OPTIMIZATION STRATEGY OF CLOUD COMPUTING SERVICE COMPOSITION RESEARCH BASED ON ANP

Xing Xu
School of Automation
Huazhong University of Science and Technology
Wuhan 430074, P.R.China
E-mail: xuxin19901201@126.com

Hexi Guo*
School of Automation
Huazhong University of Science and Technology
Wuhan 430074, P.R.China
E-mail: guohexi@163.com

ABSTRACT

Optimization strategy of service composition is the kernel content of ensuring and improving service level in Service-Oriented Architecture. However, it is difficult to realize the optimization of services in an open environment, in which there are a lot of randomicity disturbances and uncertain factors. The key is to accurately and efficiently select and composite the services. In this paper, we proposed the comprehensive evaluation metrics using ANP, an algorithm of generating all alternative atomic services and an algorithm of generating the ANP alternatives set. The metrics is for measuring the service composition quality, and combines the key performance indicators of service in cloud. The algorithm of generating all alternative atomic services is brought forward to identify the fundamental services satisfying users' requests. The second algorithm, which composites atomic services, generates feasible cloud service compositions, namely the ANP alternatives set. Consequently, ANP optimization method has been developed to find the optimal one from the feasible service compositions.

Keywords: Analytic Network Process (ANP); optimization strategy; algorithm; service composition; quality of service; atomic service

1. Introduction

The optimization selection of cloud service composition is a necessary condition for ensuring that a cloud provider guarantees the quality of service provided to the service requesters. In addition, measuring the superiority of service compositions is the key to develop optimization strategy of service composition. There are many factors need to be considered during the optimization selection of cloud service composition. This is a typical multiple attribute decision making problem due to the variety of factors, which include service price, service performance, and etc. These factors also involve the interaction and dependence relationships between them, and can not be categorized as a typical decision problem with structured hierarchy. ANP considers the dependency between different criteria, which is closer to the decision problem and is often neglected in other approaches (Saaty, 2005). Hence, ANP may facilitate the ability to measure the superiority of services residing within Clouds.

* Corresponding author
Project supported by the National Nature Science Foundation of China (No. 61273207)

ANP is used widely as an effective method to solve multiple attribute decision making problems. However, the study on optimization strategy of service composition is still limited. The current research cannot always guarantee the feasibility of the selected service composition and the satisfaction of user requirements achieved by the selected composition at the same time. The evaluation of superiority of cloud service only measures the quality of service, for example, in (Li, Wu, and Feng, 2012), the metrics for measuring the service composition quality contains only four independent indicators without considering the influence relationship between indicators. In（Huang, Zeng, Fan, and Huang, 2011), the metrics only contains the indicators for QoS, and it applies genetic algorithm to measure the superiority of service composition. The evaluation process is relatively complicated so that the method may have low practicality, especially for users who do not have the relevant knowledge. ANP model is used in (Godse, Sonar, and Mulik, 2008), which groups QoS attributes into three clusters: runtime, security and configuration management related. By following the ANP steps, the alternative services can be ranked and selected, but the optimal service composition, which is selected using the method proposed in the article, may be infeasible.

Using ANP to measure service compositions and service compositions is an alternative of ANP model. In current research, the service composition in the ANP alternatives is not always feasible and the number of service compositions in the ANP alternatives is limited to the number of atomic services. To a large extent, these shortfalls increase the difficulty in measuring the quality of service composition. Due to the characteristics of Petri net with asynchronous and concurrent (Yong, 2005), Petri net is used to distributed software systems like cloud computing, and then is applied to find the services compositions in many articles. Using Petri net, and web services is described in (Narayanan, and McIlwraith, 2002). It also analyzes the services simulation, verification and automatic composition. However, it can not guarantee the feasibility of service composition and doesn't meet users' demands.

In this paper, ANP is innovatively applied to select cloud service composition. In order to prevent selecting the unavailable service composition and ensure alternatives to satisfy the user requirement, we focus on developing algorithm that generates feasible ANP alternatives set. Finally, ANP measuring system is used to complexly measure alternatives. By doing that, we can obtain the ranking of feasible service compositions and select the optimal one among them.

2. Selection of optimal service composition using ANP

Selection of optimal service composition using ANP is an application process of ANP from receiving the requests from service requesters to accomplish the selection of optimal service composition with considering decision-maker preference.(Menascéa, Casalicchiob, Dubey, 2009) The optimization selection process includes not only the ANP evaluation process, but also the establishment of metrics for measuring service composition quality, optimizing ANP alternative set according to the requirements of users, assignment to alternatives using interactive simulation and other important processes. We summarized a series of necessary and important steps to get the service composition optimization selection process, as described in the following procedure.

2.1 ANP-based metrics for measuring service composition quality

To select the optimal service composition, the goal of ANP-based metrics is for measuring the superiority of service composition. Considering the service composition quality from the three aspects of service composition cost, hardware resource constraints and service composition quality, we proposed eleven indicators under the three criteria namely, the goal, control hierarchy and network elements are obtained in ANP-based metrics. Definitions and formulas of indicators are as follows:

(1) Service composition cost

The indicators under the criterion of service composition cost, not only contains the necessary cost for running services, but also the SLA violation cost because of the inconsistent with SLA items (Tao, and Zhang, 2011).

1) service cost: service cost is equal to the sum of hardware cost and software cost.

hardware cost =(Q+ nc*Qc)*s

where Q is the energy consumption of running the services composition, nc is the number of cooling equipment required when running the services composition, Qc is energy consumption of every cooling equipment is the average price per Kwh.

software cost = $\sum_{i=1}^{3} S_i * N_i * C_i + \sum_{i=1}^{n} S_t * N_t * C_t$

where S_i is the cost percentage of software which the running service composition requires, N_i is the number of the required software, C_i is the price of the software; S_t is the cost percentage for maintenance and updating of the required software, N_t is the times of maintenance and updating of the required software, C_t is the unit price of maintenance and updating of the required software.

2) SLA violation: the expense of violating SLA depends on the rate of service level agreement violations. Higher the rate of service level agreement violations, more different between the services quality provided and the one stipulated in SLA, so the higher costs need to pay (Yonggen, Wei, and Jie, 2012). SLA violation is measured by the average rate of service level agreement violations.

(2) Hardware resource constraints

The indicators ,under the criterion of hardware resource constraints, consist of CPU resource constraints , memory resource constraints and I/O resource constraints because service can be divided into three kinds which are services using CPU intensively, services using memory intensively and services using I/O intensively , according to the utilization of hardware resources to run the service composition.

Table 1. Indicators under the criterion of hardware resource constraints

Criterion	**Indicator**	**Formula**	**Interpretation formula**
Hardware resource constraints	CPU resource constraints	CPU resource constraints = f* Sf*T	f is the average workload, Sf is the frequency of invoking service, T is execution time of service process.
	memory resource constraints	memory resource constraints = $\sum_{i=0}^{n} D_b * O_i * K_i$	D_b is the size of data entity, O_i is the number of service process, K_i is the services number of each service process.
	I/O resource constraints	I/O resource constraints = Un* $\sum_{i=0}^{n} D_i * N_i$	D_i is the size of exchanging data entity, N_i is necessary number of data transmission, Un is the number of concurrent users' requests.

(3) QoS of service composition

Cloud computing is divided into 3 modes according to the level of services which are Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS).

Table 2. Indicators under QoS of service composition

Indicator	**Description**
Network delay	the time taking for a bit of data to travel across the network from one node or endpoint to another
Service response time	the time elapsed between the time when user makes a request to the time when user receives the response result
Throughout	the average rate of successful message delivery over a communication channel
Expandability	the ability to deal with a large number of online users
Error rate of service request	the average occurring rate of unexpected handling in per unit time
Service delivery stability	to measure the ability of offering the stable service to users

The performance of SaaS is closely related with the underlying infrastructure, in many situations, the indicator value of higher hierarchy would be improved by using the software optimization method with the support from infrastructures and platforms (Luo, Jin, Song, and Dong, 2011). So we put forward the indicators measuring the service quality of IaaS and PaaS to evaluate the QoS of service composition.

By analyzing the goal and criteria of ANP, the control hierarchy in ANP had already obtained. We also got all the elements in ANP network, and it is obvious that from the definitions and formulas of metric elements, indicators are not strictly independent. By analyzing the relationship between metric elements, the relationship between service composition cost and QoS of service composition is contradictory, because it is well known that we pursuit low cost while require high quality service, but the fact is that higher the quality, more the cost. The relationship between hardware resource constraints and QoS of service composition is contradictory in the same way. In addition, the indicators under the criterion of QoS influence each other. in other words, the QoS cluster has inner dependence.

Through the analyses above, the metrics for measuring the superiority of service compositions using ANP is obtained, as shown in Figure1.

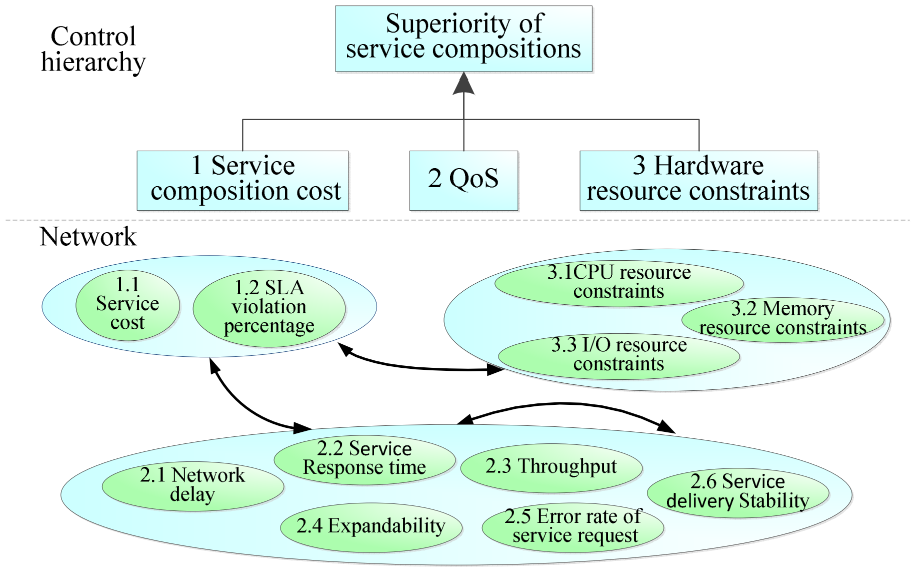

Figure1. The ANP-based metrics for measuring the superiority of service compositions

2.2 Generating of ANP alternative set

The alternative service compositions in ANP alternative set consist of permutation and combination of the atomic service. If ANP alterative set contains all the results of permutation and combination of the atomic services indiscriminately, not only the number of alternatives would soar with the increase of the service complexity, but also the following comparison work would be too large and the judgment process would be too cumbersome for the decision maker.

Besides, there are some infeasible alternatives in all the permutations and combinations of atomic services which contain all possible conditions and corresponding results of the service composition. If the infeasible alternatives and feasible ones are compared together, on the one hand it leads a waste of measuring, work, on the other hand the infeasible alternatives might get higher priority than the feasible ones so that it leads to the failure of evaluation.

To optimize the ANP alternative set, we put forward algorithms of generating all alternative atomic services and generating ANP alternative set. These two algorithms would be applied to filter out the infeasible alternatives. Therefore, on the one hand it would reduce the number of service compositions in ANP alternative set. on the other hand it ensures alternatives are all feasible. The detailed content of these two algorithms would be stated in the third chapter.

2.3 Comprehensive measurement

To build the service composition CPN models, we applied CPN and CloudSim to interactive simulate the models, then the assignment of metric element attributes of ANP alternatives, especially the objective attributes, is accomplished. In other words, the indicator values of alternatives are obtained.

Then we used ANP method to measure the superiority of the alternative service composition. Through pairwise comparing and judging the influence degree between indicators and utilizing a scale of 1 to 9 for the pairwise comparison judgments, the unweighted supermatrix W is obtained. Then we compared the influence degree between clusters in the same way and got the cluster matrix A. Next we applied the formula $W' = a_{ij}W_{ij}$ to obtain the weighted supermatrix W'. At last, the limit super matrix W^{∞} is obtained, that is to say the optimal service composition is got (Satty, 2004). The result is also the service composition that the ANP-based metrics advises service provider to use the service composition to achieve the optimal service effect.

3. Algorithms of generating ANP alternative set

In order to solve the problem of ANP alternatives selection, This section studied how to obtain the available and indivisible services by analyzing cloud service architecture CPN model, put forward an algorithm of generating all alternative atomic services and an algorithm which composites atomic services of generating feasible cloud service compositions, namely ANP alternative set. The overall thought of generating all alternative atomic services and the ANP alternative set by the method of Petri net structural analysis is showed below.

The S invariant set is obtained by analyzing Petri net model of cloud service application system. Then, complete information subnet is got using algorithm of generating all alternative atomic services. As a result, transitions in the complete information subnet can constitute the atomic services set. The T invariant set can be obtained by analyzing complete information flow subnet. Then, applying algorithms of generating ANP alternative set, the set of T invariants and corresponding service compositions which satisfied the user requirement, can be got, namely the ANP alternatives. As shown in the Figure2.

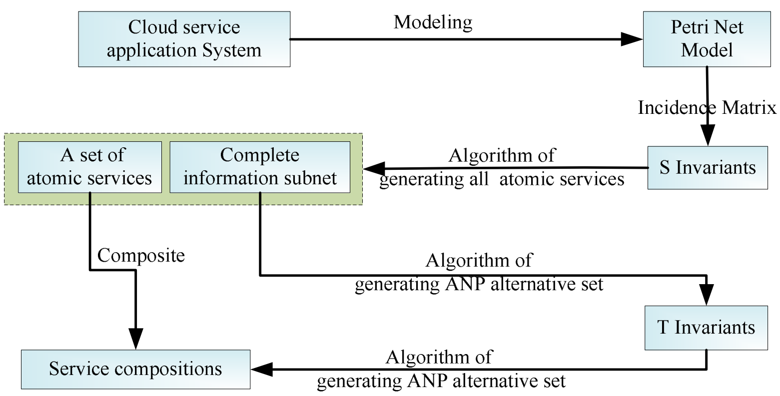

Figure2. Overall thought of generating the ANP alternative set

3.1 Algorithm of generating all alternative atomic services

We definite $PL = \{pl_k \mid k=1 \ldots h\}$ to indicate complete information flow path from the source node to the sink node in Petri net. *AS* is a set of atomic services.

1. Transform the CPN model into OPN model, and then obtain S invariant of the OPN model using the improved FM algorithm. A set of S invariant is indicated as $X=\{x_i \mid i=1 \ldots n\}$ and the S invariant components is indicated as $[x_i]$.
2. Select a sink node P_i.
3. Select all the S invariant components $[x_i]$ which contains the sink node P_i , and then these S invariant components constitute a subnet of the OPN net $[x]=\cup[x_i]$. After that, add the corresponding S invariant into *PL*.
4. $\forall t_j \in [x]$, If $^*t_j \subset [x]$, Then the sink node absorbs all Flow of information, and [x] is the complete information flow subnet; otherwise, go to step 5.
5. Find place P_k which satisfies $(P_k \notin [x]) \wedge (P_k \in {}^*t_j)$. Select all the S invariant components $[x_j]$ which contains the sink node P_k, and then these S invariant components constitute a new subnet $[x] = \cup [x_k]$. After that, add the corresponding S invariant into *PL*, and return to step 4.
6. $\forall p \in P$, If $p \notin [x_i]$ then Find all the paths from the source node to the place *p*, and then add them to the [x], add the corresponding S invariant into *PL*.
7. The transitions in the set of *PL* are added into the set AS, namely atomic service set.

3.2 Algorithm of generating ANP alternative set using Petri net

In the complete information flow subnet which satisfied the user request, there are a variety of possible outcomes of service compositions. We analyze the subnet by the structural analysis of Petri net to obtain the service compositions satisfied the user requirement, Achieving the purpose of getting the feasible ANP alternatives.

Part1:

1. Complete information flow subnet is analyzed to get the T invariants, which indicated by $Y=\{y_i \mid i=1 \ldots m\}$.
2. t_i is indicated to atomic cloud service which is the user input, and t_j is indicated to atomic cloud services which is the users expected output. $YC=\Phi$.
3. If $Y=\Phi$ Then go to step 6. Else go to step 4.
4. $i=1$, $\{y_i \mid i=1 \ldots m\}$. If $(t_i \in y_i) \wedge (t_j \in y_i)$ Then $YC=YC+y_i$ and go to step 5, Else go to step 5.
5. If $i=m$ Then go to step 6, Else $i=i+1$ and go to step 4.
6. End, return *YC*;

Part 2:

We definite *Pr* to indicate a set of source nodes of the Petri net model of cloud service applicant system, *Pc* is a set of the sink nodes and *PT* is a set of the occurrence net.

1. If $YC=\Phi$ Then go to step 8 Else $i=1$ and go to step 2.
2. If $yc_i \subseteq pt_j$ Then $YC=YC-yc_i$ and go to step3 Else go to step 4.
3. If $YC=\Phi$ Then step 8 Else $i=i+1$ and step 2.
4. $\forall t_i \in yc_i$, If $(t_i^* \subseteq yc_i) \cap ({}^*t_i \subseteq yc_i)$ Then $YC=PL-yc_i$ and $PT=PT+yc_i$ and go to step 2 Else go to step 5.

5. $\forall t_i \in pl_i$, If $(t_i^* \not\subset yc_i) \cup (^*t_i \not\subset yc_i)$ Then add places which satisfied $p_i \subset (t_i^* \not\subset yc_i \cup {}^*t_i \not\subset yc_i)$ and arcs between transition t_i and place p_i into $yc_{i.}$
6. $\forall p \in yc_i$, If $(t \subset {}^*P) \cap (t \subset yc_i)$ Then go to step 6 Else add the transitions which satisfied $t \subset {}^*p$ and the places which satisfied t^* and arcs between them into yc_i, constituted $yc_{i1} \ldots yc_{in}$, and go to step 2.
7. $\forall p \in yc_i$ and $p \notin Pc$, If $(t \in P^*) \cap (t \in yc_i)$ Then go to step 2 Else add the transitions which satisfied $t \subset {}^*p$ and the places which satisfied t^* and arcs between them into yc_i, constituted $yc_{i1} \ldots yc_{in}$, and go to step 2.
8. End and Return *PT*. The *PT* is the set of service compositions satisfied the user requirement, namely the ANP alternatives set.

4. Case study

In this section, we will present a case study for applying ANP to find the optimal one from the feasible service compositions to illustrate ANP optimization method selecting high quality service composition and algorithms of generating all alternative atomic services and to generating feasible cloud service composition paths.

.

We had applied our approach to a cloud service system in visitor management of an industrial community. Assuming a scenario is that a visitor, who had reserved, makes the request of registering. Because of the variety of service compositions to register, we find the optimal one in the service compositions using the above method.

(1) Using pragmatic approach with formulated rules, The CPN models of the visitor management of an industrial community is transformed to equivalent Place/Transition nets.

(2) Using algorithm of generating all alternative atomic services, the Petri subnet and the set of all atomic services, with regard to user requirement, are obtained. As shown in Figure 3.

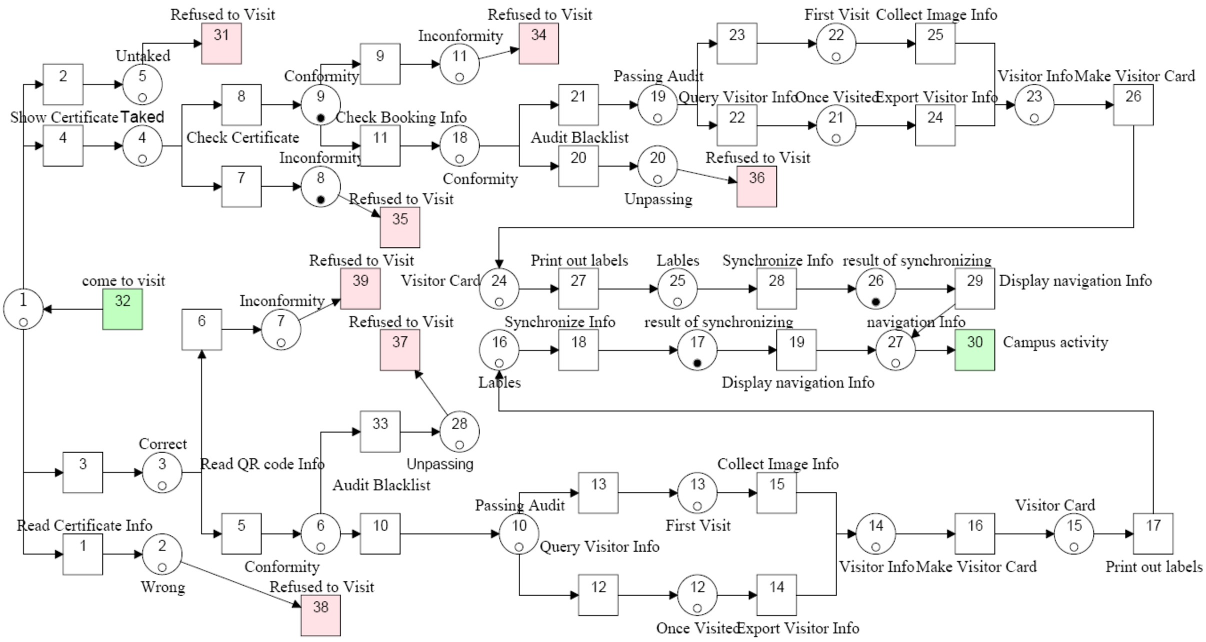

Figure 3. Complete information subnet of the registration service

Notes: In the figure above, the rectangle of Petri net represents transition and the circle of Petri net represents place. Additionality, the transitions correspond to atomic services

(3) A set of T invariants will be obtained by algorithm of generating ANP alternative set. as listed in the following table

Since the user requirement is to have activity in the industrial community, the T invariant which includes transition t1、t2、t6、t7、t9、t20、t33 don’t satisfy the user requirement from analyzing the above T invariants. Applying the algorithm of generating ANP alternative set, we can get the feasible service compositions which correspond to the T invariants like T1、T2、T8、T11.

Table 3. T invariants of registration service

No.	T invariants of registration service	
1	T1={t3, t5, t10, t13, t15, t16, t17, t18, t19, t30, t32}	T3={ t1, t32, t38}
2	T2={ t4, t8, t11, t21, t23, t25, t26, t27, t28, t29, t30, t32}	T4={ t2, t32, t31}
3	T8={ t4, t8, t11, t21, t22, t24, t26, t27, t28, t29, t30, t32 }	T5={ t4, t7, t35, t32}
4	T9={ t3, t6, t32, t39}	T6={ t4, t8, t9, t32, t34}
5	T10={ t3, t5, t32, t33, t37}	T7={ t4, t8, t11, t20, t36, t32}
6	T11={ t3, t5, t10, t12, t14, t16, t17, t18, t19, t30, t32}	

(4) Assignment to the ANP alternatives Indicators
We applied CPN and CloudSim to interactive simulate the models, and then the assignment of metric element attributes of ANP alternatives is accomplished, as shown in Table 4.

Table 4.The value of metric elements

Control Hierarchy	Metric element	Service composition 1	Service composition 2	Service composition 3	Service composition 4
QoS	Stability of service delivery	98.57	98.63	98.88	98.88
	Error rate of service request	0.1	0.09	0.09	0.08
	Throughput	216.6	218.18	216.07	219.62
	Response time of service	11.71	12.83	11.12	11.8
	Network delay	232.68	276.68	121.3	214.8
	Expandability of service	1062	1161	1366	1263
Hardware resource constraints	I/O resource constraints	202440	159840	121464	95904
	Memory resource constraints	43380	39960	48200	44400
	CPU resource constraints	72600000	81840000	91476000	97416000
Service composition cost	SLA violation percentage	0.73	0.65	0.62	0.63
	Service cost	2.088	2.201	1.787	1.957

PS: Service composition1——self-service registering for the regular visitors;

Service composition2——self-service registering for the regular visitors;
Service composition3——registering at the front desk for the new visitors;
Service composition4——registering at the front desk for the new visitors.

(5) Comprehensive measurement
Through pairwise comparing and judging the influence degree between indicators and utilizing a scale of 1 to 9 for the pairwise comparison judgments, we obtained the unweighted supermatrix W and cluster matrix A. The limit supermatrix is obtained by using Super Decisions software, as shown in Table 5. Namely, the indicator weights are obtained. At last, we put the indicator values of corresponding alternatives into the metrics and got the results of comprehensive measurement, as shown in Table 6.

Table 5. The limit supermatrix

E1	**E2**	**E3**	**E4**	**E5**	**E6**	**E7**	**E8**	**E9**	**E10**	**E11**
0.177077	0.177077	0.177077	0.177077	0.177077	0.177077	0.177077	0.17708	0.177077	0.177077	0.177077
0.026276	0.026276	0.026276	0.026276	0.026276	0.026276	0.026276	0.02628	0.026276	0.026276	0.026276
0.080445	0.080445	0.080445	0.080445	0.080445	0.080445	0.080445	0.08045	0.080445	0.080445	0.080445
0.089334	0.089334	0.089334	0.089334	0.089334	0.089334	0.089334	0.08933	0.089334	0.089334	0.089334
0.079276	0.079276	0.079276	0.079276	0.079276	0.079276	0.079276	0.07928	0.079276	0.079276	0.079276
0.050761	0.050761	0.050761	0.050761	0.050761	0.050761	0.050761	0.05076	0.050761	0.050761	0.050761
0.223981	0.223981	0.223981	0.223981	0.223981	0.223981	0.223981	0.22398	0.223981	0.223981	0.223981
0.217646	0.217646	0.217646	0.217646	0.217646	0.217646	0.217646	0.21765	0.217646	0.217646	0.217646
0.010425	0.010425	0.010425	0.010425	0.010425	0.010425	0.010425	0.01043	0.010425	0.010425	0.010425
0.008689	0.008689	0.008689	0.008689	0.008689	0.008689	0.008689	0.00869	0.008689	0.008689	0.008689
0.03609	0.03609	0.03609	0.03609	0.03609	0.03609	0.03609	0.03609	0.03609	0.03609	0.03609

PS: E1: Expandability of service; E2: Throughput; E3: Stability of service delivery; E4: Response time of service; E5: Error rate of service request; E6: Network delay; E7: SLA violation percentage; E8: Service cost; E9: CPU resource constraints; E10: I/O resource constraints; E11: Memory resource constraints.

Table 6. Comprehensive Value of service compositions

Service composition	1	2	3	4
Comprehensive Value	0.240102	0.232556	0.278727	0.248615

As we can see from Table3, the comprehensive value of Service composition 3 is highest, namely the optimal service composition in the ANP alternative set. Service composition 3 corresponds to the service of registering at the front desk for the new visitors and the service composition cost of the alternative is the least. And because the weight value of service composition cost indicator is highest, as we can see from Table2, it is appropriate to select service composition3 as the optimal one.

5. Conclusion

Optimization strategy of cloud computing service composition is an essential part of Cloud computing. This paper addresses Optimization strategy of cloud computing service composition from ANP aspects. It proposes an ANP optimization method to select high quality service composition, which contains algorithms to generate alternative atomic services and feasible cloud service composition paths.

Experimental results show that the algorithms and the proposed method are effective and efficient. Our future work will focus on the stability of Cloud composition system.

REFERENCES

Yong, C.Y. (2005). *Theory and Application of Petri Net.* Beijing, Publishing House of Electronics Industry.

Tao, F., Zhang, L., Guo, H., Luo, Y.l., & Ren, L. (2011). Typical characteristics of cloud manufacturing and several key issues of cloud service composition. *Computer Integrated Manufacturing Systems(03),* 477-486.

Godse, M., Sonar, R., & Mulik, S. (2008). *Web Service Selection Based on Analytical Network Process Approach.* IEEE Asia-Pacific Services Computing Conference, January, 2008.

Huang, S. X., Zeng, S., Fan, Y. S., & Huang, G. Q. (2011). Optimal service selection and composition for service-oriented manufacturing network. *International Journal of Computer Integrated ,Manufacturing, 24(5SI),* 416-430

Luo, J.Z., Jin, J.H., Song, A.B., & Dong, F. (2011). Cloud computing: architecture and key technologies. *Journal on Communications(07),* 3-21

Menascé, D. A., Casalicchio, E., & Dubey, V. (2010). On optimal service selection in Service Oriented Architectures. *Performance Evaluation, 67(8),* 659-675.

Narayanan,S., & McIlwraith. (2002). *Simulation, verification and automated composition of Web services*. 11th International World Wide Web Conference (WWW2002) , 37, 77-88

Saaty, T.L. (2005). The Analytic Hierarchy and Analytic Network Processes for the Measurement of Intangible Criteria and for Decision-Making *Multiple Criteria Decision Analysis: State of the Art Surveys, 78,* 345-405

Saaty, T.L. (2004). Decision making — the Analytic Hierarchy and Network Processes (AHP/ANP). *Journal of Systems Science and Systems Engineering, 13(1),* 1-35.

Li, S., Wu, L.F., & Feng, W.L. (2012).The research on Web composite service based on AHP. *Jiangsu Commercial Forum, 08,* 87-89

Yonggen, G., Wei, Z., & Jie, T. (2012). *A Study of SLA Violation Compensation Mechanismin Complex Cloud Computing Environment.* the Instrumentation, Measurement, Computer, Communication and Control (IMCCC 2012).