An Efficient QoS Preference Oriented Recommendation Scheme for the Internet of Services

Le Xin, Yushun Fan, Hongbo Lai

An Efficient QoS Preference Oriented Recommendation Scheme for the Internet of Services

Le Xin, Yushun Fan, and Hongbo Lai

Department of Automation, Tsinghua University,
100084 Beijing, China
xin-l09@mails.tsinghua.edu.cn, fanyus@tsinghua.edu.cn,
laihongbo_001@163.com

Abstract. Internet of services provides the services capabilities to interact and collaborate. Users are able to create, share and access services by means of heterogeneous devices and interaction channels, in extremely personalized ways. However, in the dynamic and ever-growing scenario, lacks of customization and flexibility have become more and more crucial. In this paper, a QoS preference oriented service recommendation scheme is proposed. First, the customer's preference is well described by a matrix model, which can be updated dynamically with customer's feedback evaluation. Entropy evaluation method is introduced to measure the weight coefficient of each QoS attribute. Furthermore, ant colony optimization algorithm is used to seek optimal service compositions. This recommending procedure will be repeated until the customer's requirement is satisfied, also can be restarted whenever the service process is blocked in execution. Theoretical analysis and numerical simulation indicate that the proposed scheme can satisfy the customer's requirements effectively and flexibly.

Keywords: Service recommendation; Customer's preference; Internet of Service.

1 Introduction

Internet of Services (IoS)[1, 2] is a relatively new paradigm, enabling the operation and reorganization of business services, in which the enterprises package their business units into business modules and provide services on users' demand. The services, either IT or business, are arranged under standard protocol [3]. Because of the unified ports of service units, the services can be used as "plug and play" mode on customers' demand [4], supported by the technologies as cloud computing [5]. The business collaboration is achieved by providing and consuming services, which are managed according to the service level agreement between users and providers.

Based on this environment, complex business application can be dynamically composed by existing service components from different providers. The composite services may have different performances with the same functionalities, such as price, response time, availability, reputation, security level and so on. These quality-of-service (QoS) factors, as the measurements of non-functional features, are important to the service selection and composition. After functional matching based on UDDI, users can compare candidate services from the respects of QoS. However, for both single task service discovery and aggregative service composition, requesters still lack an efficient method to solve the global optimization because of the following reasons.

1) Different users value different characters of services. For example, patients require fast medical services, while banks need stable data services. The researches on service selection mostly handle different QoS factors equally [6, 7], which lead to absence of customization in service selection. How to differ and determine the importance of each QoS attributes, dynamically corresponding with user's online feedback is still a weakness to identify and satisfy requesters' demand in IoS.

2) Under the users' global constraints to QoS [8], searching the optimal composition has been proved to be an NP-hard problem [9]. Exhaustive and evolutionary algorithms are two kinds of approaches of QoS-based optimization computation. Some of the selection efficiency has close relationship with the increase of candidate services; and some must restart from the initial state when interrupted by unavailable services, which leads to lack of flexibility [10, 11].

In IoS, personalized service recommendation has become more and more crucial. Considering the large number of services, requesters also need flexible and effective composition. In this paper, we study an efficient service recommendation method based on customer's preferences of QoS. First, the degree of customer's preference to each QoS attribute is well described by a value similar to membership in fuzzy space. Then services' matching level can be computed by customer's preference and QoS value of service. On these bases, services are recommended dynamically to customer, and customer's feedback evaluations of the services are obtained. Preference matrix is re-computed, and then goes to the next iteration. The ant colony optimization algorithm is used to optimize the searching in service selection.

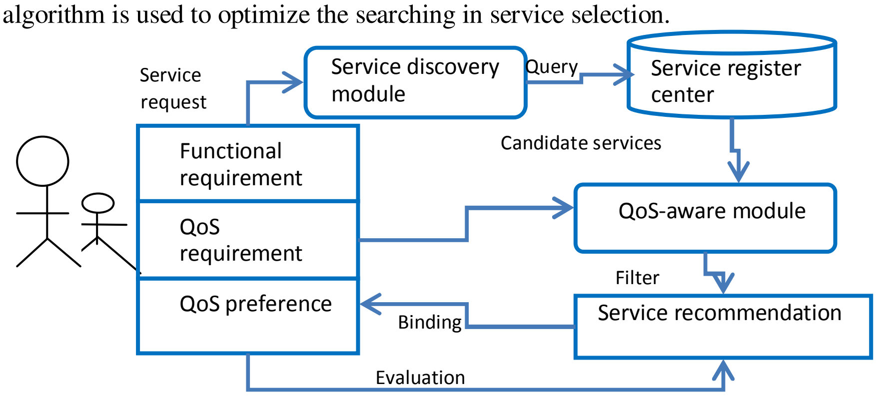

Fig. 1. QoS preference Oriented recommendation model

2 QoS Preference Oriented Service Recommendation Model

In this section we designed a service selection method through a process as illustrated in Fig.1. When customer raises request in IoS, service discovery is carried out. Through querying in register center, the candidate services which fit functional requirement can be obtained. Then customer's restrict to QoS is checked to filter out those not matching constraints, detailed computation method of which can refer to Huang [6]. After these, the list of services can be sorted and recommended, at the same time, customer's preference to QoS can be perceived by his evaluation to services on line. With these feedbacks, service recommendation result is able to be optimized closer and closer to customer's requirements. In the following models, Section 2.1 pays attention to measure customer's preference and sort the services by match degree.

2.1 Customer's Preference Model

In the model proposed above, it is difficult to measure customers' preference, which leads to lack of customization. In this section, we give the QoS model and customers' preference model based on fuzzy logic matrix first. Entropy evaluation method is introduced to measure the weight coefficient of each QoS attribute. And then services' matching level can be calculated by these models.

After normalization, the i th QoS factor can be divided into n level, e.g. if $n=5$, it means {very high, high, medium, low, very low}. For each level there is a median value $\{m_{i1}, m_{i2}, m_{i3}, \cdots, m_{in}\}$, and $1 \ge m_{i1} \ge m_{i2} \ge m_{i3} \ge \cdots \ge m_{in} \ge 0$ .For q_i', euclidean distance is used to determine which level it belongs to.

$D_{ij} = | q_i' - m_{ij} |$, if $D_{ik} = \min_{1 \le j \le n}\{D_{ij}\}$, then the i th QoS factor belongs to the k th level.

Definition 1 Quality of service is denoted by a $m \times n$ matrix Y. The matrix element is $y_{ij} = \begin{cases} 0, & \text{if } q_i' \text{ does not belong to the } j\text{th level} \\ 1, & \text{if } q_i' \text{ belongs to the } j\text{th level} \end{cases}$.

Definition 2 Customers' evaluation is denoted by a, and $a \in [0,1]$. The greater a is, the higher evaluation is.

Definition 3 Customers' preference is denoted by a $m \times n$ matrix P. The element of which is $p_{ij}(i = 1, 2, \cdots m;\ j = 1, 2, \cdots n) \in [0,1]$ is the customer preference degree to i th QoS factor belongs to j th level, and $\sum_{j=1}^{n} p_{ij} = 1$.

Customer's preference can be obtained by records of evaluation. After N times judging on the services, preference matrix is, $P(N) = (1/N)\sum_{k=1}^{N} a_k Y_k$. In the following section, it will be proved that $\lim_{N\to\infty} P(N) = \lim_{N\to\infty}(1/N)\sum_{k=1}^{N} a_k Y_k = E(a \cdot Y)$ can present the customer's true preference consistently.

Definition 4 match degree between customer's preference and quality of service can be calculated as $M = m^{-1}\sum_{i=1}^{m} c_i \sum_{j=1}^{n} y_{ij} p_{ij}$, Where c_i is the weight coefficient corresponding to the i th QoS attribute, and $c_i \in [0,1]$, $\sum_{i=1}^{m} c_i = 1$.

To determine the weight of each attribute, entropy evaluation method can be used. Shannon entropy [12] is a measure of the average information content. The entropy H of a discrete random variable Y_i with possible values $\{y_{i1}, \cdots, y_{in}\}$ and probability mass function $p_{ij}(j = 1, \cdots, n)$ is denoted as, $H(Y_i) = E(I(Y_i)) = k\sum_{j=1}^{n} p(y_{ij}) I(y_{ij})$ $= k\sum_{j=1}^{n} p(y_{ij}) \ln[p(y_{ij})]^{-1} = -k\sum_{j=1}^{n} p_{ij} \ln p_{ij}$.

The entropy H should be normalized to $H \in [0,1]$. This indicates that $k = (\ln n)^{-1}$.

In the information system, entropy is a measure of disorder, or more precisely unpredictability. The higher entropy means the lower utility value. Hence we use $1-H(Y_i)$ to measure the weight coefficient of the i th QoS attribute. After normalization, it can be denoted as $c_i = [1-H(Y_i)] / \sum_{i=1}^{m} (1-H(Y_i))$. Thus,

$$M = m^{-1} \cdot \sum_{i=1}^{m} \{[1+(\ln n)^{-1} \sum_{j=1}^{n} p_{ij} \ln p_{ij}] / \sum_{i=1}^{m} (1+(\ln n)^{-1} \sum_{j=1}^{n} p_{ij} \ln p_{ij})\} \cdot \sum_{j=1}^{n} y_{ij} p_{ij}$$

Here, if an element of the preference matrix is $p_{uv} = 0$, to avoid $\ln p_{uv} = -\infty$, let $p_{uv} = 10^{-\lambda} \min_{1 \le j \le n, j \ne v} \{p_{uj}\}$, $p_{uj} > 0$, where λ is a positive integer.

For a given service, the values of QoS attributes are fixed. The customer's evaluation to services can be found in the historical information register and transferred to preference matrix. So the match degree between customer's preference and quality of service can be obtained, which is the basis for sorting of recommending services.

3 QoS Preference Based ACO Algorithm

The ant colony optimization algorithm (ACO) is a bionics probabilistic technique for solving optimization problems which can be transferred to finding good paths through graphs, initially proposed by M. Dorigo in 1992 [13]. The Convergence of ACO algorithm has been proved [14], which means it is able to find the global optimum in finite time, though difficult to estimate the theoretical speed of convergence.

In the service recommendation scheme, ACO can be used to optimize the service selection for each task node. The service composition based on process model is illustrated in Fig. 2.

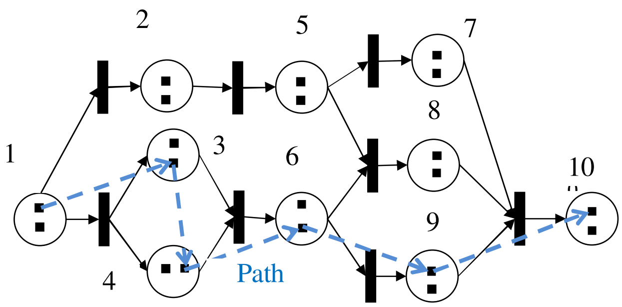

Fig. 2. An example of service composition

When service process is aggregated, there are several services for a task to choose. From the start to the end, it can be presented as a path, e.g. Path 1 is an executing sequence, which can be denoted as a linked list, liking the node presented by $(taskNo, serviceNo)$. Thus service selection problem can be transferred to seeking an optimal path in the directed digraph. In this example, No. 2 and No. 3 tasks are in parallel split pattern, while No. 5 and No. 6 mean exclusive choice. As shown in the

Fig. 2, we handle the parallel task as sequence pattern to seek the path, so as to guarantee them being executed both; however, the aggregative QoS value is still calculated as parallel pattern.

The ACO algorithm is suitable for service composition selection because of its flexibility. Seeking an optimal path can starts from any task node. Thus if service process is interrupted by an unavailable service, the optimization algorithm can be restarted immediately.

3.1 The Algorithm for Service Recommendation Oriented to QoS Preference

The ACO algorithm for service recommendation based on customer's preference is illustrated as following.

Algorithm 1 QoS preference oriented ACO algorithm for service recommendation

1) Initialization: set iteration count $N_s = 0$, the max iteration count $N_{s\max}$, customer's preference matrix $P(N_{eva})$, initially $N_{eva} = 0$, and pheromone $\tau_{ij}(0)$. Put M ants in the start task node, the initial number of ant is $k = 1$;

2) The iteration count $N_s = N_s + 1$;

3) The ant number $k = k + 1$;

4) When the k th ant is in the $taskNo$ th task, update the $\{allowed\}$ list, which contains the services in next task node able to link directly to the $taskNo$ th ant;

5) Choose the next service according to the state transition rule. The ant staying in service i moves to service j .

If $q \le Q_0$, j is determined by $\zeta_{ij}(t) = Max(\tau_{ij}(t)^{\alpha} * \eta_{ij}^{\beta}),\ j \in \{allowed\}$

Else if $q > Q_0$, the probability to select service j is defined by

$$\rho_{ij}(t) = \begin{cases} \tau_{ij}(t)^{\alpha} * \eta_{ij}^{\beta} / (\sum_{u \in allowed} \tau_{iu}(t)^{\alpha} * \eta_{iu}^{\beta}),\ j \in \{allowed\} \\ 0, \text{otherwise} \end{cases}$$

Where q is a uniform distribution random number in [0,1]; $Q_0 \in [0,1]$ is a predefined parameter which determine the probability of following the optimal result; $\zeta_{ij}(t)$ is transition rule from service i to service j ; $\tau_{ij}(t)$ denotes the density of pheromone; α presents the importance of pheromone while β stands for importance of expected heuristic information; $\{allowed\}$ means the collection of services in the task node which can be moved to at next step.

η_{ij} is the local expected factor, presenting the heuristic information in QoS-based service selection, and measuring the short term benefit, which has the same function with match degree of service j , thereby, $\eta_{ij} = M(\text{service}\, j, P(N_{eva}))$

To avoid the probability to select service j being 0 when $q \ge Q_0$, the initial pheromone $\tau_{ij}(0)$ should not be 0. Thus it can be defined as the average local expected factor, $\tau_{ij}(0) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \eta_{ij} / 2$.

6) If there is no service can be searched, go to 4) and move to another task node. Otherwise, go to 7);

7) If $k < M$, update the allowed list and go to 3), otherwise, go to 8);

8) Update the global pheromone.

During Δt, M ants have completed food seeking, so called an iteration. Thus the pheromone between service i and service j is updated as,

$$\tau_{ij}(t+\Delta t) = R\cdot\tau_{ij}(t) + \Delta\tau_{ij}(\Delta t)$$

$R \in [0,1]$, and $(1-R)$ means the evaporation rate of pheromone over time, aiming to reduce the influence of previous cases.

Here $\Delta\tau_{ij}(\Delta t) = \sum_{k=1}^{M} \Delta\tau_{ij}^{k}(\Delta t)$, and $\Delta\tau_{ij}^{k}(\Delta t)$ means in this iteration, the increment of pheromone left by the k th ant, qualified by,

$$\Delta\tau_{ij}^{k}(\Delta t) = \begin{cases} Q/L_k, & \text{if ant } k \text{ go through } i \text{ and } j \\ 0, & \text{otherwise} \end{cases}$$

Q is a constant. L_k means the length of path ant k go through, which is the measurement of difficulty to find food. Thus in service recommendation, let $L_k = 1/M(\text{comp } k, P(N_{eva}))$

$M(\text{comp } k, P(N_{eva}))$ denotes the match degree between customer's preference and service composition presented by the path of ant k in this iteration.

9) Recommend the service compositions sorted by $M(\text{comp } k, P(N_{eva}))$ to customer;

10) If customer doesn't evaluate any services or compositions, go to 11);

Otherwise, when the customer evaluate N_k new services or compositions, the preference matrix is recalculated as,

$$P(N_{eva}+N_k) = (1/(N_{eva}+N_k))\sum_{k=1}^{N_{eva}+N_k} a_k Y_k = P(N_{eva}) + N_k(P(N_k) - P(N_{eva}))/(N_{eva}+N_k)$$

Where $P(N_k)$ denotes the average evaluation to the new N_k services.

$P(N_k) = (1/N_k)\sum_{k=N_{eva}+1}^{N_{eva}+N_k} a_k Y_k$. Then let $N_{eva} = N_{eva} + N_k$;

11) If customer's requirement is satisfied, or $N_s > N_{s\max}$, stop. Otherwise, go to 2).

4 Experiment and Analysis

To evaluate the performance of proposed service composition recommendation method from the convergence rate and computation complexity, experiments have been conducted in the example process model illustrated as Fig.2. For each task there are 500 candidate services. Every service has 3 QoS attributes which are generated randomly in $[0,1]$ according to uniform distribution. Each attribute has 5 levels. Thus we can skip the normalization steps and calculate the QoS matrix directly. The parameters are set as follows: the amount of ants $M = 50$, pheromone factor $\alpha = 0.9$, expected heuristic factor $\beta = 1$.

4.1 The Influence of Pheromone Persistence Factor

The pheromone evaporation rate $(1-R)$ determines the global search ability and convergence rate. If the pheromone evaporates too quickly, the initial pheromone trails never been searched will decrease almost to zero, which reduces the global search ability and stochastic property. However, if the evaporate rate is too low, the convergence rate will be lessened. Figure 3 describes the variances of R and iteration times. The number is average value of 10 experiment results. The stop rule of iterations is that difference between the last two maximum match degree values is less than 0.0001.

These results indicate that 0.6~0.8 is a proper scope of R. In this scope, the performance of algorithm is stable, and it gains balance in convergence and global optimization.

4.2 The Process of Recommendation Optimization

According to the result above, let $R = 0.7$, which means the evaporation rate of pheromone is 0.3. An optimal path is preset, denoted as Path1 {(1,5),(3,8),(4,3),(6,12),(9,6),(10,9)}. Let No.5 service of No.1 task, No.8 service of No.3 task,…, No.9 service of No.10 task, all have the highest level QoS values. The customer's preference matrix is defined as $y_{i1} = 1, i = 1,2,3$. Thereby the match degree of service composition in Path1 is 100%. Other services' QoS values are generated randomly. Figure 4 describes the variances of match degree and iteration times, which are the average values of 10 experiment results.

Figure 4 shows that based on ACO algorithm, the match degree value of service composition increases rapidly and gets to the optimal value, which means ACO algorithm has good global convergence property. This recommending method can improve the agility and flexibility of service composition, even if the process is complex.

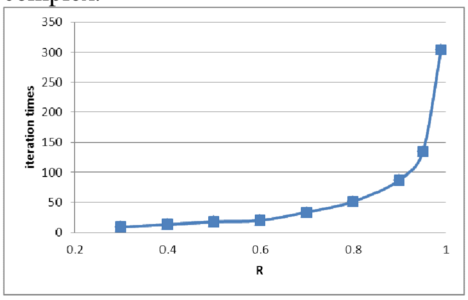

Fig. 3. Variances of R and iteration times

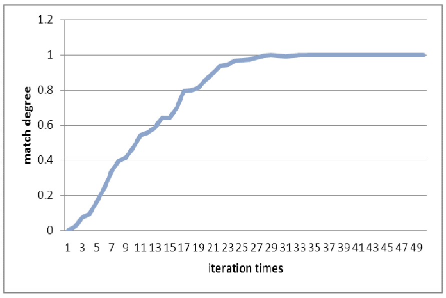

Fig.4. Variances of iteration times and match degree

5 Conclusion

In this paper, a QoS preference oriented service recommendation method is proposed. Ant colony optimization algorithm is used to optimize the service composition selection. Supported by the matrix model, customer's preference to QoS can be obtained and expressed dynamically. Furthermore, ant colony optimization algorithm is used to seek the optimal service compositions. After each iteration, several compositions are recommended according to the match degree with customer's requirement. During this period, with customer's evaluation feedback of services, the preference matrix will be recalculated and updated. The recommendation results are optimized gradually with the ACO algorithm iterating. This recommending procedure will be repeated until the customer's requirement is satisfied. If the service process is stopped in execution by some unavailable services, the recommendation procedure can restart immediately to search from the blocked task node. Theoretical analysis and numerical simulation indicate that the service composition recommendation algorithm proposed in this paper can satisfy the customer’s requirements effectively and flexibly.

Acknowledgments. This work is supported by National Natural Science Foundation of China under grant No. 61033005 and 61174169, National Key Technology R&D Program under Grant No. 2010BAH56B01.

References

1. Schroth, C., Janner, T.: Web 2.0 and SOA: Converging concepts enabling the internet of services. IT Professional **9** (2007) 36-41
2. European Commission ICT Research in FP7, http://cordis.europa.eu/fp7/ict/home_en.html
3. Lijun, M., Chan, W.K., Tse, T.H.: An Adaptive Service Selection Approach to Service Composition.: IEEE International Conference on Web Services (ICWS '08) (2008) 70-77
4. Ying, L., Fangyan, R., Ying, C., Dong, L., Li, T.: Services ecosystem: towards a resilient infrastructure for on demand services provisioning in grid.: IEEE International Conference on Web Services (ICWS'04) (2004) 394-401
5. Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., Zaharia, M.: A view of cloud computing. Commun Acm **53** (2010) 50-58
6. Huang, A., Lan, C.W., Yang, S.: An optimal QoS-based Web service selection scheme. Inform Sciences **179** (2009) 3309-3322
7. Fan, X.Q., Fang, X.W., Jiang, C.J.: Research on Web service selection based on cooperative evolution. Expert Syst Appl 38 (2011) 9736-9743
8. Chen, Z., Liang-Tien, C., Bu-Sung, L.: DAML-QoS ontology for Web services.: Web Services, 2004. Proceedings. IEEE International Conference on (2004) 472-479
9. Canfora, G., Di Penta, M., Esposito, R., Villani, M.L.: A Lightweight Approach For QoS-aware Service Composition.: IEEE international conference on service-oriented computing (ICSOC'04) (2004)
10. LIU, S.L., LIU, Y.X., ZHANG, F., TANG, G.F., JING, N.: A dynamic web service selection algorithm with qos glabal optimal in web services composition. Journal of Software 18 646-656
11. Zhang, C.W., Su, S., Chen, J.L.: Genetic Algorithm on Web Services Selection Supporting QoS. Chinese Journal of Computer 29 (2006) 1029-1037
12. Shannon, C.E.: A Mathematical Theory of Communication.: (1948)
13. Dorigo, M.: Optimization, learning and natural algorithms. Ph. D. Thesis, Politecnico di Milano, Italy (1992)
14. Zlochin, M., Birattari, M., Meuleau, N., Dorigo, M.: Model-based search for combinatorial optimization: A critical survey. Ann Oper Res 131 (2004) 373-395