Egyptian Informatics Journal xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Egyptian Informatics Journal

journal homepage: www.sciencedirect.com

Full length article

An effective collaborative movie recommender system with cuckoo search

Rahul Katarya *, Om Prakash Verma [1]

Department of Computer Science & Engineering, Delhi Technological University, Delhi, India

ARTICLE INFO

Article history:
Received 21 November 2015
Revised 8 October 2016
Accepted 18 October 2016
Available online xxxx

Keywords:
Recommender system
Collaborative filtering
k-mean
Cuckoo search optimization
Movie

ABSTRACT

Recommender systems are information filtering tools that aspire to predict the rating for users and items, predominantly from big data to recommend their likes. Movie recommendation systems provide a mechanism to assist users in classifying users with similar interests. This makes recommender systems essentially a central part of websites and e-commerce applications. This article focuses on the movie recommendation systems whose primary objective is to suggest a recommender system through data clustering and computational intelligence. In this research article, a novel recommender system has been discussed which makes use of k-means clustering by adopting cuckoo search optimization algorithm applied on the Movielens dataset. Our approach has been explained systematically, and the subsequent results have been discussed. It is also compared with existing approaches, and the results have been analyzed and interpreted. Evaluation metrics such as mean absolute error (MAE), standard deviation (SD), root mean square error (RMSE) and t-value for the movie recommender system delivers better results as our approach offers lesser value of the mean absolute error, standard deviation, and root mean square error. The experiment results obtained on Movielens dataset stipulate that the proposed approach may provide high performance regarding reliability, efficiency and delivers accurate personalized movie recommendations when compared with existing methods. Our proposed system (K-mean Cuckoo) has 0.68 MAE, which is superior to existing work (0.78 MAE) [1] and also has improvement of our previous work (0.75 MAE) [2].

1. Introduction

A recommendation system has become an indispensable component in various e-commerce applications. Recommender systems collect information about the user's preferences of different items (e.g. movies, shopping, tourism, TV, taxi) by two ways, either implicitly or explicitly [3–7]. An implicit acquisition of user information typically involves observing the user's behavior such as watched movies, purchased products, downloaded applications. On the other hand, a direct procurement of information typically involves collecting the user's previous ratings or history. Collaborative filtering (CF) is the way of filtering or calculating items through the sentiments of other people [8–10]. It first gathers the movie ratings given by individuals and then recommends movies to the target user based on like-minded people with similar tastes and interests in the past. Additional impression on which some recommender systems are based is clustering. Clustering is a popular unsupervised data mining tool that is used for partitioning a given dataset into homogeneous groups based on some similarity or dissimilarity metric [11–14]. Collaborative filtering and clustering have been discussed in detail in the next section. Hybrid cluster and optimization approach is applied to improve movie prediction accuracy. Such a hybrid approach has been used to overcome the limitations of typical content-based and collaborative recommender systems. For clustering, k-means algorithm is applied and for optimization, cuckoo search optimization is implemented. K-means algorithm is an enormously greater clustering algorithm when compared to other clustering methods in relations of time, complexity or effectiveness for a particular number of

Peer review under responsibility of Faculty of Computers and Information, Cairo University.

* Corresponding author at: Department of Computer Science & Engineering, Delhi Technological University (Formerly Delhi College of Engineering), Shabad Daulatpur, Main, Main Bawana Road, Delhi 110042, India.
E-mail address: rahulkatarya@dtu.ac.in (R. Katarya).
URLs: http://www.dtu.ac.in/Web/Departments/CSE/faculty/ (R. Katarya), http://www.dtu.ac.in/Web/Departments/CSE/faculty/ (O.P. Verma).
[1] Department of Computer Science & Engineering, Delhi Technological University (Formerly Delhi College of Engineering), Shabad Daulatpur, Main, Main Bawana Road, Delhi 110042, India.

http://dx.doi.org/10.1016/j.eij.2016.10.002
1110-8665/

clusters [12,15]. Clustering algorithm with a bio-inspired algorithm such as cuckoo search [16–21] delivers optimize results. The cuckoo search has shown best performance when compared with other algorithms such as genetic algorithms and particle swarm optimization. Simulations and comparison of the cuckoo search were greater to these existing algorithms for multimodal objective functions. To find the best results we have to find the most suitable weight among all possible ones. Cuckoo search was also performed well and showed good results that found the appropriate weights [22]. That is why cuckoo optimization algorithm is also used to obtain optimized weight in our work. Besides being one of the most efficient algorithms, it was found that it takes less time than other algorithms applied to the same dataset. The approach of k-means and cuckoo has been applied to the dataset, and the results have been observed regarding evaluation metrics such as mean absolute error (MAE), standard deviation (SD), root mean square error (RMSE) and t-value. These parameters examined and discussed to evaluate the performance of movie recommendation system. Regarding accuracy and precision, the experiment results reflect that the proposed approach is capable of providing more reliable movie recommendations as compared to the existing cluster-based CF methods. In numerous research, the clustering approaches are conducted with the entire dimensions of data which might lead to somewhat inaccuracy and results in more computation time. In general, designing expert movie recommendations is still a challenge, and discovering effective clustering method is a critical problem in this condition. To address aforementioned, a hybrid model- based movie recommendation approach is proposed to alleviate the issues of both extraordinary dimensionality and data sparsity. That is the reason we selected a cuckoo algorithm with k-means for optimization. On the comparison with some other optimization algorithms, the cuckoo was found to perform better than others. The major contributions of this research work are:

- We proposed a novel recommender system with K-means & cuckoo search optimization.
- Our system is innovative and efficient so far, as it employed Cuckoo search algorithm for excellent recommendations for Movielens Dataset.
- Our hybrid model has 0.68 MAE, which is superior to existing work (0.78 MAE).
- Our model also has excellent improvement of our previous work (0.75 MAE).
- The performance with respect to time is also better as compared to already existing systems.
- We used well known Movielens dataset (http://grouplens.org/datasets/movielens/100k/) to analyze the behavior of our proposed system.

The remainder of this article is planned as follows: Section 2 gives a brief explanation of the related work that was carried out on collaborative recommendation systems and clustering-based collaborative recommendation. The proposed approach called as a k-mean-cuckoo approach for movie recommender system is explained in Section 3. In Section 4, experiment results performed on Movielens dataset are described, and finally summarization of this article with future work are highlighted in Section 5.

2. Background and related work

Recommender systems are based on a variety of approaches such as content based [23,24], collaborative approach [9,25–27], hybrid [28,29]. Furthermost movie recommendation systems are centered on collaborative filtering and clustering. In movie recommender systems the user is asked to rate the movies which user has already seen then these ratings are applied to recommend other movies to the user that user has not perceived by utilizing collaborative filtering that is based on similar ratings. Collaborative filtering [9,10,30–32] is tremendously spreading in such a way that this approach influences most of the recommender systems. Collaborative filtering majorly classified into two principal classes such as memory-based collaborative filtering and model based collaborative filtering. Memory-based collaborative filtering [4,5,33] explores for nearest neighbors in the user space for an active user and dynamically recommend the movies. The shortcomings related to this method are computation complexity and data sparsity. Many authors [34] tried to reduce this computational complexity and memory bottleneck issues such as in item based collaborative filtering technique, in which relations between items were computed for neighborhood region around a target object. They showed in their empirical studies that item-based method could decrease the time of computation as well as deliver rationally correct prediction and accurateness. Model-based collaborative technique [3,4,8,33,35] produces a prebuilt model to collect rating patterns based on the database of users and ratings that can treat the issues of data sparsity and scalability. Model-based collaborative filtering is time-consuming and its offline in nature. Clustering based techniques are broadly used in movie recommendation systems to reduce the problem of scalability. Various researchers applied clustering-based methods on recommender systems that delivered expert recommendations [36–41]. The purpose of clustering is to partition objects into groups known as clusters in such a way that two objects within the same cluster have a minimum distance between them to identify similar objects then clustering process is performed offline to build the model. When a target user arrived, the online module allocates a cluster with a substantial similarity weight to the user, and the prediction rating of a specified item is computed based on the same cluster members instead of searching whole user space. The k-nearest neighbor (kNN) algorithm is the orientation algorithm in collaborative filtering recommendation process which is applied in recommendation process [42–45]. kNN based recommender systems for collaborative filtering recommendation process are reliable and with precise recommendations. A bio-inspired algorithm such as cuckoo search has exclusive background sensing abilities and employ a special method to facilitate the evolution of continuing resolutions into novel and quality recommendations by generating clusters with reduced time as discussed in next section.

3. K-means-cuckoo based collaborative filtering framework

To overcome the limitations of a collaborative recommender system, we propose a hybrid cluster and optimization based technique to improve movie prediction accuracy. Our motive is to design a unified model solution that incorporates user ratings from the Movielens dataset for predictions. We use K-means as clustering algorithm and cuckoo search as optimization algorithm and then apply to Movielens dataset for improved efficient recommender systems Initially k-means clustering algorithm is applied to Movielens dataset for clustering of users into different clusters. The clusters are selected randomly at first then users are inspected one by one by calculating the differences in their ratings and the centroid of the clusters, and if their difference is smallest, then the user gets allocated to the cluster to which they are closest. However, at this moment not assure that each user has been assigned to the real cluster with a minimum difference of centroid. So each user's distance is compared to its cluster mean and with other clusters mean and relocate the users according to the smallest distance from any cluster's mean. Now this iterative relocation

Please cite this article in press as: Katarya R, Verma OP. An effective collaborative movie recommender system with cuckoo search. Egyptian Informatics J (2016), http://dx.doi.org/10.1016/j.eij.2016.10.002