A New Expectation Trust Benefit Driven Algorithm for Grid Environments

Yongsheng Hao[a], Guanfeng Liu[b], Yuebin Xu[b] and Junyan Wang[b]
[a] *Network Center, Nanjing University of Information Science & Technology*
hyslove@163.com
[b] *Computer Software Lab, Qingdao University*
guanfeng.liu@yahoo.com, xyb21@qdu.edu.cn

Abstract

Since the resource of nodes can access and leave the Grid system freely, it is possible some malicious nodes exist in the Grid environment. In addition, the numbers and qualities of nodes' resource can be changed dramatically and optionally, so these malicious nodes must be affecting the utilization of Grid resource. Hence, the trust management mechanisms are developed and widely used in the resource scheduling for Grid environments. However, most of the trust management models are mainly focus on the nodes historic behaviors and few of them consider the future trust of the node based on its previous activities. This paper proposes a new method to detect the nodes behaviors of resource providers in Grid environments based on D-S theory. Through simulating experiments, the new mechanism could record the behaviors exactly and prevent the malicious ones accessing the Grid system effectively. It would give a strong support for the success of resource scheduling in Grid environments.

1. Introduction

In the nowadays research, the Web Service Resource Framework [1] (WSRF) comes out with the combination of Service-Oriented Architecture [2] (SOA) and conventional Grid techniques. The platform based on WSRF [3] tries to confirm all of resource on the internet, which includes computing, storage, information and knowledge. It virtualizes them and makes them easy to be used at anytime in anywhere. However, it also makes a difficulty thing to manage the resource scheduling in Grid environments because of the dynamic, heterogeneous, distribution and open of Grid environments [4]. Therefore, trust models are good approach used to solve these problems.

There have been a lot of significant research achievements in trust management in Grid environments. For example, A.Dogan [6] and S.Song [11] proposed a minimization failed rate resource scheduling architecture and algorithm. In addition, G.Liu [12] [13] presented some reputation architectures in the Grid economy. Those researchers give the way of resource scheduling and trust management mechanisms from different views. In this paper, we propose the representation and updating mechanism of trust function and likelihood function. They could calculate the nodes' trust through detecting the cost of nodes' receiving trust in Grid environments. In addition, we have proved the speculation trust function is sensitive and timely in simulating experiments, which could detect the trust situation of resource providers in Grid environments.

The rest of this paper is structured in the following sections. The related work is briefly discussed in the section 2.Section 3 gives the presentation of trust function and the likelihood function. In the Section 4, we make an analysis and give the trust detection mechanism. The algorithm of resource scheduling is presented in the section 5. Section 6 is the simulation implementation and results. Section 7 concludes the paper and suggests future direction for improvement.

2. Related work

Different approaches, policies and architectures are developed in the trust management for Grid environments. M.Blaze proposed the Trust Management (TM) conception [5] in 1996, which gave us a new thinking to solve the security problems in distributed environments. W.H.Winsborough [7] defined these TM systems as authority systems based on capability. However, in these systems the resource providers still need award the certificate of authority to the resource consumers. So it is difficulty to construct the dynamic trust relationships among strangers. It is a good way to build this relationship based on the host authority [8]. N.H.Li designed a role-based trust management framework (RT) [9] [10]. The nodes were divided into different roles according to the different functions of different behaviors in the strategies. E.Damiani [14] adopted distributed polling algorithm to share the trust information. However this system still has disadvantages because of the randomicity of polling and the huge expense of the system. M.Gupta [15] calculated the reputation by using reputation computation agents. He adopts debit-credit reputation computation and credit-only reputation computation. The dynamic trust metric [16] adopted a method named forget factor to punish malicious peer. These models and polices indeed push the development of the trust management in Grid computing, However, nodes future behaviors can not be detected exactly and the resource scheduling can not be instructed by the nodes future behaviors in these research.

In this paper, we propose a new method to detect the future behaviors of resource providers in Grid environments based on D-S theory. In addition, we present an algorithm based on the nodes behaviors detection method.

3. Uncertain inference and its representation

In this section, the trust function will be given to forecast the Grid nodes' behaviors. Then we can get the forecast trust function base on this.

3.1. Uncertain inference theory

In Grid environments, the trust relationship among nodes in Grid is uncompleted, inaccurate and unreliable. The theory of D-S can be used in the field of uncertain inference. So we analyze the jobs scheduling in Grid according to functions in D-S theory. Usually we use section [Bel(A),Pl(A)] to describe the unknown extensions of A.

3.2. Trust function and likelihood function

In Grid environments, it is a hard thing to confirm the trust level of Grid nodes. The trust level of node GP_i in the Grid is determined by the job which is executed on it. Before the resources scheduling, we assume that the GP_i promises to provide the number of services is S_{ij} which include the number of promised credible services is R_{ij} and the services that actually provided number is T_{ij}. Since we always hope every node in the Grid can notice the trustworthy of local services before execution. Then we will make a discussion according to the next three situations.

1. $R_{ij} > T_{ij}$

In this situation, the number of trustworthy services provided actually is less than the number that promised before the execution. We define the trust function and the likelihood function in this scheduling as follows,

$$Bel(A) = \frac{T_{ij}}{R_{ij}}, \quad Pl(A) = Bel(A).$$

2. $R_{ij} < T_{ij}$

$$Bel(A) = \frac{R_{ij}}{T_{ij}}; \; Pl(A) = 1.$$

3. $R_{ij} = T_{ij}$

This is the ideal situation. It can guarantee the deadline of jobs and the utilization of the resources in the Grid completely.

$$Bel(A) = 1; \; Pl(A) = Bel(A).$$

3.3. Update strategies

We hope get the trust function and likelihood function of GP_i from the history of its transactions. Then we can use these functions to predict the trustworthy of Grid nodes. In the history of Grid nodes' transactions, we deem the latest data has a better reference value. So we give different weights of transaction records in different times. The latest, the higher value of weight will be $x_i(x>1)$. Provided that the node GP_i gives the trust function as above policy is $Bel_1(A)$, $Bel_2(A)$, $Bel_3(A)\ldots Bel_n(A)$ and the likelihood function $Pl_1(A)$, $Pl_2(A)$, $Pl_3(A)\ldots Pl_n(A)$.Then the affection factor is $x_i(x>1)$, the prediction trust function $Bel^n{}_i(A)$ and the prediction likelihood function $Pl^n{}_i(A)$ of GP_i can be served as the basis of the n+1 times transaction. It is shown as follows.

$$Bel_i^n(A) = \frac{\sum_{t=1}^{n} x^t Bel_t(A)}{\sum_{t=1}^{n} x^t} = \frac{Bel_i^{n-1}(A) + \frac{x^n Bel_n(A)}{\sum_{t=1}^{n-1} x^t}}{1 + \frac{x^n}{\sum_{i=1}^{n-1} x^t}} \tag{1}$$

The formula of above shows that we can predict the nodes' trust function from its history data and use the function to predict the actual situations of Grid nodes. Actually not all history data have reference value. We only deem the data of n latest times as the judge standard. At the same time, we should prevent the Grid node change its trustworthy behavior for its interest. We inspect the situation that the trustworthy of Grid node reduce from Bel_1 to Bel_2 $(Bel_1 > Bel_2)$. If the node makes n-1 times transactions based on the trustworthy is Bel_2, then the permissible error of prediction trust function is ζ and the trustworthy estimation of node is Bel_2.

$$Bel_2(A) + \zeta = \frac{(xBel_1(A) + \sum_{t=2}^{n} x^t Bel_2(A))}{\sum_{t=1}^{n} x^t}$$

$$\sum_{t=0}^{t=n-1} x^t = \frac{Bel_1 - Bel_2}{\zeta} \qquad (2)$$

Suppose the value of factor ζ can fluctuate in the range of $[0,0.05]$. Because of $(Bel_1 > Bel_2)$, then, $\max(Bel_1 - Bel_2) = 1$, and because of $x > 1$, we let n=5 and then $x > 1.75$. In the mean while, we must prevent the situation that the Grid node improves the value of trust function greatly from Bel_1 to Bel_2 $(Bel_1 < Bel_2)$ in once transaction. We set the value of trust function is not higher than $(Bel_1 + Bel_2)/2 + \delta, \delta \in (-0.05, 0)$. That is to say the value of trust function of one Grid node can not be totally changed based on one times improvement of its trustworthy. Then,

$$\frac{(Bel_1 + Bel_2)}{2} + \delta \le \frac{(Bel_1 \sum_{1}^{n-1} x^i + x^n Bel_2)}{\sum_{1}^{n} x^i}$$

$$\frac{\sum_{1}^{n} x^i}{\sum_{1}^{n-1} x^i} \ge \frac{2(Bel_1 - Bel_2)}{(Bel_1 - Bel_2 - 2\delta)}$$

$$\frac{\sum_{1}^{n} x^i}{\sum_{1}^{n-1} x^i} \ge 2 \qquad (3)$$

Set $n = 5$, then $x \le 2$.
So $1.75 \le x \le 2.0$

In the same method, the prediction likelihood function is as follows,

$$Pl_i^n(A) = \frac{\sum_{t=1}^{n} x^t Pl_t(A)}{\sum_{t=1}^{n} x^t} = \frac{Pl_i^{n-1}(A) + \frac{x^n Pl_n(A)}{\sum_{t=1}^{n-1} x^t}}{1 + \frac{x^n}{\sum_{t=1}^{n-1} x^t}} \tag{4}$$

4. Trust detection mechanisms

For predicting trust function $Bel_i^n(A)$ and trust function $Bel_n(A)$:

$$|(Bel_i^n(A) - Bel_n(A)| < \partial \tag{5}$$

∂ is the variation range of Grid credibility function. Grid node changes in the permitted range and there are not suspicious activities.

$$|(Bel_i^n(A) - Bel_n(A)| > \partial \tag{6}$$

Grid node may change its trust function with strategy and there may be suspicious activities. Under job scheduling, we don't know the number of credible services in Grid node GP_i which supplies actually and we can only confer by trust function of GP_i. If the Grid node has dealt n times and received the prediction trust function is $Bel^n_i(A)$, it may confer credible services amount in Grid nodes:

$$T_{ij}^{'} = Bel_i^n * R_{ij} \tag{7}$$

Base on this information, we can define a parameter to get trust loss GBL_j for every Grid nodes:

$$GBL_j = \sum_{t=1}^{j} T_{it} - T_{it}^{'} \tag{8}$$

This value reflects Grid node which pays out superfluity loss for the sake of receiving trust, if this value less than zero, this Grid node may exist shadiness action. Then we may define this Grid node trust function $Bel_j(A)=0$.

The value reflects the extra losses of the Grid node paid for getting trust. If the value is less than zero, the Grid node must have suspicious activity and the value of the trust function of the Grid node can be set zero. If Tit>Tit', the scheduling will be successful. Sucn indicates the quantity of successful trades. $Sucm = \sum_{t=1}^{l} T_{it}^{'}$ is the quantity of resources in l trades. The expression below indicates the efficiency of resource utilization in all scheduling.

$$Ratr_i = \frac{Sucm}{\sum_{t=1}^{l} T_{it}} \tag{9}$$

5. Expectation trust benefit driven algorithm (ETBDA)

Expectation trust benefit driven Grid job scheduling heuristic arithmetic has two stages. Firstly, we calculate the max expectation benefit function value of every demand-for-trust jobs in Grid nodes, then map jobs and resources which has max expectation benefit when reliable service is not wasted. After that, we can update the trust function and supervise the reliability of Grid node by adopting the parameter GBL_j. If the GBL_j is negative, the Grid node has to regain the trust certification. If there is no job which can satisfy the trust requirement, then we schedule the jobs which have no trust requirement by Min-min algorithm until a Grid node can provide the adequate and reliable service number for a certain job. The algorithm is described as follows:

Input: the trust information of jobs and resources, CT matrix

Output: jobs mapping precept

1. **Initialization**: $GPS = (GP_{1,} GP_{2,} ... GP_i ... GP_m)$ $Jobs = (Job_{1,} Job_{2,} Job_{3,} ... Job_j ... Job_n)$, CT is used to save the pairs of job-resource;

2. **Repeat**

3. Set CT=null;

4. **For** all $Job_i \in Jobs$ and $Job_i \in$ Strict demand-for-reliability jobs

5. SetGP(Jop_i);

6. **End For**;

7. **For** all $Job_i \in Jobs$ and $Job_i \in$ demand-for-reliability jobs

8. SetGP(Jop_i);

9. **End For**;

10. Execute according to Min-min algorithm until finding a resource which can satisfy the reliability of some job;

11. **Until** (Jobs is empty);

The function of SetGP(Jop_i) is given as follows,

Function: SetGP(Jop_i)

1. **For** all $GP_i \in GPS$

2. **If** resource GP_i can satisfy the job's reliability requirement according to speculating trust function

3. Calculate Jop_j expectation trust benefit $p(i, j)$ of GP_i;

4. **End If**;

5. **End For**;

6. Get the maximum of expectation benefits of Jop_j;

7. Find the resources whose expectation benefit equals to the maximum, if there are more than one Grid nodes satisfying the trust requirement for maximum, the Grid node which offers the minimum reliable services will be chose and the job-resource will be saved into CT;

8. Execute (GP_i, Job) and update trust function of GP_i ;

9. **End For**;

6. Simulation experiments and results

6.1. The rationality of influence factor

We suppose x=1.2 and make a study of the two situations: the Grid node from Bel(A)=0.3, Pl=0.4 to Bel(A)=0.8, Pl(A)=0.9 and another Grid node from Bel(A)=0.5, Pl=0.6 to Bel(A)=0.2, Pl=0.3 (Figure 1). It can be seen from the figure that the prediction function can well reflect the change of sensitivity of Grid nodes.

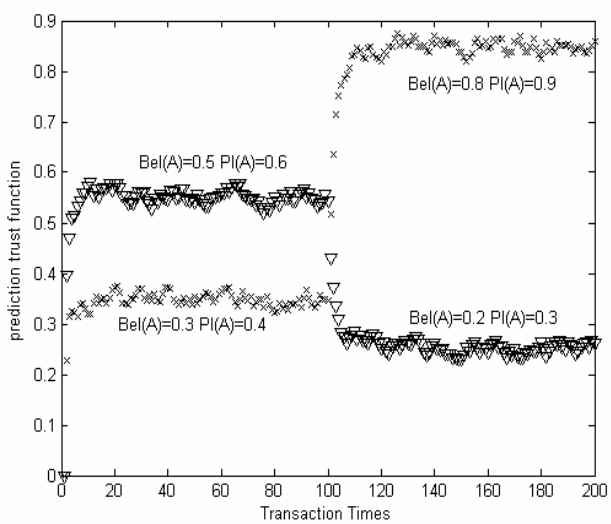


Figure 1: Sensitivity of prediction function to the Grid nodes' behaviors

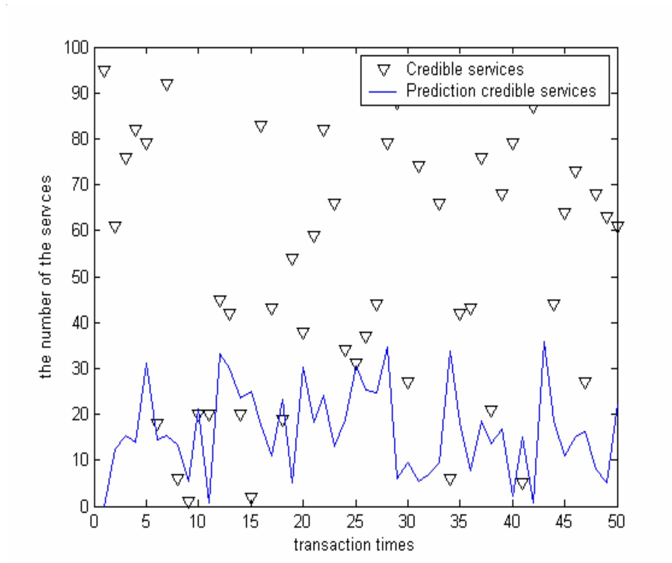


Figure 2: Sensitivity of prediction credibility function to random behaviors

The special situation is considered. Supposing that certain node is not responsible for its behavior at all, in another word, it has randomness. We study the accuracy of prediction trust function. The note can provide 100 Grid services. If the number of prediction reliable services is smaller than the number of credible services provided, Grid resource can be used. We can know from the experiment and the analysis in section 3 that only 6 resources scheduling in 50 fail. But in this case, the efficiency of resource utilization is very low and it is 22.90% (Figure 2).

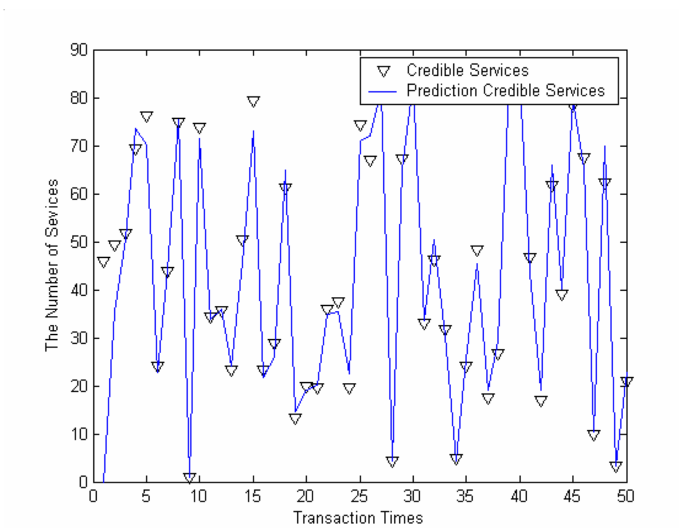


Figure 3: Sensitivity of prediction credibility function to random behaviors

When it comes to the prediction ability of Bel(A)=0.8, Pl(A)=0.9 to the quantity of the resources, it can be seen from the Figure 3 that only four resource scheduling fail. The efficiency of resource utilization is 81.45%.

6.2. Comparison of the ability of detection

We study the detection ability of Grid resource in 5 cases which are as follows, $Bel(A)=0.3$ $Pl(A)=0.4$, $Bel(A)=0.5$ $Pl(A)=0.6$, $Bel(A)=0.8$ $Pl(A)=0.9$, $Bel(A)=Pl(A)=1$. According to the section 3, if $T_{it} > T_{it}^{'}$, the scheduling will be successful. *Sucn* is the times of successful transaction and $Sucm=\sum_{i=1}^{l} T_{it}^{'}$ is the resources quantity of successful transactions in l transactions. The expression below is the utilization ratio of resources in all scheduling.

Table1: The comparison of different values of trust function

Bel and *Pl*	*Sucn*	*Sucm*	Total Resource	$Ratr_l$
Random	809	13798	49433	27.9 %
$Bel(A)=0.3, Pl(A)=0.4$	794	9298	15149	62.9 %
$Bel(A)=0.5, Pl(A)=0.6$	802	20891	29733	70.2 %
$Bel(A)=0.8, Pl(A)=0.9$	911	32706	45795	71.4 %
$Bel(A)=1, Pl(A)=1$	1000	45601	45672	99.8 %

We suppose that a Grid note makes 1000 times of Grid resource scheduling. It is shown that despite of the difference of the trust function, success rates of resources scheduling are all comparatively high. But the utilization efficiency of resources are different. If trust function of a Grid note has a higher value, the utilization efficiency of resources will be higher. The reason is that the higher credibility of a Grid note has the more exact estimation of the quantity of a Grid node.

6. 3. Comparison of two algorithms

This trial inspects the situation that the Grid system is composed of 1000 computation services to the work connection consisted by 1000, 2000, 3000, 4000, 5000, 6000, 7000 jobs (Figure 4). Each Grid node provides reliable services by some random number (an integer from 1 to 100), then the number of reliable services is allowed to change in [-0.05, 0.05], and it supposed that 5% Grid node's behaviors are random. It supposed that half of the job hasn't trust request by contrast to the other half, the trust request of which is a random numeral (an integer from 1 to 100). For demand-for-trust job, if the confidence level satisfied each job's execution time is 1; the time will lengthen (computing the ratio of the requested reliable services to provided reliable services).

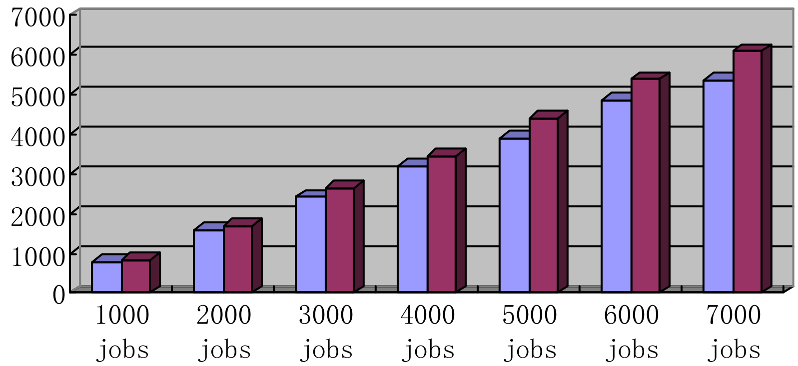


Figure 4: Benefits of two algorithms

We find that the (ETBDA) can enhance the benefit (Figure 4) along with the job number increasing, the enhancement will be more apparent. The reason is that we can avoid the benefit losing partly and reliable services' wasting. Along with the job number increases, the prediction trust function can provide accurate forecast. Then more enhancements will be seen in the work benefit.

7. Conclusion

The paper proposes a method to detect the supply situation of Grid resources based on D-S theory which is monitored by trust function $Bel_i^n(A)$ and trust lost GBL_j. In addition, we propose the representation and updating mechanism of trust function and likelihood function, which calculate the nodes' trust through detecting the cost of receiving trust of nodes in Grid environments. What's more, we have proved the speculation trust function is sensitive and timely in simulating experiments. In the Grid environment, how to deal with fault-tolerant of the unreliable Grid service and enhances the use factor of the entire Grid system is a future problem waiting for research.

References

[1] A.Oram, "Peer-to-Peer: Harnessing the Power of Disruptive Technologies", *New York: O'Reilly & Associates, Inc Press*, 2001.

[2]F.Azzedin and M.Maheswaran, "Integrating trust into grid resource management systems, International Conference on Parallel Processing, IEEE Press, Canada 2002, pp.47-54.

[3] M.P.Papazoglou and D.Georgakopoulos, "Service-oriented computing", Communications of the ACM, 2003, 46(10), pp.24-28.

[4] I. Foster, "the Physiology of the Grid-An Open Grid Service Architecture for Distributed Systems Integration", Open Grid Service Infrastructure WG Global Grid Forum, 2002.

[5] M.Blaze, J.Feigenbaum and J.Lacy, "Decentralized trust management", Proceedings of the IEEE Symposium on Security and Privacy, Washington, 1996, pp.164-173.

[6] A.Dogan and F.Ozguer, "Matching and scheduling algorithms for minimizing execution time and failure probability of applications in heterogeneous computing", IEEE Transactions on Parallel and Distributed Systems, 2002, 13(3): pp.308-323.

[7] W.H.Winsborough, K.E Seamons and V.E. Jones, "Automated trust negotiation", Proceedings of DARPA Information Survivability Cone and Exposition, New York, 2000, pp.88-102.

[8] W.Johnson, S.Mudumbai and M.Thompson, "Authorization and attribute certificates for widely distributed access control", Proceedings of the 7th Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises, Washington, 1998, pp.340-345.

[9] N.H.Li, J.C.Mitchell and W.H.Winsborough, "Design of a role-based trust management framework", Proceedings of the IEEE Symposium on Security and Privacy, 2002, pp.114-130.

[10] N.H.Li, W.H.Winsborough and J.C.Mitchell, "Distributed credential chain discovery in trust management", Proceedings of the 8th ACM Conference on Computer and Communications Security, New York, 2001, pp.156-165.

[11] S.Song, Y.K.Kwok and K.Hwang, "Trusted job scheduling in open computational grids: security-driven heuristics and a fast genetic algorithm", Proceedings of the 19th IEEE International Parallel & Distributed processing Symposium, Denver, USA, 2005, pp.33-40.

[12] G.Liu, Y.Xu and Y.Hao, "Resource Management Using New Reputation Model in Computational Economy for Grid", Proceedings of International Conference on Convergence Information Technology, Gyeongju, Korea, November 2007, pp.2056-2061.

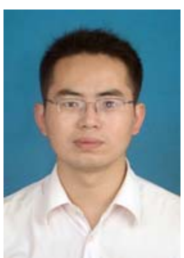
[13]G.Liu, Y.Xu and Z.Pan, "New Resource Selection Approach Based on Reputation Driven Min-min Algorithm in the Grid Economy", Proceedings of ATIP's 3rd workshop on Chinese High Performance Computing, November 10-11, 2007, Nevada, USA, pp.180-187.

[14] E.Damiani, D.Vimercati and S.Paraboschi, "A Reputation-Based Approach for Choosing Reliable Resource in Peer-to-Peer Networks", Proceedings of the 9th ACM conference on Computer and Communications Security, Washington, USA, 2002, pp.207-216.

[15] M.Gupta, P.Judge and M.Ammar, "A Reputation System for Peer-to-Peer Networks", Proceedings of the 13th International workshop on Network and Operating System Support for Digital Audio and Video, USA, 2003, pp.144-152.

[16] J.Chang, H.Wang and Y.Gang, "A Dynamic Trust Metric for P2P System", Proceedings of the 5th International conference of Grid and Cooperative Computing Workshops, Changsha, China, 2006, pp.117-120.

Authors



Yongsheng Hao received a MS Degree of Engineering from Qingdao University in 2008. Now, he is a research assistant of Network Center, Nanjing University of Information Science & Technology. His current research interests include distributed & parallel computing, mobile computing, Grid computing, Web Service, particle swarm optimization algorithm and genetic algorithm.

Guanfeng Liu received a MS Degree of Engineering from Qingdao University in 2008. His current research interests include mobile computing, Grid computing, Mobile agent system and Web-based Social Networks.

Yuebin Xu is a professor of Computer Software Lab in Qingdao University now. He has edited many books about Operation system and published several papers related to Grid computing and parallel computing. His current research interests include distributed & parallel computing, mobile computing, Grid computing, Operation system, Web Service, particle swarm optimization algorithm and genetic algorithm.

Junyan Wang is a postgraduate student of Qingdao University now. His current research interests include mobile computing and Grid computing.