other Approaches

Proposed approach [50] optimizes the weights according to historical rating using GA. The historical rating data can be used for discovering some of the valuable attributes of user and item that are called latent attributes. The recommendation is generated by the similarity of the user in each cluster. In order to check the performance of the proposed algorithm, authors used Each Movie dataset. Experimental results in this paper showed that the proposed approach improves the quality of recommendations. In its extended version [51] the proposed approach consists of two modules: explicit attribute-based recommender and implicit attribute-based recommender. Chromosomes in the genetic algorithm are

\(^2\text{http://www.mymedialite.net/}\)
used to find weights of implicit or latent attributes of materials for learner. In this approach the interests of learner based on explicit attributes of learning materials can be modeled in a multidimensional information model by preference matrix. Metadata for Architectural Contents in Europe (MACE) dataset is used for experiment. The proposed model outperforms several content-based filtering and nCo techniques and probabilistic latent semantic analysis model (PLSA) [52] on accuracy measures and can reduce some problems such as cold-start and sparsity.

In the [53], a hybrid genetic algorithm and graph theory-based online recommendation algorithm has been proposed. This approach used Singular Value Decomposition (SVD) [54] matrix factorization and stochastic gradient descent [55] to model user, item feature vector and improve parameter and accuracy. Experiments performed on KDD Cup 2012 Track 1 dataset and showed that the proposed model in compare with various nCo and Model-based collaborative filtering (mCo) baselines is able to recommend online friends with superior accuracy.

A framework to generate ranking functions for items recommendation based on genetic programming proposed in [56], which called GUARD (Genetic Unified Approach for Recommendation). As shown in Figure 7, this approach included six main phases: 1) Feature extraction, 2) Definition of genetic programming functions and terminals, 3) Terminals caching, 4) Evolutionary Process, 5) Choosing of the best solution and, 6) Recommendation. The proposed framework was tested in two movie datasets namely Movielens 100 k and Movielens 1M. The framework compared to a memory-based method (based on the cosine distance) [57] and a model-based (PureSVD) method.

The proposed model [58] is a multi-objective EC-based recommendation

![Figure 7. The architecture of GUARD [57].](image-url)
model which maximizes two performance metrics termed as accuracy and novelty. Proposed method compared with nCo, mCo and hybrid CF techniques on the Movie Lens data and demonstrate that the proposed algorithm can make more diverse yet accurate recommendations the probabilistic.

The paper [59] proposed a framework and used the multi-objective evolutionary algorithm based on decomposition to improve the novelty and the diversity for top-k recommendations of an item-based collaborative filtering technique without reducing the accuracy. The datasets which are used to test the framework were Movie Lens and Jester. It is compared with nCo approaches in diversity and novelty of the recommendation list. Proposed approach outperforms nCo approaches. This approach is further refined by using a multi-objective immune algorithm [60].

6. Conclusions
The rapid growth of information on the web and volume of information and also increasing the number of users demonstrate the requirement of the recommender systems. This review first, introduced the three main classical and traditional filtering methods: 1) Collaborative Filtering, 2) Content-based Filtering and 3) Context-Aware. Second, the evolutionary computing and the ways which they can solve the problems were explained. Third, it introduced and evaluated the number of recent research papers of evolutionary computing approach in the recommender system. The approaches were analyzed from four viewpoints: a) recommendation techniques, b) the evaluation methods, c) datasets which are used in the approaches and, d) the methods which compare with proposed approaches. Depending on considering viewpoints, approaches which utilize evolutionary computing in recommendation systems were divided into three main groups: 1) approaches in which evolutionary computing is used to optimize weights of recommendation techniques or different component, 2) approaches utilizing evolutionary computing for clustering of items or users and 3) hybrid approaches.

As for future research in EC-based recommendation, the authors suggest using the EC-based recommendation (especially genetic algorithm) for the complex recommendation scenarios, to optimize and improve some problems of these recommendations.

References


on Computer and Information Technology (WCCIT), Sousse, 22-24 June 2013, 1-6.


M. Sadeghi, S. A. Asghari


Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.
A wide selection of journals (inclusive of 9 subjects, more than 200 journals)
Providing 24-hour high-quality service
User-friendly online submission system
Fair and swift peer-review system
Efficient typesetting and proofreading procedure
Display of the result of downloads and visits, as well as the number of cited articles
Maximum dissemination of your research work

Submit your manuscript at: http://papersubmission.scirp.org/
Or contact jsea@scirp.org