International Journal of Modern Trends in Engineering and Research
www.ijmter.com
e-ISSN No.:2349-9745, Date: 2-4 July, 2015

CCPLSA: A Hybrid Approach for Similarity Computation to Rank Cloud services

R.Anupriya[1],S.Kirthica[2], Rajeswari Sridhar[3]

[1]*Department of Computer Science and Engineering, Anna University, Chennai, priyaanu2206@gmail.com*
[2] *Department of Computer Science and Engineering, Anna University, Chennai, s.kirthica@gmail.com*
[3] *Department of Computer Science and Engineering, Anna University, Chennai, rajisridhar@gmail.com*

Abstract—Due to the vast diversity of available cloud services, it is difficult for a customer to decide which cloud service satisfies his/her requirements at the most. To ease the customer's cloud service selection process an efficient ranking algorithm is required. QoS helps in identifying the customer's requirements and thereby aiding in optimal cloud service selection. We propose to analyze three QoS ranking algorithms and identify the issues in these algorithms. In order to overcome the identified constraints, we have designed a hybrid QoS rank identification framework for cloud services; Correlation Coefficient based Probabilistic Latent Semantic Analysis (CCPLSA), by combining two similarity computation methods. CCPLSA takes advantage of the past service usage experiences of other consumers and QoS attributes. A comparison of the various similarity computation approaches is performed to rank the clouds.

Keywords-cloud computing; rank identification; quality of service; similarity computation

I. INTRODUCTION

The number of cloud customers that use cloud services increase exponentially every day, and hence it has become increasingly difficult for cloud users to decide which provider can fulfill their Quality of Service (QoS) requirements. Every cloud service provider offers similar services with varied prices and performance levels by providing a variety of features. For example, a cloud provider might provide a cheaper service for storage, but it may be expensive for computation. Therefore, given the assortment of cloud service offerings, a major challenge for customers is to discover the apt cloud providers that can satisfy an individual user requirement [1], [2]. In the context of cloud computing, an active user is requesting ranking information from the Cloud Rank framework. A Cloud Rank framework is one which ranks services by using the identified QoS values of various currently available cloud services. A user can obtain service ranking prediction of all available cloud services from the Orchestration layer. More accurate rank identification results can be achieved by providing QoS values of more cloud services, by analyzing the requirement of the active user based on user's data.

Currently, there is no intelligent cloud framework to evaluate cloud offerings and rank them based on their ability to meet the user's QoS requirements. Here, we propose an intelligent cloud framework and a rank computation algorithm that measures the quality which is used to prioritize cloud services. Rank computation is based on QoS attributes selection and similarity computation methods. We propose to compare four similarity computation methods that are used to predict service rank. Kendall Rank Correlation Coefficient (KRCC) evaluates the degree of similarity by considering the response time and throughput between the current user and other users, Probabilistic Latent Semantic Analysis (PLSA) estimate the requested service availability from the set of available services based on resource availability. Fuzzy Multi Attribute Decision Making uses linear fuzzy (FMADM) membership function which is used to find the best cloud service among a given set of services using the features access time and availability, CCPLSA computation combine two different similarity methods for different QoS attributes. By evaluating these four methods user can select an optimal service from a set.

II. RELATED WORKS

QoS is defined as a set of properties including response time, throughput, availability, accountability, failure probability, etc. These QoS values of some properties (e.g., response time, user-observed availability, etc.) are essential to be measured at the customer-side. It is difficult to get QoS details and related information from cloud service providers, as these QoS values depend, to a greater extent, on the quality of Internet connection. Hence, the QoS values of the same cloud service would vary for different users.

Zheng et al [1] have proposed a personalized ranking framework that predicts the QoS rank of a set of cloud services. In their work, two attributes (response time, throughput) were used to compute similarity. For similarity computation they proposed KRCC. However, using only two attributes will greatly affect the prediction accuracy. To address this problem, we use set of QoS attributes to predict the service rank. Personalized QoS Prediction [1] approach evaluates all the service candidates equally so the rank prediction result is not providing better accuracy. To overcome this problem we consider both user experience and QoS attribute to predict the service rank.

Shivakumar et al [2] proposed Rank Cloud Services using FMADM, to rank cloud services in terms of QoS attributes. The limitation of this approach that it uses sharp cut off range for attributes selection based on a rectangular membership function. Kim et al [3] defined PLSA based service allocation. It describes a trust model which analyzes the history information of each node and allocates resources according to customer requests.

Saurabh Kumar Garg et al. [4] have presented the first framework –SMI Cloud to compute all the QoS attributes proposed by Cloud Service Measurement Index Consortium (CSMIC). They have focused on some key challenges in designing metrics for each quantifiable QoS attribute for measuring the service level of each Cloud provider. Garg et al. [5] have proposed a framework to measure the quality and prioritize Cloud services. This framework makes major impact and creates healthy competition among Cloud providers to satisfy their Service Level Agreement (SLA) and improve their Quality of Services (QoS).

ZibinZheng et al. [6] have proposed Cloud Rank approaches to rank the cloud services using a greedy algorithm. It ranks the component, which considers the explicitly rated items and the unrated items equally. Jilian Wu et al. [7] defined QoS prediction for service selection based on collaborative filtering. It predicts the unknown values for QoS-based selection using neighborhood based collaborative filtering approach.

We aim to examine three QoS ranking algorithms namely, KRCC, PLSA, and FMADM. This paper also proposes a new QoS ranking prediction framework for cloud services by combining KRCC and PLSA named “CCPLSA”.

III. SYSTEM DESIGN

System architecture of our orchestration layer, which provides QoS rank identification for cloud services, is given in fig. 1. The target users of the orchestration layer are the cloud applications, which need personalized cloud service ranking in order to choose an optimal cloud service. Within the Cloud Rank framework, there are several modules. First, based on the user-provided QoS values, similarities between the active user and training users are calculated. Second, based on the similarity values, a set of similar users can be identified. After that, two algorithms are used (i.e., CloudRank1 and CloudRank2) [1] to make personalized service ranking by taking advantages of the past service usage experiences of similar users along with the QoS values obtained from cloud. Finally, the ranked results are provided to the active user. The training data in the Cloud Rank framework can be obtained from the QoS values provided by other users based on previous experience and the QoS values collected by monitoring cloud service.

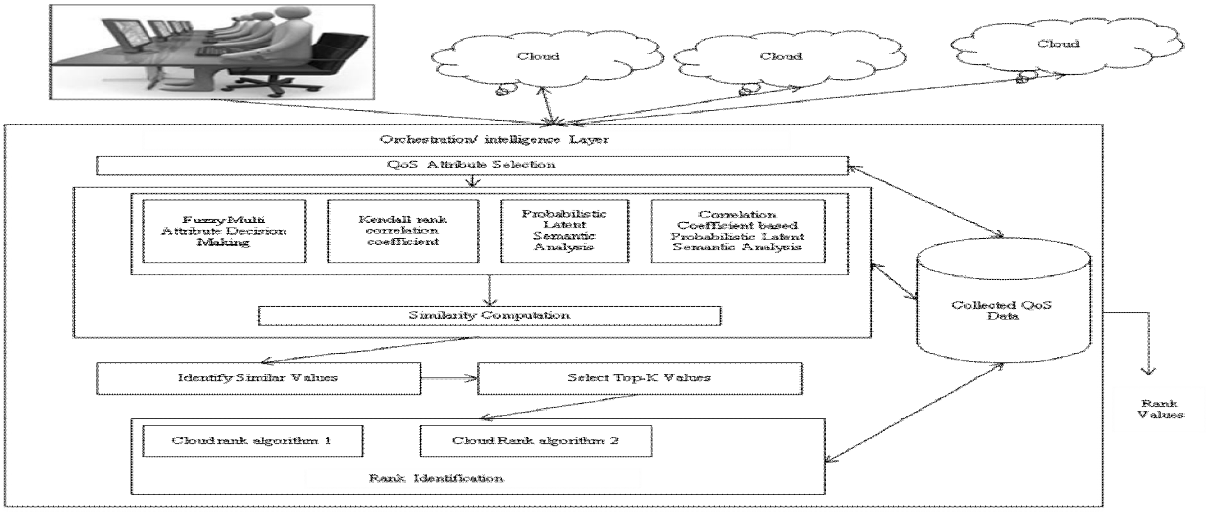


Fig 1. Detailed system architecture of cloud rank

A. Identify QoS attributes

There are two types of QoS requirements which a user can have: functional and non-functional. Some of the requirements cannot be measured easily given the nature of the Cloud. Attributes like security and user experience are not easy to quantify. Here we are considering only few attributes to perform similarity computation and use these to rank the cloud services. QoS attributes are used to compute similarity values using the four approaches.

B. Similarity computation

After identifying features, in this module, ranking similarity computation performs a comparison on the individual users' requested QoS values. The ranking similarity between requested user and training users is in the interval of [-1,1], where -1 is obtained when the order of active user is the exact reverse of training user , and 1 is obtained when order of active user is equal to the order of training user. Here four computation methods with different set of QoS attributes are used to compute similarity. The four methods are KRCC, PLSA, FMADM, and the proposed CCPLSA Approach. The following sub-sections discuss the approaches in detail.

KRCC. Kendall Rank Correlation Coefficient (KRCC) evaluates the degree of similarity by considering the response time and throughput between the current user and other users. Compute concordant and discordant values based on response time and throughput. Both attributes are moving in same direction (Increase or decrease) then the values are called concordant values. If the values are moves in different direction then it is called discordant values.

PLSA. Probabilistic Latent Semantic Analysis (PLSA) is used to estimate the requested service availability from the set of available services. In PLSA first we have to find the latent (hidden) variable. Then using Baye's theorem compute a specific attributes probability for a given set of QoS attributes. The attributes used in PLSA are resource availability.

FMADM. Fuzzy Multi Attribute Decision Making (FMADM) calculates rectangular function which is used to find the best cloud service among a given set of services. Two features access time and service availability are used in the rectangular function.

CCPLSA. It combines two different similarity methods of different QoS attributes which is shown in fig. 2. Here we are combining KRCC and PLSA. It is a two step approach, where we first compute KRCC for user requested service. Then perform PLSA for values obtained from KRCC.

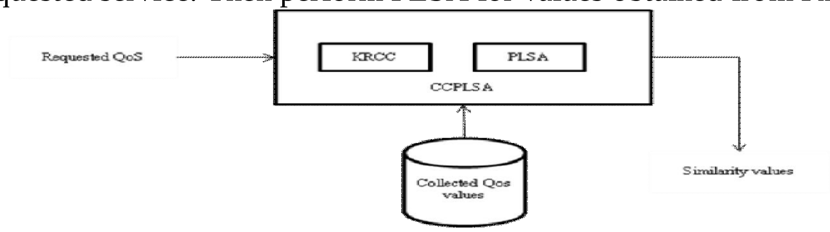


Fig2.CCPLSAComputation

C. Identify similar values

In similarity computation, similar values are obtained from different methods. QoS values of dissimilar users will greatly affect the rank identification accuracy. To address this problem, we exclude the users with negative correlations (negative similarity values i.e. -1 to 0) and only pick the Top-K similar users for making QoS ranking identification.

$$N(u) = \{v \mid v \in Tu, Sim(u, v) > 0, v \neq u\} \quad 2.1$$

Where Tu is a set of the Top-K similar users to the user u and sim (u,v) > 0 excludes the dissimilar users with negative similarity values. Equation 2.1 is used to eliminate negatively correlated value.

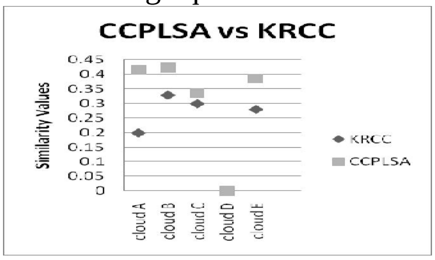
D. Rank identification

.A positively correlated value of Top-K user is used for ranking. Ranking of cloud services is one of the most important features of the cloud framework. Rank identification system computes the relative ranking values of various available cloud services based on the QoS requirements of the customer. Here two ranking oriented approaches, named Cloud rank algorithm1 and Cloud rank algorithm 2, are used [1]. Ranking-oriented approaches predict the QoS ranking directly without predicting the corresponding QoS values.

IV. EXPERIMENTS

To compare the performance of different similarity computation methods, we implement cloud rank algorithm using the KRCC, PLSA, FMADM and CCPLSA similarity computation methods. Here, we set Top-K to non-Zero Positive values. Similarity values for each method based on QoS attributes were computed. We proposed CCPLSA approach, in that the features incorporated with CCPLSA and KRCC are same and also with PLSA. Similarity values between CCPLSA and KRCC are shown in fig. 3. Both methods incorporated same set of attributes, so we compare these two, but computation is different. Hence the similarity values obtained by the two algorithms are different. In fig. 4, similarity values between CCPLSA and PLSA are shown.

After similarity computation, identify similar user will eliminate the dissimilar cloud users. Here cloud D get zero so it is eliminated. The rank relationships between CCPLSA and KRCC are shown in fig. 5 and CCPLSA and PLSA are shown in fig. 6. It has been observed that for a particular scenario the CCPLSA and KRCC provided all the clouds with same rank. However CCPLSA ranked cloud A with 3 and cloud B with 4 and PLSA rank cloud A with 4 and cloud B with 3. In order to identify the accuracy of the algorithm that has ranked correctly, we used the Normalized Discounted Cumulative Gain (NDCG) measure which is computed from the Discounted Cumulative Gain (DCG) as explained in the following equations.



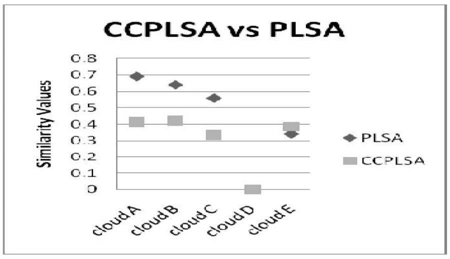
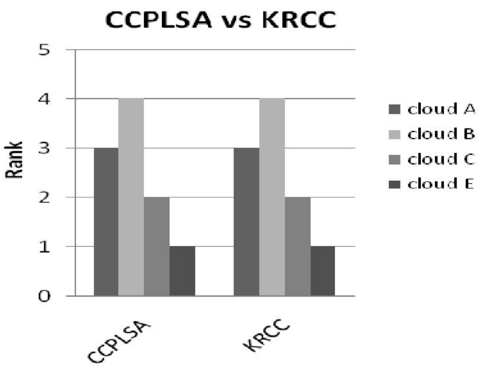


Fig3:CCPLSA vs KRCC *Fig:CCPLSA vs PLSA*



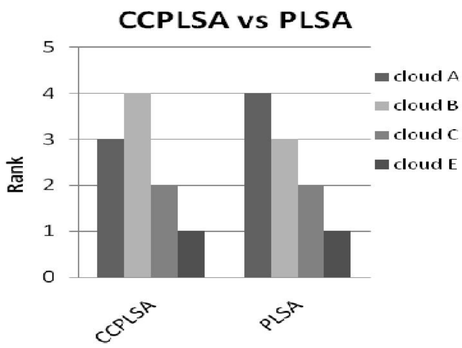


Fig5:CCPLSAvs KRCC *Fig6:CCPLSA vs PLSA*

Discounted Cumulative Gain:

$$DCG_K = \mathrm{rel}_1 + \sum_{i=2}^{K}\left(\frac{\mathrm{rel}_i}{\log_2(i)}\right). \quad 4.1$$

Normalized DCG:

$$NDCG = \frac{\text{Discounted Cumulative Gain (DCG)}}{\text{Ideal Discounted Cumulative Gain (IDCG)}} \quad 4.2$$

Where reli - Relevant Service,P- Particular rank position (K).Using equations 4.1 and 4.2 we determine the DCG, manually compute the IDCG and determine NDCG. The results are tabulated in table 1.

Table 1.NDCG Comparison

Similarity methods	DCG	IDCG	NDCG
CCPLSA	8.64	8.79	0.982
KRCC	8.62	8.79	0.98
PLSA	8.22	8.79	0.93

From the table 1, the rank identification accuracies of employing CCPLSA and KRCC are better than PLSA as the NDCG is higher for these two approaches. With the increase of attributes, CCPLSA algorithm would yield better results in ranking clouds when compared with KRCC and PLSA. When the attributes are sparse, the similarity computation methods do not have enough information for making accurate rank identification. From the above Fig. 5 the performance of KRCC and CCPLSA is same for the sample experiment conducted.

CONCLUSION AND FUTURE WORK

An intelligent cloud framework to evaluate cloud offerings and rank them based on their user requested Quality of service requirements was implemented by using four similarity computational approaches namely Kendall Rank Correlation Coefficient, Probabilistic Latent Semantic Analysis, Fuzzy Multi Attribute Decision Making and a CCPLSA approach. Also the rank identification algorithm as described in cloud rank algorithm1 and cloud rank algorithm 2 were implemented by using different set of attributes. The overall performance efficiency of the cloud rank framework was studied. It can be concluded that with the limited set of features used, CCPLSA provided better cloud ranking when compared to PLSA while CCPLSA could not be distinguished when compared to KRCC in ranking cloud. The inconclusive nature of CCPLSA approach in ranking is due to sparse use of features. The features need to be increased to perform a thorough analysis. In addition, the choice of features needs to be analyzed before using them as parameters to rank cloud services. As our current approaches did not incorporate all attributes into a single similarity method the choice of algorithm still remain ambiguous. We will conduct more investigation on the combination of different QoS attributes, so that the users can obtain better accuracy in rank identification.

REFERENCES

[1] Zheng, Z., Wu, X., Zhang, Y., Lyu, M.R., Wang, J. QoS Ranking Prediction for Cloud Services. IEEE Transaction on Parallel and Distributed Systems, 1213-1222, 2013.

[2] Shivakumar, U., Ravi, V., and Gangadharan, G.R. Ranking Cloud Services Using Fuzzy Multi-Attribute Decision Making. IEEE International Conference on Fuzzy Systems (FUZZ), 1-8, 2013.

[3] Kim, H., Lee. H., Kim, W., Kim, Y. A Trust Evaluation Model for QoS Guarantee in Cloud Systems. International Journal of Grid and Distributed Computing, 2010.

[4] Garg, S.K., Versteeg, S., Buyya, R. A framework for ranking of Cloud computing services. Future Generation Computer Systems (2013)1012–1023, 2013.

[5] Garg, S, K., Versteeg, S., Buyya, R. SMICloud: A Framework for Comparing and Ranking Cloud Services. International Conference on Utility and Cloud Computing, 210-218,2011

[6] Z., Zhang, Y., Lyu, M.R. A QoS-Driven Component Ranking Framework for Cloud Computing. International Symposium on Reliable Distributed Systems (SRDS'10), 184-193, 2010.

[7] Wu, J., Chen, L., Feng, Y., Zheng, Z., Zhou, M.C., Wu, Z. Predicting QoS for Service Selection by Neighborhood - Based Collaborative Filtering. IEEE Transactions on systems, man, and cybernetics, 428-439, 2013.

[8] Chen, X., Zheng, Z., Qi Yu., Lyu, M.R. Web Service Recommendation via Exploiting Location and QoS Information. IEEE Transactions on parallel and distributed systems, 1913 – 1924, 2014.

