Handling Free Riders in Peer-to-Peer Systems

Loubna Mekouar, Youssef Iraqi, and Raouf Boutaba

University of Waterloo, Waterloo, Canada
{lmekouar, iraqi, rboutaba}@bbcr.uwaterloo.ca

Abstract. In reputation-based peer-to-peer systems, reputation is used to build trust between peers and help selecting the right peers to download from. In this paper, we argue that reputation should not be used for service differentiation among the peers. To provide the right incentives for peers to share files and contribute to the system, the new concept of *Contribution Behavior* is introduced for partially decentralized peer-to-peer systems. Service differentiation is achieved based on the *Contribution Behavior* of the peers rather than their reputations. Simulation results assess the ability of the proposed algorithm to effectively identify free riders and malicious peers that upload malicious content, hence reducing the level of service provided to these peers and preserving network resources. On the other hand, good peers that contribute to the system receive better services which increases their satisfaction significantly.

1 Introduction

In a Peer-to-Peer (P2P) file sharing system, peers communicate directly with each other to exchange information and share files. In an open P2P system, peers often have to interact with unknown peers (i.e. strangers) and need to manage the risks involved with the interactions. For example, if a user wants to download a file, the user is given a list of peers that can provide the requested file. The user has then to choose one peer from which the download will be performed. Since the open and anonymous nature of Peer-to-Peer systems open the door to misuses (by malicious peers) and abuses (by free riders), peers need to be able to reason about trust in order to avoid untrustworthy peers.

Trust management is any mechanism that allows to establish mutual trust which will motivate peers to cooperate. Building trust is difficult especially when we are dealing with strangers in virtual communities. In such interactions, risk is involved and in order to minimize this risk and get advantage from these interactions, trust is needed. Several reputation-based P2P systems [1, 2, 3, 4, 5] were introduced to build trust among peers. These systems are used to attribute a value to a peer based on its past transactions. The higher the reputation score, the more confident we are that this peer will upload an authentic file. When people interact with each other over time, the history of past transactions will help inform them about their real behavior. In addition, peers are motivated to display good behavior as it will have an impact on their future interactions. Political scientist Robert Axelrod refers to this phenomenon as the *shadow of future* [6].

Z. Despotovic, S. Joseph, and C. Sartori (Eds.): AP2PC 2005, LNAI 4118, pp. 58–69, 2006.

1.1 Motivation and Contribution

Almost all of the proposed reputation management schemes try to achieve one or more of the following goals:

1. Isolate malicious peers from the network by downloading files from the reputable peers, hence reducing malicious uploads
2. Increase the users satisfaction
3. Use the network resources more efficiently
4. Motivate peers to share files and contribute to the system
5. Reward the reputable peers by providing better services to them

Goals 1, 2 and 3 have been more or less addressed by most reputation management schemes. Goals 4 and 5 are mostly related to providing incentives and service differentiation. Few works have addressed service differentiation. Section 6 presents the most important works.

Most proposed reputation management schemes help reduce malicious uploads by choosing the high reputable peers for downloads. They also help increase the peers satisfaction. However, they do not provide incentives for peers to have a high reputation value and hence share. Indeed, the reputation considered in the proposed schemes is for trust (i.e. maliciousness of peers), based on the accuracy and quality of the files uploaded.

In eBay, members have interest in building trust and get a high reputation value in case they want to become "sellers". The higher is the reputation of a member, the higher is the chance that buyers will trust to deal with him.

In a P2P file sharing system, the situation is different. What is the interest that a peer can gain from having a high reputation value? This peer will be more and more requested for uploads which is not a gain for this peer, but more for the peers that download from it. This is why service differentiation is needed.

Few reputation schemes proposed service differentiation among the peers (cf. Section 6). However, these schemes considered peers' reputation as a guideline for service differentiation. This means that a peer with a high reputation, will receive better service than a peer with a lower reputation.

This however does not address the problem of free riders. Free riders are peers that take advantage of the system without contributing to it[1]. Providing a mechanism to detect free riders is an important issue since in [7], it has been found that most of the shared content in Gnutella is provided by only 30% of the peers. This means that 70% of the peers are free riders. There should be a mechanism to reward the contributing peers and encourage other peers to share their content.

However, free riders can have a high reputation[2], but this only means that the files that they are providing are authentic. If the reputation is used as a guideline for service differentiation, then free riders will also receive the same

[1] Or with a very small contribution.

[2] E.g. a free rider may upload few authentic files and get a high reputation. Then, the free rider starts taking advantage of the system thanks to its high reputation. In the literature, this phenomenon is called "milking".

service as the participating peers. Using reputation for service differentiation, will not allow detecting free riders. It will however provide better service to high reputable peers and low or no service to low reputable peers.

In this paper, we argue that a good scheme for service differentiation should be able to detect free riders and malicious peers and lower the service provided to them. This will have a double effect. On one hand, this will encourage free riders and malicious peers to change their behavior. And, on the other hand, good peers will receive a better service and will be motivated to continue providing good service. In this paper, we propose such a scheme and show that it is able to detect free riders and malicious peers and reduce the services provided to them while providing good peers with a better service.

The paper is organized as follows. Section 2, describes briefly the reputation management scheme considered in this work. Section 3 presents the proposed new contribution management scheme while, section 4 discusses service differentiation issues for partially decentralized P2P systems. Section 5 presents the performance evaluation of the new scheme and Section 6 describes the related works. Finally, section 7 concludes the paper.

2 Reputation Management

In this section, we describe briefly the reputation management scheme considered in this paper. For more details, please refer to [8].

2.1 Notations and Assumptions

In this paper, we consider partially decentralized P2P systems. In these systems, supernodes index the files shared by peers connected to them, and proxy search requests on behalf of these peers. Queries are therefore sent to supernodes, not to other peers. In the remaining of the paper, the following notations are used:

- Let P_i denotes peer i
- Let $D_{i,j}$ denotes the size of downloads performed by peer P_i from peer P_j
- Let $D_{i,*}$ denotes the size of downloads performed by peer P_i
- Let $D_{*,j}$ denotes the size of uploads by peer P_j
- Let $A^F_{i,j}$ be the appreciation of peer P_i of downloading the file F from P_j
- Let $Sup(i)$ denotes the supernode of peer i

2.2 The Reputation Management Scheme

After downloading a file F from peer P_j, peer P_i will evaluate this download. If the file received corresponds to the requested file, then we set $A^F_{i,j} = 1$. If not, we set $A^F_{i,j} = -1$. In the latter case, either the file has the same title as the requested file but different content, or that its quality is not acceptable. Each peer P_i in the system has four values, called *reputation data* (REP_{P_i}), stored by its supernode:

1. $D^+_{i,*}$: Satisfied downloads of peer P_i from other peers,
2. $D^-_{i,*}$: Unsatisfied downloads of peer P_i from other peers,
3. $D^+_{*,i}$: Satisfied uploads from peer P_i to other peers,
4. $D^-_{*,i}$: Unsatisfied uploads from peer P_i to other peers

Note that we have: $D^+_{i,*} + D^-_{i,*} = D_{i,*}$ and $D^+_{*,i} + D^-_{*,i} = D_{*,i} \forall i$.

When a peer P_i joins the system for the first time, all values of its *reputation data* REP_{P_i} are initialized to zero[3].

When receiving the appreciation (i.e. $A^F_{i,j}$) of peer P_i, its supernode $Sup(i)$ will perform the following operation:

If $A^F_{i,j} = 1$ then $D^+_{i,*} = D^+_{i,*} + Size(F)$,
else $D^-_{i,*} = D^-_{i,*} + Size(F)$.

Then, the appreciation is sent to $Sup(j)$ that will perform the following operation:

If $A^F_{i,j} = 1$ then $D^+_{*,j} = D^+_{*,j} + Size(F)$,
else $D^-_{*,j} = D^-_{*,j} + Size(F)$.

We compute the *Authentic Behavior* of a peer P_j as:

$$\begin{array}{ll} AB_j = \frac{D^+_{*,j} - D^-_{*,j}}{D^+_{*,j} + D^-_{*,j}} = \frac{D^+_{*,j} - D^-_{*,j}}{D_{*,j}} & \text{if } D_{*,j} \neq 0 \\ AB_j = 0 & \text{otherwise} \end{array} \qquad (1)$$

Note that AB_i is a real number between -1 (if $D^+_{*,j} = 0$) and 1 (if $D^-_{*,j} = 0$).

3 Contribution Management

We believe that trust in a peer-to-peer system should be addressed according to the following dimensions: 1) *Authentic Behavior*, 2) *Credibility Behavior*, and 3) *Contribution Behavior*

Authentic Behavior (AB): this is the reliability of a peer in providing accurate and good quality files. Good peers have usually a high *authentic behavior* value, while malicious peers usually get lower values since they are providing malicious content. This value represents the reputation of a peer. It allows to differentiate between good and malicious peers.

Credibility Behavior (CB): this represents the sincerity of a peer in providing a honest feedback. The *credibility behavior* is an important indicator that allows to identify liar peers and reduce their effect on the reputation system. In [5], the concept of *Suspicious Transaction* was introduced to compute the *credibility behavior*.

Contribution Behavior (CTB): in this paper, we introduce the new concept of *Contribution Behavior* that allows to distinguish between peers that contribute positively[4] to the system (i.e. altruistic) and the free riders (i.e. egoistic).

[3] This is a neutral reputation value.

[4] We do not consider uploading malicious content as a contribution. Only authentic uploads are taken into consideration.

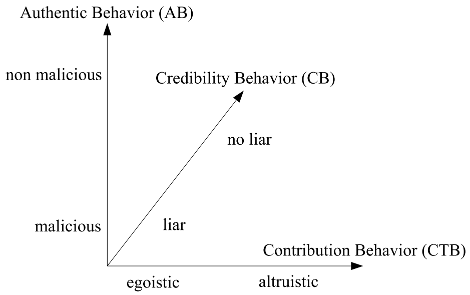


Fig. 1. Peer Behavior Dimensions

The behavior of a peer P_i is characterized by the triplet (AB_i, CB_i, CTB_i) (cf. Figure 1) which characterizes the behavior of the peer in terms of *Authentic Behavior* (sending authentic or inauthentic files), *Credibility Behavior* (lying or not in the feedback) and *Contribution Behavior* (contributing positively or not to the system). Good peers will have high values along the three defined dimensions.

We compute the *Contribution Behavior (CTB)* of a peer P_j as follows:

$$\begin{array}{ll} CTB_j = \frac{D^+_{*,j} - D^-_{*,j}}{D^+_{j,*} + D^-_{j,*}} = \frac{D^+_{*,j} - D^-_{*,j}}{D_{j,*}} & \text{if } D_{j,*} \neq 0 \\ CTB_j = D^+_{*,j} - D^-_{*,j} & \text{otherwise} \end{array} \qquad (2)$$

The intuition behind equation 2 is as follows. While the reputation value is based only on the uploads of a peer to reflect its authentic behavior (cf. equation 1), the contribution behavior should be based on both the uploads and the downloads of the peer.

The contribution of a peer is the ratio between what the peer has provided to the system and what it has consumed from it. The term $D^+_{*,j} - D^-_{*,j}$ means that the contribution value is sensitive to the maliciousness of the peer. This term allows to affect both free riders and malicious peers.

Ideally, a peer should be charged only for its authentic downloads since it is not responsible for the malicious content that it received from other peers. However, some malicious peers may rate all their downloads as inauthentic so that these downloads will not be counted in the contribution value. To avoid this situation, the total downloads is used for computing the contribution value. This will motivate the peers to deal only with the high reputable peers.

4 Service Differentiation

We divide service differentiation into two categories: *implicit* and *explicit.*

Implicit service differentiation, is the service differentiation that results from the normal evolution of the system. For example, when a peer has a low reputation, this peer will have a low probability of being selected for uploads, which will not allow it to increase its contribution value nor its reputation.

Explicit service differentiation, is the one that results from the explicit decision of system entities. For example, a supernode may decide to enforce service differentiation policies on the peers it manages. *Explicit* service differentiation can also be enforced at the level of the peer. For example, a peer may decide not to upload a file to a peer with a low credibility value (along the *Credibility Behavior* dimension), since the later peer may wrongfully send negative feedback and affect badly the reputation of the peer performing the upload. A peer may also decide not to upload a file to a peer with a low contribution value (along the *Contribution Behavior* dimension), since the peer requesting the upload may be a free rider.

The new concept of *Contribution Behavior* can be used to enforce service differentiation at any level (i.e. supernode or peer). To show its effectiveness, in this paper we enforce service differentiation policies at the supernode level. When a peer P_i sends a request to its supernode $Sup(i)$, this later will associate to the request a probability $prob_i$ according to the contribution level of peer P_i. This is the probability of performing the requested service by $Sup(i)$. The higher the contribution value is, the more chances the supernode will execute the requests for this peer[5]. This probability is computed as follows:

```
if $D_{i,*} \leq MinDownload$  $prob_i = 1$
else
    if $CBT_i \leq 0$  $prob_i = 0$
    else  $prob_i = Min\{CBT_i, 1\}$
```

Since a new peer that joins the system will have its contribution value set to 0, we allow these new peers to download a minimum amount set to a parameter $MinDownload$. In this case, the probability used by the supernode is 1. After exceeding this minimum amount of download, the probability used by the supernode will be computed according to the contribution value of the peer. The value of $MinDownload$ should be carefully chosen not to encourage peers to change identities and benefit from free downloads. Note that in case that $CBT_i \geq 1$, $prob_i$ is set to 1. This means that the peer is contributing to the system more than what it is consuming from it.

5 Performance Evaluation

In the performance evaluation section, we will compare the following schemes:

1. The reputation management scheme with no service differentiation ($NOSD$). This is the same scheme presented in [8]. This is to show the importance of service differentiation among the peers.

[5] To prevent peers from repeatedly sending the same request to the supernode over and over until the request is handled, a time period can be associated with each request. This will motivate peers to contribute if they want their requests to be processed by the system.

2. The reputation management scheme with the reputation value as a guideline for service differentiation. We will call this scheme the Reputation-Based Service Differentiation ($RBSD$). Since the reputation values (i.e. AB_i) are between -1 and 1, in this scheme, the probability $prob_i$ is computed as follows: $prob_i = (1 + AB_i)/2$, where AB_i is computed as in Eq. 1.
3. The reputation management scheme with the *Contribution Behavior* as a guideline for service differentiation. We will call this scheme the Contribution-Based Service Differentiation ($CBSD$).

To assess the effectiveness of the considered schemes in identifying free riders, a high percentage of free riders is assumed. In this section, we do not consider peers that lie in their feedbacks. This issue has been addressed in [8].

5.1 Simulation Parameters

We use the following simulation parameters:

- We simulate a system with 1000 peers and 1000 files.
- File sizes are uniformly distributed between 10MB and 150MB.
- At the beginning of the simulation, each peer has at most 45 randomly chosen files and each file has at least one owner.
- As observed by [9], KaZaA files' requests do not follow the Zipf's law distribution. In our simulations, file requests follow the real life distribution observed in [9]. This means that each peer can ask for a file with a Zipf distribution over all the files that the peer does not already have. The Zipf distribution parameter is chosen close to 1
- Peers are divided into two categories: Contributors and Free riders. Free riders constitute 70% of the peers. From each category, 30% of the peers are malicious peers that send inauthentic content. Peers behavior and distribution are summarized in table 1.
- To assess the performance of the considered schemes in a highly dynamic environment, only 40% of all peers with the requested file are found in each search request. This is due to the partial search results obtained in partially decentralized P2P systems with supernodes.
- Free riders share files with a probability of 5%. In addition, 250 of the non malicious free rider peers will accept uploading the first file to get a high reputation.
- $MinDownload$ is set to the average file size (i.e. 70MB).
- We simulate 90000 requests.

According to table 1, peers with indices from 1 to 700 belong to the category of free riders, peers with indices from 701 to 1000 belong to the category of contributor peers. Accordingly, peers with indices from 491 to 700 are malicious peers that provide malicious content in addition of being free riders. Peers with indices from 701 to 790 provide malicious content but still participate in uploading files to other peers. We have considered a situation where we have a high percentage of free riders as observed by [7] to show the effectiveness of our proposed scheme in identifying and isolating free riders and malicious peers.

Table 1. Peer Behavior and Distribution

		Probability of sending inauthentic files	
Category of peers	Percentage	Malicious (30%)	Non malicious (70%)
Contributors	30%	0.9	0.01
Free Riders	70%	0.9	0.01

5.2 Performance Parameters

In these simulations, we will focus on the following performance parameters:

- Percentage of successful requests: computed as the total number of requests that have been performed for the peer during the simulation over the total number of all submitted requests by this peer.
- Peer contribution level: shows the contribution behavior of each peer which is computed using equation 2.
- Peer load share: this parameter is computed as the normalized load supported by the peer. This is computed as the sum of the uploads performed by the peer over the total uploads in the system.

5.3 Simulation Results

No Service Differentiation Case:

Figure 2 depicts the peer load share in the case of the $NOSD$ scheme. The X axis represents the number of requests while the Y axis represents the peer load share. From the figure, it is clear that the reputation management scheme is able to isolate malicious peers (i.e. peer id 491 to 790), as they are not requested for uploads. It is also clear that the free riders do not contribute significantly

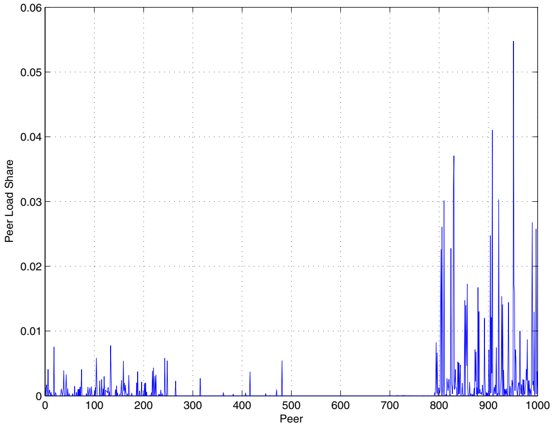


Fig. 2. Peer Load Share for $NOSD$

to the system. All the load is almost supported exclusively by non malicious contributor peers (i.e. peer id 791 to 1000).

Since there is no service differentiation, all the requests sent to the supernode will be performed regardless of the contribution of the peers. This is obviously unfair to the peers that contribute to the system.

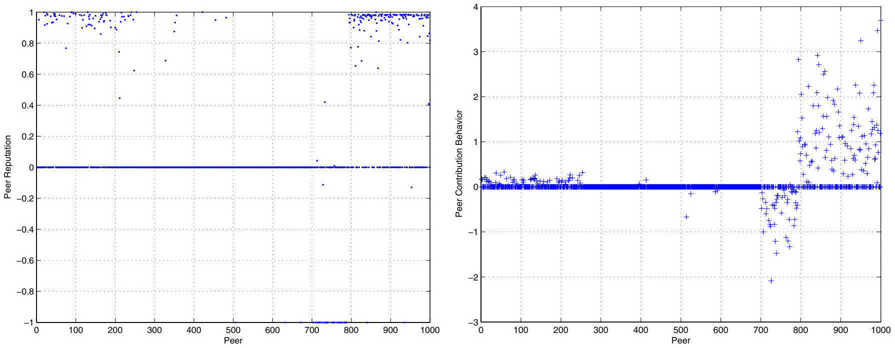


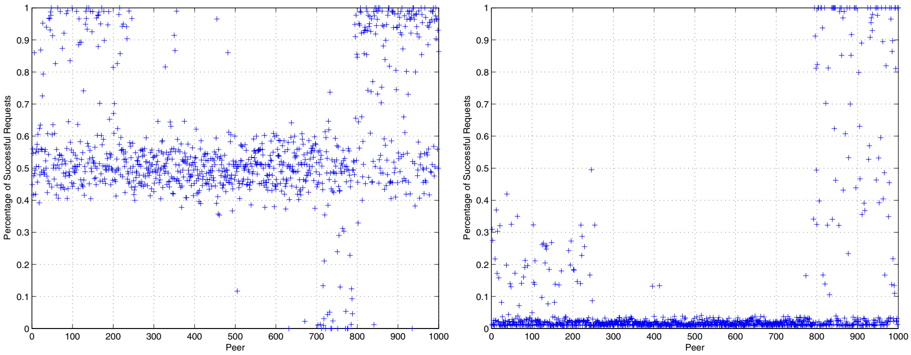
Fig. 3. (a) Peers Reputation in $RBSD$, (b) Peers Contribution Behavior in $CBSD$

Service Differentiation Case:

Figure 3.a depicts the reputation values of the peers in the case of the Reputation Based Service Differentiation ($RBSD$) scheme. It is clear that the scheme is able to identify malicious peers. However, the scheme is not able to differentiate between free riders and contributor peers. Reputation is not a good indicator of the contribution of the peer as we can see from comparing figure 2 and figure 3.a.

Figure 3.b depicts the *Contribution Behavior* value in the case of the Contribution Based Service Differentiation ($CBSD$) scheme. By comparing this figure with figure 2, we can notice that the Contribution Behavior value is a good indicator of the peer load share. In other words, a peer with a high contribution level will support more load than a peer with a low contribution level. Note that the Contribution Behavior values of malicious peers (i.e. peer id 491 to 790) are negative. This is because malicious peers are harming the system by uploading malicious files. This means that the Contribution Behavior value can be used for service differentiation which will effectively reward good peers and punish both free riders and malicious peers.

Figure 4 shows the percentage of successful requests for (a) $RBSD$ and for (b) $CBSD$. From figure 4.a, we can notice that free riders have about 50% chance to have their request processed by the supernode. Free riders with high reputation values (i.e. peer id 1 to 250) have almost the same percentage of successful requests as non malicious contributor peers. However, free riders did not contribute at the same level. In figure 4.b, free riders with IDs from 1 to 250, have a lower percentage of successful requests since they uploaded



(a) (b)

Fig. 4. Percentage of Successful Requests (a) *RBSD*, (b) *CBSD*

only few files compared to non malicious contributor peers. The later peers receive a high percentage of successful requests since they have supported almost all the load. They contributed significantly and positively to the system. The supernode processed their requests with a high probability. Some of the malicious peers uploaded more malicious content than good one, hence their percentage of successful requests is very low. This is because their contribution is negative as shown in figure 3.b.

Note that in these simulations, we assumed a static peer behavior. This means that peers do not change their behavior over time. This is to assess the capability of the proposed scheme in detecting malicious and free rider peers and preventing them from obtaining good service. In a real life system, however, peers will tend to change their behavior and we expect free rider peers with rational behavior to change from free riding to contributing to the system.

6 Related Work

The authors in [10] proposed a service differentiation protocol (SDP) for completely decentralized unstructured P2P networks. This protocol works by sending the *reputation score* of the requesting peer to other peers. These peers will map the reputation score to a Level of Service. These peers will provide service to the requesting peer according to this level. In addition of being proposed for completely decentralized P2P systems, this scheme does not take into account the maliciousness of the peers.

In [11], the authors introduce a reputation-based mechanism that assigns a better service to higher performing peers. The proposed scheme provides incentives for peers to improve their performance. The reputation is classified into two categories: provider selection and contention resolution. In provider selection, a peer among the peers offering a service is chosen to provide the service. In contention resolution, a peer among the peers requesting a service is selected by the provider peer. This scheme uses the reputation value as a guideline for service differentiation. In this paper, we have shown that this does not lead to a useful service differentiation. In addition, it proposes providing the peer requesting a file from the peers with a similar reputation value (i.e. concept of "Layered Communities"). This approach will most probably incur an important increase of malicious uploads. Indeed, if a peer receives a service from a low reputation peer, it will most probably receive bad service (e.g. malicious file) and hence does not help the peer in providing good service to others. In this paper, we propose to provide only eligible peers with the requested service. Once the request is approved, peers will receive the service from the most reputable providers. Receiving malicious content will just pollute the P2P file sharing system and waste network's resources.

In [12], the authors analyze the effectiveness of different incentives mechanisms to motivate peers to share files. The paper proposes the *reputation-based peer-approved* that uses a reputation mechanism based on rating peers according to the number of files they are advertising. Peers are allowed to download files only from peers with lower or equal rating. However, rating peers according to the number of files they are advertising is not efficient. Malicious peers can advertise a high number of malicious files. These peers will still receive good services since they will be able to upload from other peers that have a high rating value. Even non malicious peers may advertise a large number of useless files and still benefit from the system.

KaZaA, a proprietary partially-decentralized P2P system, has introduced the *participation level* for rating peers. In KaZaA, the participation level is computed as follows: (Uploads in MB/Downloads in MB)*100. Priority is given to peers with high participation level, however the exact process of how this priority is given is not known. In KaZaA, malicious peers that upload malicious content will still have a high value of participation level. As shown in [8], KaZaA is not able to detect malicious peers.

7 Conclusion

In this paper, we propose a contribution management scheme for partially decentralized peer-to-peer systems. We introduce the new concept of "*Contribution Behavior*" which is used for service differentiation rather than the use of reputation. The use of contribution behavior as the basis for service differentiation, provides the right incentives for peers to share files and contribute positively to the system. Simulation results have shown the ability of the proposed scheme to effectively identify free riders and malicious peers and prevent them from using fully the system. The use of *Contribution Behavior* for service differentiation along with the use of the *Authentic Behavior* for reputation management solve the main problems of peer-to-peer systems; free riders and malicious peers. This will provide good peers with higher satisfaction and will achieve better network resource utilization.

References

1. Aberer, K., Despotovic, Z.: Managing Trust in a Peer-2-Peer Information System. In: The 9th International Conference on Information and Knowledge Management, Atlanta, USA (2001) 310–317
2. Cornelli, F., Damiani, E., di Vimercati, S.D.C., Paraboschi, S., Samarati, P.: Choosing Reputable Servents in a P2P Network. In: The 11th International World Wide Web Conference, Honolulu, USA (2002) 376–386
3. Kamvar, S.D., Schlosser, M.T., Garcia-Molina, H.: The EigenTrust Algorithm for Reputation Management in P2P Networks. In: The 12th International World Wide Web Conference, Budapest, Hungary (2003) 640–651
4. Gupta, M., Judge, P., Ammar, M.: A Reputation System for Peer-to-Peer Networks. In: ACM 13th International Workshop on Network and Operating Systems Support for Digital Audio and Video, Monterey, USA (2003) 144–152
5. Mekouar, L., Iraqi, Y., Boutaba, R.: Peer-to-peer most wanted: Malicious peers. to appear in the Computer Networks Journal (2005)
6. Axelrod, R. In: The Evolution of Cooperation. Basic Books, New York (1984)
7. Adar, E., Huberman, B.A.: Free Riding on Gnutella. Technical report, HP (2000) http://www.hpl.hp.com/research/idl/papers/gnutella/.
8. Mekouar, L., Iraqi, Y., Boutaba, R.: Detecting Malicious Peers in A Reputation-Based Peer-to-Peer System. In: The IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas, USA (2005)
9. Gummadi, K., Dunn, R.J., Saroiu, S., Gribble, S.D., Levy, H.M., Zahorjan, J.: Measurement, Modeling, and analysis of a Peer-to-Peer File Sharing Workload. In: The 19th ACM Symposium on Operating Systems Principles, New York, USA (2003) 314–329
10. Gupta, M., Ammar, M.: Service Differentiation in Peer-to-Peer Networks Utilizing Reputations. In: ACM Fifth International Workshop on Networked Group Communications, Munich, Germany (2003)
11. Papaioannou, T.G., Stamoulis, G.D.: Effective use of reputation in peer-to-peer environments. In: Proceedings of IEEE/ACM CCGrid: International Symposium on Cluster Computing and the Grid. (2004)
12. Ranganathan, K., Ripeanu, M., Sarin, A., Foster, I.: Incentive mechanisms for large collaborative resource sharing. In: Proceedings of IEEE/ACM CCGrid: International Symposium on Cluster Computing and the Grid. (2004)