Journal of Network and Computer Applications 41 (2014) 47–55

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

Incentive mechanism for P2P file sharing based on social network and game theory

Tin-Yu Wu [b], Wei-Tsong Lee [a], Nadra Guizani [c,*], Tzu-Ming Wang [a]

[a] Department of Electrical Engineering, TamKang University, Taiwan, R.O.C
[b] Department of Computer Science and Information Engineering, National Ilan University, Taiwan, R.O.C
[c] Purdue University, USA

ARTICLE INFO

Available online 13 November 2013

Keywords:
Peer-to-peer (P2P) file sharing
Reward/penalty system
Social network
Game theory

ABSTRACT

According to statistics, most P2P applications use Gnutella or BitTorrent protocols to combat free riders. However, BitTorrent restrains free riders using the choking algorithm that chokes free riders only. Whereas, Gnutella identifies and prevents malicious nodes using the EigenTrust algorithm that has been proven to be imperfect. Therefore, both of these schemes are inefficient. According to a research conducted in 2005, 85% of Gnutella network users are free riders and only 1% of the users share new files and resources voluntarily. In this paper, by considering users' bandwidth, computing power and energy, our proposed system architecture gives users corresponding counters, which are stored and managed by the server. Moreover, the file-sharing model of our system can be divided into real-time streaming media sharing and file sharing. As for real-time streaming media sharing, users can use their counters to participate in the auction and bid on the admission of high-quality real-time streaming. As for file sharing, users have to pay a certain number of counters for every unit of download bandwidth. That is to say that all system users must use the counters to bid on or purchase services, which further enables users' spontaneous resource sharing.

1. Introduction

Recently, P2P (Peer-to-peer) technology has been extensively applied to the fields of file sharing, real-time streaming media sharing and communication. Also, because of the advancement of networking devices all over the world and the development of developing countries, P2P gradually becomes one important technology over the Internet. Up to now, many famous P2P-based file sharing platforms have been presented, like BitTorrent (Obele et al., 2009) and Gnutella (Rozario et al., 2011). However, free-riding is a well-known problem in P2P networks. Free riders in a P2P network refer to peers who keep downloading resources but are reluctant to contribute resources in return. Owing to such a free-riding phenomenon, the system performance might be degraded, those who have contributed resources to the network spontaneously cannot get worthwhile files, and more and more peers stop sharing resources. Therefore, a novel resource-sharing mechanism in P2P networks is necessary. A study conducted in 2005 showed that 85% of all Gnutella users are free riders and only 1% of the users spontaneously share new files and resources (Hughes et al., 2005).

In our proposed system architecture, we give corresponding counters to users according to their bandwidth, computing power and energy. The counters will be stored and managed by the server and users can use the counters for strategic decision-making: to bid on or purchase services. Also, based on social network and game theory, we propose a novel incentive mechanism that stimulates users to contribute resources spontaneously. Because resources in P2P systems is as if public resources that everyone can access, most users are reluctant to share resources and the system performance might be degraded. Therefore, based on the social network platform, this paper designs a Novel Incentive Mechanism (NIM), that encourages peers to share resources and excludes free riders by the relationship between peers. Finally, with the aim to combat free riders effectively and achieve fair resource allocation, we use game theory to analyze users' decision-making and corresponding rewards in different situations (Wu and Chan, 2010; Wu et al., 2011).

The rest of this paper is structured as follows: Section 2 outlines the background and introduces P2P platforms, including BitTorrent, Gnutella and Private tracker. Also, social networking habits and features of users, basic concepts and applications of game theory, and the model of Pareto Efficiency are stated. Section 3 describes our proposed social-network-based NIM and the system architecture. Simulation results and performance analysis are given in Section 4, which introduces the simulated scenarios and analyzes the weighted values and variations of the system

* Corresponding author.
E-mail addresses: tyw@niu.edu.tw (T.-Y. Wu), wtlee@mail.tku.edu.tw (W.-T. Lee), nguizani@purdue.edu (N. Guizani), roy11226@gmail.com (T.-M. Wang).

1084-8045/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jnca.2013.10.006

parameters in different conditions. By using game theory, we can define users' strategy space and relative rewards in different situations and find out users' Nash equilibrium to prove our system performance and the efficiency in combatting free riders. Section 5 concludes this paper and gives objectives for the future.

2. Background and related works

2.1. P2P file sharing architecture

Different from the general centralized system architecture that relies on the computing power and bandwidth of the central core server, P2P technology depends on the computing power and bandwidth of system users to avoid the network congestion caused by the increased user traffic that servers cannot handle immediately in the centralized architecture. P2P allows peers to connect with each others for file sharing and users can connect with several nodes simultaneously to access files. P2P provides high data rates and high scalability because all P2P users can contribute hardware and bandwidth resources to the system. Whenever a peer node joins a P2P system and requests data from other nodes, the capacity of the entire network increases as well. To replicate data over peers, P2P architecture increases robustness because there will be no single point of failure in pure P2P systems (Kim et al., 2009; Wang et al., 2012).

In recent years, P2P technology has been widely applied to various file-sharing and communication programs, like BitTorrent, Gnutella and Skype. According to the presence of a central server, P2P systems can be classified into pure peer-to-peer systems, centralized peer-to-peer systems and hybrid peer-to-peer systems. Depending on the network topology, we can classify the existing P2P architecture into structured, unstructured and decentralized networks. Detailed description of BitTorrent architecture, Gnutella architecture and Private tracker architecture (Chen et al., 2010) will be given below.

2.1.1. BitTorrent

As the most popular centralized P2P architecture for distributing large amounts of data over the Internet, BitTorrent (BT) includes a central server to supply system operation. In BT systems, all peers must install BT software first. For file-sharing, a peer converts a file to a torrent file, which contains metadata, encryption and authentication of the file. Saved by the tracker, the torrent file is recorded in a tracker list that is kept in a global registry of all the uploaders and downloaders. To download the file, peers have to connect to the tracker to obtain the torrent file. Also, with the central tracker, BT allows peers that are interested in the same file to know from which other peers to download the pieces of the file. After being responded by the tracker with a list of peers having the requested file, users are able to download the file. Each tracker can supervise several downloads of multiple files simultaneously. At the same time, the peers have to announce themselves to the tracker every 30 min. Users are authenticated by the tracker while first joining the system. Finally, after getting the torrent file from the neighboring peers, the peers start to download the pieces of the file, each of which is approximately 256 KB. When a certain chunk is completely downloaded, the peer can share the file chunk. Peers that have the full original files are called seeds. With the expansion of BT usage, the Kademlia DHT has been added to BT as a supplement of the central tracker to avoid the single-point-of-failure of a central tracker (Qi et al., 2008).

Distributed Hash Table (DHT), is a distributed system and Kademlia DHT refers to a distributed hash table for decentralized P2P networks. Kademlia specifies not only the architecture of the network, but also the exchange of data among nodes. Even in the worst case, it takes only n steps to find out the target node. By using Kademlia DHT, BT can operate without the tracker: each node is responsible for routing in a small range and storing some data. Consequently, users who cannot connect to the tracker can find out the nodes interested in the same file through DHT network and start to download the file. According to Qi et al. (2008), DHT enhances the performance of BT 5 times better.

To restrain free-rider phenomenon, BT uses a choking algorithm to control the connections. In BT systems, a client must be unchoked by the target peer first before downloading data. Only when the unchoke notification is answered can the client begin downloading the fragments of a file. Every BT client unchokes a certain number of its peers and whether to use a tit-for-tat algorithm to determine who to unchoke depends on whether the peers are seeds or not. If yes, the client chooses to unchoke the peer that contributes most resources to accelerate its download and make it acting as a seed node to enhance the system performance (Mehyar et al., 2007). If not, the client uses the tit-for-tat policy but whether it is worthwhile to unchoke the peer must be considered in advance. Supposing a client fails to download a requested data chunk from a specific peer within a period of time, BT assumes that the client is snubbed by the peer and will not answer its unchoke notification hereafter. Before optimistic unchoking finds out another peer having a better downloading rate, the client still can get a minimum downloading rate from contributors, which induces free-riding.

2.1.2. Gnutella

Gnutella is a protocol for fully decentralized P2P file-sharing systems, with an estimated market share of more than 40%. Without any central servers, the Gnutella network continues to exist as long as there are at least two clients. To this day, Gnutella, more than a piece of software, refers to an open protocol.

We briefly explain how Gnutella works. To join the network, the clients must first execute a bootstrapping function to discover other peers in the network. A new client can use different methods of bootstrapping: a pre-existing address list of other peers in the network, updated web caches of known nodes, or UDP host caches. After answered by a list of working addresses, the client tries to connect to the peers in the network. Since version 0.6 of the Gnutella protocol, peers are divided into ultra nodes (ultrapeers), each of which can connect to 32 leaf nodes, and leaf nodes, each of which can connect to approximately 3 ultrapeers only. However, most peers do not stay in the network for long and their requests for files are often discarded during the transfer, which easily leads to network congestion.

As observed in Portmann et al. (2001), the number of nodes in the Gnutella networks affects the system performance directly. Because Gnutella is a pure P2P network, when there are more than a few thousand nodes in the network, the cost of peer discovery and searching increases super-linearly and even the required bandwidth will exceed the resources of a typical client. Therefore, to enhance the performance of the Gnutella network, most applications using the Gnutella protocol restrict the size of the network to a few thousand nodes, which generates a power-law topology in file-sharing applications to share files having the same nature. Currently, owing to a power-law topology, the Gnutella network cannot get beyond a network size of a few thousand nodes and thus loses its diversity of files and global search capabilities. In the Gnutella networks, a client that has nothing to share still can request for downloading files. Thus, EigenTrust algorithm in Gnutella cannot solve the free-rider problems efficiently.

2.1.3. Private tracker

Private tracker (PT) is a P2P file-sharing software similar to BT but users have to use invitation code from friends or certain websites to register at PTs. Every PT forum asks its users to follow a share ratio, which is calculated by dividing the amount of uploaded data by the amount of downloaded data. To restrain leechers, a user whose share ratio is lower than a threshold value will be banned from the PT.

Although users in PTs can achieve faster download speed, PTs require invitations to join, which puts a limitation on the diversity of files because users may belong to the same group. Too few users in the system will reduce the number of seeders and those who cannot achieve the share ratio and be banned by the system might easily re-join the PTs or transfer to other PTs. Some users even cheat by renting seedboxes to maintain the share ratio, which results in unfair resource allocation. Most PTs sites are only available for users who use the specified PT software and this further causes the inconvenience and inconsistency of PTs (Chen et al., 2010).

2.2. Social network

Because of the rapid development in the past two years, social networks have become part of our lives and dramatically changed our living habits. At present, facebook has reached 800 million members and for every member, 68% of the contacts are family members, colleagues and close friends, who have high reliability, share similar space and time features, i.e., online time and locations, and develop the same interests in similar files. Therefore, a social-network-based P2P file-sharing network encourages the peers to contribute their resources spontaneously and exclude free-riders(Altmann and Bedane, 2009; Abboud et al., 2010; Mani et al., 2009).

Designed based on social network with large number of users, our proposed mechanism can solve the last piece/block problem for PTs. Since every social networking account that refers to a user itself is unique and requires long-term management, the user does not create new accounts easily and user behaviors continue to influence the user itself. Consequently, this results in sanctions against user behaviors because users are unwilling to run the risk of involving illegal activities. In our mechanism that integrates social network with P2P file-sharing architecture, all users have to contribute to the system and operate their own sites over a long period of time to get better resources in return. Such a method can effectively restrain free riders for social networks, exclude those who are reluctant to share their resources, causing crowding out effect.

2.3. Game theory

Game theory developed by John von Neumann in 1928 was originally a branch of applied mathematics and now has been extensively applied to different fields. Game theory is usually utilized to analyze rational players' strategic interactions and investigate competitive mathematical theories and methods. By considering the rules of the game and the players' corresponding payoff, game theory can predict a player's optimal strategy and reaction to compare with the player's actual decision-making (Ouyang et al., 2009; Wang et al., 2010). A classic game theory example is the Prisoner's dilemma, in which Player A and B can choose to surrender or keep silent. Table 1 lists their corresponding payoff after making different strategies. Based on the payoff, both players should choose to keep silent, which brings higher payoff. However, when two players are separated and cannot communicate with each other, they choose the most favorable strategy for themselves: both Players A and B surrender, which results in a bad result. Therefore, to cooperate with each other brings about not only a better result but also a great benefit to strategic decision making.

Table 1
Payoff table of the prisoner's dilemma.

		Player A	
		Silence	Surrender
Player B	Silence	−1, −1	−10,0
	Surrender	0, −10	−5, −5

According to the information about the other players, all games can be classified as complete and incomplete information games. In a game of complete information, all players are informed about the other players' characteristics, strategies and payoff information. On the contrary, in a game of incomplete information, the players do not have complete knowledge about the other players. According to the time duration, games can be finite or infinite. Finally, static games of complete information, dynamic games of complete information, static games of incomplete information, dynamic games of incomplete information respectively correspond to Nash equilibrium, subgame perfect Nash equilibrium, Bayesian Nash equilibrium and perfect Bayesian Nash equilibrium.

Therefore, based on time slots, we divide P2P file-sharing architecture as if this is an indefinitely repeated game. Assuming that all P2P users are rational players and have the common knowledge about the others, we can define the best response of a player (Wang et al., 2010; Park and van der Schaar, 2010).

2.4. Pareto efficiency

Pareto efficiency, or Pareto optimality, refers to the most efficient resource allocation in the system. A Pareto improvement is a resource reallocation that makes at least one individual better off without making anyone worse off. When no more Pareto improvements can be made, this is a Pareto efficient allocation. If a system does not satisfy the conditions of Pareto efficiency, it must be inefficient and needs to be improved. However, Pareto efficiency is not absolutely ideal because of some exceptions. Although Pareto improvements are made to make one individual better off without making anyone worse off, Utilitarianism may sacrifice individuals for the greater good. In this paper, we will use Pareto efficiency to prove whether our proposed mechanism is Pareto efficient or not.

2.5. (k,t)-robust

(k,t)-robust refers to a robust system of k-resilience and t-immunity. Because users are not absolutely rational in real life, we must consider the existence of irrational users in the networks.

In game theory, to incorporate fault tolerance, the system must tolerate more than one deviating player because the system cannot simultaneously tolerate more than two deviating players. Therefore, the Nash equilibrium of the system is 1-resilient.

Defined to be k-resilient, the game can tolerate k deviating players, even anyone going home. However, to pursue the highest possible payoff, everyone stays, which is a very fragile equilibrium because it is not immune to even one irrational player. As pointed out in Abraham et al. (2011), while considering the total number of players, k, t and Byzantine agreement, (k,t)-robustness is difficult to attain for games. Byzantine agreement refers to the malicious behaviors of the players in a game.

Because players are reluctant to pay for communications, no players will communicate with one another. Nevertheless, we build a (1,1)-robust equilibrium in our system, which means players must communicate. Though assuming Player i communicates with Player j, the two players might be irrational. That is to say that a (1,1)-robust Nash equilibrium does not exist in the game.

3. The novel P2P file-sharing mechanism

The NIM presented in this paper is a centralized P2P architecture, like BT, that has a central server. Since NIM will be deployed within social networks, we use the central server of social networks as the server core for NIM. The server core stores social files, including user identity, file authentication and social information. Before downloading files, the nodes must connect to the server to obtain the social file list to find out the source nodes having the requested file chunks. Simultaneously, to support the operation of NIM and node management, we add the concept of super nodes in the Gnutella network to NIM. The NIM estimates the parameters of each node to determine to make it a super node or not. The system architecture of NIM is displayed in Fig. 1, in which the server is the server core of NIM, the nodes around the server core are super nodes, solid lines refer to communication and control signals between nodes and dotted lines refer to P2P file transfers between nodes.

The file sharing and downloading process of NIM is similar to BT. First, the node converts the file to share to a social file and saves it in the server core of NIM. A node that wants to download this file must ask the server core for the social file. At the same time, the server core keeps track of the nodes who downloaded the file and uses the central tracker for nodes who are interested in the same file to know from which other peers to download the pieces of the file. Like BT, the node can download the file after given a list of peers that have the requested file. Besides, NIM uses DHT technology to avoid the problem of single point of failure while the super nodes in NIM enhances the performance of DHT technology also (Mani et al., 2010). But, there is one special feature about the mechanism. To increase the diversity of files, the nodes have to pay extra referral fee to NIM if the source node of the requested file is not in the social network or is indirectly related to other social networks. In this way, the nodes must enrich their own diversity of files and social networks with long term operation can obtain better payoff. Furthermore, we can use the temporal and spatial patterns between nodes to improve the performance of social networks effectively.

3.1. P2P incentive mechanism

Because BT and Gnutella cannot efficiently restrain malicious peers as well as free riders, which limits the file-sharing capabilities in a P2P system, we therefore design a novel incentive mechanism, NIM, over social networks. By using the relationships and features of social network users, our proposed NIM can effectively reduce the impact of free riders and exclude malicious peers because each social networking account is unique and irreplaceable and malicious user behaviors will continue to affect the users themselves.

3.1.1. Load estimation for access points

Counters are used in our proposed NIM. In every round, NIM considers several important factors of a node and gives the node corresponding counters based on the weighted value of each factor, as displayed in the following equation:

$$N_C = E_S \times B_C \times I_S \times D_P(aP_C + bP + cT + dS) \tag{1}$$

N_C is the number of counters assigned to the node. E_S is the performance of the P2P system that correlates closely with the condition of NIM in the previous round. Supposing the total downloaded traffic of NIM in the previous round is higher than the average downloaded traffic, as shown in Eq. (3), E_S increases and brings about better payoff to the node. B_C refers to the bandwidth contribution of the node. Eq. (2) reveals that the closer the upload bandwidth contribution of the node is to the total upload bandwidth it has, the higher the value of B_C will be. I_S refers to the social networking information of the node. D_P means the data popularity, which correlates with how many times the node's data has been downloaded in the previous round. The more times the data has been downloaded, the higher the data popularity is. P_C is the computing power of the node. P is the energy of the node. T is the authenticity of the node's data. After a node downloads a file, the node grades the source of the file to restrain the spread of malicious files and to deter malicious nodes in NIM. S, the stability of the node, is directly proportional to the average retention time

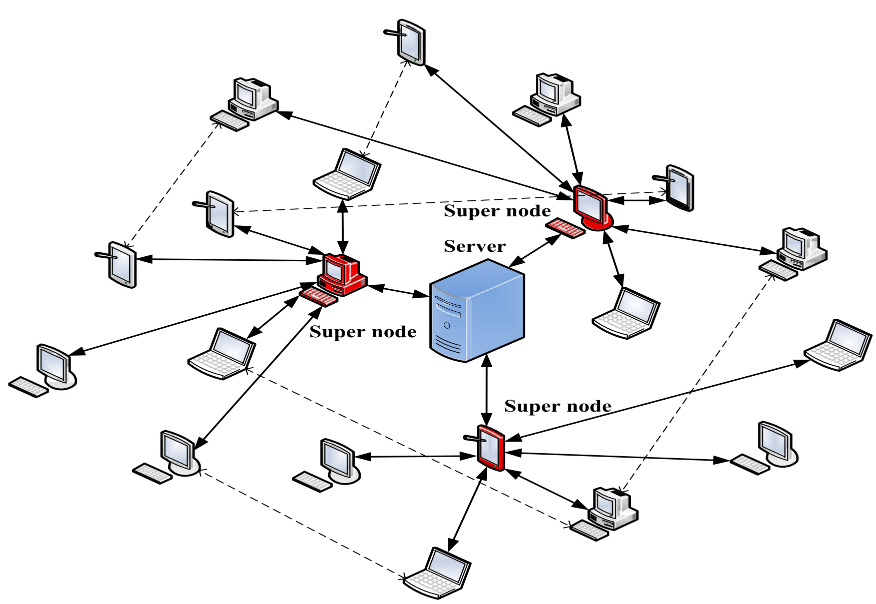

Fig. 1. System architecture of NIM.

of the node in the system and helps to improve system stability and performance. Variables a, b, c and d are the weighted values of the parameters and will be introduced in Section 4.

$$B_C = \frac{{B_{share}}^2}{B_{have}} \tag{2}$$

$$E_S = \frac{B_{now}}{B_{past}} \tag{3}$$

Recorded by the server core of NIM, the counters are managed by the server, allowing users to have different strategic decision-making in NIM. According to the bandwidth, computing power and energy of nodes, the server chooses some nodes as super nodes to manage the counters in the system and reduce the burden on the server. Since super nodes have to contribute their own resources, corresponding percentage increases in counters will be given by the server to super nodes.

3.1.2. Auction and purchase of service

In our proposed NIM, nodes get more counters by contributing resources to the system so that they can use the counters to purchase or bid on services. In the P2P system that has global search capabilities, in addition to file sharing, real-time streaming media sharing is another significant application that is often adopted to improve the quality of streaming of real-time international sports games. Real-time streaming media sharing is greatly different from common file-sharing in the P2P system. As for common file-sharing, reliable data transfer relies on the TCP protocol. As for real-time streaming media sharing, the UDP protocol that emphasizes on timeliness is utilized.

4. Simulation results and performance analysis

4.1. Simulated scenario

In our simulated scenario, we assume that all system users are rational players and the game is infinite. Because our proposed NIM is based on social networks, packet loss between nodes or between nodes and the server is not taken into account. Therefore, all nodes exist in the network since the beginning and will stay even after downloads are completed. In the analysis model in this section, the payoff for each player is relative return, instead of absolute return.

4.1.1. Evolutionary game theory model

According to traditional game theory model, all players in a game have complete knowledge of details of the game and the strategies of others. Impervious to external forces or thoughts, the players simplify and conceptualize the rules, strategies and payoff to find equilibrium in the game. To attain the above-mentioned premises, information exchange among the players must be complete and factors including moral, religion and habits must be dismissed. Under the premises, in traditional game theory based on rationality, a player's strategy based on the strategies of the others is not absolutely the best response.

In this section, we attempt to use evolutionary game theory model to investigate the Evolutionary Stable Strategy (ESS) in NIM. ESS, a strategy that resists the invasion of new strategies through time, can be a strategy set from many players. In other words, we may have the proportion or features of the strategy set, but we cannot accurately define the choice of any one specific player. Eq. (4) is denoted as an evolutionary game theory model, where I mean the players in the game, S means the strategy set of the players and π means the utility function of the players. Because the utility function directly corresponds to the satisfaction degree of the players, we regard π as the corresponding payoff of the strategy (Feng et al., 2008); (Cho and Nguyen, 2009).

$$G = \{I, S, \pi\} \tag{4}$$

If $Y \neq X \in S$ and $\overline{\varepsilon_v} \in (0{,}1)$ exists, the players choose either strategy X or Y and the equation can be stratified as

$$\mu[x, \varepsilon y + (1-\varepsilon)x] > \mu[y, \varepsilon y + (1-\varepsilon)x] \tag{5}$$

u, also called a layer, refers to a stratified sample obtained from different layers of the population. When the multi-layer vector of strategy X is higher than strategy Y, we get $\varepsilon \in (0, \varepsilon_v)$, which means that all players choose strategy X in the end and X is the ESS of the model.

As shown in Feng et al. (2008), we assume that at time t, N nodes in the P2P system are randomly distributed in each round and the strategy space of each node is [0, 1]. Strategy 0 means that when downloading resources from others, the node simultaneously contributes his own resources to others. Strategy 1 means that the node downloads resources from others but contributes no resources. Assuming that a node chooses strategy 0 and uploads D_s units of resources, the payoff of every unit of resource contribution will be c. No matter which strategy the node chooses, he downloads D_d and the cost of downloading every unit of resources will be g. In this way, we can get the payoff matrix for nodes playing with strategy 0 as (6) and the payoff matrix for nodes playing with strategy 1 as $B = A$^T. Next, we discuss the strategic changes of a node during a small time interval ε. Other researches have shown that the decision making process for each node is only related to D_s, upload resources, and c, extra payoff. As long as D_s and c are both bigger than 0, the number of nodes choosing strategy 0 increases over time and reaches 100% in the end. This solution is the ESS of the game. The bigger the values of D_s and c are, the faster the system can reach the ESS.

$$A = \begin{bmatrix} -gD_d + cD_s & cD_s \\ -gD_d & 0 \end{bmatrix} \tag{6}$$

In NIM, we lay great stress on excluding free riders spontaneously by the features of social networks and therefore one important parameter for calculating counters and payoff is social networking information. For this reason, we modify the payoff matrix to prove that NIM can enhance the performance of P2P network by referring to social networking information. To further define a novel payoff as the social networking information, I_s, we can get the payoff matrix for nodes playing with strategy 0 in NIM as (7) and the payoff matrix for nodes playing strategy $B_{NIM} = {A_{NIM}}^T$.

$$A_{NIM} = \begin{bmatrix} -gD_d + cI_sD_s & cI_sD_s \\ -gD_d & 0 \end{bmatrix} \tag{7}$$

Let $\gamma_0(t)$ denote the number of nodes playing with strategy 0 at time t, $\gamma_1(t)$ denotes the number of nodes playing with strategy 1 at time t, and $\gamma(t)$ denotes the total number of nodes in the system, we can get

$$\gamma(t) = \gamma_0(t) + \gamma_1(t) \tag{8}$$

(9) is the fraction of nodes playing with strategy 0. Then, $1 - x(t)$is the fraction of nodes playing with strategy 1.

$$x(t) = \frac{\gamma_0(t)}{\gamma(t)} \tag{9}$$

We assume that every stage game starts at time $kt(k \in N)$ and ends at time $t(k+1)$ and every node gets the average payoff with regard to other possible players. Moreover, during a small time interval ε, only an ε fraction of the total nodes participate in the game. Consequently, the number of nodes playing with strategy i at time $t+\varepsilon$ can be given by Eq. (10), which is directly proportional to the corresponding payoff of the strategy.

$$\gamma_i(t+\varepsilon)=(1-\varepsilon)\gamma_i(t)+\varepsilon\gamma_i(t)U_i(t);\ i=0,1 \quad (10)$$

(11) and (12) stand for the average payoff for nodes playing with strategy 0 and 1, respectively.

$$U_0(t)=-gD_d+cI_sD_sx(t) \quad (11)$$

$$U_1(t)=-gD_d \quad (12)$$

The number of nodes in the system at time $t+\varepsilon$ can be given by Eq. (13). The average payoff for nodes at time t can be given by Eq. (14).

$$\gamma(t+\varepsilon)=(t+\varepsilon)\gamma(t)+\varepsilon\gamma(t)\overline{U}(t) \quad (13)$$

$$\overline{U}(t)=(cI_sD_s-gD_d)x(t) \quad (14)$$

Eq. (15) shows the number of nodes choosing strategy 0 during a small time interval ε.

$$x(t+\varepsilon)-x(t)=\varepsilon\frac{x(t)[U_0(t)-\overline{U}(t)]}{1-\varepsilon+\varepsilon\overline{U}(t)} \quad (15)$$

To divide both sides of Eq. (15) by ε and substitute $\lim \varepsilon\to 0$ into the equation, we can get the replicator equation as

$$\frac{dx(t)}{dt}=x(t)[U_0(t)-\overline{U}(t)]=cI_sD_sx(t)(1-x(t)) \quad (16)$$

Next, we compare the performance of traditional P2P presented in Feng et al. (2008) with our proposed social-network-based NIM. We assume that at the beginning, the number of nodes playing with strategy 0 and 1 is half and half, $c=0.5$, $D=1$, $I_s=1.5$. Fig. 2 displays the proportion of nodes playing with strategy 0 in both systems.

In Fig. 2, the lower line refers to a traditional P2P system while the upper line means the NIM integrated with social networking information. It reveals that nodes in the NIM reach the ESS more quickly because of extra payoff from social networks. The value of I_s for NIM users is determined by their social networking information. Supposing a node is a member of a good social network, its social networking information will bring about higher payoff. On the contrary, the payoff will be greatly reduced. To obtain higher payoff based on I_s, nodes in the NIM must keep long term operation of their social networks and enrich their own diversity of files. When all nodes endeavor to improve their I_s, the average I_s in the system increases and the system will reach the ESS more quickly.

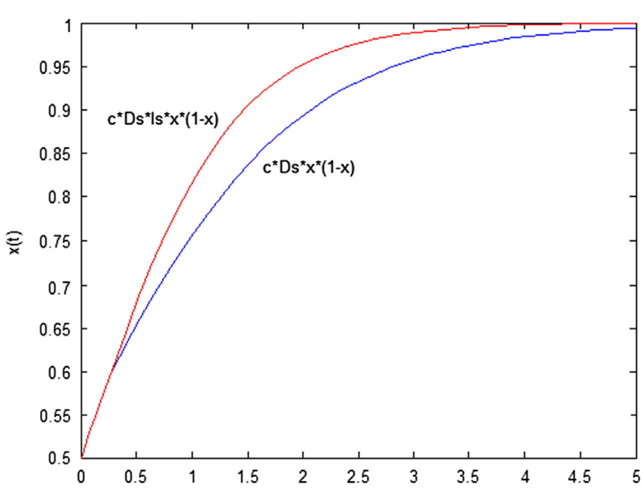

Fig. 2. The proportion of nodes playing with strategy 0 in both systems.

Fig. 3. System fairness in NIM.

4.1.2. System fairness

This section aims to make the reader understand the relationship between fairness and performance in P2P networks. To attain the optimal performance, P2P must retain the resourceful nodes that are inclined to contribute resources as long as possible to accelerate the data rate. Therefore, the upload bandwidth is a significant parameter for the system to judge the condition of a node. Supposing a user contributes much upload bandwidth but receives bad service, he will contribute less resource. For this reason, the numerical threshold between fairness and performance in P2P systems is important.

Bin et al. (2009) has pointed out that BT peers usually use two strategies: non-discriminative uploading and selective uploading, as known as the "tit-for-tat" strategy in the BT protocol. The higher proportion of peers that use the non-discriminative uploading in the system results in better system performance but worse system fairness because these peers do not exactly get corresponding payoff in return. Oppositely, the higher proportion of peers that use the selective uploading in the system leads to higher system fairness but worse system performance. However, in our proposed NIM, users simultaneously receive counters while contributing their own resources. A good reward mechanism encourages peers to share resources and helps to improve the performance of NIM. Therefore, in NIM, fairness is directly proportional to system performance. As displayed in Fig. 3, the vertical axis refers to fairness, 1 the fairest and 0.1 the most unfair, and the horizontal axis refers to reward rate, Nc/Bc, where N_c means the number of counters a node can receive and B_c means the bandwidth contribution of a node.

4.2. Parameter analysis

This section analyzes NIM by game theory and creates diagrams for game analysis in NIM to show the corresponding payoff for players playing with different strategies. Moreover, Heuristic methods are used to find out the Nash equilibrium for the players in NIM. Observing the corresponding payoff for players playing with different strategies and inferring the advantageous strategy, we figure out why the players are willing to contribute their resources and prove that the dominant strategy for players in NIM is to share their resources to the best of their abilities.

The algorithms used in Gnutella and BT to prevent free riders are proven to be imperfect and ineffective. Assume that all players are representative agents, have the same strategy space, and receive the same payoff while selecting the same strategy. Fig. 4 describes a traditional P2P-based game with many players, in which free riders that contribute no resources still have the minimum download bandwidth and thus receive higher payoff compared with normal peers. However, when the proportion of free riders in the game keeps increasing, not only the performance of the P2P system but also the payoff for all players decreases. As a result, the dominant strategy in this game is that everyone becomes a free rider, which is quite similar to the current situation.

Table 2 is the payoff matrix of Player A in NIM. Because players must contribute their resources to exchange counters for deposit and service purchase, free riders cannot get any resources in NIM. Consequently, the dominant strategy for all players in this game is to share resources and the Nash equilibrium of this game is that all players cooperate by sharing their resources. In this way, free riders do not exist in NIM.

Next, based on several important parameters in NIM, we analyze the strategy space and payoff for players in different situations, and find out the dominant strategy and Nash equilibrium in the game. Assume that all players have the same resources. First, we analyze I_s, social networking information. There are players A, B, C and D in the game and their social relationship is displayed as shown in Fig. 5: Players A, B and C are in the same social network and Player C and D are in another one. Each player can choose to cooperate or betray and the payoff matrix is given in Table 3. To download a file, a player has to pay extra referral fee to NIM if the source node of the requested file is not in the social network or is not directly related to the social network. The table reveals a Nash equilibrium: all players choose to cooperate to get the highest payoff. Once one of the players chooses to betray, the social network excludes the player and records such a betrayal that continues to affect the player's social development in the future. The payoff to the player who betrays is therefore the lowest. One more point to notice is that Player D is only socially related to Player C, so Player D must develop a social relationship with other players in the game in order to earn better payoff.

Assume that the consideration is only given to the efficacy of NIM, E_S. The players in the game can choose to increase or decrease upload/download bandwidth. The payoff analysis displayed in Fig. 6 shows that the dominant strategy for all players is to increase upload bandwidth when the system efficacy is higher than 1. When the system efficacy is lower than 1, the dominant strategy for all players is to increase download bandwidth, not only to stabilize the system, but also to propose good reasons for players to contribute their resources while sharing streaming of real-time sports games.

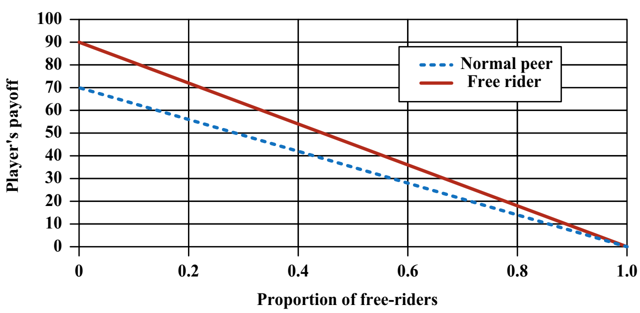

Fig. 4. Payoff analysis of traditional P2P file-sharing.

Table 2
Payoff matrix of player A in NIM.

		The other peer	
		Be a free rider	Be a normal peer
Player A	Be a free rider	0, 0	0, 100
	Be a normal peer	100, 0	100, 100

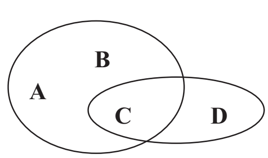

Fig. 5. Social relationship between Players' A, B, C and D.

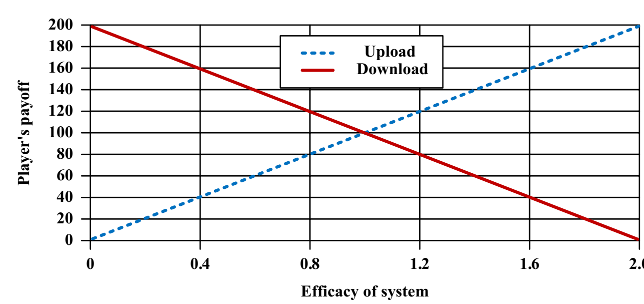

Fig. 6. Payoff analysis based on the system efficacy.

4.3. Game simulation for NIM

Finally, we simulate the variation of counters for Players A, B, C, D, E and F playing with different strategies in 20 rounds, as shown in Fig. 7. Table 4 lists the simulation parameters of the six players. Player A is a new participant whose resources to share is limited but parameters including D_P, T and S continue to increase with their stay in NIM. Player B shares all upload bandwidth he has. Player C is a normal peer whose income is equal to his expenditure.

Table 3
Payoff matrix based on social network.

				Player C			
				Cooperate		Betray	
				Player D		Player D	
				Cooperate	Betray	Cooperate	Betray
Player A	Cooperate	Player B	Cooperate	10,10,10,10	8,8,8,1	6,6,1,2	6,6,1,1
			Betray	8,1,8,8	6,1,6,1	2,1,1,2	2,1,1,1
	Betray	Player B	Cooperate	1,8,8,8	1,6,6,1	1,2,1,2	1,2,1,1
			Betray	1,1,6,6	1,1,2,1	1,1,1,2	1,1,1,1

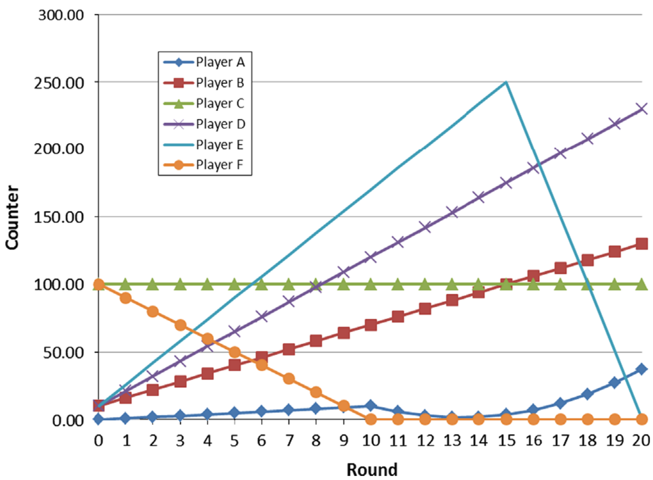

Fig. 7. Simulation result.

Table 4
Simulation parameters

Parameter	Player					
	A	B	C	D	E	F
N_C	0	10	100	10	10	100
E_C	1	1	1	1	1	1
B_C	1	2	1	2	2	0
I_S	1	1	1	1	1	1
D_P	$0.1 \rightarrow 2$	1	1	1	1	1
P_C	2	2	2	4	2	2
P	2	2	2	2	2	2
T	$0.1 \rightarrow 2$	2	2	2	2	2
S	$0.1 \rightarrow 2$	2	2	2	2	2
Bonus (%)	0	0	0	30	0	0
Spend/Round	$0 \rightarrow 5$	10	8	15	$0 \rightarrow 50$	10

Player D is a super node who has better hardware resources, higher value of P_C, and additional bonus from NIM. Player E chooses to save his counters to view high-quality streaming after the 15th round. Player F is a free rider. As a new participant, Player A receives preferential treatment from the system and the weighted values of P_C and P are raised. Player A will use his counters to download files of high data popularity in the 11th round to earn better payoff in the following rounds. Player E chooses to save his counters strategically for purchasing high-quality real-time streaming media in the 15th round. Accordingly, Player E's counters to spend in each round will be increased to 50 and he stops sharing his own resources.

Fig. 7 shows how the six players spend their counters in 20 rounds. Because NIM reduces the parameters as preferential treatment, Player A uses the counter to download files of high data popularity to share and to earn more counters. Player B shares all his upload bandwidth to get as many counters as possible. Player C chooses to earn as many counters as he spends. As a super node that contributes resources to help the server to manage system, Player D can get 30% additional bonus. To download high-quality streaming, Player E saves his counters for 15 rounds and bids on the admission of high-quality real-time streaming in the 16th round and stops sharing to maintain his own resources. Therefore, Player E has to spend 50 counters per round after the 16th round. Without contributing any resources, Player F exhausts his resources and will be excluded by the system soon.

5. Conclusion and future work

To solve the free-rider phenomenon in the current P2P file-sharing applications, this paper integrates the pros and cons of various P2P architectures and presents a Novel Incentive Mechanism (NIM) based on social network and game theory for P2P file-sharing. Our simulation results and game analysis show that by considering different conditions and contributions of all players and providing players with more strategy space by counters, our proposed NIM can restrain the number of free riders efficiently, encourage the nodes to contribute resources as much as possible, and distribute the resources to all nodes more fairly via different compensation mechanisms. Besides, NIM is immune to betraying and irrational players because of utilizing the space and time features of nodes in social networks.

Nevertheless, to implement the social-network-based NIM on the current social networks, we have to further quantify the parameters concretely and determine the range of the parameters for counters in case one specific parameter affects the system excessively. Moreover, users' habits of making friends and utilization status in social networks both contribute to the performance of NIM. Our future research orientations will focus on the security issues and coding techniques in P2P systems. Because social networks include lots of private information about users, how to enhance the security for P2P file-sharing systems is a key point. In addition, with the increase of files and control signals in P2P systems, to improve the throughput of P2P systems by coding techniques will be our future research direction.

Reference

Abraham Ittai, Alvisi Lorenzo, Halpern Joseph Y. Distributed computing meets game theory:combining insights from two fields. ACM SIGACT News 2011;42 (2):69–76.

Altmann J., Bedane ZB. A P2P file sharing network topology formation algorithm based on social network information. IEEE INFOCOM Workshops; 19–25 April 2009. p. 1–6.

Bin Fan, Lui JCS, Chiu Dah-Ming. The Design Trade-off of Bittorrent-like File Sharing Protocols. IEEE/ACM Trans Netw 2009;17(2):365–76.

Brownson O. Obele, Augustine I Ukaegbu. KangMinho. On tackling free-riders in P2P networks. In: Proceedings of the 11th international conference on advanced communication technology, vol. 03; 15–18 February 2009. p. 2084–9.

Chen Xiaowei, Chu Xiaowen, Jiang Yixin, Ren Fengyuan, Measurements, analysis and modeling of private tracker sites. In: Proceedings of 2010 18th international workshop on quality of service, China; 16–18 June 2010. p. 1–2.

Huibin Feng, Shunyi Zhang, Chao Liu, Junrong Yan, Ming Zhang. P2P incentive model on evolutionary game theory. In Proceedings of the 4th international conference on wireless communications, networking and mobile computing; 12–14 October 2008. p.1–4.

Hughes Daniel, Coulson Geoff, Walkerdine James. Free riding on Gnutella revisited: the bell tolls? IEEE Distributed Syst. Online 2005;6(6).

Jongtack Kim, Yugyung Lee, Saewoong Bahk. SALSA: Super-Peer Assisted Live Streaming Architecture. In: IEEE International conference on communications, South Korea; 14–18 June 2009. p. 1–5.

Mani M, Ngyuen A-M, Crespi N. What's up: P2P spontaneous social networking. In: IEEE international conference on pervasive computing and communications; 9–13 March 2009. p.1–2.

Mani M, Nguyen Anh-Minh, Crespi N. SCOPE: a prototype for spontaneous P2P social networking. In: Proceedings of the 2010 8th IEEE international conference on pervasive computing and communications workshops, France; March 29–April 2 2010. p. 220–5.

Mehyar M, Gu WeiHsin, Low SH, Effros M, Ho T. Optimal strategies for efficient peer-to-peer file sharing. In: 2007 IEEE international conference on acoustics, speech and signal processing, vol 4; 15–20 April 2007. p. IV-1337–40.

Osama Abboud, Thomas Zinner, Eduardo Lidanski,; Konstantin Pussep, Ralf Steinmetz, StreamSocial: A P2P Streaming System with Social Incentives. In: 2010 IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks; 14–17 June 2010. p. 1–2.

Ouyang Jing-cheng, Wang Yue-bin, Hu Xu-huai, Lin Ya-ping. An incentive mechanism using game theory for P2P networks. NSWCTC '09. In: International conference on networks security, wireless communications and trusted computing, vol. 2; 25–26 April 2009. p.715–8.

Park Jaeok, van der Schaar M. A game theoretic analysis of incentives in content production and sharing over peer-to-peer networks. IEEE J Selected Topics Signal Process 2010;4:704–17.

Portmann M, Sookavatana P, Ardon S, Seneviratne A. The cost of peer discovery and searching in the Gnutella peer-to-peer file sharing protocol. In: Proceedings of the ninth IEEE international conference on networks, Australia; 10–12 October 2001. p.263–8.
Qi Jiayin, Zhang Hongli, Ji Zhenzhou, Yun Liu. Analyzing BitTorrent traffic across large network. In: 2008 International conference on cyberworlds, Harbin; 22–24 September 2008. p.759–64.
Rozario Fabian, Han Zhu, Niyato Dusit. Optimization of non-cooperative P2P network from the game theory point of view. In: 2011 IEEE wireless communications and networking conference, USA; 28–31 March 2011. p.868–73.
Seungyun Cho, Thinh Nguyen. Modeling and dynamics analysis of P2P networks based on evolutionary games. In: Proceedings of the First international conference on advances in P2P systems, USA; 11–16 October 2009. p. 34–8.
Wang Chunzhi, Zhou Ke, Chen Hongwei, Chen Li. A game model of P2P based on punishment strategies. In: Proceedings of 2010 2nd international symposium on information engineering and electronic commerce, China; 23–25 July 2010. p. 1–3.
Wang Tzu-Ming, Lee Wei-Tsong, Wu Tin-Yu, Wei Hsin-Wen, Lin Yu-San. New P2P sharing incentive mechanism based on social network and game theory. In: Proceedings of the 2012 26th international conference on advanced information networking and applications workshops; 26–29 March 2012. p. 915–90.
Wu Tin-Yu, Chan Hung- Lin. Integrate airtime metric and geocast over P2P-based VoD streaming cache. Tamkang J Sci Eng 2010;13(1):99–106.
Wu Tin-Yu, Kuo Kai-Hua, Cheng Hua-Pu, Ding Jen-Wen, Lee Wei-Tsong. Increasing the lifetime of ad hoc networks using hierarchical cluster-based power management. KSII Trans Internet Inf Syst 2011;5(1):5–23.