

# A Survey on QoS Ranking in Cloud Computing

M. Subha<sup>1</sup>, M. Uthaya Banu<sup>2</sup>

<sup>1,2</sup>Student & Department of Computer Science and Engineering, Regional Centre of Anna University, Tirunelveli (T.N)

**Abstract**— QoS (Quality-of-Service) is an important topic in cloud computing. It is very difficult to make a decision on choosing the cloud services depending on QoS requirements. These requirements have to be satisfied by both cloud service providers and cloud users. So, Optimal Service Selection is needed to obtain high quality cloud applications. With the increasing number of Cloud services, Quality-of-Service (QoS) is usually employed for describing non-functional characteristics of Cloud services. The QoS performance of cloud applications becomes low due to unreliable Internet connections. In this paper, we have presented a widespread survey on QoS Ranking in Cloud Computing with respect to their Limitations and Inferences.

**Keywords**— Cloud Applications, Cloud Services, Optimal Service Selection, Prediction, Quality-of-Service.

## I. INTRODUCTION

Cloud computing is a new paradigm for delivering on-demand resources (e.g., infrastructure, platform, software, etc.) for customers similar to other utilities (e.g., water, electricity and gas). The current Cloud computing architecture enables three layers of services [1]. The cloud removes the need for you to be in the same physical location as the hardware that stores your data. There are number of functionally equivalent services in the cloud. Due to unreliable internet connections different cloud applications may receive different levels of quality for same cloud services so that we need to select the optimal services.

Cloud computing provides three main services, namely Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). In Software as a Service (SaaS), Clients can use the software to provide by the provider, which usually need not to install and it is usually a one of many services. Like Gmail, search engine. In Platform as a Service (PaaS), Clients can run their own applications on the platform provided; General platforms are Linux and Windows. In Infrastructure as a Service (IaaS), Client can put their own operating system on cloud. With the rapid growth of cloud computing, a number of service providers have appeared who offer similar services at different prices and performance levels.

Furthermore, due to the dynamic nature of cloud services which results from the elasticity and on-demand provision of computing resources, there are considerable fluctuations in the Quality of Service (QoS) levels of each service [2].

QoS is defined as a set of properties including response time, throughput, availability, reputation, failure probability, etc. Among these QoS properties, values of some properties (e.g., response time, user-observed availability, etc.) essential to be measured at the client-side [3]. It is impractical to get such QoS information from service providers, since these QoS values are susceptible to the uncertain Internet environment. Therefore, different users may observe quite different QoS values of the same cloud service.

Optimal Service Selection is also unrealistic for users to acquire QoS information by evaluating all service candidates by themselves, since conducting real world web service invocations is time-consuming and resource-consuming. Moreover, some QoS properties (e.g., reliability) are difficult to be evaluated as long-duration observation is required. QoS Ranking provides scalable services and adaptive to the diversity of end users. The active user and Training users are identified. The Similarity between those two users are calculated either Kendall Rank Correlation Coefficient or Pearson Correlation Coefficient.

The Kendall Rank Correlation Coefficient (KRCC) [4] evaluates the degree of similarity by considering the number of inversions of service pairs which would be needed to transform one rank order into the other. Pearson Correlation Coefficient (PCC) has been introduced in a number of recommender systems for similarity computation, since it can be easily implemented and can achieve high accuracy. The rating-based collaborative filtering approaches try to predict the missing QoS values in the user-item matrix as accurately as possible.

It considered only real time invocation of service candidates, but ranking-based prediction approach considers the past usage experience of cloud services. The rating-based approaches include dissimilar users and it provides low ranking accuracy. The former approach treated the rated and unrated items equally so the QoS Ranking approach is used to achieve better ranking accuracy of cloud services.

## II. QoS RANKING TECHNIQUES

Quality-of-Service (QoS) is describing non-functional characteristics of Cloud services. Among different QoS properties of Cloud services, some QoS properties are user-dependent and have different values for different users (e.g., response time, invocation failure probability, etc.). Ranking-based QoS prediction approaches aim at predicting the quality ranking of the target Cloud services instead of the detailed QoS values.

Bonatti et al. [5] have proposed a best way of identifying optimal Service Selection problems based on three criteria. It provides an Exact and approximated algorithms for optimal service selection based on a given set of service requests, a set of service users, the result of the matchmaking process, and a numeric preference measure. It identifies the Service Selection Problem (SSP). The high computational complexity of the service selection problem is caused by the one-time costs related to service users (e.g., Initialization and registration costs). In the absence of one-time costs, the optimal selection problem can be solved in polynomial time by applying a greedy approach. The heuristic algorithm seems to be faster, but it has no guarantees on the quality of the solution.

Breese et al. [6] have proposed collaborative filtering algorithms that predict the utility of items to a particular user (the active user) based on a database of user votes from a sample or population of other users (the user database). We use two basic classes of evaluation metrics. The first characterizes accuracy over a set of individual predictions in terms of average absolute deviation. The second estimates the utility of a ranked list of suggested items. Bayesian networks typically have smaller memory requirements and allow for faster predictions than a memory-based technique such as correlation, but Bayesian methods examined here require a learning phase that can take up to several hours and results are reflected in the recommendations.

Deshpande et al. [7] have proposed an Item-Based Top-N Recommendation Algorithms that determines the similarities between the various items from the set of items to be recommended. The goal of top-N recommendation algorithm was to classify the items purchased by an individual user into two classes: like and dislike. This algorithm is faster than the traditional user-neighborhood based recommender systems and it provides recommendations with comparable or better quality. The proposed algorithms are independent of the size of the user-item matrix.

Linden et al. [8] proposed Recommendation Algorithm which determines a set of customers whose purchased and rated items overlap the user's purchased and rated items. The algorithm aggregates items from these similar customers, eliminates items the user has already purchased or rated, and recommends the remaining items to the user. It generates high quality recommendations and the algorithm must respond immediately to new information. It is used to distinguish the online store for each customer, but it needs to apply recommendation algorithms for targeted marketing, both online and offline.

Zibin Zheng et al. [9] proposed Hybrid collaborative filtering method that to improve performance of Recommender System. It comprises a user-contribution mechanism for Web service QoS information collection and a novel hybrid collaborative filtering algorithm for Web service QoS value prediction. It collects efficient QoS information and it provides better feasibility of the web service recommender system, but it has to monitor more real-world web services and it has to investigate more QoS properties of Web services.

Performance evaluation of server farms is an important aspect of cloud computing. Hamzeh Khazaei et al. [10] have proposed an analytical technique based on an approximate Markov chain model for performance evaluation of a Cloud computing. It is considered only response time as a major factor.

A common principle of previous research is that the QoS values of services to target users are supposed to be known to all. However, many of QoS values are unknown in reality. Wu et al. [11] have proposed a neighborhood-based collaborative filtering approach to predict such unknown values for QoS-based selection. It removes the impact of different QoS scale. The prediction accuracy is improved by using a data smoothing process. It reduces the data sparsity problem using a similarity fusion approach.

Yu et al. [12] have proposed a broker-based architecture to assist the selection of QoS-based services. The objective of service selection is to maximize an application-specific utility function under the end-to-end QoS constraints. The combinatorial model defines the problem as a multidimensional multichoice 0-1 knapsack problem. The graph model defines the problem as a multiconstraint optimal path problem. QoS for web services refers to various nonfunctional characteristics such as response time, throughput, availability, and reliability and failure probability.

Broker-based architecture provides an end-to-end QoS management for distributed cloud services. The main functions of QBroker include service discovery, planning, selection, and adaptation. The efficiency of QBroker is dominated by the running time of the service selection algorithm. In both models, a user-defined utility function of system parameters may be specified to optimize application-specific objectives. The heuristic algorithm finds near-optimal solutions in polynomial time, which is more suitable for making runtime decisions. It handles rich composition structures, including sequential, parallel, conditional, and loops of services.

Saurabh Kumar Garg et al. [13] 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). They proposed an Analytical Hierarchical Process (AHP) based ranking mechanism which can estimate the cloud services based on different applications depending on QoS requirements. This technique is used **only for quantifiable QoS attributes** such as Accountability, Agility, Assurance of Service, Cost, Performance, Security, Privacy, and Usability. It is **not suitable for non-quantifiable QoS attributes** such as Service Response-time, Sustainability, Suitability, Accuracy, Transparency, Interoperability, Availability, Reliability and Stability.

Alexandru Iosup et al. [14] have analyzed the performance of many-task applications on Clouds. Similarly, many performance monitoring and analysis tools are also proposed. By utilizing these tools the authors can use the data to rank and measure the QoS of various Cloud services according to consumer's cloud services. There are three main differences between scientific computing workloads and the initial target workload of clouds: in required system size, in performance demand, and in the job execution model. The reason for this selection is threefold.

First, not all the clouds on the market are still accepting clients; **Flexi Scale** puts new customers on a **waiting list** for over two weeks due to system overload. Second, not all the clouds on the market are **large enough to accommodate** requests for even 16 or 32 coallocated resources. Third, our selection already covers a wide range of quantitative and qualitative cloud characteristics of cloud. The main feature is that the compute performance of the tested clouds is low. It provides a good solution for the scientists who need resources instantly and temporarily.

Saurabh Kumar Garg et al. [15] have presented the first framework, SMICloud 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. They also have proposed an Analytical Hierarchical Process (AHP) based ranking mechanism which can assess the Cloud services based on various applications depending on QoS requirements. Their proposed mechanism also addressed the challenge of different dimensional units of various QoS attributes by providing a constant way to evaluate the relative ranking of cloud services for each type of QoS attribute.

Zibin Zheng et al. [17] have proposed CloudRank approaches to rank the cloud services in an optimal way using a greedy algorithm. It ranks the component instead of service, but this algorithm is used to rank a set of items, which treats the **explicitly rated items and the unrated items equally**. It does not guarantee that the explicitly rated items will be ranked correctly.

Component quality ranking approaches [22] are vital for making an optimal component selection from a set of functionally equivalent components candidates. CloudRank Framework ranks the component by taking advantage of past component usage experiences of different component users.

The rating-based QoS prediction approaches aim at predicting QoS values for different service users. Zibin Zheng et al. [18] have proposed a CloudRank Prediction Framework that predicts the QoS Ranking directly instead of predicting the experimental QoS values. The predicted QoS values can be employed to rank the target Cloud services. The major challenge of making QoS-driven Cloud service quality ranking is that the Cloud service quality ranking of a user cannot be transferred directly to another user, since the user locations are different. So, the author proposed a personalized QoS Ranking for cloud services. It evaluates all the Cloud services at the user-side and rank the Cloud services based on the observed QoS performance. Moreover, it is difficult for the service users to evaluate all the Cloud services themselves, since there may exist a huge number of Cloud services in the Internet.

Xi Chen et al. [19] have proposed a novel collaborative filtering based web service recommender system to help users select services with optimal Quality-of-Service (QoS) performance. It provides a best way for a user to select an optimal web service among a large amount of service candidates.

## International Journal of Emerging Technology and Advanced Engineering

Website: [www.ijetae.com](http://www.ijetae.com) (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 4, Issue 2, February 2014)

QoS values evaluated by one user cannot be employed directly by another for service selection. It uses the location information and QoS values to cluster users and services, and makes personalized service recommendation for users based on the clustering service results. Moreover, some QoS properties (e.g., reliability) are difficult to be evaluated as long-duration observation is required.

Location-aware Web service recommender system (named LoRec), which employs both Web service QoS values and user locations for making personalized QoS prediction. Users of LoRec share their past usage experience of web services, and the system provides personalized service recommendations to them. Therefore, different users may observe quite different QoS values of the same web service. It is also unfeasible for users to acquire QoS information by evaluating all service candidates by themselves, since conducting real world Web service invocations is time-consuming and resource-consuming.

LoRec first collects user observed QoS records of different Web services and then users who have similar QoS observations together to generate recommendations.

Location information is also considered when clustering users and services. It selects the optimal cloud service based on historical QoS records of web services. The basic idea is to predict Web service QoS values and recommend the best one for active users based on historical Web service QoS records. It has to improve the scalability of LoRec (Location-aware Web service recommender system).

### III. CONCLUSIONS

Cloud computing aim is to provide scalable and adaptive to the diversity of end-users. Optimal service selection is important to obtain high quality cloud applications. A greedy algorithm treats rated and unrated items equally so it provides low quality cloud applications. CloudRank Framework provides the same quality in both algorithms. So, we suggest an optimal VM allocation is used to improve the quality of cloud applications. In this paper, we have carried out a significant review in QoS Ranking of cloud services. This survey paper will hopefully motivate future researchers to come up with high quality cloud applications using QoS ranking techniques.

**TABLE I**  
**SUMMARY OF LIMITATIONS AND INFERENCES FROM EXISTING TECHNIQUES**

Title	Techniques/ Approaches Used	Description	Limitations	Inferences
Item-Based Top-n Recommendation	Model-based Recommendation Algorithms	It determines the similarities between the various items from a set of items to be recommended.	(i) It has no guarantees on the quality of the solution. (ii) The proposed algorithms are independent of the size of user-item matrix.	(i) Recommendations are computed based on either use item-to-item or item set –to-item. (ii) Results showed that a recommender system provides better accurate recommendation than Collaborative Filtering techniques. (iii) It is faster than other rating based algorithms.
Amazon.com Recommendations: Item-to-Item Collaborative Filtering	Collaborative Filtering Algorithms	It predicts the utility of items to a particular based on a database of user votes from a sample or population of other users.	(i) Need to apply recommendation algorithms for targeted marketing, both online and offline. (ii) Need to support for offline retailers like, postal mailings, coupons, and other forms of customer communication.	(i) It generates high quality Recommendations. (ii) Results showed that the Recommendation System does little offline computation than Traditional collaborative filtering techniques.

QoS-Aware Web Service Recommendation by Collaborative Filtering	User-Collaborative Mechanism	It predicts the QoS values of web services by taking the past usage experience of service users	(i) Need to predict more real-world web services from more services. (ii) Need to investigate more QoS attributes of web services.	(i) It is implemented with JDK, Eclipse, Axis2, <sup>6</sup> and Apache Tomcat. (ii) Results showed that the item-based approach achieves better prediction accuracy than the user-based approach in our Web service QoS data set. (iii) Optimal service selection process reduces the effect of the web service invocations to the real-world Web services.
Performance Analysis of Cloud Computing Centers Using M/G/m/m+r Queuing Systems	An approximate Markov chain model	It allows us to obtain a complete probability distribution of response time and the number of tasks in the system.	(i) Not suitable for burst arrivals. (ii) Assumed general service time for requests as well as a large number of servers, which makes our model flexible in terms of scalability and diversity of service time.	(i) The performance evaluation is based on the response time of web services. (ii) Results showed that the proposed approximate method provides results with a high degree of accuracy for the mean number of tasks in the system.
Predicting QoS for Service Selection by Neighborhood - Based Collaborative Filtering	A Neighborhood based collaborative filtering approach	It predicts the unknown values for QoS-based selection.	(i) Need to use CF-based service selection architecture. (ii) Need to use two phase neighbor selection (TNS) to improve the scalability of ADF.	(i) Results showed that the WSRec approach is the best one for QoS prediction at present. (ii) A-cosine equation is used to compute the service-based similarity.
Efficient Algorithms for Web Services Selection with End-to-End QoS Constraints	The Combinational model and graph model	It maximizes an application specific utility function under the end-to-end QoS constraints	(i) Need more algorithms to obtain an optimal service selection. (ii) Need to provide a practical solution to the end-to-end QoS guarantee on service processes.	(i) The performance evaluation based on four QoS attributes and utility values for service candidates: response time, cost, availability, and reliability. (ii) Results showed that the heuristic algorithms reduce the time complexity
<b>Title</b>	<b>Techniques/ Approaches Used</b>	<b>Description</b>	<b>Limitations</b>	<b>Inferences</b>
SMICloud: A Framework for Comparing and Ranking Cloud Services	AHP based ranking mechanism	It measures the quality and prioritizes Cloud services.	(i) Non quantifiable QoS attributes are not used. (ii) It is not compatible with various QoS attributes.	(i) It provides a uniform way to evaluate the relative ranking of Cloud services for each type of QoS attribute. (ii) Performance monitoring and analysis tools are used to rank and measure the QoS of various Cloud services according to user's applications



## International Journal of Emerging Technology and Advanced Engineering

Website: [www.ijetae.com](http://www.ijetae.com) (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 4, Issue 2, February 2014)

Performance analysis of Cloud computing services for many-tasks scientific computing	A framework for performance monitoring and analysis tools	It analyzes the performance of many task applications on Clouds.	(i) After analyzing the services can be ranked. (ii) The compute performance of the tested clouds is low. (iii) It is not compatible with large resources.	(i) It achieves high performance than scientific computing alternatives such as grids and parallel production infrastructures. (ii) It focuses on a different platform (that is, clouds) and that we target a broader scientific computing community.
A framework for ranking of Cloud computing services	AHP hierarchy using SMI architecture	It computes all the QoS attributes proposed by Cloud Service Measurement Index Consortium (CSMIC)	(i) Ranking Algorithms can be deployed to rank infrastructures. (ii) It has not fully met all QoS requirements.	(i) Results showed that the cloud environment has high performance than source environment. (iii) The current clouds need an order of magnitude in performance improvement to be useful to the scientific community
CloudRank: A QoS-Driven Component Ranking Framework for Cloud Computing	CloudRank Framework	It performs component ranking by taking the advantages of service candidates	(i) Greedy algorithm treats the explicitly rated items and the unrated items equally. (ii) Personalized approach evaluates all the service candidates equally. (iii) Personalized approach is not compatible with large invocations of service users.	(i) CloudSim Tool is used to evaluate the ranking of cloud components. (ii) Results showed that the high performance of components using Normalized Discounted Cumulative gain (NDCG) metric. (iii) It aggregates the preference between a pair of components to produce a ranking of components.
QoS ranking prediction for cloud services	QoS Ranking Prediction Framework	It performs service ranking by taking the advantages of service candidates	(i) CloudRank algorithms have provided the same quality of cloud services. (ii) Personalized QoS Prediction approach evaluates all the service candidates equally. (iii) Need to use optimal VM allocation for each service.	(i) CloudSim Tool is used to evaluate the ranking of cloud services. (ii) Results showed that the high performance of components using Normalized Discounted Cumulative gain (NDCG) metric. (iii) It predicts QoS Ranking directly instead of predicting QoS values.
Web service recommendation via exploiting Location and QoS information	Collaborative filtering based Web service recommender system approach	It provides an optimal way for a user to select a proper Web service among a large amount of service candidates.	(i) Need to improve the scalability of LoRec. (ii) It provides an error when the prediction generated from both user region and web service region.	(i) Java classes are generated using the WSDL2Java tool of Axis2 package (ii) Results showed that the LoRec have achieved better prediction with more QoS data.

## International Journal of Emerging Technology and Advanced Engineering

Website: [www.ijetae.com](http://www.ijetae.com) (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 4, Issue 2, February 2014)

### REFERENCES

- [1] Buyya et al., "Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility," *Future Generation Computer Systems*, vol. 25, no. 6, pp. 599–616, 2009.
- [2] S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, "A novel ultrathin elevated channel low-temperature poly-Si TFT," *IEEE Electron Device Lett.*, vol. 20, pp. 569–571, Nov. 1999.
- [3] N. Thio and S. Karunasekera, "Automatic measurement of a QoS metric for Web service recommendation," in *Proceedings Australian Software Engineering Conference*, 2005, pp. 202–211.
- [4] J. Marden, *Analyzing and Modeling Ranking Data*. Chapman & Hall, 1995.
- [5] P.A. Bonatti and P. Festa, "On Optimal Service Selection," *Proc. 14th Int'l Conf. World Wide Web (WWW '05)*, pp. 530-538, 2005.
- [6] J.S. Breese, D. Heckerman, and C. Kadie, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," *Proc. 14th Ann. Conf. Uncertainty in Artificial Intelligence (UAI '98)*, pp. 43-52, 1998.
- [7] M. Deshpande and G. Karypis, "Item-Based Top-n Recommendation," *ACM Trans. Information System*, vol. 22, no. 1, pp. 143-177, 2004.
- [8] G. Linden, B. Smith, and J. York, "Amazon.com Recommendations: Item-to-Item Collaborative Filtering," *IEEE Internet Computing*, vol. 7, no. 1, pp. 76-80, Jan. /Feb. 2003.
- [9] Z. Zheng, H. Ma, M.R. Lyu, and I. King, "QoS-Aware Web Service Recommendation by Collaborative Filtering," *IEEE Trans. Service Computing*, vol. 4, no. 2, pp. 140-152, Apr.-June 2011.
- [10] Hamzeh Khazaei, Jelena and Vojislav, "Performance Analysis of Cloud Computing Centers Using M/G/m/m+r Queuing Systems", Published in the *IEEE Transactions on parallel and Distributed systems*, Vol. 23, No. 5, may 2012.
- [11] J. Wu, L. Chen, Y. Feng, Z. Zheng, M. Zhou, and Z. Wu, "Predicting QoS for Service Selection by Neighborhood -Based Collaborative Filtering," *IEEE Trans. System, Man, and Cybernetics, Part A*, to appear.
- [12] T. Yu, Y. Zhang, and K.-J. Lin, "Efficient Algorithms for Web Services Selection with End-to-End QoS Constraints," *ACM Trans. Web*, vol. 1, no. 1, pp. 1-26, 2007.
- [13] Saurabh Kumar Garg, Steve Versteeg and Rajkumar Buyya, "SMICloud: A Framework for Comparing and Ranking Cloud Services", Published in *Fourth IEEE International Conference on Utility and Cloud Computing* in 2011.
- [14] A.Iosup, S. Ostermann, N. Yigitbasi, R. Prodan, T. Fahringer and D.Epema, "Performance analysis of Cloud computing services for many-tasks scientific computing", Published in the *IEEE Transactions on Parallel and Distributed Systems* 22 (6) (2011) 931–945.
- [15] J. Saurabh Kumar Garg, Steve Versteeg and Rajkumar Buyya, "A framework for ranking of Cloud computing services", Published in the *ELSEVIER of Future Generation Computer Systems* 29 (2013) 1012–1023.
- [16] (2011) The IEEE website. [Online]. Available: <http://www.ieee.org/>
- [17] Z. Zheng, Y. Zhang, and M.R. Lyu, "CloudRank: A QoS-Driven Component Ranking Framework for Cloud Computing," *Proc. Int'l Symp. Reliable Distributed Systems (SRDS '10)*, pp. 184-193, 2010.
- [18] Z. Zheng, X. Wu, Y. Zhang, M. R. Lyu, and J. Wang, "QoS ranking prediction for cloud services," *IEEE Transactions on Parallel and Distributed Systems*, vol. 99, no. Preprints, p. 1, 2012.
- [19] Xi Chen, Zibin Zheng, Qi Yu and Michael R. Lyu, "Web service recommendation via exploiting Location and QoS information," Published in the *IEEE Transactions on parallel and distributed systems*, Dec 2013.
- [20] H. Ma, I. King, and M. R. Lyu, "Effective missing data prediction for Collaborative filtering," in *Proceedings of the 30th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'07)*, 2007, pp. 39–46.
- [21] R. Buyya, C. Yeo, S. Venugopal, J. Broberg, I. Brandic, "Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility", Published in *ELSEVIER of Future Generation Computer Systems* 25 (6) (2009) 599–616.
- [22] N. Ani Brown Mary and K. Jayapriya, "An Extensive Survey on QoS in Cloud computing", published in the *International Journal of Computer Science and Information Technologies*, Vol. 5 Issue 1- 2014.