

Research Article

Services Recommendation System based on Heterogeneous Network Analysis in Cloud Computing

¹Junping Dong, ¹Qingyu Xiong, ¹Junhao Wen and ²Peng Li

¹School of Software Engineering,

²College of Computer Science, Chongqing University, Chongqing 400030, China

Abstract: Resources are provided mainly in the form of services in cloud computing. In the distribute environment of cloud computing, how to find the needed services efficiently and accurately is the most urgent problem in cloud computing. In cloud computing, services are the intermediary of cloud platform, services are connected by lots of service providers and requesters and construct the complex heterogeneous network. The traditional recommendation systems only consider the functional and non-functional requirements of services but ignore the links between providers and requesters of service, which result to the service position is not accurate. Focus on the problems, this study intends to model the relationship of the cloud services participants with the format of heterogeneous information network, which intend to mine the hidden relationships between services participants in the cloud computing environment. In theoretical research, we proposed a cloud service heterogeneous network extraction and automatic maintenance model, proposed a new service recommendation system based on heterogeneous service network ranking and clustering.

Keywords: Cloud computing, clustering, heterogeneous network analysis, ranking, service recommendation system

INTRODUCTION

Cloud computing provides a new kind of computing model in which computing resources, storage resources, online applications and network infrastructures are provided by the format of services. In cloud computing, the ideas of SaaS (Software as a Service), PaaS (Platform as a Service) and IaaS (Infrastructures as a Service) have come true (Armbrust *et al.*, 2010). With the increase implement of cloud computing based applications, more and more cloud services are developed and applied in real systems, the relations between services, service providers and service requesters have become more and more complex, how to find the most suitable service in cloud computing become the key problem to be solved (Buyya *et al.*, 2008). Recommendation system is one of the best methods which can find the needed suitable items from large number of resources. In the current recommendation system researches, content-based recommendation system, collaborative filtering based recommendation system, hybrid recommendation system and network-based recommendation system are the mainly popular methods.

In a content-based recommendation system, it needs to build configure files for users and items; recommend items based on the similar of configure files. The advantages of this kind of recommendation systems are simple and effective; however, this method is not

available to cloud computing. For one thing, the cloud services are described by WSDL but there is no specific description format for users; for another, the precise of recommendation is not high enough (Pazzani and Billsus, 2007). One approach to the design of recommender systems that has seen wide use is collaborative filtering (Herlocker *et al.*, 2004). Collaborative filtering methods are based on collecting and analyzing a large amount of information on users' behaviors, activities or preferences and predicting what users will like based on their similarity to other users. User-based collaborative filtering attempts to model the social process of asking a friend for a recommendation (Burke, 2002; Zhang and Han, 2007). A key advantage of the collaborative filtering approach is that it does not rely on machine analyzable content and therefore it is capable of accurately recommending complex items such as movies without requiring an "understanding" of the item itself. Recent research has demonstrated that a hybrid approach, combining collaborative filtering and content-based filtering could be more effective in some cases. Hybrid approaches can be implemented in several ways: by making content-based and collaborative-based predictions separately and then combining them; by adding content-based capabilities to a collaborative-based approach (and vice versa); or by unifying the approaches into one model. Several studies empirically compare the performance of the hybrid with the pure collaborative and content-based methods and

demonstrate that the hybrid methods can provide more accurate recommendations than pure approaches (Sarwar *et al.*, 2000). These methods can also be used to overcome some of the common problems in recommender systems such as cold start and the sparsely problem.

The current recommendation methods can't meet the requirements of cloud computing, most of them only consider the functional and non-functional attributes of services but ignore the internships between services. The hidden network relations between services, service providers and service requesters are more important in clouding service platform.

In this study, based on the cloud service features and information network structure, we propose service recommendation system based on heterogeneous service network analysis. The main contributions of our research are shown as follow:

- We propose heterogeneous cloud service network extraction and automatic maintenance model in cloud computing.
- We propose service ranking model based on heterogeneous service network ranking and clustering.
- We propose service recommendation model based on heterogeneous service network analysis.

LITERATURE REVIEW

Among the researches of service recommendation in cloud computing, in study (Han *et al.*, 2009) authors proposed QoS network and virtual machine based cloud service recommendation framework which can quickly discovery and make selection on services. The main idea of QoS network based service ranking is a novel new innovation of service recommendation, however, the ranking method only consider services' inside attributes and ignore the relations between context providers and requesters.

Information retrieval and information network have developed a lot recently, some new researches can help solving the problems in cloud computing. PageRank (Brin and Page, 1998) has been proved to be the most successful ranking method with the rapid development of Google, in which web pages are ranked by the links between them. HITS (Kleinberg, 1999) is another famous ranking method which consider both the authority of pages and the links between pages.

According to the complex structures of nowadays information network, some heterogeneous network based analysis methods are proposed, among them, Rank Clus (Sun *et al.*, 2009a) and Net Clus (Sun *et al.*, 2009b) are two representational researches. In Rank Clus, the authors proposed one new algorithm for information network rank which considers the links between objects in network and gives some ranked clusters results; Net Clus extends Rank Clus to heterogeneous network which supports more than two types of objects in network.

In recommendation system researches, in Xue and Ling (2011), the authors proposed network analysis based recommendation system which imported the network ranking and heterogeneous network analysis to recommendation systems.

In our research, we analysis the heterogeneous network structure of cloud services and propose service network analysis based service ranking and clustering method which makes up the main components of our cloud service recommendation system.

METHODOLOGY

Cloud service network model:

Cloud service network: In our research we call the network of cloud services in cloud computing platform is Cloud Service Network which contains the connected objects of services, service providers and service requesters, the definition of Cloud Service Network is given as Definition 1.

Definition 1: Cloud Service Network (CSN). We define $CSN = \langle CS, E, W \rangle$ as a cloud service network where:

CS is the cloud service node set and $CS = SUPUR$ where S, P and R are objects sets of Services, Providers and Requesters, for convenient, we use lowercase s, p and r to denote the instance objects of three type nodes in CS, uppercase S, P and R denote both the object sets and object types (Services, Providers and Requesters).

E is the edges set of nodes, the edges of CS are the real relations in cloud service context which contains two kinds of relations: "Provide-Provided" relation between P and S types of objects, "Request-Requested" relation between R and S types. For each pair of objects, for example, one service object $s \in S$, another provider object $p \in P$, if service s is provided by p, there is an edge $e = \langle s, p \rangle \in E$, the relations between S and R is similar.

W is the weight set of edges. For the example of E, if service s is provided by p, there is an edge $e = \langle s, p \rangle \in E$, $W(s, p) = 1$, otherwise, $W(s, p) = 0$.

CSN extraction model: How to extract service network from complex service registers and service application context from cloud computing environment is the core work which we focus in this study. In our research, we propose a CSN network Extraction Model to finish the network extraction process which includes: service relation analysis, network building and network mapping sub-processes:

- Service relation analysis process analysis the relations between registered services, service providers and requesters, extract the service description from WSDL
- Network building process mainly to build the topological structure of service network
- Network mapping process focus on the transfer from cloud services to cloud service network, in the

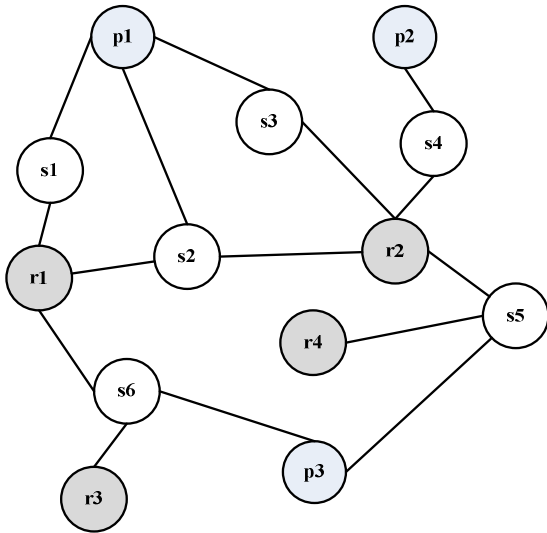


Fig. 1: CSN example

mapped CSN, services are linked with providers and requesters in the form of heterogeneous network

In order to give a more visual description of CSN, now we give a CSN example shown as Fig. 1.

From Fig. 1 we can see, in the CSN example, all the services (s), service providers (p) and service requesters (r) are linked by relations, for each provider can provide several services, every service is provided by one provider and can be called by any number of requesters.

After network extraction of CSN, how to maintain the network is another problem that should be considered in our research. We will introduce our network maintenance model in the next section.

CSN automatic maintenance model: In cloud service environment, the contexts of services are changing from

time to time, so, the mapped CSN network should consider the maintenance according to real context changes.

In order to monitor the real-time changes of services in cloud, we design a component called Service Monitor Centre to be in charge of the job. There are many Local Monitors under the center monitor, the local monitors report service changes to center and the center will do some related update operations on network structure or content.

In our research, in order to improve the effect of maintenance we use cache-based update in our model; the detail of our maintenance model is shown as Fig. 2.

Service ranking model: In cloud computing service network, services and all other partners are linked together by relations. Different with the traditional service ranking, the ranking method we proposed in this study is based on the heterogeneous cloud service network, the service ranking and clustering are running contemporaneously, the ranking results are the input of clustering and the new clustering results are the input of new ranking, all of the process is based on iteration.

Ranking probability model: The ranking algorithm we use here is based on the relations between CSN objects, for a service object, if more high ranked providers and requesters linked to it, the service object will have a more high rank.

In a $CSN = \langle CS, E, W \rangle$ network with Service (S), Provider (P) and Requester (R) types of objects, we use probability distribution $P(s)$ to denote the rank of one object s in a service space, the ranking function of service object is defined as Eq. (1):

$$P(s) = \prod_{x \in N(s)} P(x)W(s, x) \tag{1}$$

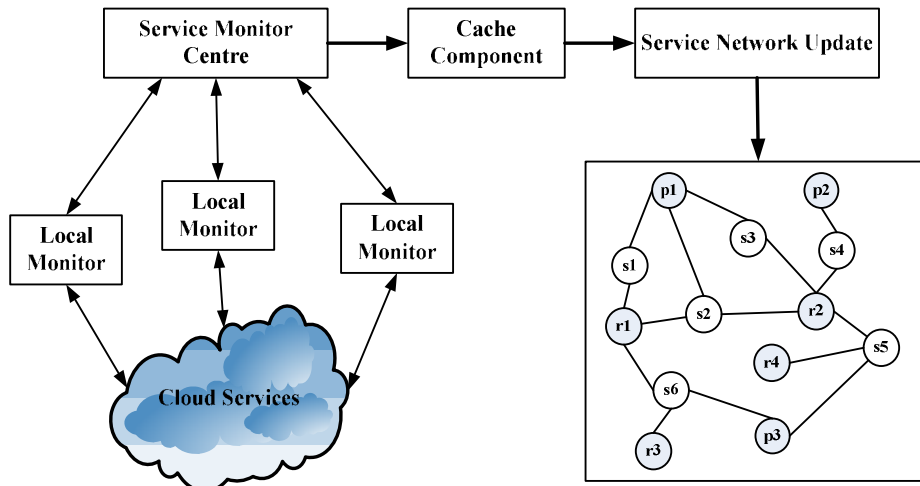


Fig. 2: CSN maintenance model

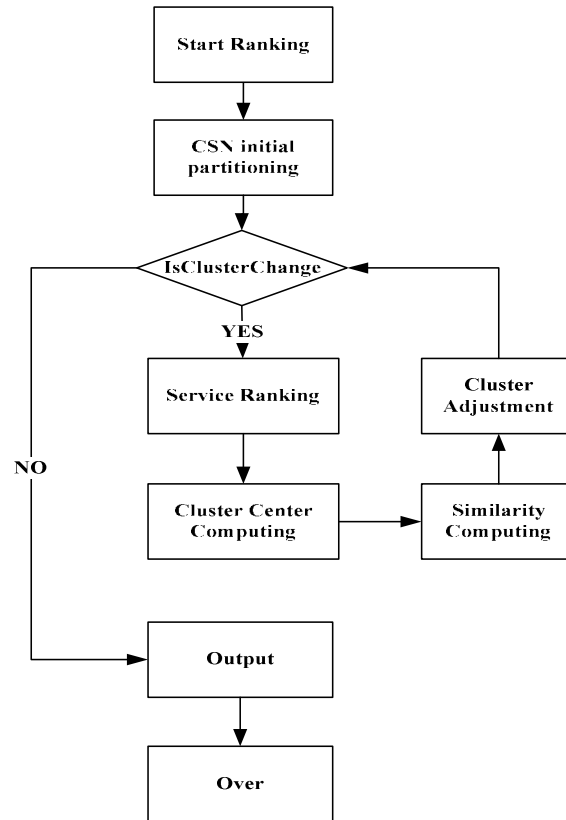


Fig. 3: Ranking with clustering

where,

$N(s)$ = The linked providers and requesters with service s

$P(x)$ = The rank score of the linked object

$W(s, x)$ = The weight of the link

Ranking based on clustering: In our service ranking process, we use clustering to get several ranked clusters. In this study, based on the probability distribution of services clustering results are described, that is to say, the service here according to the probabilistic numerical description, may belong to any cluster; therefore, the service cluster belongs to soft clustering. In a specific cluster study, service similarity, clustering convergence speed and number of clustering decisions are key problems to be researched. The main framework is shown as Fig. 3.

From Fig. 3 we can see, the ranking and clustering process includes sub-process as follows:

1. Initial partitioning on the input network CSN to K clusters
2. Judge the clustering stop condition, if the clusters are not changing or meet the maximum iteration times, if YES, continue, or jump to END
3. Ranking the service based on service ranking function

4. Computing the current cluster centers based on the ranking results
5. Computing the similarity between services and every center to get the nearest distance
6. Adjusting the clustering results based on the similarity computing
7. Repeat (2)-(6) until clustering stop
8. Output the K ranked clusters result and end the ranking process

SERVICE RECOMMENDATION SYSTEM FRAMEWORK

According to the research above, we propose service network based service recommendation system in cloud computing which contains network extraction component, network ranking and clustering component; the framework is shown as Fig. 4.

In Fig. 4, the recommendation system contains three important components: Service Network Extraction, Service Network Ranking and Clustering, the ranked clusters are recommended to users.

Compare to the traditional service recommendation system, we focused on the cloud service and proposed network based ranking and clustering method to make a better recommendation for cloud services.

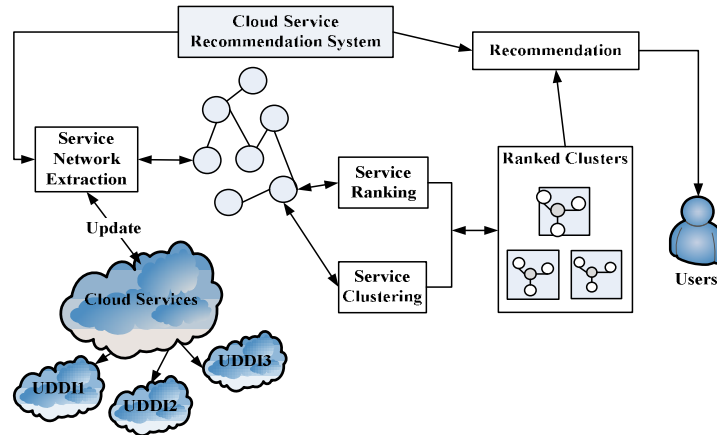


Fig. 4: Service recommendation system framework

CONCLUSION

In order to enhance the re-useable of services in cloud computing environment, we proposed heterogeneous service network based ranking and clustering model, based on the network analysis model, we proposed a novel service recommendation system which extracting the potential network structure, ranking and clustering on the cloud service network. The recommendation system provides users the ranked clusters which are conveniently for user selection. The research we proposed now only focus on the high level framework, the detail of algorithms and experiments will be finished in the later study. In future, we will find real cloud services to validate our recommendation method and development a prototype system for it.

ACKNOWLEDGMENT

This study is supported by Natural Science Foundation Project No. 61075053 and supported by Natural Science Foundation Project of CQ CSTC2010BB2244 and supported by the Fundamental Research Funds for the Central Universities Project No. CDJZR10090001.

REFERENCES

Armbrust, M., A. Fox and R. Griffith, 2010. A view of cloud computing. *Commun. ACM*, 53(4): 50-58.
 Brin, S. and L. Page, 1998. The anatomy of a large-scale hypertextual web search engine. *Comput. Networks ISDN*, 30(1998): 107-117.
 Burke, R., 2002. Hybrid recommender systems: Survey and experiments. *User Model. User-Adap.*, 12(4): 331-370.
 Buyya, R., C.S. Yeo and S. Venugopal, 2008. Market-oriented cloud computing: Vision, hype and reality for delivering IT services as computing utilities. *Proceeding of the 10th IEEE International Conference on High Performance Computing and Communications (HPCC)*, pp: 5-13.

Han, S.M., M. Mehedi and C.W. Yoon, 2009. Efficient service recommendation system for cloud computing market. *Proceeding of the 2nd International Conference on Interaction Sciences*, pp: 839-845.
 Herlocker, J.L., J.A. Konstan and L.G. Terveen, 2004. Evaluating collaborative filtering recommender systems. *ACM T. Inform. Syst.*, 22(1): 5-53.
 Kleinberg, J.M., 1999. Authoritative sources in a hyperlinked environment. *J. ACM*, 46(5): 604-632.
 Pazzani, M. and D. Billsus, 2007. Content-based recommendation systems. *Lect. Notes Comput. Sc.*, 432: 325-341.
 Sarwar, B., G. Karypis and J. Konstan, 2000. Analysis of recommendation algorithms for e-commerce. *Proceeding of the 2nd ACM Conference on Electronic Commerce*, pp: 158-167.
 Sun, Y., J. Han and P. Zhao, 2009a. RankClus: Integrating clustering with ranking for heterogeneous information network analysis. *Proceeding of the 12th International Conference on Extending Database Technology: Advances in Database Technology*, pp: 565-576.
 Sun, Y.Z., Y.T. Yu and J.W. Han, 2009b. Ranking-based clustering of heterogeneous information networks with star network schema. *Proceeding of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp: 97-805.
 Xue, L. and C. Ling, 2011. Recommendations based on network analysis. *Proceeding of International Conference on Advanced Computer Science and Information Systems*, pp: 9-15.
 Zhang, C. and Y. Han, 2007. Service recommendation with adaptive user interests modeling. *Lect. Notes Comput. Sc.*, 4882: 265-270.