International Journal of Grid and Distributed Computing
Vol. 9, No. 12 (2016), pp.213-222
http://dx.doi.org/10.14257/ijgdc.2016.9.12.19

A Study on Big Data Reliable Combination Evaluation Method based on the Cloud Service Qos Model

Li Xiating and Song Rong

Jiangxi Vocational and Technical College of Communication, Nanchang，Jiangxi Province, china, 330000
281838225@qq. com

Abstract

The vigorous development of various technologies in the cloud computation field promotes the development and progress of cloud services. But with the large accumulation of information data in the cloud services, in order to meet user's needs of cloud services multi-function, the operation capacity of a single cloud service is not enough. Now, combining two or many cloud services is a hot research problem. The other hot research problem is that the task of data construction is assigned to different cloud service platforms. However, the combination schemes are too much, so how to ensure the optimal scheme is credible? It has become the key to solve the problem. After reading a large number of literatures, the paper proposes the big data reliable combination evaluation method based on the cloud service QoS model. By getting the contribution of each of the services in the combination scheme set, the method chooses the optimal cloud service combination scheme. Then the optimal scheme is evaluated its reliability.

Keywords*: cloud service, QoS model, big data, reliable combination evaluation*

1. The Research Background

The cloud service, which is a fast-growing emerging technology, is more and more popular in the different fields. There is a great advantage about cloud computation in two parts: cost budget and performance. It uses IT resource usage pattern. It is in accordance with the demand distribution and usage payment model. It not only can reduce the hardware cost when building big data services and maintenance cost, but also can effectively achieve the low cost of big data services. There are some characters about cloud computation: on-demand access, extensible elasticity, usage charging. For the demand of computation and storage resources in the process of big data service component, these characters can be satisfied, and they play an important role in exploring big data services and improving the deployment efficiency. With the fast development of the cloud computation, many applications are built on cloud computation in many enterprises and institutions. The cloud computation is used for data analysis and calculation. Because the fast-growing cloud services have accumulated a lot of resources, so facing a huge of information resources, only two even multiple cloud services together can meet the needs for the application of multi-function. If the application is built on the single cloud service platform, the load requirement of cloud computation is too high. The computation needed by the big data need many satisfactory cloud service platforms to run at the same time. How to choose the best cloud service combination of QoS(Quality of Service) from the cloud service pool of the same function to participate in structure building of big data multi-services, it is one of the hottest issues. Based on the above problem, the paper carries out the research about big data reliable combination services based on cloud service QoS model. It needs the contribution of each subset from cloud service combination set, and then each contribution need to be ranked. The highest

ISSN: 2005-4262 IJGDC

contribution is QoS value optimal cloud service combination scheme. At the end, it completes the reliable evaluation for the combination scheme.

2. Status Quo at Home and Abroad

At present, there is a great breakthrough in the research on Web service combination at abroad. In order to meet the diversified demand, there are lots of achievements on the user-centered research of Web service combination technology. QoS of the non-functional attribute plays a vital role in the Web service combination. For example, Xia Yamei proposes a service combination method about research on several key technologies in dynamic service composition. This method can shorten the candidate service search scope that is corresponding with the abstract service search. The paper proposes an advanced concept of virtual service about cloud service combination research. These services are all real physical connection. Through such a connection, it can increase the success rate of cloud service connection way. There is a big success in the cloud service combination of QoS, when people solve the way of cloud service combination optimizations of QoS perception constraint. People improve the social cognitive algorithm, and propose learning method based on discrete optimization problems. People also do the related experiments to verify its rationality and effectiveness. Based on the above research, it can find that at present the method about cloud service combination is single. The method only research the users' certain attribute, load, the other relevant information of cloud. The research that considers two or even more methods is seldom.

3. QoS Model of Cloud Service

As the similar or the same number of cloud service in the cloud service market increase rapidly, just functional service has not already satisfied the current requirements. So non-functional indexes are needed, and these indexes are used to distinguish the decision. QoS is a evaluation index in the process of selecting cloud service, it is one of the cores of modern network technology, it can measure one service and it is the satisfaction for this service. QoS is a indispensable element in the process of selecting cloud service. It also is an important condition of trade group that among the enterprises and among the enterprises and consumers. And it plays an important role in distinguishing the success rate of service providers. The attribute of QoS is divided into two parts: relevant field attribute and general attribute. The relevant field attribute is that in the different fields of business and services, Qos has the unique attribute of this field. General attribute is that most of the fields will have some of attributes, such as time, price and so on. Because the cloud environment is complex, so QoS attribute is affected by many factors. In this paper, the study on the attribute of QoS model is basis of general attribute and relevant field attribute. Now, many scholars and institutions don't consider the load ability of cloud service platform and two attributes. So this paper will research two attributes.

3. 1. General QoS Model

First of all, the paper researches the general attribute of cloud service QoS model. Research model is shown in figure 1.

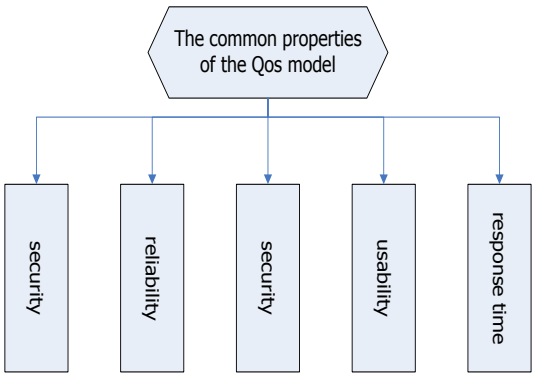

Figure 1.The General Attribute of QoS Model

(1) Time

For users, although there are many ways to measure the performance of cloud services, the most fundamental index, which can meet the demand of users, is use delay and throughput rate of cloud service. Throughput rate represents the service capacity, and it is measured by the number of service requests per time. If the throughput rate is big and delay is small, it represents the service performance is good. Service response time is from start to the end for service uses. Service response time is usually affected by some factors, such as the cloud computation capacity, information transmission time and so on. Cloud computation capacity is affect by the load capacity of server and by information transmission. If the service response is faster, it means that the server running ability is stronger.

(2) Price

The service price is usually set by the developers. The price set by the developers is usually basis of server consumption (If the cloud service platform is rented, the rental fee will be considered). The service price can also reflect that the server performance is good or bad.

(3) Reliability

Reliability is the important evaluation standard of cloud services, especially the reliability of exchange information between service providers and users and the reliability of the service provided by service providers. The service reliability is the success rate of server users. Within the maximum expected time, if the server can respond to behavior and at the same time it can give a correct result, it means it is successful.

(4) Usability

The service usability is base on the status quo, and it is probability that the server can be used or not. The principle of high usability is that the fault is zero. The service usability is usually affected by the virus, network overload and so on. In order to reduce the worry of users for cloud computation, the existing feasibility measure research always uses robustness and accuracy. Service usability is one of important symbols to measure the server performance.

(5) Safety

With the development of cloud computation, there are many crucial tasks in finance and the core of the departments of institutions, so the safety of cloud computation is more and more important. The safety of service is that developers ensure the safety of users' practical information. There are fundamental characters for cloud computation: nothingness, liquidity, and borderless. So it faces many risks: data disclosure, misuse, evil and so on. Because cloud service is public, so ensuring the confidentiality of user information is important.

3. 2. The Load Capacity of Cloud Service

The above mentioned, QoS not only has general attribute but also has the relevant field attribute. In the field of cloud service, QoS has not only above mentioned general attribute such as time and price, but also has fielded attributed such as service system load capacity. Because people has the requirements of cloud service function diversification and the requirements of high performance, so cloud computation need lots of information resources and a lot of running memories. The system load capacity has already become the hottest attribute in the cloud service field. Load capacity is also one of the important standards to measure the service level of service system. The strength of the load capacity can directly affect the service reliability and even the credibility of big data combination based on cloud service.

Based on the current study, server load capacity can be affected by three factors: CPU load, network load, and memory load. CPU load is the load degree caused by the CPU usage for CPU. Network load is the load degree caused by the network usage for network. Memory load is the degree caused by memory usage for memory.

3. 3. The Dependence of the Cloud Service

From the abstract perspective, the cloud computation is mainly made up of three levels of abstraction. Usually, users directly see the top level. The current way about cloud service mainly includes two kinds, as shown in figure 2.

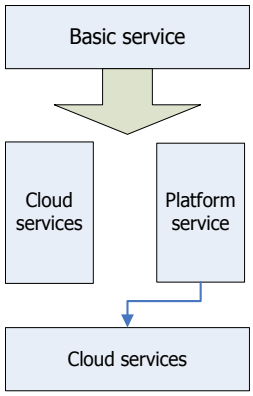

Figure 2.The Realization Form of Cloud Service

From picture 2, the first realization form of cloud service is cloud services-platform service-basic service. This realization form is that developer's rent cloud service platform in the cloud environment to develop application. Such developed cloud services, its QoS attribute not only has a relationship with service itself, but also has a relationship with the rent cloud services platform, with a certain degree of dependency. In this situation, for developers, the price must takes account into the costs of service itself consumption and the costs of renting platform. For users, when they use it, they not only consider the time of cloud computation consumption, but also consider the time delay caused by renting cloud services platform. To sum up, QoS will be affected by platform dependency.

The other form is relatively simple, developers directly research on their own cloud service platform. The QoS attribute is affected by the infrastructure. Such realization form is similar to well-known Web service realization form. The paper builds a complete model based on general attribute and relevant field attribute, as shown in figure 3.

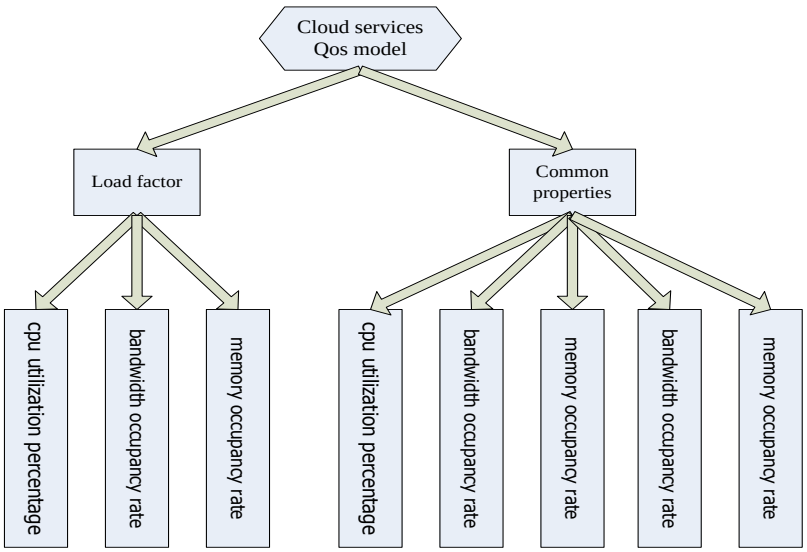

Figure 3. QoS Complete Model

3. 4. The Utility Function in the Model

Assuming the mission planning P of big data service, d dimension vector represents the corresponding QoS constraint.

$$E = \{E_1, E_2, \cdots, En, \cdots, Ed\}, \ (1 \le n \le d).$$

$En(1 \le n \le d)$ is value of expectation of the application requirements for service on the dimension QoS. The paper will give the definition of feasible combination cloud service scheme:

Defining a big data mission planning:

$$P = \{P_1, P_2, \cdots, P_n\}$$

A group of big data service QoS constraints:

$$E = \{E_1, E_2, \cdots, En, \cdots, Ed\}, \ (1 \le n \le d)$$

A feasible cloud service scheme:

$$AS = \{S_1, S_2, \cdots, S_n\}, \ s_m \in S_m (1 < m < n)$$

If the aggregate value of AS satisfies $q_i(AS) \le C_i$, so AS is a feasible combination scheme.

Generally, the given task planning r has many possible combinations. In order to pick out a optimal combination scheme QoS, it need the utility function of combination cloud service scheme to select in large number of feasible schemes, to participate in the big data mapping and deploy. As we all know, the combination evaluation of big data service is the process of selecting QoS optimal scheme in lots of combination schemes.

Combination evaluation is that big data service is evaluated by the utility function, and then evaluates the QoS combination scheme. Then these combination schemes are reasonably ranked by the utility function, the top one is a optimal combination scheme.

4. Big Data Reliable Combination Evaluation Method

4. 1. Reliable Combination Evaluation Method based on QoS

First of all, people need to define the combination of history to analyze the credibility of the evaluation. The definition of combination is as follows:

a big data mission planning

$$P = \{P_1, P_2, \cdots, P_n\}$$

A combination scheme

$$SCPLAN - K_J = \{S_1 \cdot K_1, S_2 \cdot K_{2,\cdots}, S_m \cdot K_m, \cdots S_n \cdot K_n\},$$

And $s_m \in S_m (1 < m < n)$. $S_m \cdot K_m$ is one historical record of S_m.

The above definition is the foundation; we suppose that there is n_m service set that is corresponding of any task P_m of people planning P. For S_m, each service $S_{mj}(1 \le j \le n_m)$ contains b_{mj} historical records. So, in the S_m, the total number of historical record is $\psi_m = b_{m1} + b_{m2} + \cdots + b_{mn}$. To sum up, based on QoS record, the number of combination scheme is $\psi = \psi_1 \times \psi_2 \times \cdots \times \psi_n$.

Through the above definition and analysis, then building model for combination evaluation is used by 0-1 integer programming method:

$$U(VG\_AS) = \sum_{r=1}^{d} \frac{W \max(r) - \sum_{m=1}^{n} \sum_{j=1}^{n_m} \sum_{v=1}^{b_{mj}} q_d(s_{mj} \cdot G_v) \times x_{mj-v}}{W_{\max}(r) - W_{\min}(r)} \times Q_r$$

$$\sum_{m=1}^{n} \sum_{j=1}^{lm} \sum_{v}^{b_y} q_r(s_{mj} \cdot G_v) \times x_{mj-h} \le C_r, 1 \le v \le d$$

$$\sum_{j=1}^{n_m} \sum_{v=1}^{b_{mj}} x_{mj-h} = 1, x_{mj-h} \in \{0,1\}$$

$$\sum_{r=1}^{d} Q_r = 1, Q_r \in [0,1]$$

The quality of combination scheme is evaluated by the definition of the combination utility function to show the value of Qos optimal combination scheme. The definition is following: a big data mission planning $P = \{P_1, P_2, \cdots, P_n\}$, each service combination $AS = \{s_1, s_2, \cdots, s_n\}$, $s_m \in S_m (1 < m < n)$. The utility function value of AS is that AS satisfies the average value of the constraint condition of the utility function.

Above all, the definition of reliable combination evaluation method based on QoS historical record is got. The definition is following:
a big data mission planning: $P = \{P_1, P_2, \cdots, P_n\}$.

Fist of all, list all cloud service combination schemes, then for each combination scheme, and calculate its utility function value. The scheme, its value of utility function is the highest, is the optimal combination of Qos values, and is the final evaluation result of credible combination service.

4. 2. Reliable Combination Evaluation Method based on Contribution Degree

In the actual situation, usually there are large numbers to meet cloud services, and each cloud service can accumulate a large number of QoS historical records. Based on the method in the section 4. 1, it will cost a lot of calculation in the implementation process, and it will lead to greatly increasing of cloud platform load. In order to improve the work efficiency of cloud computation, the paper proposes the concept named contribution and it can't affect the evaluation result. Based on the parts of QoS historical records combination scheme the reliable combination is evaluated, this result is contribution degree. In contrast, if its contribution degree is highest, it is final reliable combination evaluation result. First, people should calculate the combination scheme of the highest utility function value. There are Z of function values. Then, in these combination schemes, people can choose a scheme that its contribution degree is the highest. The definition of contribution degree of combination scheme is: a big data mission planning $P = \{P_1, P_2, \cdots, P_n\}$, each possible combination cloud service $P = \{P_1, P_2, \cdots, P_n\}$, each service combination $AS = \{s_1, s_2, \cdots, s_n\}$, $s_m \in S_m (1 < m < n)$. The contribution degree of AS is the average value of utility function of combination scheme based on QoS historical records combination schemes. First, people should find the method of combination scheme of the highest value of utility function, shown

as figure 4:

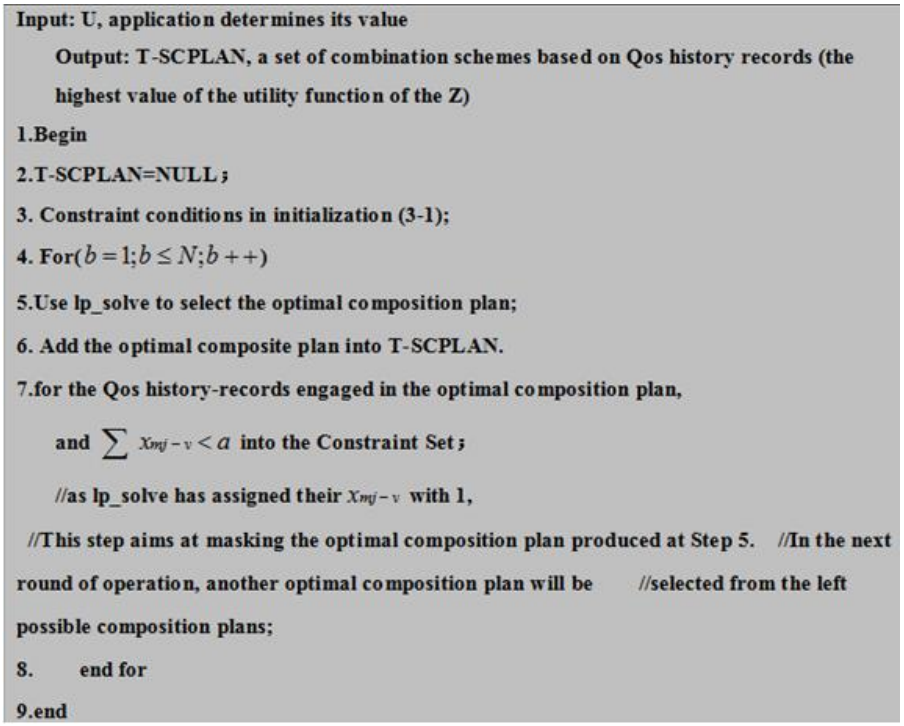

Figure 4. Find out Z Most Efficient use of the Value of the Algorithm

In combination with the contribution degree of cloud service combination scheme, the implementation of this method includes three basic steps:

(1)Using the algorithm 3.2, get the combination scheme based on Z QoS historical records that is corresponding of big data mission planning, and record their average utility function valence.

(2)Calculate the contribution degree of each cloud service combination schemes, based on Z combination schemes of QoS historical records.

(3)From all the contribution degrees of combination schemes select the highest one, and take it as QoS optimal combination cloud service scheme. The optimal combination scheme is participate in the construction and deploy of big data.

Method 1 is compared with method 2, people can find that although the process is not significantly streamline, LP_SOLVE branch and bound algorithms used in the method 2 can not only reduce the number of data records, but also can improve the calculation efficiency.

5. The Experiment Evaluation

5. 1. The Experiment Configuration

In order to verify the reliable combination method based on the contribution degree, described in 4.2, now we use an experiment to verify from complex rate and optimal rate. In building the cloud environment, the big data mission planning $P=\{P_1,P_2,P_3,P_4\}$ is taken as an example, the each corresponding task $P_m=\{1\le m\le 4\}$, the combination scheme number set is $\{3,4,5,6\}$, its corresponding historical records number set is $\{15,20,25,30,35\}$, the result is the scope of QoS general attributes(price, time, safety, reliable, usable).

Table 1. General Attribute Values Range

General attribute	price	time	safety	reliable	usable
Values range	[40，45]	[5，10]	[0. 2，1. 0]	[0. 5，1. 0]	[0. 3，1. 0]

In order to reach the goal of experiment evaluation, the needed historical records is selected within specify QoS internal value, according to the random principle, and QoS historical record set is got by the synthetic way. In order to simulate the cloud service QoS value distribution in the real cloud environment, we take the random method to generate the required QoS history records, for example table 1 lists each dimension QoS attribute and the value range generated randomly.

5. 2. The Experiment Results Analysis

To sum up, on the basis of the given the number of cloud services, the consumption time of combination evaluation is proportional of the growth of the number of QoS historical records. But method 4.1 needs to list all the combination schemes, so its consumption time exponential level growth. For method 4.2, because it can eliminate most of data that is not in conformity with the conditions before the calculation, so it greatly reduces the amount of calculation, and reduces a lot of the time consumption. In comparison, the method 4.2 can save more time and is more efficient.

Reliable combination evaluation scheme based on the contribution degree is named the most feasible technical scheme. So in the cloud environment, this method can not only provide the reliable combination evaluation, which can run the data with the high efficient, but also greatly shorten the wasting time of reliable combination evaluation based on QoS historical records, and won the high quality cloud service combination scheme at the same time.

5. 3. Comparison and Analysis

For the reliable combination evaluation scheme based on the contribution degree and the traditional service combination, the focus of the former one is that the complex calculation is ran by the loosely coupled way based on the existing cloud services. It not only can achieve faster and new development, but also can provide the reliable calculation model for the deploy service application. The latter one is that it takes the combination evaluation as the typical problem to research. How to select the optimal QoS or the optimal scheme in the lots of sets that meet all the requirement of service function? This is the focus attention for the latter one method.

With the development of the cloud technology, the traditional services face the increasingly serious problems, such as dynamic network environment and commercial that is possible made by service providers. It will appear that the QoS value provided by service providers is not conforming to the actual value. So the combination evaluation scheme, based on the QoS value provided by the service providers, can not be completely convinced. Combining the service resources through the different platforms can meet the requirement of data computation resources and storage resources. Big data service reliable combination evaluation method based on QoS historical records not only can reduce the cost of building big data and improve efficient, but also provides the new calculation model and technology.

6. Research Results

The paper proposes big data service reliable combination evaluation method based on QoS historical records to improve the credibility of the big data service combination scheme. Because this method may consume lots of time in the process of combining, we propose a reliable combination evaluation method based on the contribution degree. It means that we select the parts of the more reliable and optimal schemes in the all QoS historical records to evaluate, to shorten the calculation time of combination evaluation and improve the work efficiency. The above experiment results show that comparing with the traditional method, the method based on the contribution degree has the obviously advantage on the execution time, the reliable combination evaluation method based on the contribution degree can get the high optimal degree in most cases at the same time.

7. Conclusion

The cloud service is in the better tendency development, the data of the more application service will be not only limited to distribution in a single cloud service platform. Multiple cloud service platforms combination is not simple superposition. Because each platform has its own capacity, it will show the biggest advantages, to maximize the application performance.

There are many cloud service platforms with the same function and performance, but the performances are different. The one of the development trend is selecting the optimal cloud service combination scheme from the cloud service pool of having huge number of cloud services. Evaluating the reliable of the selected combination schemes is an essential tache for every company or institution before developing. This paper proposes big data reliable combination evaluation method based on cloud service QoS model, and it meet the requirement of combination scheme evaluation method for the current developers.

References

[1] A. H. Lzeng, H. Ngu, B. Benatallah, R. Podorozhny and H. Lei, “Dynamic composition and optimization of Web services”, Distributed and Parallel Databases, vol. 24, no. l-3, **(2008)**, pp. 45-72.
[2] J. Cardoso and C. Bussler, “Semantic Web Services and Processes: Semantic Composition and Quality of Service”, On the Move to Meaningful Internet Computing and ubiquitous Computer, Rvine CA, **(2002)**.
[3] D. A. Menase, “QoS Issues in Web Services”, IEEE Internet Coinputing, vol. 6, no. 6, **(2002)**, pp. 72-75.
[4] M. C. Jaeger, G. J. Rojec-Goldmann and G. Muhl, “QoS aggregation for Web service composition using workflow pattems”, Proc. of the 8th IEEE Int. Conf. on Enterprise Distributed objects Computing (EDOC), Caledonia: Monterey, **(2004)**.
[5] A. F. M. Huang, C.-W. Lan and S. J. H Yang, “An optimal QoS-based Web service selection scheme”, Information Science, vol. 179, no. 19, **(2009)**, pp. 3309-3322.
[6] C.-F. Liao, Y.-W. Jong and L.-C. Fu, “Toward reliable service management in message-oriented pervasive systems”, IEEE transactions on service computing, vol. 4, no. 3, **(2010)**, pp. 183-195.
[7] D. A. Menasce, E. Casalicchio and V. Dubey, “On optimal service selection in Service Oriented Architectures”, Performance Evaluation, vol. 67, no. 8, **(2010)**, pp. 659-675.
[8] A. Dempster, “Upper and lower probabilities induced by multivalued mapping”, Annals of Mathematical Statistics, vol. 38, no. 2, **(1967)**, pp. 325-339.
[9] D. Yong, W.-K. Shi and Z.-F. Zhu, “Combining belief functions based on distance of evidence”, Dicision Support Systems, vol. 38, no. 3, **(2004)**, pp. 489-493.
[10] X.-H. Gu and K. Nahrstedt, “On composing stream on parallel and distributed systems”, IEEE Transactions on Parallel and Distributed Systems, vol. 17, no. 8, **(2006)**, pp. 824-837.

Authors

Li Xiating, She is born in June, 1979. She is an associate professor. Her research direction is computer science. She works at Jiangxi Vocational and Technical College of Communication in Nanchang City, Jiangxi Province now. She host a project named a study on the application of association rules mining in Sakai platform personalized service, colleges and universities in Hu'nan Province education department scientific research project subject in general, 15C0646.