

Research Article

CC-PSM: A Preference-Aware Selection Model for Cloud Service Based on Consumer Community

Yan Wang,¹ Jian-tao Zhou,¹ and Hong-yan Tan²

¹College of Computer Science, Inner Mongolia University, Hohhot, Inner Mongolia 010021, China

²Institute of Information Engineering, Chinese Academy of Sciences, State Key Laboratory of Information Security, Beijing 100093, China

Correspondence should be addressed to Jian-tao Zhou; cszjtao@imu.edu.cn

Received 28 May 2015; Accepted 16 July 2015

Academic Editor: Krishnaiyan Thulasiraman

Copyright © 2015 Yan Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to give full consideration to the consumer's personal preference in cloud service selection strategies and improve the credibility of service prediction, a preference-aware cloud service selection model based on consumer community (CC-PSM) is presented in this work. The objective of CC-PSM is to select a service meeting a target consumer's demands and preference. Firstly, the correlation between cloud consumers from a bipartite network for service selection is mined to compute the preference similarity between them. Secondly, an improved hierarchical clustering algorithm is designed to discover the consumer community with similar preferences so as to form the trusted groups for service recommendation. In the clustering process, a quantization function called community degree is given to evaluate the quality of community structure. Thirdly, a prediction model based on consumer community is built to predict a consumer's evaluation on an unknown service. The experimental results show that CC-PSM can effectively partition the consumers based on their preferences and has good effectiveness in service selection applications.

1. Introduction

Cloud services, which take resource virtualization as the core technology and on-demand resource provisioning as the main characteristics, provide strong processing capacity to complete the users' all kinds of applications, while saving the cost of software installation and maintenance. With the development of cloud computing technology, more and more IT enterprises package their existing applications to cloud services for the users' consumptions, which form a pool with mass services. In the cloud service market, a three-tier service framework is formed, including cloud platform, service provider, and consumer [1, 2]. A service provider provides the consumers with services by renting virtual resources of the cloud platform. Different service providers can adopt different technologies, so they will provide a number of services with different quality. Facing the competing services with the same or similar function, service selection has become a core question concerned by the cloud consumers [3].

How to make a consumer select a satisfactory service from the service pool so as to improve the customer's

loyalty and *attract* more consumers is significant for the cloud platform [4]. Therefore, more and more scholars have attached great importance to the study on the methods of cloud service selection. At present, main research directions of cloud service selection can be grossly divided into two categories. One is the service selection method based on objective QoS, which transforms the problem of service selection to the problem of QoS calculation, and then recommends an optimal service or service composition to a consumer by building a mathematical model based on the service's QoS and the consumer's preference [5–9]. The other is the service selection method based on subjective recommendation, which regards service as “commodity,” and then copies reputation mechanism based on user feedbacks in the electronic commerce to solve the problem of service selection [10–14].

Both of the above categories have their virtues and faults. On the one hand, the service selection method based on objective QoS provides a prediction of the consumer's satisfaction to an unknown service. But the premise of such a situation is that all the information, including the service's QoS

and the consumer's preference, needs to be clearly described. In fact, it is very difficult for the consumer to accurately describe his/her preference. Besides that, due to the economic interests, the service provider may give a false QoS description. On the other hand, the service selection method based on subjective recommendation builds a model for selecting a credible service through reputation mechanism, but it still needs to distinguish the consumers with different preferences. If not, it will make a wrong recommendation. Then how to find the trusted users for recommending services has become a key problem of the method. The traditional way is to choose the nearest N neighbors, such as collaborative filtering (CF). But it needs to search the similar users at each round, which is unfavorable for the efficiency and stability of recommender system. Because of the consumers' inherent characteristics, such as the differences in knowledge structure, professional requirements, and social factors, the consumers' preferences would show some group features. It is a natural form of some service preference-oriented consumer communities in the cloud. If a stable group structure based on the consumers' preferences is found, we can recommend services through the group structure. For the service requester, the recommendations from the consumers similar with him are trustworthy [14].

Based on the idea, a service selection model based on community discovery for consumer preference is proposed, which breaks the limitation of traditional service selection methods. It does not depend on the description of the services' QoS parameters and the consumers' preferences but helps to improve the efficiency and stability of service recommendation. Our main ideas and contributions are as follows:

- (1) The concept of consumer preference similarity is proposed to identify the differences of the consumers' preferences.
- (2) The consumer community structure is discovered based on consumer preference similarity, which contains some stable groups for service recommendation.
- (3) A novel service selection model based on consumer community is presented, which explores a method of predicting the consumers' evaluations through the stable and subjective information.

We conducted a lot of experiments with the simulated data and the data sets from Epinions.com, and the results showed the effectiveness of our work. In addition, we obtained many interesting and useful findings from the experiments.

The remainder of the paper is organized as follows. In Section 2, the related work is introduced and discussed. In Section 3, a community discovery model based on consumer preference is presented, and a consumer community discovery algorithm for the model is described. Then in Section 4, a service selection model based on consumer community (CC-PSM) is presented, and the corresponding algorithm is given, while, in Section 5, the experimental results are reported and analyzed to evaluate the effectiveness and feasibility of CC-PSM. Finally, the paper is concluded and possible directions of future work are highlighted in Section 6.

2. Related Work

In the cloud environment, one of main obstacles to choose a cloud service for the consumer is the diversity and agility of cloud service, which leads to the difficulty of comparing different cloud services. To solve the problem, some scholars are devoted to the study on the parameter matching approach between services and consumers. Ding et al. [5] designed a service recommendation method based on the historical data of the service performance, which can regulate multi-attribute matching between provider solutions and customer demands. However the consumer's demand description and the service provider's performance description often have no unified standard. For this, Karim et al. [15] explored a mechanism to map the users' QoS requirements to the right QoS specifications of SaaS and then proposed a set of rules to perform the mapping process.

Godse and Mulik [16] thought that cloud service selection was a multicriteria decision-making problem (MCDM), so they made use of Analytic Hierarchy Process (AHP) technique for prioritizing the services features and also for expert-led scoring of the services. In [6], the authors gave full consideration to different criteria of the cloud services and different demands of the consumers and proposed a multicriteria cloud service selection methodology. Zhao et al. [7] proposed an innovative service selection algorithm based on the service's response time, trust degree, and monetary cost to find the appropriate services with satisfying the users' multiple QoS requirements. Nevertheless, in order to improve the accuracy and efficiency of MCDM and reduce the uncertainty of decision problem, it must ensure that the cloud service provider not only offers real description of the service, but never changes the service parameters during the period of decision-making, which is unfavorable to resist malicious or low-quality service providers in the cloud environment. So a cloud service selection method based on parallel multicriteria decision analysis was put forward to rank all cloud services in different time period in [8].

In addition, due to the fuzziness and uncertainty of user preference, some scholars have studied the quantification of user preference. In [17], considering the uncertainty of user subjective and objective weight preferences, the authors obtained the user subjective weight by intuitionistic fuzzy set and objective weight by attribute significance of rough set. Finally, the uncertain user QoS preference-aware cloud service selection was transformed to a multiple attribute decision-making problem to select a best service for the user. In [18], the authors transformed a consumer's qualitative and semiquantitative personalized preferences into quantitative numeric weights based on AHP.

To improve the consumer's satisfaction is the ultimate goal of service selection method. In [9], the authors focused on QoS-satisfied predictions about the composition of cloud service components and presented a QoS-satisfied prediction model based on a hidden Markov model. Zheng et al. [10] pointed out that a pure QoS prediction was not enough for the consumer to reflect accurate order of the candidate services. So they designed a personalized QoS ranking prediction framework. The framework discovers the users with similar

preference through comparing the users' ranking similarity of the commonly invoked services and then makes QoS ranking prediction of a set of cloud services for the target user based on the similar users' ranking with him.

To eliminate a sharp trust crisis between cloud consumers and cloud service providers, trust is introduced in the process of service selection. In [19], a novel cloud service-composition method based on the trust span tree was proposed. Through the credible relationship evolution, the trust union of service providers with same or similar function is formed, which helps to exclude uncertain or malicious service from the trust span tree. But trust is a subjective cognition judgment. Therefore, the trust evaluation needs not only objective measurement but also subjective perception. Ding et al. [11] designed a novel framework named CSTrust for conducting cloud service trustworthiness evaluation by combining QoS prediction and customer satisfaction estimation. To reduce the bias caused by unreasonable feedback from unprofessional or malicious cloud users, a method was proposed for filtering the feedback from such users in [12]. After processing, the aggregated result can quantitatively reflect the overall quality of a cloud service.

Whether a service is credible to satisfy the consumer's personality preference or not is the main consideration in [20]. This paper utilized fuzzy clustering technology to classify services by the requesters' preferences and gave a service selection algorithm to select a closest classification with the requester's preference. With the integration of trust evaluation method and CF technique, Wang and Zhang [13] presented a trustworthy service selection model based on collaborative filtering. This model introduces consumer correlation to embody the impact of the requester's personal characteristics on selection process, computes creditability of recommendation, and employs analytic hierarchy process to decide the weight of each factor in service reputation. Abedinzadeh and Sadaoui [14] presented ScubAA, a novel generic agent trust management framework based on the theory of human plausible reasoning. ScubAA recommends to the user a list of the most trusted services in terms of a single personalized value derived from several types of evidences such as user's feedback, history of user's interactions, context of the submitted request, references from third party users as well as from third party service agents, and structure of the society of agents.

As seen from the above literatures, selecting the appropriate recommendation users is the key to the service selection based on trust. In [21], an innovative idea for selecting the reliable recommendation users was explored. The authors thought that sparsity, cold-start and trustworthiness were major issues challenging service recommendation in adopting similarity-based approaches. With the prevalence of social networks, to a certain extent, the user's characteristics and preference were exposed from the data in blogs and social-networking sites. So a social network-based service recommendation method with trust enhancement was proposed in the paper. The method assesses the degree of trust between users in social network by a matrix factorization, and then recommendation results are obtained by an extended random walk algorithm.

The above literatures used different technologies to realize an intelligent service selection for consumers from different angles, including quantifying the consumers' demands or preferences, predicting the quality of service, matching the consumer's demands and the service's performance, and recommending the services based on trustworthy consumers. Different methods have different merits and limitations. Firstly, the method based on the parameters matching between the consumer's demands and the service's performance is relatively simple and intuitive. But this type of the method mostly requires the consumers to explicitly give his/her QoS demands, which is too difficult for consumers to provide these parameters. So its feasibility is poor in practice. Secondly, the method by predicting the quality of service only considers the general performance of the service but ignores the differences of QoS demands. Thirdly, the fuzziness and uncertainty of user preference imply that it is a difficult thing to mine the users' preferences. The methods of quantifying the consumers' preferences are still under study. In addition, there are still no efficient ways to validate the quantitative results. Finally, the method based on trust recommendation is now widely accepted in the field of electronic commerce and social networking service. Thus the method has good application prospects in terms of cloud services selection. The key problem of the method is the choice of trustworthy customers for service recommendation. But the existing methods for selecting the trustworthy customers are mostly unstable and inefficient, which affect the accuracy of the selection results and the feasibility of the selection process.

From what has been discussed above, we can see clearly that the core problem of service selection is to find a feasible method to identify different consumers' preferences. To resolve the problem, a cloud service selection model based on simple evaluation information and the idea of the trust in community network will be built in the paper. At first, we need to consider how to mine the consumers' preferences based on available information in the cloud so as to improve the feasibility of the method in practical application. Inspired by the literature [21], a consumer is a mapping for a real person in real society, whose behavior and preference are relatively stable. Due to different profession and background, the preferences of the consumers would show some group features. So we can cluster the consumers with similar preference to form a stable community, which is the basis of service selection. According to the theory of human plausible reasoning [14], the consumer would trust the consumers who are similar with him/her more than the others. Then we design a service selection model based on community trust.

3. Community Discovery Model Based on Consumer Preference

Predicting a consumer's evaluation on an unknown service mainly relies on the recommendations from other consumers. So whether other consumers' evaluations on a target service are worthy to trust or not will have a large impact on the accuracy of prediction. Therefore it is a key step to find the trusted recommendation users in the process of service

selection. Much large-scale distributed complex system can be described in the form of a complex network, in which some modules or communities can be observed. A community may be a group of nodes based on some certain concept such as node similarity. Community structure can serve as a bridge between a single node and the macro network system. Although there is no direct interaction or contact between cloud consumers, a complex network based on the inner relationships between consumers is still formed through their selection and evaluation for cloud services. The network also tends to be obvious characteristics, community structure, as other complex networks. The study of consumer community structure in the cloud, which shows consumer groups with common features, has the important theoretical significance for analysis and prediction of consumer behavior. Based on the idea of community discovery in the complex network, this paper discovers consumer community based on preference similarity between consumers. In the community structure, the users in one community are similar in preference. In this section, we will model the process of community discovery through the following definitions and give an algorithm to perform the model.

The data foundation of our study is the service evaluation information in cloud, which includes three parts: service, consumer, and service evaluation. If a consumer buys a service according to his/her business requirements, he/she will give a comprehensive evaluation to the service at response time, security, reliability, availability, and other performances through his/her service experience, which is called service evaluation. This paper sets service evaluation on a scale of 1 to 5. The value 1 is for a very poor service and 5 for a completely satisfactory service. Assume that each entity is abstracted as a node; then a concept of bipartite network for service selection is given as follows.

Definition 1 (bipartite network for service selection). A bipartite network for service selection, W , is a three-tuple, $W = (U, S, G)$.

- (1) $U = \{u_1, u_2, \dots, u_n\}$ is a finite set of consumers, $n > 0$.
- (2) $S = \{s_1, s_2, \dots, s_m\}$ is a finite set of services, $m > 0$.
- (3) The set of consumers and the set of services are disjoint, $U \cap S = \emptyset$.
- (4) $G = \{(u_x, s_i, G_{xi}) \mid u_x \in U \wedge s_i \in S \wedge 1 \leq G_{xi} \leq 5\}$ is a mapping from consumers to services. $\forall u_x \in U, s_i \in S$, if u_x has bought s_i , there exists an edge between u_x and s_i , whose weight is G_{xi} . G_{xi} represents the evaluation of u_i on s_j .

3.1. Consumer Preference Similarity. In current research results, there exist a lot of similarity measure methods based on network structure, such as CN, Salton, and Jaccard [22]. It can be seen from these results that common neighbor is an important indicator to describe the similarity between nodes. Obviously, the consumers who are interested in similar business tend to select the same service, so their similarity is much higher. However, it is not enough to judge the similarity between consumers only based on the same services chosen

by them. Due to losing their evaluations on the services, it is likely to cause misjudgments. In this paper, a factor of evaluation difference [13] is introduced to compute the preference similarity between consumers. The related definitions are given as follows.

Definition 2 (service selection similarity). Let $N(u_x) = \{s_i \mid \exists (u_x, s_i, G_{xi}) \in G\}$ be a set of neighbor nodes of u_x . $\forall u_x, u_y \in U, u_x \neq u_y$, the service selection similarity between u_x and u_y is a function of $N(u_x)$ and $N(u_y)$, which is defined by

$$\begin{aligned} \text{Sim}_s(x, y) &= \begin{cases} \frac{|N(u_x) \cap N(u_y)|}{\min\{|N(u_x)|, |N(u_y)|\}} & N(u_x) \neq \emptyset \wedge N(u_y) \neq \emptyset \\ 0 & \text{otherwise,} \end{cases} \quad (1) \end{aligned}$$

where $|N(u_x)|$ is for the number of elements in $N(u_x)$; $|N(u_x) \cap N(u_y)|$ for the number of elements in the intersection of $N(u_x)$ and $N(u_y)$.

Definition 3 (service evaluation difference). $\forall u_x, u_y \in U, u_x \neq u_y$, the service evaluation difference between u_x and u_y is a function of $N(u_x)$, $N(u_y)$, and G , which is defined by

$$\begin{aligned} \text{GD}(x, y) &= \begin{cases} \frac{\sum_{s_i \in N(u_x) \cap N(u_y)} [G_{xi} - G_{yi}]}{|N(u_x) \cap N(u_y)| * \alpha} & |N(u_x) \cap N(u_y)| \neq 0 \\ 0 & \text{otherwise,} \end{cases} \quad (2) \end{aligned}$$

where $[G_{xi} - G_{yi}]$ is for the absolute difference of the evaluations on s_i between u_x and u_y ; α for the normalizing factor (α is set to 4 because it is the maximum of the evaluation difference between consumers).

Definition 4 (consumer preference similarity). $\forall u_x, u_y \in U, u_x \neq u_y$, the consumer preference similarity between u_x and u_y is a function of $\text{Sim}_s(x, y)$ and $\text{GD}(x, y)$, which is defined by

$$\text{Sim}_p(x, y) = \frac{\text{Sim}_s(x, y)}{e^{\text{GD}(x, y)}}. \quad (3)$$

$\text{Sim}_p(x, y)$ is directly proportional to $\text{Sim}_s(x, y)$ and inversely proportional to $\text{GD}(x, y)$. When $\text{Sim}_s(x, y) = 0$, $\text{Sim}_p(x, y) = 0$; that is, both of u_x and u_y are completely different in service preference. When $\text{Sim}_s(x, y) \neq 0$ but $\text{GD}(x, y) = 0$, $\text{Sim}_p(x, y)$ lies completely with $\text{Sim}_s(x, y)$. When $\text{Sim}_s(x, y) \neq 0$ and $\text{GD}(x, y) \neq 0$, $\text{Sim}_p(x, y)$ is decided by $\text{Sim}_s(x, y)$ and $\text{GD}(x, y)$.

3.2. Community Discovery. The general idea of community discovery based on network structure is to translate the problem of community discovery into a clustering problem by using node similarity in the network. Regarding the center of cluster as an abstract node, the following part will present the definitions of similarity between node and cluster or between clusters in the clustering process.

Definition 5 (similarity between node and cluster). $\forall u_x \in U, \forall C_y \subseteq U, C_y \neq \emptyset \wedge u_x \notin C_y$, the similarity between u_x and C_y is defined by

$$D_{\text{sim}}(u_x, C_y) = \frac{\sum_{\forall u_i \in C_y} \text{Sim}_p(x, i)}{|C_y|}, \quad (4)$$

where $|C_y|$ is for the number of elements in C_y .

Definition 6 (similarity between clusters). $\forall C_x, C_y \subseteq U, C_x \neq \emptyset \wedge C_y \neq \emptyset \wedge C_x \cap C_y = \emptyset$, the similarity between C_x and C_y is defined by

$$D_{\text{sim}}(C_x, C_y) = \frac{\sum_{\forall u_i \in C_x, \forall u_j \in C_y} \text{Sim}_p(i, j)}{|C_x| * |C_y|}. \quad (5)$$

A measurement index is needed to evaluate how well the consumer community structure matches the golden standard groups. So an evaluation function for community structure, as the fitness function of the optimization process, is designed to compare the quality of different community partitions and validate different community discovery algorithms. The mostly used evaluation function for community structure, Modularity Q [23], does not apply to evaluate the quality of consumer community structure in the cloud because of only considering node degree, not edge weight. This paper built a quantization function suitable for consumer community structure in the cloud, called community degree. The index is used to look for a rigid network structure with strong ties in a highly dynamic environment. The related definitions of community degree are given as follows.

Definition 7 (community cohesion). $\forall C_k \subseteq U, C_k \neq \emptyset$, the community cohesion of C_k is defined by

$$AS_{\text{in}}(k) = \begin{cases} 0 & |C_k| = 1 \\ \frac{\sum_{\forall u_x, u_y \in C_k} \text{Sim}_p(x, y)}{|G_{\text{in}}|} & \text{otherwise.} \end{cases} \quad (6)$$

Of these, $|G_{\text{in}}|$ is the total number of pairs of consumers in C_k .

Definition 8 (community separation). $\forall C_k \subseteq U, C_k \neq \emptyset$, the community separation of C_k is defined by

$$AS_{\text{out}}(k) = \begin{cases} AS_{\text{in}}(k) & |C_k| = n \\ \frac{\sum_{\forall u_x \in C_k, \forall u_y \notin C_k} \text{Sim}_p(x, y)}{|G_{\text{out}}|} & \text{otherwise.} \end{cases} \quad (7)$$

Of these, $|G_{\text{out}}|$ is the total number of pairs of consumers, respectively, from inside and outside C_k .

The definition reflects the interdependence between the nodes in a community and other nodes outside it. For each C_k , the higher its community cohesion is and the lower its community separation is, the stronger its independence is. In other words, a community is a group, in which the nodes have strong ties. The definition of community independence is given by normalization as follows.

Definition 9 (community independence). $\forall C_k \subseteq U, C_k \neq \emptyset$, the community independence of C_k is defined by

$$CI(k) = \begin{cases} 0 & |C_k| = 1 \\ 1 - \frac{AS_{\text{out}}(k)}{AS_{\text{in}}(k)} & \text{otherwise.} \end{cases} \quad (8)$$

The closer the value is to 1, the better the independence of the community is.

Definition 10 (community structure). A structure C is called a community structure, if and only if C satisfies the following conditions:

- (1) $C = \{C_1, C_2, \dots, C_c\}$ ($1 \leq c \leq n$);
- (2) $C_1 \subseteq U \wedge C_2 \subseteq U \wedge \dots \wedge C_c \subseteq U$;
- (3) $\forall C_i, C_i \neq \emptyset$;
- (4) $\forall C_i, C_j, C_i \cap C_j = \emptyset$;
- (5) $C_1 \cup C_2 \cup \dots \cup C_c = U$.

Definition 11 (community degree). For a community structure C , the community degree of C is defined by

$$M(C) = \frac{1}{c} \sum_{k=1}^c CI(k) * \frac{1}{c} \sum_{k=1}^c AS_{\text{in}}(k) * \ln \left(\sqrt{\frac{U_n}{c}} \right). \quad (9)$$

Of these, c is the number of communities in community structure C ; U_n is the number of consumers.

The definition is a typical objective function in clustering to formalize the goal of attaining high intracommunity similarity (consumers within the same community are similar) and low intercommunity similarity (consumers from different communities are dissimilar). The community degree of a community structure C is inevitably affected by the community independence and community cohesion of every community in C . Besides that, after lots of trials and errors, the community size should also be considered as a factor. Introducing the community size can help to avoid generating too much small communities. For the different community structures $C_{(1)}$ and $C_{(2)}$ of W , if $M(C_{(1)}) > M(C_{(2)})$, it shows that $C_{(1)}$ is superior to $C_{(2)}$. Typically, after calculating $M(C)$ of every community structure in the clustering process, the community structure corresponding to its peak will be optimal that represents the best division for the nodes in U .

3.3. Algorithm Description. Through the definitions and equations above, we have completed the modeling for the whole process of consumer community discovery in the cloud. Then we will give an algorithm for the implementation of the model. There are many clustering methods based on node similarity, including partitioning cluster, hierarchy cluster, and density-based cluster. Because it is quite difficult to estimate clustering radius or observe some proper cluster centers in the cloud, this paper adopts a hierarchy cluster algorithm to discover consumer community based on consumer preference similarity. Three primary reasons for this are as follows: (1) In the cloud, there may be some consumer communities with different size and even complex shape.

Input: W (bipartite network for service selection),
 $(u_x, u_y, \text{Sim}_p(x, y))$ (triple collection of consumer preference similarity)
Output: $C_M = \{C_1, C_2, \dots, C_c\}$ (community structure which Maximize $M(C)$)
(1) Initialize community structure in W , let $C = \{C_1, C_2, \dots, C_n\}$, where $C_i = \{u_i\}$, n for the number of consumers.
(2) Initialize similarity between clusters to output a triple collection $(C_x, C_y, D_{\text{sim}}(C_x, C_y))$, where
 $(C_x, C_y, D_{\text{sim}}(C_x, C_y)) = (u_x, u_y, \text{Sim}_p(x, y))$.
(3) Select one or more pairs of clusters from the triple collection $(C_x, C_y, D_{\text{sim}}(C_x, C_y))$, whose similarity satisfies the conditions of multi-step merge or transitive merge in this round, and merge them by making use of the two strategies.
(4) Update partition C , $C = \frac{C}{\{C_{i1}\}\{C_{j1}\}} \cdots \{C_{ik}\}\{C_{jk}\} \cup \{C_{i1} \cup C_{j1}\} \cdots \{C_{ik} \cup C_{jk}\}$, where k for the number of merged pairs of clusters in this round.
(5) Update the similarities of these merged clusters and other clusters.
(6) Calculate community degree $M(C)$.
(7) When $|C| > 1$, repeat Step 2.
(8) Form a clustering tree, $C_{(\text{tree})} = \{C_{(1)}, C_{(2)}, \dots, C_{(n')}\}$
(9) Select community structure $C_{(M)}$ which maximizes $M(C)$ from $C_{(\text{tree})}$, $C_{(M)} = \{C_{(i)} \mid C_{(i)} \in C_{(\text{tree})} \wedge M(C_{(i)}) = \text{Max}_{C \in C_{(\text{tree})}}(M(C))\}$.

ALGORITHM 1: Consumer community discovery algorithm.

The hierarchy clustering method has stronger recognition ability in this respect [22]. (2) For a completely unknown environment, the hierarchy clustering method need not estimate the number of clusters and clustering radius in advance, which can control different levels of clustering size flexibly. (3) The hierarchy clustering method is not dependent on the choice of empirical parameters and not sensitive to noise data, so it can effectively filter isolated point.

In this paper, an improved agglomerative hierarchical clustering method is proposed, which maps preference similarity to Euclidean distance, and then iteratively merges the most similar pairs until meeting some preset ultimate condition. According to the definitions and equations above, once consumer preference similarities in bipartite network are provided, the clusters can be iteratively merged by the hierarchical cluster until all the consumers are into one community. The whole process forms a clustering tree, in which each layer corresponds to a community structure. The community structure C , which maximizes $M(C)$, is the most optimal community structure in the cloud.

In order to improve the efficiency and accuracy of the traditional hierarchical clustering methods, the algorithm introduces multistep merge and transitive merge. In the multistep merge strategy, there may be one or more pairs of clusters to be merged during each round. A threshold for merge deviation ε is set up. For every round in the clustering process, select multiple pairs of clusters, whose similarity is not lower than the difference between the highest similarity between clusters and ε , and then merge them separately. The task of transitive merge is that when a consumer and other multiple consumers have the same preference similarity, such as $\text{Sim}_p(1, 2) = 0.8$, $\text{Sim}_p(2, 3) = 0.8$, merge them into one cluster. A detailed description of the algorithm is given by Algorithm 1.

4. Service Selection Model Based on Consumer Community

According to human psychological cognitive habits, a consumer's trust in other consumers who are in the same

community with him is higher than the consumers who are in other communities. And a consumer's trust in the consumers from different community has a certain difference because of different similarity between communities. The formal definition of community trust is given below.

Definition 12 (community trust). The symbol $T(x, y)$ represents that how much consumer u_x trusts the service evaluation of consumer u_y . Let $u_i \in C(u_i)$; the following property is satisfied:

$$\forall u_x, u_y, u_z \in U : (C(u_x) = C(u_y) \wedge C(u_x) \neq C(u_z)) \implies T(x, y) > T(x, z). \quad (10)$$

In this section, we will model the process of community trust-driven service selection and give an algorithm for the model.

4.1. Prediction of Service Evaluation. The consumer community structure forms a stable partition of the consumers in the cloud, which embodies different consumer groups with similar preferences. If the service has been chosen and evaluated by a certain amount of consumers who are in the same community with the requester, the requester's evaluation on the service can be accurately predicted only through the evaluations of the consumers in the same community with him. However, if the bipartite network for service selection is sparse and a predicted service is rarely or has never been chosen by the consumers in the same community with the requester, it is not enough to predict his/her evaluation on the service only through the recommendations from the consumers in his/her community. Thus we need some evaluations from a wider range of the consumers who have interacted with the target service. When predicting the requester's evaluation on unknown through the consumers in other communities, the similarities between the requester and other communities are regarded as the weights of service evaluations from these communities. For a new cloud service, an initial evaluation, which is less than half of the highest

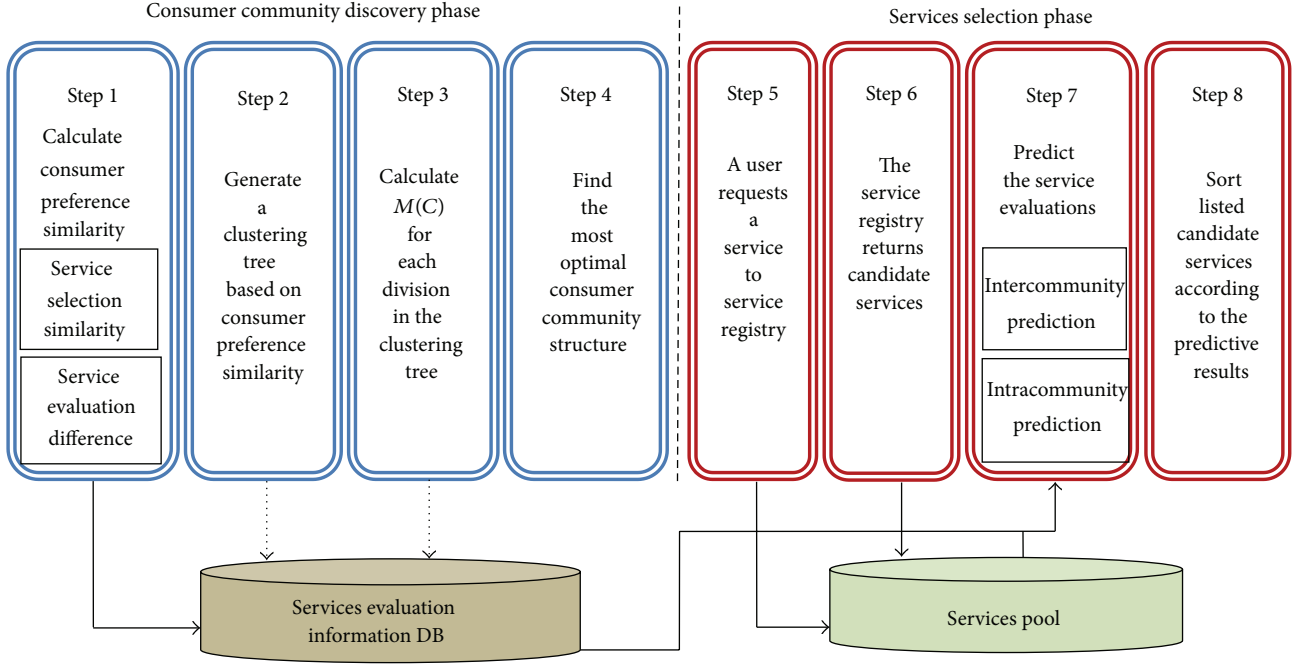


FIGURE 1: Service selection process driven by community trust.

evaluation of this kind of service, is assigned. This is because higher initial value can lead to the attacks of malicious service nodes, while lower initial value can make the service never be chosen. The definitions involved in the process are as follows.

Definition 13 (community public service evaluation). $\forall C_x \in U, \forall s_i \in S$, the community public service evaluation of community C_x on service s_i is defined by

$$CG_{xi} = \begin{cases} \frac{1}{|C_{xi}|} \sum_{u_y \in C_x} G_{yi} & (u_y, s_i, G_{yi}) \in G \\ V_{int} & \text{otherwise,} \end{cases} \quad (11)$$

where $C_{xi} = \{u_y \mid u_y \in C_x \wedge (u_y, s_i, G_{yi}) \in G\}$ is for the consumers in C_x who have interacted with service s_i , $|C_{xi}|$ for the number of the consumers in C_{xi} , and V_{int} for the initial evaluation.

Definition 14 (intracommunity predictive evaluation). $\forall u_x \in U, \forall s_i \in S$, the intracommunity predictive evaluation of consumer u_x on service s_i is defined by

$$CP(x, i) = CG_{yi} \quad (u_x \in C_y). \quad (12)$$

Definition 15 (intercommunity predictive evaluation). Let $C = \{C_1, C_2, \dots, C_c\}$ be a community structure. $\forall u_x \in U, \forall s_i \in S$, the intercommunity predictive evaluation of consumer u_x on service s_i is defined by

$$EP(x, i) = \sum_{y=1}^c w_y * CG_{yi}, \quad (13)$$

where $w_y = D_{sim}(x, C_y) / \sum_{z=1}^c D_{sim}(x, C_z)$, which is normalized similarity weight between consumer u_x and community C_y .

In order to determine when to use the intracommunity prediction or when to use the intercommunity prediction, the definition of predictive decision parameter is given. If the parameter is greater than a given threshold value, we use the intracommunity prediction; otherwise, we use the intercommunity prediction.

Definition 16 (predictive decision parameter). $\forall u_x \in U, \forall s_i \in S$, the predictive decision parameter of user u_x to service s_i is defined by

$$PD_{xi} = \frac{|N(s_i) \cap C_y|}{|N(s_i)|} * \frac{1}{c} \sum_{k=1}^c AS_{in}(k), \quad (14)$$

where $N(s_i)$ is for a set of consumers who has chosen service s_i , C_y for a set of consumers in the community that consumer u_x belongs to, and $C_y = \{u_z \mid u_x \in C_y \wedge u_z \in C_y\}$.

4.2. Service Selection Process and Algorithm. The overall process of service selection model consists of two phases, as shown in Figure 1. The first is consumer community discovery phase, which is the basic of the second phase. In this phase, the system collects service evaluation information to construct a bipartite network for service selection. Then the consumer preference similarity is integrated by calculating the service selection similarity and the service evaluation difference between consumers based on network structure and edge-weight (service evaluation) of the bipartite network for service selection, corresponding to Step 1. Finally, the system utilizes an improved hierarchical clustering algorithm to discover consumer community structure and chooses the optimal community structure by community degree, corresponding to Step 2, Step 3, and Step 4.

```

Input:  $W$  (bipartite network for service selection),  $C_M$  (consumer community structure)
       $(u_x, u_y, \text{Sim}_p(x, y))$  (triple collection of consumer preference similarity), requester  $u_x$ , the candidate service list SC
Output: the sorted candidate services
Begin
for each service  $s_i$  in SC
{ If ( $s_i$  is a new service)
  {Assign an initial evaluation to service  $s_i$ ,  $PG_{xi} = G_{\max}/2$ }
  Else
  {Calculate the predictive decision parameter  $PD_{xj}$ ;
  If ( $PD_{xi} > T$ ) //  $T$  is a threshold of predictive decision parameter
    Select intra-community prediction to calculate  $PG_{xi}$ 
  Else
    Select inter-community prediction to calculate  $PG_{xi}$ ;}}
Sort the candidate services according to predictive evaluation  $PG_{xi}$ 
End

```

ALGORITHM 2: Service selection algorithm.

The second is service selection phase, which recommends a most satisfactory service to a requester according to his/her preference. The phase includes the following steps. (1) Return the candidate services which meet the requester's functional requirements; (2) compute each service's predictive decision parameter. If the parameter is higher than a predefined threshold, intracommunity prediction will be chosen; oppositely, intercommunity prediction will be chosen. (3) Sort listed candidate services according to the predictive evaluations, and then select a service with the highest value to recommend to the requester. The implementation of the process is given by Algorithm 2.

5. Experiments and Analysis

The algorithms for the model were implemented by Matlab. The simulation data and public data sets were used to test and analyze the algorithms. The experiments on simulation data sets were used to validate the effectiveness and accuracy of the community discovery algorithm and the service selection algorithm. The experiments on public data sets were used to compare CC-PSM with other two kinds of methods.

5.1. Data Sets. Simulation data set is generated through emulating the consumers' service selection and evaluation in the cloud environments. Assume that the number of services meeting essential functional requirements of a requester is S_n in the experiments. For the requester, service selection is actually to choose a most satisfactory service from these services. Each service contains four QoS parameters ranged from 1 to 5, whose values are dynamically generated by the random function. Corresponding to different QoS parameters, the weightings of the requester's preference are, respectively, set to w_j ($1 \leq j \leq 4$ and $\sum_{j=1}^4 w_j = 1$). All the consumers are divided into four groups, and the consumers in the same group have the same preference. Initially, each consumer randomly selects n services and evaluates the selected services according to his/her preference. Then a bipartite network for service selection is gradually formed. Specific parameters in simulation data sets are shown in Table 1.

TABLE 1: Description of simulation data set.

Symbol	Description	Value range
S_n	The number of cloud services	40
U_n	The number of cloud consumers	40~400
C_n	The number of consumer groups	4
Q_n	The number of QoS parameters	4
q_i	The value of QoS parameters	1~5
w_i	The weighting of consumer preference	0~1
WD	The average of preference difference between groups	0.1~0.5
SS_n	The maximum number of services chosen by a consumer	0~30
G_s	service evaluation	1~5
p	The probability of a consumer selecting a service that has been selected by other consumers in the same group	0.1~1

Public data set, which is the same in structure as service evaluation data in the cloud, is selected to test the feasibility of the service selection algorithm in the cloud. We adopted Epinions dataset which was collected by Paolo Massa in a 5-week crawl. The dataset contains 49,290 consumers who have rated a total of 139,738 different items at least once.

5.2. Experimental Results and Analysis. Three standard metrics were adopted to evaluate CC-PSM, including purity [24], $F1$ -measure [25], and MAE [26]. Purity and $F1$ -measure are criteria for the quality of community structure and MAE is a criterion for predicting errors.

5.2.1. Validation of Consumer Community Discovery Model. In this section, the consumer community discovery algorithm was, respectively, applied under different consumer scale. First of all, we randomly generated four kinds of consumer preferences, which made the cloud consumers divided into four groups at the same size. The weightings of consumer preference in each group are enumerated in Table 2.

TABLE 2: Setting of consumer preference weightings.

Consumer group	(w_1, w_2, w_3, w_4)
1	(0.3, 0.3, 0.2, 0.1)
2	(0.4, 0.1, 0.2, 0.3)
3	(0.2, 0.5, 0.3, 0.1)
4	(0.1, 0.1, 0.2, 0.6)

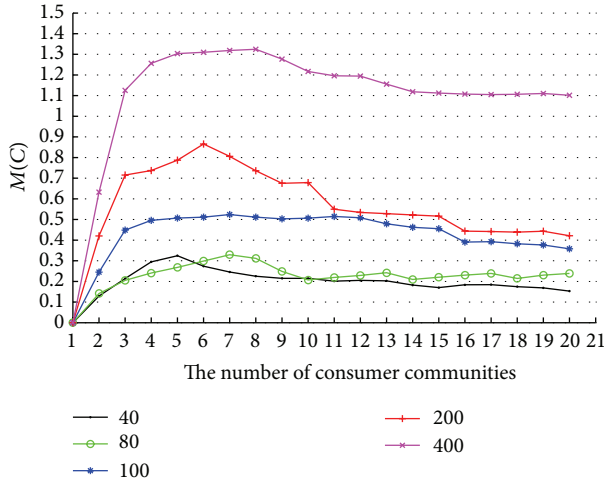


FIGURE 2: The relation between community structure C and $M(C)$ under different consumer scale.

At the same time, set the parameter $p = 0.6$, $SS_n \leq 10$, and the others are shown in Table 1.

The experimental results are as shown in Figure 2 and Tables 3 and 4. Figure 2 shows how $M(C)$ changes with the changing of the number of consumer communities. As a result of the limitation of display area, only the values of $M(C)$ corresponding to the clustering number from 1 to 20 are given. As seen in Figure 2, the value of $M(C)$ slowly increases with the reduction of the number of consumer communities in the clustering process, which reflects the fact that some consumers with high similarity have not been clustered into one community. When the value of $M(C)$ achieves the peak at a certain clustering number, the consumer community structure is the best optimal. And when the clustering process continues, it can partition dissimilar consumers into one community, which inevitably leads to a rapid decline of $M(C)$. In addition, the experimental results have proved from another aspect that community degree can accurately evaluate the consumer community structure and is of great significance for the effect of consumer community discovery algorithm.

The specific situation of the optimal community structure corresponding to different consumer scale is shown in Table 3. Among them, the comparison between the members in the optimal community structure and the members in actual groups under 100 consumers is given in Table 4. As you can see from the experimental data in Tables 3 and 4, the clustering algorithm for consumer community has a good effect on community detection. The accuracy of the algorithm is not affected by consumer scale, and the purity of the discovered community structure is high.

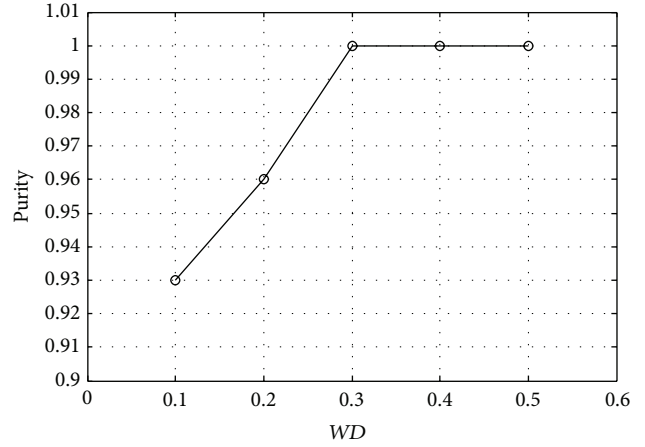


FIGURE 3: The relation between WD and purity.

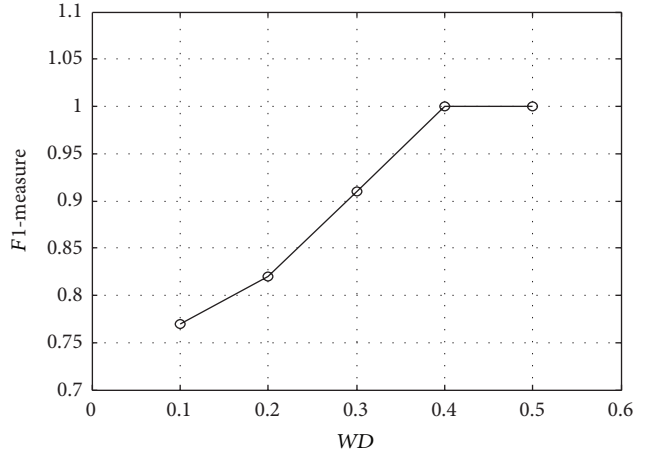


FIGURE 4: The relation between WD and $F1$ -measure.

5.2.2. Analysis on Community Discovery Model

(1) *Relation between WD and the Accuracy of Community Structure.* In order to observe the sensitivity of the algorithm for preference difference between consumers, the experiments adjusted the weightings of consumer preference in Table 2 to make the preference difference of four groups gradually strengthen. In this group of experiments, set $U_n = 200$, $p = 0.8$, the value of WD varying from 0.1 to 0.5, and the others the same as in Table 1. The experimental results are as shown in Figures 3, 4, and 5.

In Figures 3 and 4, the purity and $F1$ -measure of the optimal community, respectively, achieve 93% and 77% when $WD = 0.1$. As the increasing of WD , the accuracy of the community discovery algorithm becomes higher and higher. When $WD \geq 0.3$, the purity of the community structure achieves 100%. When $WD \geq 0.4$, the algorithm can detect the original consumer group division. Figure 5 shows how $M(C)$ changes with the changing of the clustering number of consumer communities under different WD . As can be seen from Figure 5, for the simulation data at different WD , the clustering numbers of the optimal community structure are

TABLE 3: The optimal consumer community structure.

Number of consumers	Number of communities	Number of consumers in the smallest community	Number of consumers in the biggest community	Purity	F1-measure
40	5	4	10	1	0.93
80	7	3	20	0.97	0.88
100	7	4	29	0.96	0.84
200	6	4	51	0.97	0.89
400	8	5	102	0.98	0.92

TABLE 4: The predicted community structure and actual groups under 100 consumers.

ID	The predicted consumer sets	The true consumer sets
1	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25
2	26, 27, 28, 29, 30, 32, 33, 34, 35, 36, 37, 38, 41, 42, 43, 45, 46, 48, 49, 50, 51, 54, 57, 58,	26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50
3	76, 77, 78, 79, 80, 81, 82, 83, 85, 86, 87, 88, 89, 92, 93, 94, 95, 96, 97, 99, 100	76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100
4	52, 55, 56, 59, 61, 62, 63, 66, 67, 69, 70, 71, 72, 73, 75	51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75
5	53, 60, 64, 65, 68, 74	
6	31, 39, 40, 44, 47	
7	84, 90, 91, 98	

different. When the preference difference between consumer groups is relatively large, namely, $WD \geq 0.4$, the value is the same as that in actual situation, while when the difference is not obvious, the value is slightly higher than the number of the original consumer groups because of the existence of some small communities. For example, the number of clusters is 5 when $WD = 0.3$, while the number of clusters is 7 when $WD = 0.2$.

(2) *Relation between p and the Accuracy of Community Structure.* In order to observe the way of the probability p affecting the accuracy of the final community structure, the experiment about the relation between p and the accuracy of community structure was finished. In the experiment, set $Un = 200$, $WD = 0.3$, the value of p varying from 0.5 to 0.9, and the others the same as in Table I. The experimental results are shown in Figures 6, 7, and 8.

Figures 6 and 7 show the purity and F1-measure of the optimal community structure under different p . As shown in Figures 6 and 7, when $p = 0.5$, 16% of consumers are clustered with dissimilar consumers and F1-measure is 0.72. With the growth of p , the purity and F1-measure of the optimal community structure become higher and higher. When p

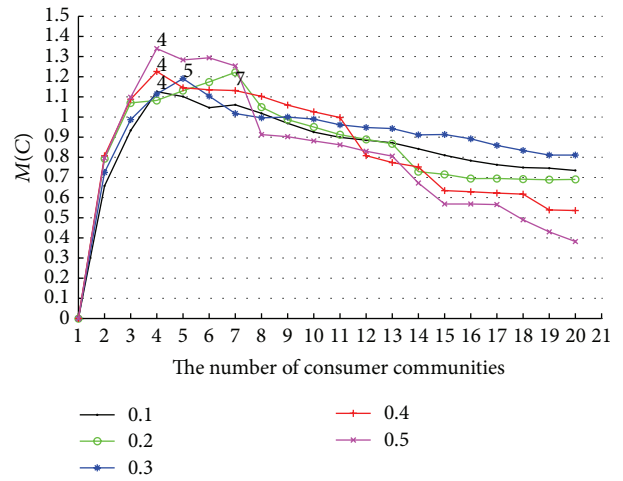


FIGURE 5: The relation between different community structure C and $M(C)$ under different WD .

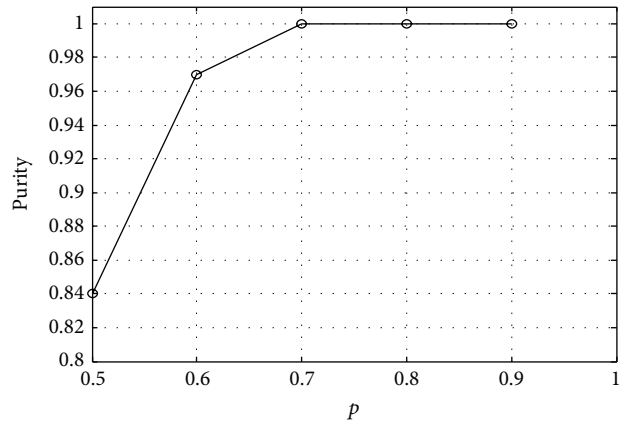
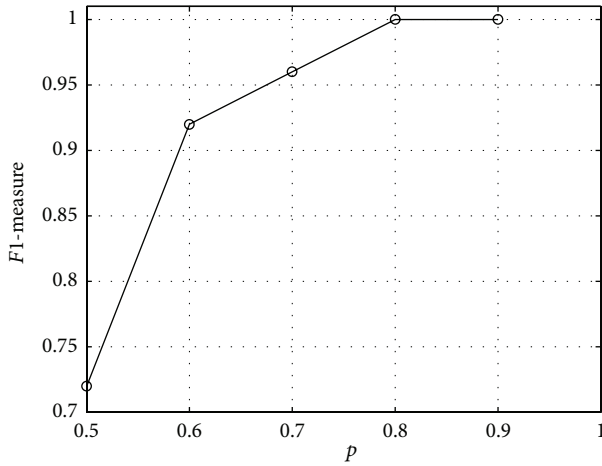
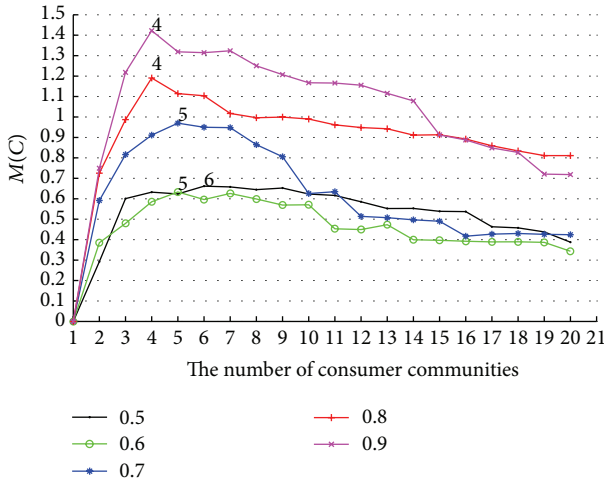


FIGURE 6: The relationship between p and purity.

achieves 0.8, the recognition rate of the algorithm reaches 100%. Figure 8 also shows the algorithm can discover the optimal consumer community structure at that time. Similar to Figure 5, when the consumers in the same group had chosen less same services, the clustering number of the optimal community structure could be higher than that of the original consumer groups because of the existence of some small communities or some wrong partitions.

Through comprehensive analysis on experimental results, the following conclusions can be drawn. (a) The preference


 FIGURE 7: The relationship between p and $F1$ -measure.

 FIGURE 8: The relation between different community structure C and $M(C)$ under different p .

difference between consumers will affect the accuracy of the consumer community discovery algorithm. However, on the whole, the recognition rate of the algorithm still remains at a high level. (b) Whether a consumer's behavior can reflect his/her preference or not, it will have a greater impact on the accurate of the community discovery algorithm than the consumers' preference differences. In common sense, the consumer should not choose some services which deviate from his/her interests. So CC-PSM is effective when the value of p keeps at a reasonable level.

5.2.3. Analysis on Service Selection Model. In this section, the accuracy of CC-PSM is compared with the other two service selection methods. The two methods, respectively, are service selection based on public evaluation and service selection based on nearest neighbors.

Service Selection Based on Public Evaluation (SSPE). The public evaluation of service s_i is the average value of

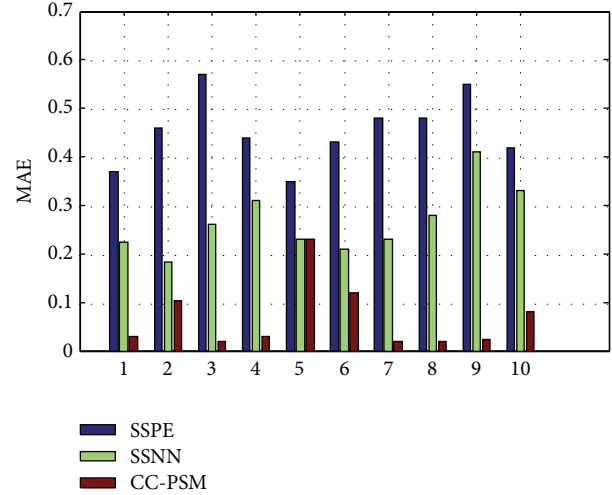


FIGURE 9: Tests of three algorithms on the simulation dataset.

the evaluations from the consumers who have interacted with it. Formally,

$$G_i = \frac{\sum_{u_x \in N(s_i)} G_{xi}}{|N(s_i)|}, \quad (15)$$

where $N(s_i)$ is for a set of consumers who have chosen s_i ; $|N(s_i)|$ for the number of the elements in $N(s_i)$.

Service Selection Based on Nearest Neighbors (SSNN). The method computes the consumer's predictive evaluation on a service through his/her n -nearest neighbors. The predictive evaluation on service s_i of consumer u_x can be expressed as follows:

$$PN_{xi} = \sum_{k=1}^n w_k * G_{ki}, \quad (16)$$

where n is for the number of the nearest neighbors; $w_k = \text{Sim}_p(k, x) / \sum_{j=1}^n \text{Sim}_p(j, x)$.

(1) **Analysis Based on Simulation Data Set.** Based on the data set of 400 consumers in Section 5.2.1, the experiment ran 10 simulations on different candidate services with random quality and evaluated the accuracy of the service selection algorithm by using MAE. The threshold value of predictive decision parameter is set to 0.1. The comparison results are shown in Figure 9. Seen from Figure 9, the MAE of CC-PSM is lower than that of the other two algorithms. And for SSNN, its MAE can be equal with that of CC-PSM, but in most cases, its MAE is significantly higher than that of CC-PSM. In general, prediction error of CC-PSM is still at a lower level.

(2) **Analysis Based on Public Data Set.** Due to the influence of uncertainty factors, it is inevitable that there is greater error when the algorithm is applied in real data than in simulation data. For example, whether a consumer's preference is constant or not and whether a consumer's service evaluation can accurately reflect his/her preference or not

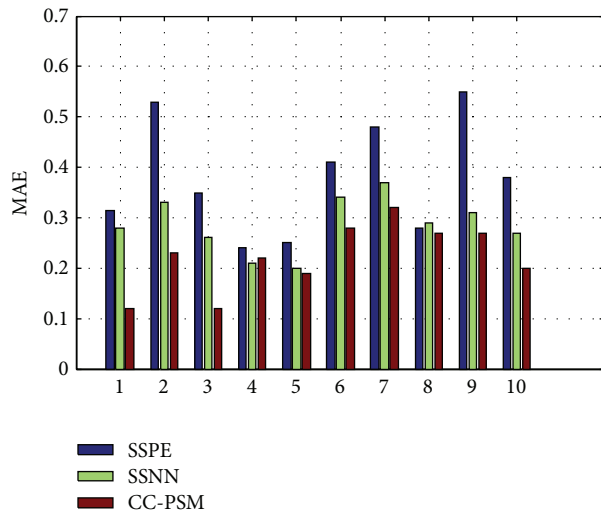


FIGURE 10: Tests of three algorithms on Epinions dataset.

will have great influence on the accuracy of selection model. In the experiment, CC-PSM was compared with the above two methods on the public data set; the threshold value of predictive decision parameter is the same as the above. The experiment randomly selected five datasets with 500 consumers and then randomly divided each dataset into two parts of the training set and the test set with partition ratio of {80%, 20%}. The predictive results based on the training set of the three methods were compared, respectively, with the actual evaluation in the test set. The experimental results are shown in Figure 10. Through Figure 10, you can see that the error of CC-PSM is lower than that of SSPE and SSNN in most cases, so CC-PSM is feasible in application.

6. Conclusions

The cloud service selection is an active research topic in cloud computing. The main drawback of current work is the inability to accurately understand the consumers' preferences. In the actual service environment, the consumers' selection and evaluation information for the services objectively reflect the inner correlation between them. If a service is recommended by a stable group in which the consumers have high similarity with a requester, it will win the requester's higher trust than other services. This viewpoint conforms with our cognition to real situation and has extensive explanatory power. Based on this idea, CC-PSM is built, and the corresponding algorithms are designed. The experimental results have shown that this model and its algorithms achieve the purpose of effectively detecting the consumer community structure and accurately predicting the requester's service evaluation, which have certain extensibility and adaptive ability.

But, this method still has many insufficiencies. The future research can be done from the following aspects: (1) the efficiency problem of the model; (2) cold start problem of the model; (3) other factors that affect community trust.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors wish to thank Natural Science Foundation of China under Grants nos. 61262082, 61262017, and 61462066, Natural Science Foundation of Inner Mongolia under Grant no. 2015MS0608, and Research Program of Science and Technology at Universities of Inner Mongolia Autonomous Region Grant no. NJZY008.

References

- [1] Y. C. Lee, C. Wang, A. Y. Zomaya, and B. B. Zhou, "Profit-driven scheduling for cloud services with data access awareness," *Journal of Parallel & Distributed Computing*, vol. 72, no. 4, pp. 591–602, 2012.
- [2] J. Chen, C. Wang, B. B. Zhou, L. Sun, Y. C. Lee, and A. Y. Zomaya, "Tradeoffs between profit and customer satisfaction for service provisioning in the cloud," in *Proceedings of the 20th ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC '11)*, pp. 229–238, ACM, June 2011.
- [3] M. Armbrust, A. Fox, R. Griffith et al., *Above the Clouds: A Berkeley View of Cloud Computing*, University of California at Berkeley, 2009.
- [4] S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, and A. Ghalsasi, "Cloud computing—the business perspective," *Decision Support Systems*, vol. 51, no. 1, pp. 176–189, 2011.
- [5] S. Ding, C. Xia, Q. Cai, K. Zhou, and S. Yang, "QoS-aware resource matching and recommendation for cloud computing systems," *Applied Mathematics and Computation*, vol. 247, pp. 941–950, 2014.
- [6] Z. U. Rehman, F. K. Hussain, and O. K. Hussain, "Towards multi-criteria cloud service selection," in *Proceedings of the 5th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS '11)*, pp. 44–48, IEEE, July 2011.
- [7] L. Zhao, Y. Ren, M. Li, and K. Sakurai, "Flexible service selection with user-specific QoS support in service-oriented architecture," *Journal of Network and Computer Applications*, vol. 35, no. 3, pp. 962–973, 2012.
- [8] Z. U. Rehman, O. K. Hussain, and F. K. Hussain, "Parallel cloud service selection and ranking based on QoS history," *International Journal of Parallel Programming*, vol. 42, no. 5, pp. 820–852, 2014.
- [9] Q. Wu, M. Zhang, R. Zheng, Y. Lou, and W. Wei, "A QoS-satisfied prediction model for cloud-service composition based on a hidden markov model," *Mathematical Problems in Engineering*, vol. 2013, Article ID 387083, 7 pages, 2013.
- [10] 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. 24, no. 6, pp. 1213–1222, 2013.
- [11] S. Ding, S. Yang, Y. Zhang, C. Liang, and C. Xia, "Combining QoS prediction and customer satisfaction estimation to solve cloud service trustworthiness evaluation problems," *Knowledge-Based Systems*, vol. 56, pp. 216–225, 2014.

- [12] L. Qu, Y. Wang, and M. A. Orgun, "Cloud service selection based on the aggregation of user feedback and quantitative performance assessment," in *Proceedings of the IEEE 10th International Conference on Services Computing (SCC '13)*, pp. 152–159, July 2013.
- [13] H.-Y. Wang and D.-Y. Zhang, "A trustworthy service selection model based on collaborative filtering," *Journal of Electronics and Information Technology*, vol. 35, no. 2, pp. 349–354, 2013.
- [14] S. Abedinzadeh and S. Sadaoui, "A trust-based service suggestion system using human plausible reasoning," *Applied Intelligence*, vol. 41, no. 1, pp. 55–75, 2014.
- [15] R. Karim, C. Ding, and A. Miri, "An end-to-end QoS mapping approach for cloud service selection," in *Proceedings of the IEEE 9th World Congress on Services (SERVICES '13)*, pp. 341–348, IEEE, July 2013.
- [16] M. Godse and S. Mulik, "An approach for selecting software-as-a-service (SaaS) product," in *Proceedings of the International Conference on Cloud Computing (CLOUD '09)*, pp. 155–158, 2009.
- [17] B. Mu, S. Li, and S. Yuan, "QoS-aware cloud service selection based on uncertain user preference," in *Rough Sets and Knowledge Technology*, vol. 8818 of *Lecture Notes in Computer Science*, pp. 589–600, Springer, 2014.
- [18] M. Sun, T. Zang, X. Xu, and R. Wang, "Consumer-centered cloud services selection using AHP," in *Proceedings of the International Conference on Service Science (ICSS '13)*, pp. 1–6, April 2013.
- [19] C.-H. Hu, J.-B. Liu, and J.-X. Liu, "Services selection based on trust evolution and union for cloud computing," *Journal on Communications*, vol. 32, no. 7, pp. 71–79, 2011.
- [20] R.-Z. Du, J.-F. Tian, and H.-G. Zhang, "Cloud service selection model based on trust and personality preferences," *Journal of Zhejiang University (Engineering Science)*, vol. 47, no. 1, pp. 53–61, 2013.
- [21] S. Deng, L. Huang, and G. Xu, "Social network-based service recommendation with trust enhancement," *Expert Systems with Applications*, vol. 41, no. 18, pp. 8075–8084, 2014.
- [22] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical Review E*, vol. 69, no. 2, Article ID 026113, 2004.
- [23] M. E. Newman, "Mixing patterns in networks," *Physical Review E. Statistical, Nonlinear, and Soft Matter Physics*, vol. 67, no. 2, 2003.
- [24] Evaluation of Clustering, April 2013, <http://nlp.stanford.edu/IR-book/html/htmledition/evaluation-of-clustering-1.html>.
- [25] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.
- [26] A.-L. Deng, Y.-Y. Zhu, and B.-L. Shi, "Collaborative filtering recommendation algorithm based on item rating prediction," *Journal of Software*, vol. 14, no. 9, pp. 1621–1628, 2003.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

