



# Cloud Service Prediction Using KCFC Approach

K. Indira<sup>1</sup>, C. Santhiya<sup>1</sup>(✉), and K. Swetha<sup>2</sup>

<sup>1</sup> Department of Information Technology,  
Thiagarajar College of Engineering, Madurai, India  
csit@tce.edu

<sup>2</sup> Thiagarajar College of Engineering, Madurai, India

**Abstract.** Cloud computing is an emerging paradigm where the user can benefit many efficient services. Since many services resemble the same functionality, the user faces relevant and un-relevant information as data burden. So, Recommender System (RS) is getting used to suggest the user only the information that suits their search. Here (CFC)-Collaborative Filtering Coefficient is used as RS which functions by analyzing user history and similar service from neighbor users. Pearson coefficient is used to calculate the association between the services. But, it works for existing users not for new users because the further user details are not sufficient to recommend a service. To overcome this, the KNN approach is utilized to classify the recommendation from a k-nearest neighbor by finding the resemblance between various client ratings using Euclidean Distance measure. Thus, the KNN-CFC hybrid novel approach can create a new efficient RS framework which supplies the client a most relevant service information with low execution time for various data densities and different users and services.

**Keywords:** Quality of service · k-Nearest Neighbour · CFC-Collaborative Filtering Coefficient · Pearson coefficient · Recommendation system

## 1 Introduction

Cloud computing is an amounting mechanism that provides services based on user demand (computing power, storage resources, applications, etc.). It has become famous now, since it allows for plenty of nature with the purpose of improves internet people on the way to struggle it, for example reliability, security, agility, performance, etc. [1]. It provides three various services, Infrastructure as a service, Platform as a service and Software as a service via the internet with pay-as-you-go pricing. Many cloud services supply similar services which are the challenge designed for the users to decide a relevant repair so as to fits precisely with their needs. Due to the existence of several websites that provide an enormous quantity of information on the same items, the clients are overcharged with irrelevant and relevant service info. To find the suitable web services based on user request, the web service suggestions are used. Recommender systems take users' interest as input and apply certain methods to sieve the

available information and deliver the user with the enormous lump of information called recommendations. But, whenever a new user drives through the website, the recommender system is not able to predict a user's interest due to the lack of the user's details known as cold start problem [2]. Regarding the recommendation system, cold start issue with new user regarding no rating of new products new; those services will not have any ratings, therefore will not be able to suggest to other users. Since the majority work in recommendation system is based on nearest neighbors' information, this work helps the cloud providers to promote their services and cloud users in the direction of recognize services so as to assemble their QoS requirements (e.g., response time, security, availability, reliability, throughput, etc.). High QoS with services satisfies the client then some additional features are to be used such as user's ratings, service completing, revenue model and business methods, etc. [3]. CF approach with QoS based RS is used to suggest the relevant services by using Pearson coefficient. The approach is used to forecast in cooperation QoS category of cloud services. Reliable ranking can be attained through this approach using the Pearson coefficient. The Collaborative Filtering Algorithm (CF) is a known technique in the recommender systems that deal with the human making decisions through their previous search. Adding to the experience, decisions may be taken based on some knowledge they may come across from a group of associates. This set of knowledge can be taken for suggestions. CF-based RS allow users to give rating about a set of elements (hotels, theatres) in such a way that where every information is stored in the database which can be used for recommendation. The suggesting quality comes from the capacity to find the common number of users made that votes or ratings [4]. The CF faces the troubles of missing of data and enlargement of data. Data sparsity refers to the situation when available data (e.g., ratings) is not such client to conclude the region of the energetic user so, nearly all and sundry of the services resolve not encompass precise ratings or one ratings at all [5]. Many available services will not be used for this reason. To overcome this KNN-CFC (K Nearest Neighbour – Collaborative Filtering Coefficient) approach is used to attain more reliable rankings than a CF approach and also attain exactness and scalability. In kNN, resemblance metric such as Euclidean distance is used to study the relation of a query to the neighboring query. The association of training dataset with 1 – D distance space based on found similarities and labeling them [6].

## 2 Related Works

Mezni et al. [7] stated that the client could face large similar operating services due to the presence of many websites. So, the recommendation system was provided to solve this issue. Then the new user may suffer from low recommendations due to few votes and service ratings. This problem was called the cold start problem. To overcome this recommendation system was combined with unclear recognized notion examination. After Implemented the unclear recognized notion Analysis-based service proposal come up to the concert and the excellence of fashioned recommendations in judgment through on hand high-tech solutions.

Hayyolalam et al. [9] provided a numerical examination of the previous approach for QoS-aware cloud service composition/selection. QoS monitoring can provide the QoS of a cloud service (Client workstation monitoring, server-side monitoring, and third-party monitoring). The quality assessment checklist had been famed, and the quality of service was monitored. The selection for the recommendation was made based on the assessment report.

Ertuğrul et al. [10] proposed the KNN approach which was the machine learning technique finds the k-Nearest Neighbour by using the resemblance metric such as Euclidean distance. Based on the gradient between the query and each sample weight was given to the query similarity for differentiation. Mapping process kNN was carried out to predict the likeliness, with a question being positioned happening a location procession and supplementary samples were mapped to a 1-D distance line based on the expanse to the uncertainty.

Ortega et al. [11] proposed to beat the data burden problem recommendation System was used. It was better to use the recommendation framework that would function as to design and implement recommendation method, increasing the execution time. Here the proposed work was CF4 Opensource library intended en route for holdout CF based RS, which allows reading RS Dataset, jam-packed and easy entrée to figures.

Kumar et al. [12] made the comparison of many various machine learning algorithms to obtain the successful extraction of internet answer instance plus throughput events. Baking was the successful method for extracting the absent answer time and sequential minimum optimization regression be the achievement technique designed for extracting throughput values. The absent principles of the real world datasets was extracted using this two models for recommendation instances.

Hadeel et al. (2014) proposed the significance of digital knowledge, social network in addition to information withdrawal to individualize the social information withdrawal application inside stipulations of the instructor's and student's requirements. information withdrawal techniques such as classification, clustering of the post, comments given in the Facebook like social networks were discovered to improve the student's usage of e-learning.

### 3 Methodology

#### 3.1 Quality of Service

QoS defines the range of execution, consistency, and availability obtainable by an submission and by the proposal [14, 15]. QoS is extremely significant designed for users, who anticipate providers to distribute the superiority and meant for providers, encompass to switch the trade-offs among QoS levels and equipped expenses. Through this approach, the service promotion and identification of quality services are done. High QoS with services satisfies the client then some additional features are to be used in business model and completion time. The work deals with the QoS-based cloud service recommendation, and propose a CF approach using the Spearman coefficient to suggest cloud services. The scheme is new to estimate in cooperation place and QoS

rating for cloud services. QoS monitoring involves Client-side monitoring, server-side monitoring, and intermediary monitoring. The a range of excellence metrics like business plan, failure time and failure recovery etc. suggestions [16, 17, 18]. CF-based RS helps the client to give a rating for a set of services (hotels, theatres)in such a way that where every information is stored in the database which can be used for recommendation. The quality of advice comes from the capacity to find the typical number of users made that votes or ratings.

### 3.2 KNN-CFC (K Nearest Neighbor - Collaborative Filtering Coefficient)

Reliability, accuracy and scalability are attained through KNN-CFC approach. In kNN, the relative of a uncertainty to the neighboring query is learned from a parallel metric, Euclidean distance [20, 21]. The association of training dataset with 1 – D distance space based (Fig. 1).

$$d(x,y) = (i = 1 \text{ to } n) \tag{1}$$

Therefore in Eq. (2)  $(x, y)$  is the distance between the samples  $x$  and  $y$ , having  $n$  dimension. The decisions can be taken from the class of a sample based on the extracted similarity. The difference among the sample. KNN is largely used in the classification problems though it is effective for both classification and regression problems. Consider a spread of blue circle(bc) and blue square(bs) have to find the class of red star(rs). Here  $k$  is the nearest neighbor we wish to take a vote from. If  $k = 3$  then by placing the rs as a center it will group the bc and it comes under the class of bc Fig. 2.a, b.

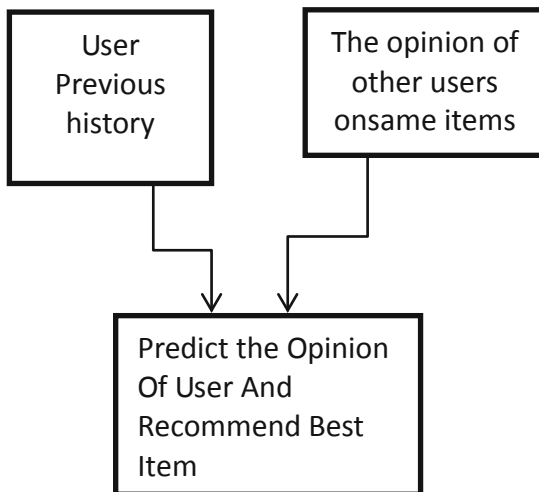
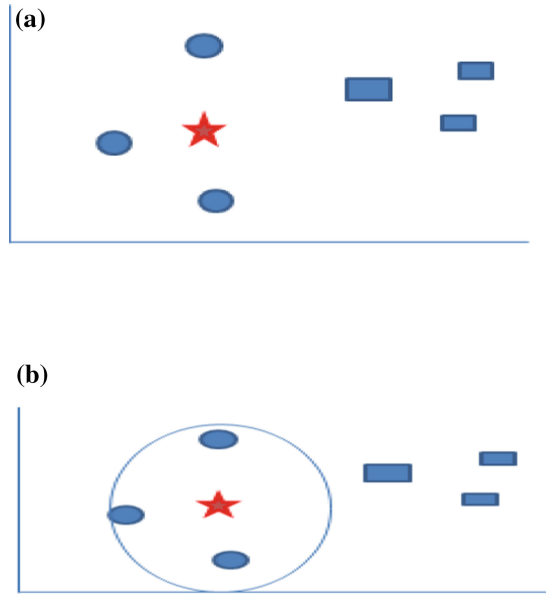


Fig. 1. Predicting recommendation



**Fig. 2.** a. The red star among the sparse particles. b. The red star belongs to the class blue circle.

#### **KNN-CFC algorithm steps:**

**Step 1:** Loading the dataset as input for preprocessing state.

**Step 2:** Set value of K for predicting number of clusters to be made.

**Step 3:** Predicted class iteration from 1 to no of the training set.

**Step 4:** Calculate Euclidean distance to predict the similarity among the dataset.

**Step 5:** Sort the distance in ascending order.

**Step 6:** Get the Top k rows that attain the most similar pattern.

**Step 7:** Get the frequent class of rows that represents the association among the dataset.

**Step 8:** Return predicted class from the observation.

### **3.3 System Architecture**

Figure 3 says that the user asks for a cloud service and the server recommends the user's search-related items by predicting the users' history and other users similar search topics. The recommendation system using Collaborative Filtering Coefficient filter and cluster the similar information and the result is listed with KNN classification approaches that predict the user service similarity from the k amount of adjacent nodes, the similarity is calculated from the Euclidean distance then predict the top frequently occurring services as a class to attain accuracy [22]. Then the novel recommendation system suggests the user the relevant information. Here Cloud provider may provide movie recommendation as a cloud service that is controlled by cloud administrator. This cloud service which contains multiple movie recommendation but for the perfect

service to the potential user, the recommendation engine will perform KNN classification which has the more impact on CFC. This Core CFC and KNN will provide the perfect service to the cloud users. Distance between the nearest cloud service users is calculated by various distance metrics and finally which has evaluated with real time cloud service providers. Here history of the users is very important essential keyword for the perfect services.

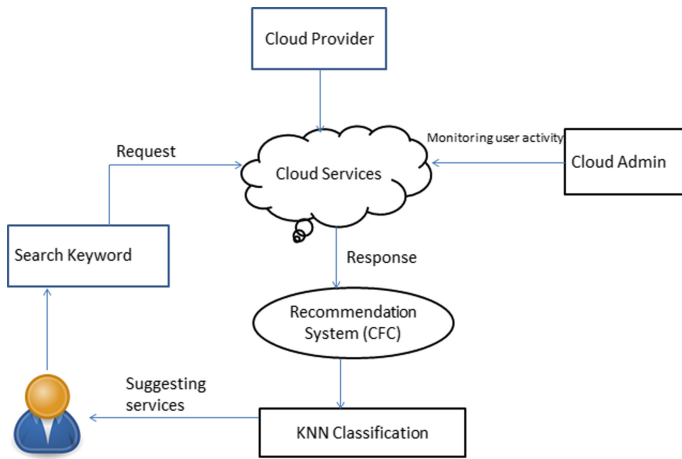


Fig. 3. System Architecture

## 4 Results and Discussion

The movie lens dataset is analyzed with 6000 users with 400 tags for timesharing nature in the hadoop framework environment to process the user information in a parallel manner. This service is deployed in the open source cloud platform (IaaS) to movie recommendation as a cloud service. With this setup the effective execution time and response time of the various users to be measured.

### 4.1 Precision, Recall and F-Measure

The new proposed system provides precision, recall and f-measure to evaluate the recommender system in a better way. The following table gives the values evaluation metrics (Figs. 4 and 5).

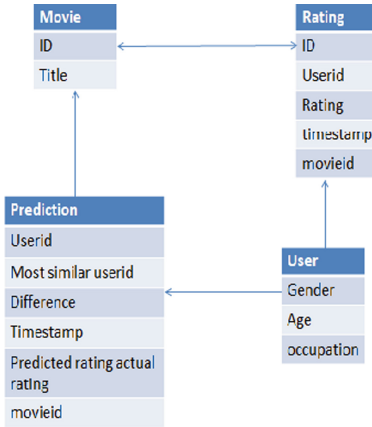


Fig. 4. Class diagram for a prediction

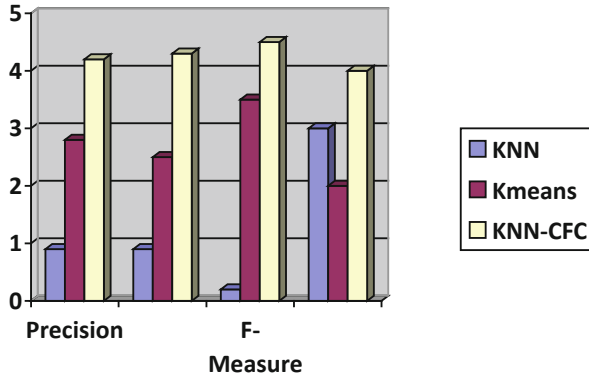


Fig. 5. Precision, Recall and F-measure

### 4.2 Evaluation Measure

We espouse the Mean Absolute Error (MAE) method and the Root Mean Squared Error (RMSE) method, which are broadly second-hand in a lot of recommendation systems, to appraise the routine of the new model. The terminology is as follows:

$$MAE = \frac{1}{N} \sum_{i=1}^N |q(i) - r(i)| \tag{2}$$

where  $N$  is the number of testing data samples and represents the prediction score according to Eq. (2).

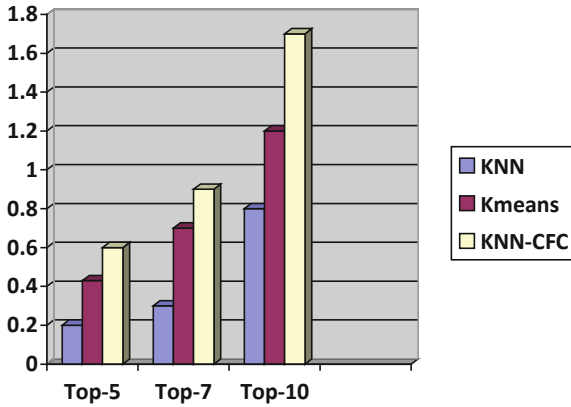


Fig. 6. Mean Absolute Error

### 4.3 Execution Time with Different no of Services

The experiment deals with the no of services and time in seconds for various data density in Fig. 6. Initially, the time is 100 s for 96 number of services for the data density 20%. And the time increases from 250–390 for the range 300–600 number of services for the density 40% and 60%. When there is a high data density, the recommendation time increases along with the improved number of services. It limits the search process to minimum data density. So, there commendation algorithm (collaborative filtering approach) can find parallel users and top-rated services.

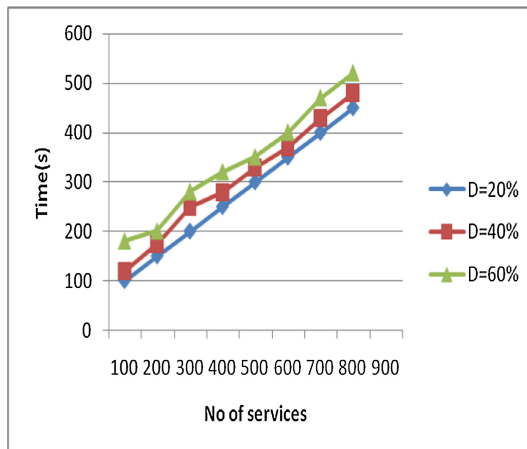


Fig. 7. Execution time comparisons



#### 4.4 Recommendation Time with Various no of Users

The experiment deals with the no of users and time in seconds for various data density in Fig. 7. Here the computation time is 110 s when the user’s count are less than 100 for rating density of 20% with top k recommendation [10]. Then the time varies from 190 to 350 for the user’s count is in the range of 400–700 for density 40% in top k = 10. So, the time increases with the user’s count and ratings as well. This helps to find the similarity in first few steps and works better with large ratings [24] (Fig. 8).

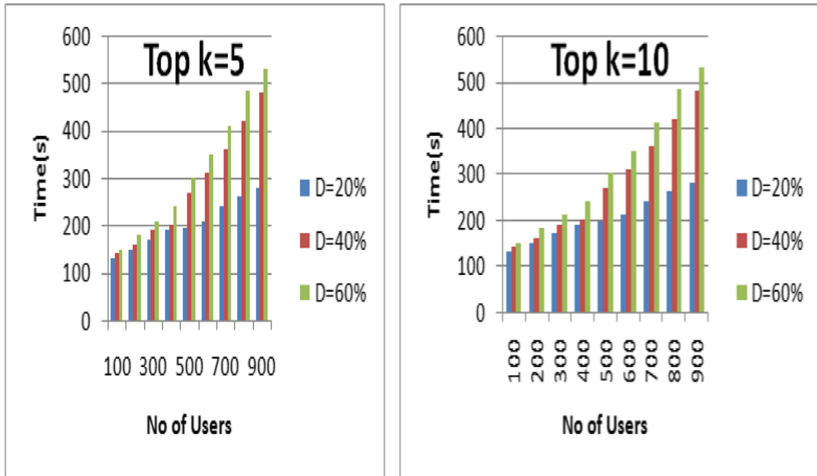


Fig. 8. Execution time comparisons

## 5 Conclusion

Recommendation system provides the service that the client looking for and overcome the challenge of correct service selection. Here we used a hybrid RS framework that combines the CFC clustering technique and the KNN technique. The CFC clustering technique is used to extract the active users relevant services based on the user history and similar neighbor user’s service votes. The KNN technique is suitable to find the recommendation even for the new user’s searching of services that reduce the service delivery time from cloud providers to the users. The user can also benefit the relevant service information. The execution time is low for predicting the large amount similar services from the cloud provider. The recommendation time is less when comparing with many cloud services even for large user rates. The recommendation framework performs with the help of QoS ranking and user rating. The research had been made with the small dataset and research on an actual social cloud will be the real test proposal in the future works.

## References

1. Jannati, H., Bahrak, B.: An improved authentication protocol for distributed mobile cloud computing services. *Int. J. Crit. Infrastruct. Prot.* **19**, 59–67 (2017)
2. Camacho, L.A.G., Alves-Souza, S.N.: Social network data to alleviate cold-start in recommender system. *Inf. Process. Manag.: Int. J.* **54**(4), 529–544 (2018)
3. Wang, S., Zhao, Y., Huang, L., Xu, J., Hsu, C.H.: QoS prediction for service recommendations in mobile edge computing. *J. Parallel Distrib. Comput.* (2017)
4. Wei, J., He, J., Chen, K., Zhou, Y., Tang, Z.: Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Syst. Appl.* **69**, 29–39 (2017)
5. Salah, A., Rogovschi, N., Nadif, M.: A dynamic collaborative filtering system via a weighted clustering approach. *Neurocomputing* **175**, 206–215 (2016)
6. Li, J., Zhang, K., Yang, X., Wei, P., Wang, J., Mitra, K., Ranjan, R.: Category preferred Canopy-K-means based collaborative filtering algorithm. *Future Gener. Comput. Syst.* **93**, 1046–1054 (2018)
7. Mezni, H., Abdeljaoued, T.: A cloud services recommendation system based on Fuzzy Formal Concept Analysis. *Data Knowl. Eng.* **116**, 100–123 (2018)
8. Bobadilla, J., Ortega, F., Hernando, A., Bernal, J.: A collaborative filtering approach to mitigate the new user cold start problem. *Knowl.-Based Syst.* **26**, 225–238 (2012)
9. Hayyolalam, V., Kazem, A.A.P.: A systematic literature review on QoS-aware service composition and selection in cloud environment. *J. Netw. Comput. Appl.* **110**, 52–74 (2018)
10. Ertuğrul, Ö.F., Tağluk, M.E.: A novel version of k nearest neighbor: dependent nearest neighbor. *Appl. Soft Comput.* **55**, 480–490 (2017)
11. Ortega, F., Zhu, B., Bobadilla, J., Hernando, A.: CF4 J: collaborative filtering for Java. *Knowl.-Based Syst.* **152**, 94–99 (2018)
12. Kumar, S., Pandey, M.K., Nath, A., Subbiah, K., Singh, M.K.: Comparative study on machine learning techniques in predicting the QoS-values for web-services recommendations. In: 2015 International Conference on Computing, Communication & Automation (ICCCA), pp. 161–167. IEEE, May 2015