A Group Based Reputation System for P2P Networks*

Huirong Tian, Shihong Zou, Wendong Wang, and Shiduan Cheng

State Key Laboratory of Networking and Switching Technology,
Beijing University of Posts and Telecommunications, Beijing, P.R. China
tianhr@bupt.edu.cn

Abstract. In large scale P2P networks, it is less likely that repeat interactions will occur between same peers for the asymmetric interests between them. So it is difficult to establish the direct trust relationship between peers and the network is vulnerable to malicious peers. A group based reputation system GroupRep for P2P networks is proposed in this paper to solve this problem. In GroupRep, the trust relationship is classified into three tiers: the trust relationship between groups, between groups and peers, and between peers. A peer evaluates the credibility of a given peer by its local trust information or the reference from the group it belonging to. A filtering cluster algorithm is proposed to filter unfair ratings provided by malicious peers. As a result, good peers are distinguished from malicious ones. Choosing the download source based on the trust value of responders can make good peers happy with the high ratio of the success query and the high satisfaction level under malicious collusive attacks with front peers.

1 Introduction

The open accessibility of Peer-to-Peer(P2P) networks attracts users to join the network to offer contents and services for others as well as try to obtain resources from the network. However, it also make the P2P network vulnerable to malicious peers which wishes to poison the network with corrupted data or harmful services for personal or commercial gain [1]. So, ascertaining the validity of the resources or services which peers want to obtain is very important to guarantee their profit.

As determining the validity of resources is a costly operation[1], reputation systems have been presented to solve this problem indirectly. Studies [2][3][4] present mechanisms that enable peers to evaluate the credibility of every file version after they get responses about different file versions. Most reputation systems focus on evaluating the credibility of resource providers. Some of them ([5-8]) calculate a global trust value for every peer. However, in large scale P2P networks, the feasibility and the necessity of establishing global trust for every peer are still doubtful. Others ([1-4], [9, 10]) allow peers to calculate a local trust

* This paper was supported by the National 973 project of China(No.2003CB314806, No2006CB701306), the NSFC (NO.90204003, NO.60472067).

L.T. Yang et al. (Eds.): ATC 2006, LNCS 4158, pp. 342–351, 2006.

value for other peers with shared information. Generally, the shared information is collected by flooding reference trust requests to peers' friends. However, in large scale P2P networks, flooding mechanism is not scalable. In addition, such reputation systems with shared information can't be adopted into the partially decentralized P2P networks where no direct management messages are allowed between peers [8].

Enlightened by the management of companies and the cooperation between them, a group based reputation system GroupRep is proposed in this paper. Peers join in a group and share their ratings by the group reputation management mechanism. The trust relationship in GroupRep can be viewed as three tiers: the trust relationship between groups, between groups and peers and between peers. A peer calculates a given peer's trust value with its local trust information or the group's reference which it belongs to. The probability that repeat interactions occur between the same groups is larger than that occur between the same peers. So in GroupRep it is easy to establish the trust relationship between peers. A cluster filtering algorithm, which is based on personalized similarity of peers and rating importance similarity of peers on providing ratings, is proposed to filter the unfair ratings provided by malicious peers. As a result, GroupRep can distinguish good peers from malicious peers effectively and make them happy with the high ratio of the success query and the high satisfaction level under malicious collusive attacks with front peers.

The rest of this paper is organized as follows: section 2 presents related work. The group based reputation system is defined in section 3. The simulation and analysis of the proposed model is followed. In the final section the conclusion is stated.

2 Related Work

Kamvar S.[5] proposed EigenTrust to calculate a global trust value for every peer based on their behavior history. Dou W.[6] presented a similar trust model where pre-trusted peers are unnecessary while these peers' existence is the basic assumption of EigenTrust. In [7] the direct trust relationship between peers is modeled as the web page link and the global trust value of every peer is calculated by the distributed pagerank algorithm. Mekouar L.[8] proposed a reputation management scheme RMS_PDN for partially decentralized P2P networks to enable every superpeer to maintain the contribution of its leaf peers and calculate their global trust value. Most of the reputation systems enable peers to calculate a local trust value for a given peer with shared information. Marti S.[1] evaluated the performance of P2P networks where peers choose the download source based on the trust value calculated by shared information. Wang Y. [9] proposed a Bayesian network trust model to evaluate a peer's credibility on different aspects according to different scenarios. NICE[10] is distinguished from other reputation systems in the intrinsic incentive that peers are willing to store the ratings which demonstrate themselves credible. PeerTrust[13] differs from other reputation systems in that trust value is computed based on

three basic trust parameters and two adaptive factors. PeerTrust is effective against malicious behaviors of peers. The main weakness of PeerTrust is that it doesn't suit highly dynamic environments as it requires peers cooperation for storing the reputation information and it can't distinguish and punish malicious peers.

3 Group Based Reputation System

Enlightened by the management of companies and the cooperation between them, we propose a group based reputation system for P2P networks. The file sharing inside the group seems as peers' providing services for this group while that between the peers which belong to different groups can be viewed as the cooperation between groups.

In GroupRep, the trust relationship is classified into three tiers: the trust relationship between groups, between groups and peers and between peers. Groups establish their direct trust relationship based on the cooperation between them. A group evaluates the credibility of members according to their behavior history of proving services for this group and cooperating with other groups. A peer stores limited local trust information based on the behavior history of other peers. When it calculates the trust value of a given peer, it firstly checks its local trust information. If there is not local trust information of the given peer, it will ask the group which it belongs to to get a reference trust value. On receiving the reference trust request, the group directly gives the trust value if the requested peer is its member; Otherwise, it gets the reference trust value by asking the group which the requested peer belonging to.

In order to clarify the idea easily, we take the P2P network as a file-sharing network, although the proposed reputation system can be adopted by many other P2P service environments.

In this section, the group based reputation system GroupRep is presented at first. Then, we explain the process of trust information updating.

3.1 GroupRep

We assume peers sharing same interests voluntarily construct logic groups. When a peer enters the network, it can use some bootstrap mechanism just as that in Gnutella 0.6[11] to find its interested group. Then it would be accepted as a member by the recommendation of other members in this group. There are one or several group managers in one group. And all the trust information inside and outside the group can be handled correctly. In addition, the communication messages are encrypted to guarantee the integrity and confidentiality.

1. The Trust relationship between groups

Groups and their direct trust relationship form the trust network which is modeled as a directed weighted graph $G_{Trust} = (V, E)$, where $V = \{G_1, ..., G_M | M \in N\}$, $G_i, 1 \leq i \leq M$ is the group in P2P networks; $E = \{e_{G_iG_j} | G_i, G_j \in V\}$, $e_{G_iG_j}$is the direct trust relationship between $G_i$and G_j . The weight of $e_{G_iG_j}$ is

the trust value $Tr_{G_iG_j}$ of G_i to G_j . We define $Trust^{path}_{G_sG_t} = (e_{G_sG_i}, ..., e_{G_jG_t})$ as the trust reference path from G_s to G_t , which is a path from G_s to G_t in G_{Trust}. Groups calculate the trust value of other groups by the following equation:

$$Tr_{G_iG_j} = \begin{cases} \frac{u_{G_iG_j} - c_{G_iG_j}}{u_{G_iG_j} + c_{G_iG_j}}, & u_{G_iG_j} + c_{G_iG_j} \neq 0 \\ Tr^{reference}_{G_iG_j}, & u_{G_iG_j} + c_{G_iG_j} = 0 \; and \; Trust^{path}_{G_iG_j} \; exists \\ Tr_{G_iG_{strange}}, & others \end{cases}, \quad (1)$$

where $u_{G_iG_j \geq 0}$ and $c_{G_iG_j \geq 0}$ are the utility and the cost respectively which peers in G_j have brought to peers in G_i . When $u_{G_iG_j} + c_{G_iG_j} = 0$, if $Trust^{path}_{G_iG_j}$ exists, $Tr_{G_iG_j}$ is defined as the reference trust value $Tr^{reference}_{G_iG_j}$ which is calculated based on the strongest path rule. Otherwise, $Tr_{G_iG_j}$ is defined by $Tr_{G_iG_{strange}}$ which is calculated according to the behavior history of strange groups that G_i has had encountered.

The strongest trust reference path. Given a set of trust reference paths from G_i to G_j , the strongest trust reference path is the path along the most credible group. We define the minimal trust value along the trust reference path as the reference trust value of this path. So $Tr^{reference}_{G_iG_j}$ is the reference trust value of the strongest trust reference path. If there are multi strongest trust reference paths, $Tr^{reference}_{G_iG_j}$ is the average of the reference trust values of these paths. As shown in Fig.1, the two strongest trust reference paths from group A to group H are "$A- > D- > C- > H$" and "$A- > F- > G- > H$". And the reference trust value of them is 0.3 and -0.5 respectively. So $Tr^{reference}_{G_AG_H}$ is -0.1.

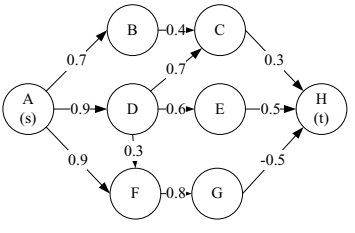


Fig. 1. The strongest trust reference path

Adaptive trust value to strange groups

$$Tr_{G_iG_{strange}} = \begin{cases} \frac{u_{G_iG_{strange}} - c_{G_iG_{strange}}}{u_{G_iG_{strange}} + c_{G_iG_{strange}}}, & u_{G_iG_{strange}} + c_{G_iG_{strange}} \neq 0 \\ 0, & others \end{cases}, \quad (2)$$

Where $u_{G_iG_{strange}} \geq 0$ and $c_{G_iG_{strange}} \geq 0$ are the utility and the cost separately which strange groups (groups that have not had transactions with G_i) have brought to G_i .

2. The trust relationship between groups and peers
We define $Tr_i^{G_i}$ as the trust value of group G_i to peer i:

$$Tr_i^{G_i} = \begin{cases} \frac{u_i^{G(i)} - c_i^{G(i)}}{u_i^{G(i)} + c_i^{G(i)}}, & u_i^{G(i)} + c_i^{G(i)} \neq 0 \ and \ i \in G_i = G(i) \\ Tr_{strange}^{G(i)}, & u_i^{G(i)} + c_i^{G(i)} = 0 \ and \ i \in G_i = G(i) \\ min\{Tr_{G_iG(i)}, Tr_i^{G(i)}\}, & i \notin G_i \end{cases} \quad (3)$$

$$Tr_{strange}^{G(i)} = \begin{cases} \frac{u_{strange}^{G(i)} - c_{strange}^{G(i)}}{u_{strange}^{G(i)} + c_{strange}^{G(i)}}, & u_{strange}^{G(i)} + c_{strange}^{G(i)} \neq 0 \\ 0, & u_{strange}^{G(i)} + c_{strange}^{G(i)} = 0 \end{cases}, \quad (4)$$

where $G(i)$ is the group that peer i belongs to, $u_i^{G(i)} \geq 0$ and $c_i^{G(i)} \geq 0$ are the utility and the cost separately that i has brought to other peers. $u_{strange}^{G(i)} \geq 0$ and $c_{strange}^{G(i)} \geq 0$ are the utility and the cost separately that peers in $G(i)$ have brought to others when they upload files as the first time. $Tr_{strange}^{G(i)}$ is adapted according to the first uploading behavior of $G(i)$ members against the malicious peers which change ID join the group.

3. The trust relationship between peers
The trust value of peer i to j is defined as follows:

$$Tr_{ij} = \begin{cases} \frac{u_{ij} - c_{ij}}{u_{ij} + c_{ij}}, & u_{ij} + c_{ij} \neq 0 \\ Tr_j^{G(i)}, & u_{ij} + c_{ij} = 0 \end{cases}, \quad (5)$$

where $u_{ij} \geq 0$ and $c_{ij} \geq 0$ are the utility and the cost separately which peer i believes what peer j has brought to it according to its local trust information. If there is no local trust information of j, defines Tr_{ij} as the reference trust value of $G(i)$.

3.2 Trust Information Updating

In GroupRep, a peer keeps limited local trust information which is updated based on its ratings to others. A group would update the trust information of members, familiar groups, strange peer or strange groups weighted by the credibility of ratings it has received.

When peer i has finished a transaction with peer j, it would firstly update its local trust information, then it reports its rating to $G(i)$. If $j \in G(i)$, $G(i)$ updates j's trust information. Otherwise, $G(i)$ updates $G(j)$'s trust information and transfers this rating to $G(j)$ to enable $G(j)$ to update j's trust information. A group would update the trust information of strange peers or strange groups if its members firstly provide service or a strange group firstly provide service to its members. Then the problem is how to determine the credibility of ratings. In order to reduce the effect of unfair ratings, we define a cluster algorithm to filter them. This filtering cluster algorithm is based on the personalized similarity of peers and the rating importance similarity of peers on providing ratings.

The peers which are similar both in personalized similarity and in the rating importance similarity are taken as similar peers. These peers associated by their similar peers within the same group form a cluster. The ratings submitted by the maximum cluster (which is noted as *the rating cluster* C_G) are seemed as credible by the group. Then we define the personalized similarity and the rating importance similarity as following:

1) The personalized similarity of peer i and j: It is measured by the root-mean-square of the ratings which they provide to same peers. It is defined by equation(6)

$$S_{PS}(i,j) = \begin{cases} 1 - \sqrt{\frac{\sum_{k \in ComSet(i,j)} (Tr_{ik} - Tr_{jk})^2}{|ComSet(i,j)|}}, & ComSet(i,j) \neq \emptyset \\ 0, & ComSet(i,j) = \emptyset \end{cases}, \tag{6}$$

where $ComSet(i,j)$ denotes the set of peers that have interacted both with peer i and with peer j in the observed period. If $S_{PS}(i,j) > S_{PS}^{threshold}$, peer i and j are personalized similar, where $S_{PS}^{threshold}$ is the threshold to determine whether two peers have the same personalized metric.

2) The rating importance similarity of peer i and j on providing ratings : Given the observation that peers in a malicious collusive group with front peers, front peers give fair ratings outside the group to increase the personalized similarity with these peers, and give high ratings selectively inside group to magnify its partners. The ratings importance similarity of i and j on providing ratings is measured by the relative importance difference of the ratings which they has given to the same peers. It is defined by the following equation:

$$S_{RIS}(i,j) = \begin{cases} 1 - \sqrt{\frac{\sum_{k \in ComSet(i,j)} \frac{|ri_{ik} - ri_{jk}|}{ri_{ik} + ri_{jk}}}{|ComSet(i,j)|}}, & ComSet(i,j) \neq \emptyset \\ 0, & ComSet(i,j) = \emptyset \end{cases}, \tag{7}$$

where $ri_{ik} = \frac{T_{ik}}{\sum_{l \in R_j} T_{jl}}$ is the relative importance of the rating which i has given to k to the total ratings which j has given in the observed period, T_{ik}, which can be defined as $u_{ik} + c_{ik}$, is the importance of the rating which i has given to k, and R_j is the set of peers to which j has given ratings in the observed period. Similarly, $ri_{jk} = \frac{T_{jk}}{\sum_{l \in R_i} T_{il}}$]. If $S_{RIS}(i,j) > S_{RIS}^{threshold}$, the ratings given by peer i and j has the same rating importance similarity, where $S_{RIS}^{threshold}$ is the threshold used to determine whether the rating importance is similar.

At the initial stage of P2P systems, there are no enough ratings to get *the rating cluster*. Group $G(i)$ would update the trust information weighted by the trust value of the peers which give the ratings. If $G(i)$ get *the rating cluster* $C_{G(i)}$ periodly, it would filter the rating based on $C_{G(i)}$. So $G(i)$ measures the rating credibility given by i by the following equation:

$$Cr_i^{G(i)} = \begin{cases} 1, & C_{G(i)} \neq \emptyset \ \ and \ i \in C_{G(i)} \\ 0, & C_{G(i)} \neq \emptyset \ \ and \ i \notin C_{G(i)} \\ 1, & C_{G(i)} = \emptyset \ \ and \ there \ is \ no \ trust \ information \ in \ the \ group \\ Tr_i^{G(i)}, & others \end{cases}, \tag{8}$$

and $G(j)(\neq G(i))$ measures the rating credibility given by i by equation (9):

$$Cr_i^{G(j)} = \begin{cases} 1, & there\ is\ no\ trust\ information\ in\ the\ group \\ min\{Cr_i^{G(i)}, Tr_{G(j)G(i)}\}, & others \end{cases} \tag{9}$$

If the rating credibility is not larger than 0, the corresponding trust information would not be updated. In addition, each group would discount all the maintained trust information periodly to make a peer's recent behavior always matters and the peer has continuing incentives to behave honestly.

4 Simulation and Analysis

GroupRep is implemented based on Query Cycle Simulator[14]. At the same time, we also implement RMS_PDN and a conventional reputation system with shared information noted as RSSI, where the trust value is calculated by the local trust information or by the reference of friends and friends' friends. We evaluate the effectiveness of GroupRep, RMS_PDN and RSSI against malicious collusive attacks with front peers. The data are collected after the 100th query cycle and the results are averaged over 5 runs.

The efficiency of the network describes how good peers can efficiently get reliable files. They are as follows:

- The Ratio of the Success Query (RSQ): if good peers issue q requests and q_s of them are satisfied with authentic files, then $RSQ = q_s/q$.
- The satisfaction level (Sat): if the size of authentic contents downloaded by i is $authentic_i$, and the size of inauthentic contents downloaded by i is $inauthentic_i$, then $Sat_i = (authentic_i - inauthentic_i)/(authentic_i + inauthentic_i)$. The overall satisfaction level of good peers is $Sat = \frac{\sum_{i \in V_g} Sat_i}{|V_g|}$, where V_g is the set of good peers.

4.1 Simulation Environment

In order to compare with RMS_PDN, the simulated network is partially decentralized with 1000 peers where the fraction of malicious peers f_m is within the range [0.1, 0.5]. Peers are constructed as 20 groups and assigned to groups with the random uniform distribution. In GroupRep, we assume a logic group corresponds with a P2P overlay group constructed with a supernode and its leaf nodes. The reference trust requests between peers in RSSI are forwarded by the manage module of our simulation. In GroupRep and RSSI, the length of the limisted trust information mainted by peers is 10, the TTL of flooded reference trust requests is 3. So in the idea situation, a peer in RSSI can get the trust value of all peers by the reference of its friends and friends' friends. In GroupRep, groups get *the rating cluster* with $S_{PS}^{threshold} = 0$ and $S_{RIS}^{threshold} = 0.5$ each 5 query cycles and discount all the maintained trust information with factor 0.8 every 3 query cycles.

There are 10000 unique files in the network. Each file is characterized by the content category c and the popularity rank r within this category. c and r both follow the uniform distribution. 20 content categories are hold by 20 groups respectively. Files are distributed with the uniform random distribution based on the content categories that peers are interested in. File sizes are randomly and uniformly between 10MB and 150MB. We assume the files owned by malicious peers are also owned by good peers and all the files can be located successfully. The utility of a success file sharing or the cost of downloading an inauthentic file is equal to the file size.

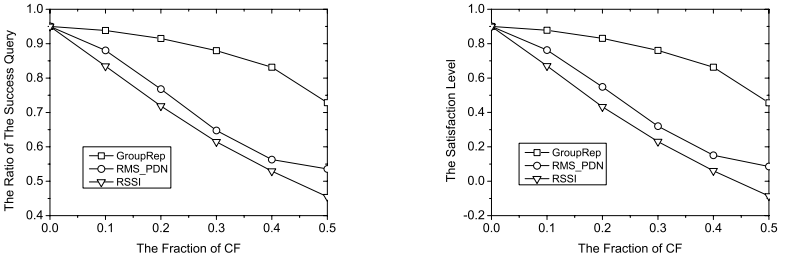
Peers are always online and issue queries. The peers in the same group share the same interest. Good peers request files randomly in their interested category with the probability 0.8. Malicious peers request files randomly to know other peers. For good peers, the probability of providing inauthentic files is 5%, while malicious peers provide inauthentic files for 80% download requests. On getting the list of file providers, peers choose the most credible provider as the download source based on their trust value.

4.2 Effectiveness Against Malicious Collusive Attacks with Front Peers

Under malicious collusive attacks with front peers, malicious peers act as a collusive group. Most of them upload inauthentic files to good peers while they upload good files to peers within their collusive group. At the same time, in GroupRep and RMS_PDN, they give negative ratings to good peers and give high positive ratings to partners which they have had a transaction with. Others named as front malicious peers act as moles and upload authentic files just as good peers and giving good peers fair ratings. These front malicious peers try to cover other malicious peers behavior by give them high positive ratings. In our simulation, the high positive rating is set as the maximum file size 150. In addition, in GroupRep and RMS_PDN we strength the attack by making malicious peers randomly select another malicious peer to give high positive rating in every query cycle except submitting a rating after a transaction. In RSSI, on receiving reference trust requests, normal malicious peers give reference trust value as 1 if the requested peer is malicious. Otherwise, it would be -1. Front malicious peers would give reference trust value based on their local trust information if the requested peer is not their partners. The peers within this collusive group are noted as *CF*. In our simulation, the fraction of front malicious peers to the whole malicious peers is 0.2. As the probability that good peers upload authentic files is 0.95, the ratio of the success query is 0.95 and the satisfaction level is 0.9 when there are no malicious peers.

As shown in Fig.2(a), RSQ decreases in GroupRep much more slowly than that does in RMS_PDN and RSSI. This is because in GroupRep the cluster filtering algorithm can filter most unfair ratings and keep still larger than 0.7 when f_m reaches 0.5. In RMS_PDN, the trust information is updated directly, so malicious peers have good reputation by giving highly positive ratings each other. Thus, RSQ of RMS_PDN decrease as f_m increasing quickly. In RSSI, a peer can't

evaluate a given peer's trust value exactly. This is more seriously in malicious collusive attacks with front peers as malicious peers always recommends their partners as trusted peers. In addition, if a peer can't calculate the trust value of the given peer by its local information or by trust friends' reference, it has no discernment on good peers and malicious peers. So in RSSI decreases dramatically as f_m increasing. in Fig.2(b), Sat always changes with the same trend as RSQ does. So it is concluded that GroupRep is more efficient than RMS_PDN and RSSI under malicious collusive attacks with front peers.



(a) The ratio of the success query (b) The satisfaction level

Fig. 2. Efficiency

5 Conclusion

Enlightened by the management of companies and the cooperation between them, we propose a group based reputation system GroupRep. The trust relationship in GroupRep is classified into three tiers: the trust relationship between groups, between groups and peers, and between peers. A peer calculates a given peer's trust value based on its local trust information or the reference of the group which it belongs to. So in GroupRep the trust relationship between peers can be established effectively because the probability that repeat transactions occur between same groups is larger than that between same peers. Personalized similarity and rating importance similarity can tell whether two peers are similar on providing ratings. The filtering cluster algorithm based on the similarity between peers can filter unfair ratings given by malicious peers. As a result, good peers can be distinguished from malicious ones. Therefore, choosing the download source based on the trust value of responders makes good peers happy with the high ratio of the success query and the high satisfaction level under malicious collusive attacks with front peers.

References

1. Marti S., and Garcia-Molina H.:Limited Reputation Sharing in P2P Systems. Proceedings of the 5th ACM conference on Electronic commerce, May 17-20, 2004, New York, NY, USA

2. Damiani E., di Vimercati D. C., Paraboschi S., et.al.:A Reputation-Based Approach for Choosing Reliable Resources in Peer-to-Peer Networks. In Proceedings of the 9th ACM conference on Computer and communications security, pages 207-216. ACM Press, 2002
3. Cornelli F., Damiani E., di Vimercati D. C., et.al.: Choosing Reputable Servents in A P2P Network. In: Lassner D, ed. Proc. of the 11th Int'l World Wide Web Conf. Hawaii: ACM Press, 2002. 441 449.
4. Selcuk A.A., Uzun E., and Pariente M.R.: A Reputation-Based Trust Management System for P2P Networks. Cluster Computing and the Grid, 2004. IEEE International Symposium on , April 19-22, 2004 Pages:251 - 258.
5. Kamvar S., and Schlosser M.:The EigenTrust Algorithm for Reputation Management in P2P Networks. WWW, Budapest, Hungary, 2003
6. Dou W., Wang H.M., Jia Y., et.al:A Recommentdation-Based Peer-to-Peer Trust Model. Journal of Software, 2004, 15(4):571 583
7. Yamamoto A., Asahara D., Itao T., et.al.:Distributed Pagerank: A Distributed Reputation Model for Open Peer-to-Peer Networks. Proceedings of the 2004 International Symposium on Applications and the Internet Workshops (SAINTW'04)
8. Mekouar L., Iraqi Y., and Boutaba R.:A Reputation Management and Selection Advisor Schemes for Peer-to-Peer Systems. in 15th IFIP/IEEE International Workshop on Distributed Systems: Operations & Management, CA, USA, 2004
9. Wang Y., and Vassileva J.:Trust and Reputation Model in Peer-to-Peer Networks. Third International Conference on Peer-to-Peer Computing (P2P'03), IEEE, September 01 - 03, 2003
10. Lee S., Sherwood R. and Bhattacharjee B.:Cooperative Peer Groups in NICE. IEEE Infocom, San Francisco, USA, 2003
11. http://rfc-gnutella.sourceforge.net
12. Lai K., Feldman M., Stoica I., et.al:Incentives for Cooperation in Peer-to-Peer Networks. Workshop on economics of p2p systems, June 2003 Berkeley, CA.
13. Xiong L., and Liu L.:PeerTrust: Supporting Reputaion-Based Trust for Peer-to-Peer Electronic Communities. IEEE Transactions on Knowledge and Data Engineering, Vol. 16, No. 7, July 2004.
14. http://p2p.stanford.edu/www/demos.htm