Journal of AI and Data Mining
Vol 8, No 4, 2020, 557-571. ***Methodological Paper*** DOI: 10.22044/JADM.2020.9063.2049

A New Incentive Mechanism to Detect and Restrict Sybil Nodes in P2P File-Sharing Networks with a Heterogeneous Bandwidth

M. Babazadeh Shareh[1], H. R. Navidi[2*], H. Haj Seyyed Javadi[2], M. HosseinZadeh[3]

1. Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.
2. Department of Mathematics and Computer Sciences, Shahed University, Tehran, Iran.
3. Iran University of Medical Sciences, Tehran, Iran.

Received 18 November 2019; Revised 01 April 2020; Accepted 24 June 2020
**Corresponding author: navidi@shahed.ac.ir (HR. Navidi).*

Abstract
In cooperative peer-to-peer (P2P) networks, there are two kinds of illegal users, namely free riders and Sybil nodes. Free riders are those who try to receive services without any cost. Sybil users are rational peers that have multiple fake identities. There are some techniques available to detect free riders and Sybil users such as the tit-for-tat and Sybil guard methods. The free riding prevention methods are easily by-passed by Sybil nodes and Sybil detection algorithms just focus on the network specifications that are nothing to do with free riders. There is no technique capable of detecting both of these attacks simultaneously. Therefore, the main objective of this research work is to propose a single mechanism to detect both kinds of these illegal users based on the game theory. In fact, an innovative incentive mechanism is designed that cannot be deceived by Sybil nodes. Obtaining new centrality and bandwidth contribution formulas for this incentive mechanism approach is the basic idea of this work. The results of this work show that as the life of the network passes, free riders are identified, and through detecting Sybil nodes, the number of services offered to them will be decreased. By means of the proposed mechanism, with up to 40% of initial defectors and 10% Sybil nodes, the network can properly deliver the services.

Keywords: *File Sharing Network, P2P Network, Free Rider, Sybil Attack, Incentive Mechanism.*

1. Introduction

There is a type of social networking site that, like the peer-to-peer (P2P) networks, allows a group of users to make their resources directly available to the other users using a software. In this case, there will be a distributed system that operates in a decentralized self-organized way [1].

The P2P architecture has revolutionized the share of huge files on the Internet. This design gives each peer the chance to present in sharing files at the same time. This contribution is self-directed. In addition, each node chooses the level of participation. The achievement of a P2P system is on the basis of the level of peer participation [2].

The flexible infrastructure of the P2P networks has multiple cons and probs. Misuse of the system by some nodes can lead to many obstacles for the network [3].

Both the provider and the consumer users share their resources and compete to get more resources [4]. The services that are provided in these networks can be the use of shared files, data storage, bandwidth, computing power, etc. [5]. So far, several well-known software programs such as Bit Torrent [6] and Gnutella [7] have been introduced based on the platform of this network. Of course, today, BitTorrent is the most successful of these products [8-10], which is quite evident with the continuous use of the users and its high popularity [11].

In the P2P networks' infrastructure, the users give or get services. The users can download a file from others or send them file chunks. Any user uses the system by receiving a file or spends a resource for servicing the rest. Based on this plan, any user attempts to consume maximum services. Free rider users (or defectors) are willing to receive their files between the available resources and flee from giving services to other nodes [5].

This treatment can destroy the network and decrease the amount of files in the system, and consequently, convert the network into a malformed one. According to a study conducted in 2005, almost 85% of the nodes of Gnutella are defectors, and just less than 1% of the nodes provide services voluntarily [12].

In a P2P network, there is no base supervisor server to monitor the performance of the peers. Therefore, it is very difficult to detect malicious behaviors. An essential issue is to manage the nodes and encourage them to cooperate. The essential methods available to create cooperative activities in such networks are reputation and trust. The researchers have conducted multiple methods by means of trust approximation, user's reputation data, and punishing the defectors [13].

Due to the danger of free riding in the P2P networks, in the recent years, some researchers have tried to use new methods to design innovative algorithms for curbing free-rider's activities [14]. As mentioned earlier, file sharing networks are also vulnerable to free riding. Hence, some researchers have been trying to tackle the free riding problem exclusively in this kind of network. Designing a new file-sharing platform from scratch is one possible solution [15].

Apart from the free riding problem and its solutions, Sybil attack is an emerging problem in the social networks. Predicting Sybil activities in order to contain them and detecting malicious profiles are already the tested methods, especially in large-scale social networks [16, 17]. By the growth of the size of the P2P networks, it is essential to employ new methods for its fundamental problems. Considering a network in the size of twitter, eliminating free riders and curbing Sybil attacks could be far more difficult. Therefore, an incentive mechanism and game theory-based approaches have been applied to the networks in the recent years. Nevertheless, using the hybrid approaches can take the full advantage of various methods simultaneously [18, 19, 20].

This paper is based on a comprehensive research work on defeating the attackers in the p2p networks. The process of this work started with the problem definition. The main objective of the work was to cope with free riders and Sybil nodes. The major question was how we could use an incentive mechanism in order to reduce the network attacks. Reviewing the previous works was the first step toward solving the problem. After specification of the problem background, a new approach was designed. The proposed method was assessed through a generally generated network. The final report showed the efficiency of the new approach. Figure 1 shows the research methodology.

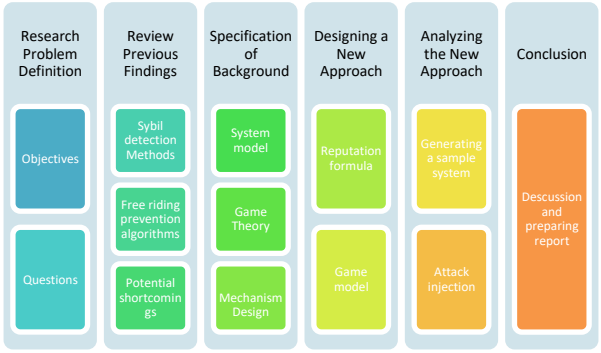

Figure 1. Research methodology.

The rest of this article has been arranged as what follows. The second section describes the previous methods. Section 3 describes the network structure on which the proposed method has been worked on. Section 4 gives a brief account of the concepts of the game theory and the design mechanism. Section 5 suggests the proposed method. Section 6 explains the implementation and evaluation processes. Finally, conclusions and summarization are given in Section 7.

2. Related Works

The trust and reputation mechanisms have been widely used in order to prevent the selfish behavior of free rides [21,22]. In this way, a record is created for the node in the system lifetime; thus the users with the best reputation will be selected for exchanging the files. So far, in most works, an incentive mechanism designed by means of the game theory (cooperative model and non-cooperative model) has been used [23,24].

At a minimum payment system, a user is rewarded or credited after each operation with a unit of virtual money by either the system or by the user who gains profit, taking the sending of a message or the uploading of a file in response to a request. The minimum payment incentive system was analyzed based on the game theory. In this investigation, the difference between the minimum payment incentive systems was evaluated in comparison with the other incentive systems [25]. Lian has proposed an incentive system based on Tit-for-Tat and Eigen Trust that generates preferences for well-behaved nodes, while correctly punishing the colluders [26]. Centeno has suggested an incentive mechanism based on Q-learning in order to persuade the users to act in a right way [27]. Cui has proposed a new probing structure based on the evolutionary game model to analyze the efficiency of the motivation

mechanism. Chunk exchange between the users are considered as the sender-receiver game model to cover both types of functionality in users [5]. GaMe-PLive is a new structure for P2P streaming based on the game theory. The main achievements of this method are avoiding free-riding and reduction of the loss of data broadcast [28]. Mahini has suggested a new collaborative streaming method by means of the game model and data coding. Interaction of the users was implemented by the Beer-Quiche game. The Nash equilibrium of the proposed method shows that the proposed game works as an incentive mechanism. This method can detect and punish malicious users [29].

The lack of base controller in the P2P systems makes them weak against several attacks such as the existence of Sybil nodes [30]. This kind of nodes tries to breach the network rules by means of collusion and gain of the maximum profit [31]. A Sybil attacker uses multiple fake names (ID) in the system. Even though it has been proven that entirely removing the Sybil attackers is not possible, many research works have been conducted to control this kind of attack [30]. Rowaihy has proposed a new method that can identify the Sybil node's ID by means of their IP address [32]. Dinger and Hartenstein have suggested an algorithm named "self-registration", where one ID registers itself through r other users [33].

The routing strategies are developed by means of a graph named Bootstrap Graph extracted from an interaction model among the users [34]. Haifeng *et al*. have suggested the Sybil Guard that is a distributed algorithm to restrict the entry of Sybils. In this paper, by making an Attack Verge among the users and setting up a track between nodes, it is shown that creating a trusting relationship is possible only between honest nodes, and although Sybil nodes have many edges, they are not capable of setting a trusting interaction. Such Sybil attackers cannot reach their purposes [35].

Xu *et al*. have suggested an algorithm called SRNC, where each attack edge runs with executive order O (1) that shows a better performance compared to the Sybil Limit algorithm with a temporal function [31]. Shareh *et al*. have proposed a mechanism that can defect free riding and Sybil attack simultaneously. In this mechanism, a new reputation formula has been used to track the peer's activities in the network. The reputation formula can penalize malicious users [36].

Liu *et al*. have developed a protocol called ANS to identify the unknown nodes with an intelligent policy and cut off its relationship with its neighbors [16].

The authors in [3] have proposed a motivation mechanism based on the reputation for the P2P networks. The proposed mechanism consists of two main features. The first one is all nodes that have a trust number calculated to analyze the reputation grade precisely. The second feature is a stable service provided to the trusted and loyal nodes. The level of granted service depends on the value of the user's reputation [3].

A few articles have focused on the comparison of the presented incentive mechanisms. In [2], some of the recent methods have been analyzed on the same network to calculate the efficiency of each. The Peer Communications Network is the base for deciding about the performance of each algorithm [2].

A new method based on the game theory has been developed that can reduce the impact of free riders on the Bit Torrent. The proposed model has considered two different modes for un-choking. The first one is the optimistic mode and the other one is a regular mode. The results obtained are analyzed in comparison with the Bit Torrent's main choking algorithm [37].

3. System Model

In this section, a brief explanation about the system used for the proposed method has been presented.

3.1. P2P Networks

The P2P system is a distributed infrastructure whose files and services are spread among nodes. The files and services are exchanged among the users without the help of any server. The P2P system permits nodes to share new files and concurrently use the other node's files [24].

The pattern of managing in this system leads it to get larger rapidly and have numerous files. Joining new nodes to the system can improve the network power and diversity. Due to the duplicated data among the users, there is no single point of failure [24].

3.2. P2P File-sharing Networks

In the recent years, sharing files has been an important usage of the P2P systems. There are three kinds of file-sharing networks. The first one is pure, the second one is centralized, and the third one is a hybrid system. These classes are based on the presence or absence of a central server. The

method proposed in this paper is based on a hybrid system [36].
Joining the users in the system is autonomously. The bandwidth contribution and service level depend on the nodes. The interaction of nodes is modeled by means of a random graph that does not have any particular topology. The edges of the graph are based on the demanded chunks. This is an active graph that updates in each interaction step. The chunks of each file will be sent to a number of users. Then these users swap the chunks. As a result, all demanders will give the file. The distributed servers have an important role to save the treatment of peers in each step of file sharing. Using this information, the peers who have been found to have a malicious function on the network can be recognized and the service delivered to them can be limited. In order to achieve scalability in the P2P network, communications among the servers are handled based on the Chord algorithm. By means of this algorithm, the downloading cost from the servers will be low. The downloaded data is the best criterion for servicing other nodes. Figure 2 shows an overview of the network model.

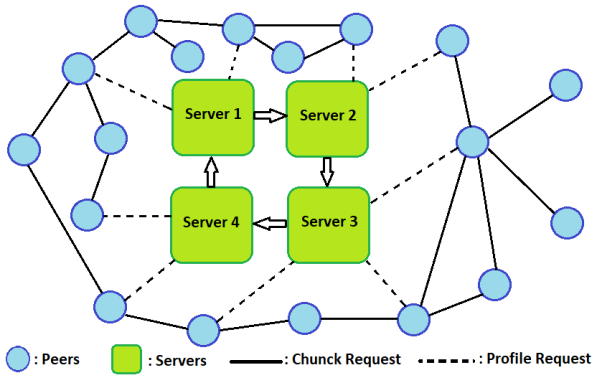

Figure 2. An overview of the P2P file-sharing network.

4. Game Theory and Mechanism Design

Since the method proposed in this paper has been developed based on the game theory and mechanism design, here, more explanations and a literatures review of both mentioned techniques are presented.

4.1. Game Theory

Game theory, a mathematical method invented in the 1950s, has attracted the attentions of many researchers working in various scopes of science from banking and finances, computer engineering, social science, and communication in computer networks. The game theory has a lot of capabilities in communication and distributed system controlling. Therefore, it has been widely employed in computer networks and decentralized systems.
Nash equilibrium and motivation are the most important tools in the game theory. By means of these tools, the user's treatment and decisions can be computed. Indeed, the game theory makes a strong mathematical structure available for analyzing treatment of the users and connections among them. Besides, it is fundamentally distributed and provides an infrastructure for designing the distributed methods in a resource allocation problem. Furthermore, many standard game equations have been extended in the recent years, some involved in the evolutionary game model [38].
The evolutionary gaming approach was first proposed by Maynard in 1972, where the analytical method was integrated with a dynamic evolution process. In order to study the complex system [39], the need for the complex calculation in the Nash Equilibrium networks has highlighted the importance of the evolutionary gaming approach as an alternative method for solving the problems. The evolutionary game theory has been argued to be superior to the Nash Equilibrium approach in the case of offering a more efficient and stable problem-solving method. The evolutionary game theory has been derived from the genetic population model. This method is different from the classic game models as it focuses more on the changes in strategy rather than balancing the strategy. This approach has attracted the attention of the economists. The evolutionary game model is based on a test-and-trial process, giving an insight on identifying the best strategies [4].

4.2. Mechanism Design

The mechanism design is an important field in economics, which leads to a mathematical viewpoint in economics. Thanks to the mechanism design, the economic researchers can design the algorithms, standards, and systems like computer scientists do. It is a strong structure for the social behavior design. Social behavior is a collection of the user's decisions among multiple strategies. The mechanism design provides a framework to persuade the users to select a good social behavior. Participants of the society are completely autonomous, and their strategy selection is in a game theoretic manner [40].
In the recent fifty years, the game theory and the mechanism design have been widely used in the engineering and computer sciences. In particular, the game theory and the mechanism design have been used as a significant structure to model and

analyze the distributed problems in the environments involving numerous users that interact rationally and intelligently. The non-cooperative and incomplete information games are used in the mechanism design. The mechanism design tries to unveil the private information of the users. It also suggests the best social response to each user's strategy. Indeed, the mechanism design can be considered as the reverse engineering of the game theory. In other words, it is the art of designing the desired game rules to achieve a specific result. Designing proper regulations that can satisfy the objectives of the system is the main concern of the mechanism design [41].

5. Proposed Method

The proposed mechanism is in such a way that free riders are removed from the system or are forced to change the strategies, and the players tend not to collude and deviate from the game. The new mechanism has the capability of detecting free riders and Sybil nodes simultaneously. In fact, the mechanism has been designed to detect free riders but the notion of curbing Sybil nodes has been embedded in the heart of the mechanism.

This mechanism is based on reputation. The users' activities are kept in the network to calculate the amount of their reputation and receive a service in accordance with the reputation value. Therefore, the main activities of each peer are effective in its future. At each step of the game, the reputation of each peer is calculated from Equation (1):

$$\theta_i = aC_i + b(S_i \times B_i) \quad (1)$$

where:

θ_i is the amount of the i[th] peer reputation,

C_i is the amount of the i[th] peer centrality,

S_i is the amount of the i[th] peer stability,

B_i is the amount of sharing the i[th] peer bandwidth, and a and b are two constant coefficients with the sum of one that determines the importance of each indicator.

The way to calculate C_i, S_i, and B_i will be explained below. Equation (2) defines when node i will be a free rider. $\bar{\theta}$ is the average peers' reputation.

$$\theta_i < \bar{\theta} \quad (2)$$

A free rider can increase the amount of θ_i with file sharing, staying in the system and bandwidth allocation. The amount of θ_i of each user will not be lost with his/her departure. However, the amount of S_i in every departure will reset. This aims to eliminate "white washers".

5.1. Calculating C_i

In order to detect a Sybil attack, the chunk exchanging graph is stored in the adjacency matrix form. It is a weighted and directed matrix. The edge weight shows the number of chunks, and the edge's direction shows the sender and the receiver. With this graph, the amount of C_i will be calculated. The nodes that have more social relationships take more C_i. In this case, if two Sybil nodes permanently send chunks to each other, their centrality is reduced, and they will lose their reputation. In fact, by using C_i in the reputation formula, the deviation in the game will be prevented. Suppose that the $M_{n\times n}$ matrix is the exchanged chunks, where n represents the number of peers. The centrality of the k^{th} peer(C_k) is calculated by Equation (3):

$$C_k = \frac{\sum_{i=1}^{n}\sum_{j=1}^{M[k,i]}\alpha^j}{\sum_{i=1}^{n} M[i,k]} \quad (3)$$

where, $\alpha < 1$ is a reduction coefficient that is close to 1. For example, using $\alpha = 0.9$, if one peer takes two chunks and gives two chunks to two peers, then its centrality will be $2/2 = 1$, while if it takes two chunks and gives two chunks to one peer, then its centrality will be $1.9/2 = 0.95$. Therefore, the amount of centrality for the peers that have conspired will gradually decrease. It can be said that centrality shows the amount of each node's generosity.

5.2. Calculating S_i

In order to create motivation to continue working in the P2P systems, a parameter called stability will affect the reputation of each node. Stability is any peer's willingness to stay in the system, and is calculated using Equation (4).

$$S_k = \min\left(\frac{t_k}{2 * t_{avg}}, 1\right) \quad (4)$$

where:

S_k is the stability of k in the system,

t_k is the time of peer k work in the current session, and t_{avg} is the mean stability of nodes in the system.

If S_k becomes more than one, it will turn into one. With each exit of peer k, its t_k amount will

be reset to zero. Thus the P2P systems can somewhat undermine the effectiveness of "white-washers" in the system.

5.3. Calculating B_i

B_i is the amount of bandwidth to be shared by peer i. This parameter causes the heterogeneity of bandwidth in a P2P network to have no effect on the system performance. Reaching equilibrium based on a low bandwidth will affect the performance of the entire system. The bandwidth contribution rate is calculated by the functions shown in Equation (5).

$$B_i = f\left(B_s, B_h\right) = \log_{B_h+1}(B_s+1) \qquad (5)$$

where:

B_s is the amount of shared bandwidth, and B_h is the maximum bandwidth of peer i.

The following equations ((6), (7), (8)) define the attributes of Eq. (6), which has led to the selection of this function.

$$f(0, B_s) = 0 \qquad (6)$$

$$f(B_h, B_h) = 1 \qquad (7)$$

$$f(kB_s, kB_h) > f(B_s, B_h) \qquad (8)$$

The third attribute that has been presented by Eq. (8) shows that a peer with less bandwidth must share a larger portion of bandwidth to increase its B_i factor.

5.4. Game Model

Here, the proposed game will be formulated. Game G is defined as $G = \langle N, S, U \rangle$, where N is the set of entire players, $S = \{Defect, Cooperation\}$ is the entire strategies, and U is the outcome function that is shown with the following matrix. Suppose that one peer gains pay-off as much as γ by receiving a chunk from the network. On the other hand, it consumes its bandwidth as much as λ fraction per each service it presents. In addition, a peer like i improves its centrality factor (C_i) with each given service. It considers this value as ΔC_i. According to the formula proposed for calculating the reputation, an outcome like $a\Delta C_i$ will go to this peer, where the a range is $[0,1]$. This value is added to the peer reputation number (θ_i) so that it receives a better service in the next rounds. Now, the one-step matrix of the file sharing game between i *and* j is figured as follows (9):

$$\begin{bmatrix} 0 & \gamma \\ a\Delta C - \lambda & \gamma + a\Delta C - \lambda \end{bmatrix} \qquad (9)$$

If $a\Delta C - \lambda$ is greater than zero, then the game equilibrium will be cooperation/cooperation. For this purpose, it is enough to calculate a as follows (10):

$$a\Delta C - \lambda > 0 \Rightarrow a\Delta C > \lambda \Rightarrow a > \frac{\lambda}{\Delta C} \qquad (10)$$

In (10), λ is a fraction of bandwidth that is used to send a part of the file, and ΔC is the value of centrality increase. The maximum value of λ and ΔC can be calculated given the general characteristics of the P2P networks, and thus the maximum value of the right hand side of the fraction will be achieved. Now, if factor a is determined to be more than this value, the game equilibrium will be cooperation/cooperation. In the following, a game mode will be explored, where a Sybil attack is also considered.

It is claimed that the proposed model has the ability to detect Sybil attacks and can prevent the nodes' collusion. In fact, it blocks the ways to cheat on the mechanism and deviation from the game. In order to prove this claim, the game is modeled into the situations in which some nodes tend not to collude. Suppose that each node in the k^{th} step of the game can choose one of the three strategies namely defect, collusion or cooperation. The strategy space will be changed as follows:

$$S = \{Defect, Collusion, Contribute\} \qquad (11)$$

Defect is the strategy of a free rider. Contribute is the strategy of a normal user, and collusion is the state that one peer gives no services. However, it reports cooperating via another fake ID to show itself as a cooperator without incurring bandwidth and improves its reputation number. In this way, despite the lack of cooperation with the system, it is known as a cooperator peer. The game matrix between two peers i and j can be constructed as follows:

$$\begin{array}{c} \\ Defect \\ Collusion \\ Contribute \end{array} \begin{array}{c} \begin{array}{ccc} Defect & Collusion & Contribute \end{array} \\ \begin{bmatrix} \frac{0}{0} & \frac{0}{b\Delta' C_j} & \frac{\gamma_i}{b\Delta C_j + \lambda_j} \\ \frac{b\Delta' C_i}{0} & \frac{b\Delta' C_i}{b\Delta' C_j} & \frac{\gamma_i + b\Delta' C_i}{b\Delta C_j + \lambda_j} \\ \frac{b\Delta C_i - \lambda_i}{\gamma_j} & \frac{b\Delta C_i - \lambda_i}{\gamma_j + b\Delta' C_j} & \frac{\gamma_i + b\Delta C_i - \lambda_i}{\gamma_j + b\Delta C_j + \lambda_j} \end{bmatrix} \end{array} \qquad (10)$$

With respect to the proposed centrality formula, it can easily show $\Delta' C \le \Delta C$. Now, given the P2P file-sharing network characteristics, the reduction coefficient (b) is determined in such a way that

for each peer as i, the equation $b\Delta' C_i < b\Delta C_i - \lambda_i$ is established, and then the game's equilibrium will be cooperation/cooperation. This means that the users will have no incentive to collude since they gradually receive less outcome if colluded. If the previous equation is not established, two states will exist. First, $b\Delta' C_i = b\Delta C_i - \lambda_i$, in which the game will have four Nash equilibriums. That is to say, peers are indifferent between choosing the cooperation or collusion strategy. Secondly $b\Delta' C_i > b\Delta C_i - \lambda_i$, where collusion/collusion will play the Nash equilibrium. This state is in no way desirable to a social planner.

5.5. Proving Convergence

In this section, it will be shown that if there is any tension in the game balance, the game will converge to the equilibrium point. You are required to make correspondence between this game and that of the prisoner's dilemma game. The file-sharing game has been designed based on rewarding. We reverse the pay-off of the game to be like game penalties in the Prisoner's Dilemma. Thus using the Taylor and Jonker equations [42], it can be shown that the game will be convergent to cooperation/cooperation.

5.6. Pseudo-code of Proposed Method

In the following, we describe the proposed method. These algorithms work according to the suggested game model and the reputation formula introduced.

Algorithm 1 : P_i *as a Server* [36]

```
1: do{
2:     if (P_j demands c_k){//c is a chunk and p is a peer
3:         if (strategy_i == Defect or strategy_i == Collude)
4:             decline(p_j)
5:         else if (strategy_i == Cooperate)
6:             deliver(c_k, P_j)
7:         else if (strategy_i == Reciprocate)
8:             fetch_Profile(P_j)
9:             if (P_j ≠ freeRider)
10:                deliver(c_k, P_j)
11:    }//endif
12:    update(strategy_i)//By means of EvolutionaryModel
13:}while(true)
```

Algorithm 2 : P_i *as a Client* [36]

```
1: do{
2:     demand(P_j, C_k) for some j, k
3:     if (recieved(c_k))
4:         announce(positive, Shared_Bandwidth_j, P_j)
5:     else
6:         announce(negative, 0, Peer_j)
7:     if (strategy_i = Collude)
8:         for each j that P_j Colludes P_i
9:             announce(positive, Bandwidth_j P_j)
10: while(true)
```

Pseudo-codes are written in three sections for the users in server or client mode and for servers. Algorithm 1 and algorithm 2 are the same as the algorithms proposed in [36] but algorithm 3 is dedicated for this article proposed method. Algorithm 1 shows the peer function as the server. If a peer requires a file and sends a request to another user, it will report the user's performance to the server. In case the file is received, a positive report, and otherwise, a negative report will be sent to the server. The colluder peers performing Sybil attacks always send a positive report to their partners. Algorithm 2 shows the function of a user as a client.

On the other hand, the servers receive the user performance reports in each round of the game and store them in the user profile. At the end of each round, the service graph of each user is modified and its centrality is calculated. Then using Formula 1, the reputation of each peer is calculated. A user whose reputation is less than the average of all users' reputation is known as a free rider. Algorithm 3 shows the server performance in each round of the game.

Algorithm 3: *Server*

```
1: do{
2:     for each fetch _Profile(P_i)
3:         if (θ_i < θ̄)
4:             send(freeRider)
5:         else
6:             send(not freeRider)
7:     wait for reports
8:     for ech P_i in P 2P file sharing network {
9:         update(C_i) //Based on proposed centrality
10:        update(S_i) //Based on proposed stability formula
11:        update(B_i) //Based on proposed bandwidth formula
12:        θ_i = aC_i + b(S_i × B_i)
13:    }//end for
14:    update(θ̄) //θ̄ is average of θ
15: while(true)
```

6. Results and Discussions

The simulation parameters of the proposed method are listed in table 1. C is the cost of providing service, C_r is the strategy inquiry cost of a node, Q is the cost-revenue ratio in a service, K is the noise or error in Fermi Equation [43, 44], and α is the reduction coefficient in Eq (3). a and b in Eq (1) are 0.6 and 0.4, respectively. The simulations were implemented in the Matlab software.

Table 1. Simulation parameters.

Parameter	Value
Population size	10000
Game rounds	200
Number of neighbors	8-16
C:Cost of providing service	1
Cr: Inquiry cost	0.1
Q: Cost/revenue ratio	0.35
K: Noise	0.1
α: Reduction coefficient	0.99

This network has 10,000 users. The bandwidth of each node is a random number between one and ten, where each node shares at its disposal a part of it. The network has been intended as Mesh, and each node in each game round is associated with eight random nodes. In each round, each user sends a request to all of its neighbors, and every neighbor will respond according to their strategy. Playing starts with equal percentages of cooperator, defector users, and reciprocator. A cooperator provides services in any condition. A defector is a free rider, and a reciprocator provides service if the request is not from a defector.

The game's evolutionary model is done based on the FERMI equation [43, 44]. Only two colluding users on the network are supposed to be randomly distributed in the network. The numerical results in figure 3 show that by increasing the rounds, the nodes proceed towards cooperating or reciprocating to achieve the greatest benefit by changing their strategy. The red, green, and blue colors are assigned to the defector, reciprocator, and co-operator, respectively.

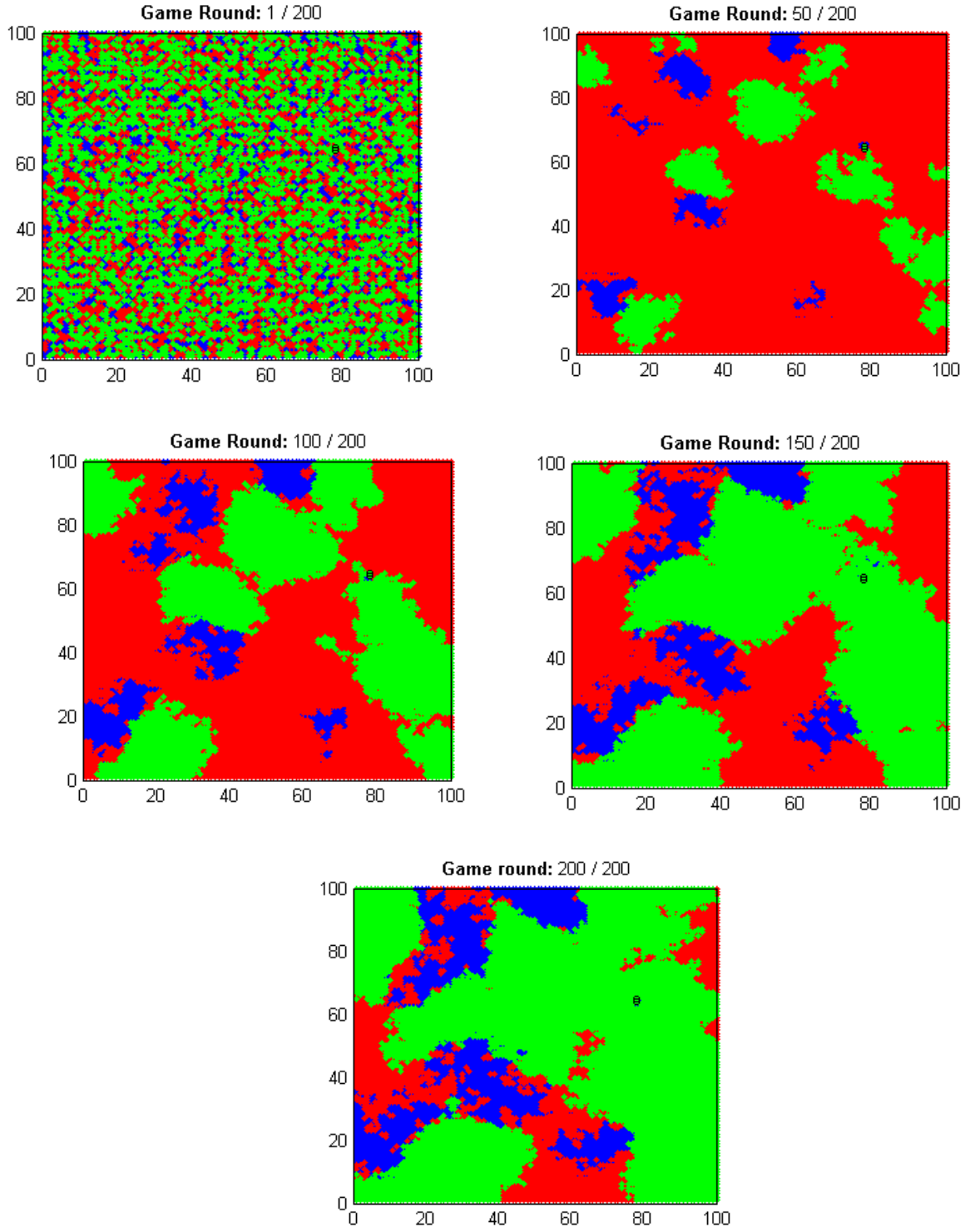

Figure 3. Simulation results for 200 rounds of the game. Red, green, and blue are assigned to the defector, reciprocator, and co-operator, respectively.

Figure 4 shows the effect of centrality relation in detecting Sybil attacks. As the game goes on, the average consumed service (ACS) of Sybil users will be decreased.

Figure 5 shows the Sybil node compared to the other nodes in the system. Gradually, more players identify Sybil nodes. That is why the growth of the total services they receive is reduced.

Figure 6 shows the strategy distribution in each game round. As the game goes on, the number of defectors (free riders) decreases gradually and the users tend to contribute.

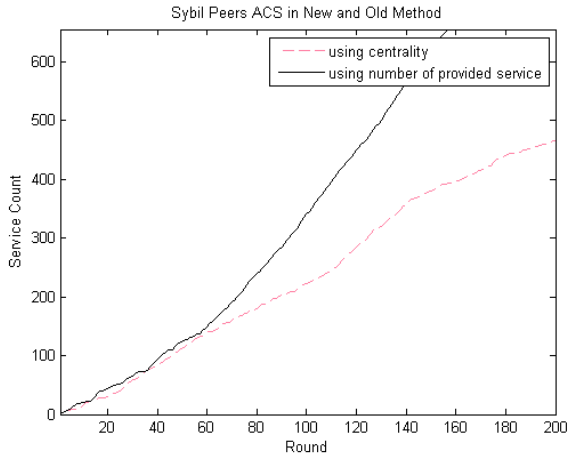

Figure 4. Effect of centrality.

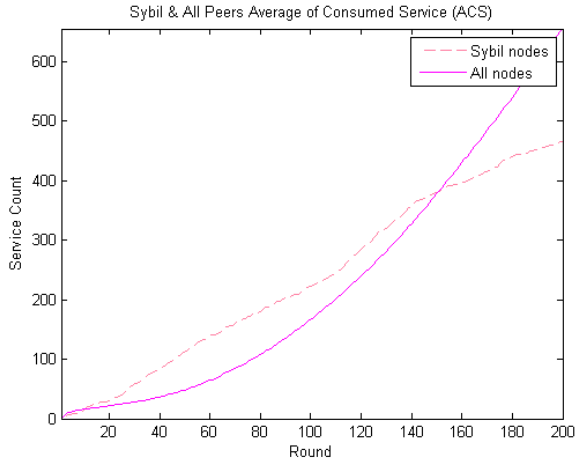

Figure 5. Identifying Sybil nodes.

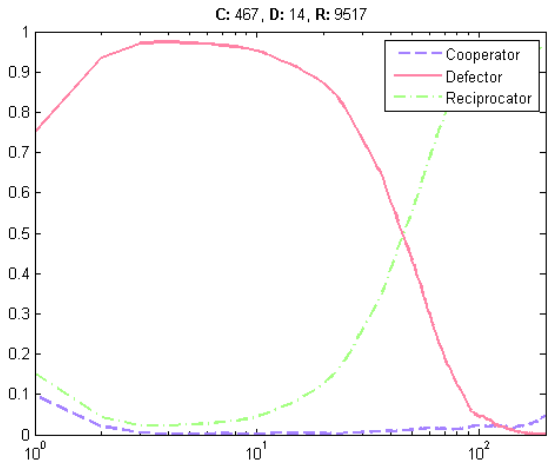

Figure 6. Strategy distribution in each game round.

The use of evolutionary games makes it easy to identify and restrict free-riders. Figure 7 shows the successful download rate of free-riders and reciprocators in different game periods. The more we approach the end of the game, the more free-riders are restricted.

In the simulations conducted, it is assumed that there are two Colluder users. Figure 8 shows the average of incoming services in the two hundredth round of the game for a higher number of Colluder users.

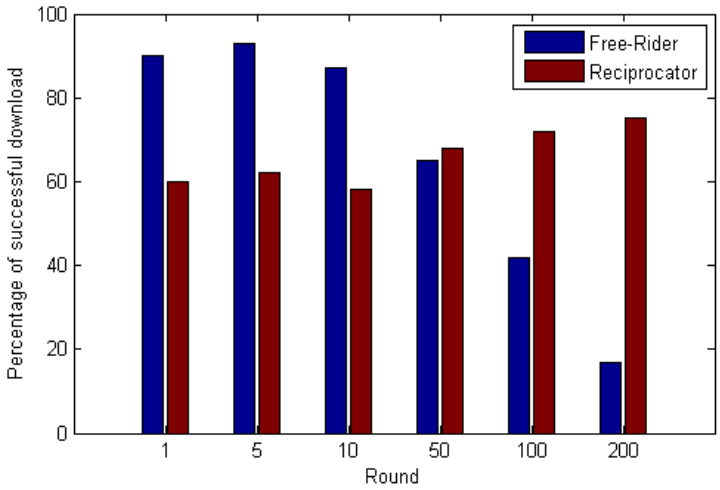

Figure 7. Percentage of successful download.

The results obtained show that as the number of Colluder users increases, the service received by them will increase accordingly. However, the provided colluders form a large part of users, the amount of service received will be reduced for all users. With up to 10% of the initial Sybil nodes (colluders), the proposed method can handle the system properly. Beyond that, Sybil nodes will not be able to receive a high quality service but because of their collusion, the entire system will face a tremendous decrease in the delivered files.

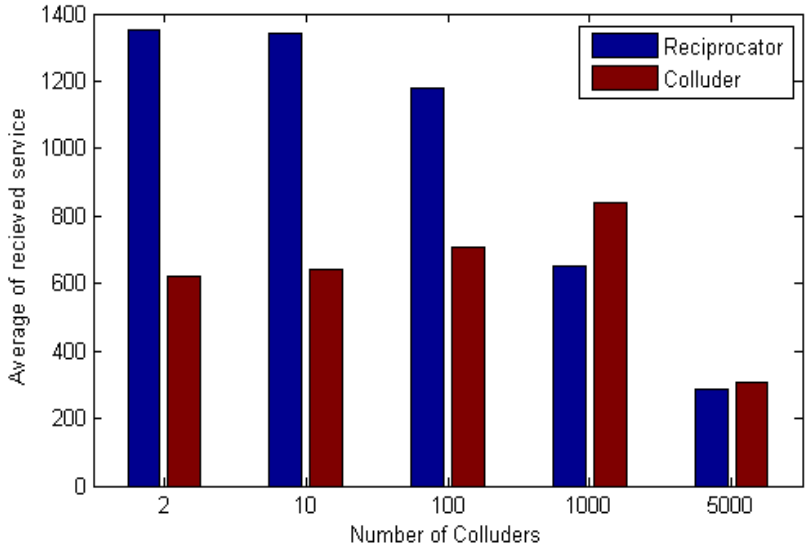

Figure 8. Average of received service in round 200.

The proposed method is examined assuming 30% free rider user at the beginning of the game. If the amount of free riders changes at the start of the game, the performance of the proposed incentive mechanism will also change. Figure 9 shows the percentage of users with each strategy in round 200 for different values of the free rider. With up to 40% of the initial free riders, the proposed method works properly. It has some positive impacts up to 60% of the free rider. However, it loses its effectiveness from 60% to 80%, and from then on, it will have a reverse effect.

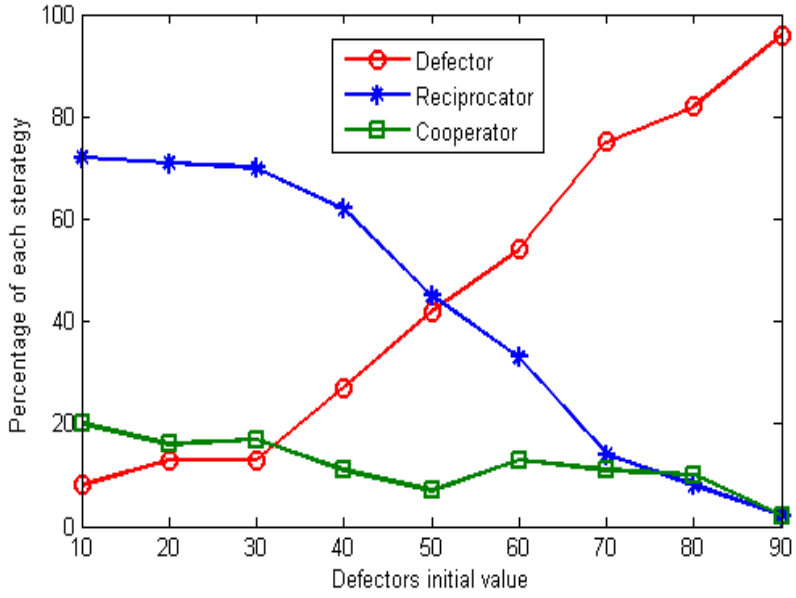

Figure 9. Effect of number of free-rider on strategy.

In order to investigate the effect of B_i on the reputation, regardless of the amount of bandwidth, the users have been classified into two different groups: users with B_i greater than 0.5 and users with B_i lower than 0.5. The simulation results by the parameters in table 1 show that the average of consumer services of users with $B_i > 0.5$ is greater by about 15%. This means that each user can improve its B_i to get more services. According to Equation (5), the users with a low bandwidth can improve their B_i factor by sharing a large portion or all bandwidth but the users with a very high bandwidth can do the same but by sharing only a small amount of bandwidth. Figure 10 shows the average of incoming services in two

modes. In one mode, a bandwidth condition is considered, and in the other mode, a bandwidth has no effect on the reputation.

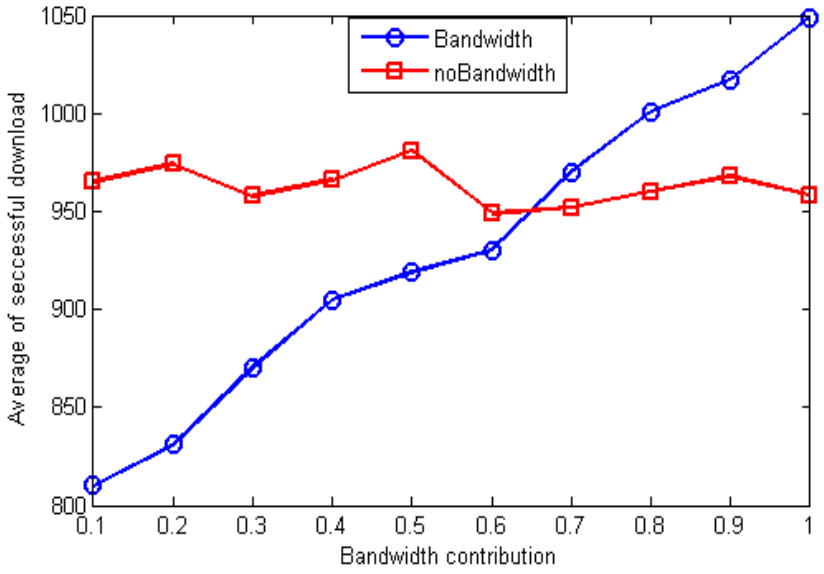

Figure 10. Effect of bandwidth contribution formula.

7. Conclusions

The peer-to-peer (P2P) file sharing networks work collaboratively, and lack a central controller. Hence, various damages threaten these networks. The free rider users try to use the system features without providing the service to the others. The other users may try to break the system rules by colluding and providing false reports.

In this paper, a reputation-based method was presented, while "Sybil attack discovery" was contributed to the file-sharing game in a heterogeneous P2P network. In this way, some limitations have been assigned to the Sybil nodes, and free riders have been blocked. One of the reputation factors is the amount of shared bandwidth.

This factor was designed to protect the users with less bandwidth, and face no penalty in case of full sharing. The proposed method gradually detects the free riders and Sybil. In the future, the users can be detected more rapidly by designing trusted nodes in the network structure. Also adding new factors to the reputation formula makes the game more complete.

References

[1] Lua, E. K., Crowcroft, J., Pias, M., Sharma, R., & Lim, S. (2005). A survey and comparison of peer-to-peer overlay network schemes. IEEE Communications Surveys & Tutorials, vol. 7, no. 2, pp. 72-93.

[2] Guo, D., Kwok, Y. K., Jin, X., & Deng, J. (2016). A performance study of incentive schemes in peer-to-peer file-sharing systems. The Journal of Supercomputing, vol. 72, no. 3, pp. 1152-1178.

[3] Chang, J., Pang, Z., Xu, W., Wang, H., & Yin, G. (2014). An incentive compatible reputation mechanism for P2P systems. The Journal of Supercomputing, vol. 69, no. 3, pp. 1382-1409.

[4] Zhang, Q., Xue, H. F., & Kou, X. D. (2007). An evolutionary game model of resources-sharing mechanism in P2P networks. In Intelligent Information Technology Application, Workshop on (pp. 282-285). IEEE.

[5] Cui, G., Li, M., Wang, Z., Ren, J., Jiao, D., & Ma, J. (2015). Analysis and evaluation of incentive mechanisms in p2p networks: a spatial evolutionary game theory perspective. Concurrency and Computation: Practice and Experience, vol. 27, no. 12, pp. 3044-3064.

[6] Obele, B. O., Ukaegbu, A. I., & Kang, M. (2009). On tackling free-riders in P2P networks. In Advanced Communication Technology, 2009. ICACT 2009. 11th International Conference on (vol. 3, pp. 2084-2089). IEEE.

[7] Rozario, F., Han, Z., & Niyato, D. (2011). Optimization of non-cooperative P2P network from the game theory point of view. In 2011 IEEE Wireless Communications and Networking Conference (pp. 868-873). IEEE.

[8] Atlidakis, V., Roussopoulos, M., & Delis, A. (2014). EnhancedBit: Unleashing the potential of the unchoking policy in the BitTorrent protocol. Journal of

Parallel and Distributed Computing, vol. 74, no. 1, pp. 1959-1970.

[9] Manoharan, S., & Ge, T. (2013). A demerit point strategy to reduce free-riding in BitTorrent. Computer Communications, vol. 36, no. 8, pp. 875-880.

[10] Izhak-Ratzin, R. (2010). Improving the BitTorrent protocol using different incentive techniques (Doctoral dissertation, University of California Los Angeles).

[11] Wang, H., Wang, F., Liu, J., Xu, K., & Wu, D. (2013). Torrents on twitter: Explore long-term social relationships in peer-to-peer systems. IEEE Transactions on Network and Service Management, vol. 10, no. 1, pp. 95-104.

[12] Hughes, D., Coulson, G., & Walkerdine, J. (2005). Free riding on Gnutella revisited: the bell tolls?. IEEE distributed systems online, vol. 6, no. 6.

[13] Ren, Y., Li, M., Xiang, Y., Cui, Y., & Sakurai, K. (2013). Evolution of cooperation in reputation system by group-based scheme. The Journal of Supercomputing, vol. 63, no. 1, pp. 171-190.

[14] Alotibi, B., Alarifi, N., Abdulghani, M., & Altoaimy, L. (2019). Overcoming Free-Riding Behavior in Peer-to-Peer Networks Using Points System Approach. Procedia Computer Science, vol. 151, pp. 1060-1065.

[15] Jain, A., & Kumar, S. (2018). FriendShare: A secure and reliable framework for file sharing on network. Journal of Network and Computer Applications, vol. 120, pp. 1-16.

[16] Al-Qurishi, M., Alrubaian, M., Rahman, S. M. M., Alamri, A., & Hassan, M. M. (2018). A prediction system of Sybil attack in social network using deep-regression model. Future Generation Computer Systems, vol. 87, pp. 743-753.

[17] Ramalingam, D., & Chinnaiah, V. (2018). Fake profile detection techniques in large-scale online social networks: A comprehensive review. Computers & Electrical Engineering, vol. 65, pp. 165-177.

[18] Chen, Z., Cheng, Y., Deng, X., Qi, Q., & Yan, X. (2019). Agent incentives of strategic behavior in resource exchange. Discrete Applied Mathematics, vol. 264, pp. 15-25.

[19] Sahoo, S. R., & Gupta, B. B. (2019). Hybrid approach for detection of malicious profiles in twitter. Computers & Electrical Engineering, vol. 76, pp. 65-81.

[20] Zarezade, M., Nourani, E., & Bouyer, A. (2020). Community detection using a new node scoring and synchronous label updating of boundary nodes in social networks. Journal of AI and Data Mining, vol. 8, no. 2, pp. 201-212.

[21] Zhou, R., & Hwang, K. (2007). Powertrust: A robust and scalable reputation system for trusted peer-to-peer computing. IEEE Transactions on parallel and distributed systems, vol. 18, no. 4, pp. 460-473.

[22] Kamvar, S. D., Schlosser, M. T., & Garcia-Molina, H. (2003). The eigentrust algorithm for reputation management in P2P networks. In Proceedings of the 12th international conference on World Wide Web (pp. 640-651). ACM.

[23] Chen, Z., Qiu, Y., Liu, J., & Xu, L. (2011). Incentive mechanism for selfish nodes in wireless sensor networks based on evolutionary game. Computers & Mathematics with Applications, vol. 62, no. 9, pp. 3378-3388.

[24] Wang, Y., Nakao, A., Vasilakos, A. V., & Ma, J. (2011). On the effectiveness of service differentiation based resource-provision incentive mechanisms in dynamic and autonomous P2P networks. Computer Networks, vol. 55, no. 17, pp.3811-3831.

[25] Figueiredo, D., Shapiro, J., & Towsley, D. (2005). Incentives to promote availability in peer-to-peer anonymity systems. In 13TH IEEE International Conference on Network Protocols (ICNP'05) (pp. 12-pp). IEEE.

[26] Lian, Q., Yu, P., Yang, M., Zhang, Z., Dai, Y., Li, X., & Yu, R. P. (2007). Robust incentives via multi-level tit-for-tat, Currency and Computation: practice and experience, vol. 20, no. 2, pp. 167-178

[27] Centeno, R., Billhardt, H., & Hermoso, R. (2013). Persuading agents to act in the right way: An incentive-based approach. Engineering Applications of Artificial Intelligence, vol. 26, no. 1, pp. 198-210.

[28] Mahini, H., Dehghan, M., Navidi, H., & Masoud Rahmani, A. (2016). GaMe-PLive: a new game theoretic mechanism for P2P live video streaming. International Journal of Communication Systems, vol. 29, no. 6, pp. 1187-1203.

[29] Mahini, H., Dehghan, M., Navidi, H., & Rahmani, A. M. (2017). Peer-assisted video streaming based on network coding and Beer-Quiche game. AEU-International Journal of Electronics and Communications, vol. 73, pp. 34-45.

[30] Douceur, J. R. (2002). The sybil attack. In International Workshop on Peer-to-Peer Systems (pp. 251-260). Springer Berlin Heidelberg.

[31] Xu, L., Chainan, S., Takizawa, H., & Kobayashi, H. (2010). Resisting sybil attack by social network and network clustering. In Applications and the Internet (SAINT), 2010 10th IEEE/IPSJ International Symposium on (pp. 15-21). IEEE.

[32] Rowaihy, H., Enck, W., McDaniel, P., & La Porta, T. (2007, May). Limiting sybil attacks in structured p2p networks. In IEEE INFOCOM 2007-26th IEEE International Conference on Computer Communications (pp. 2596-2600). IEEE.

[33] Dinger, J., & Hartenstein, H. (2006). Defending the sybil attack in P2P networks: Taxonomy, challenges, and a proposal for self-registration. In First International Conference on Availability, Reliability and Security (ARES'06) (pp. 8-pp). IEEE.

[34] Trifa, Z., & Khemakhem, M. (2012). Mitigation of Sybil Attacks in Structured P2P Overlay Networks. In Semantics, Knowledge and Grids (SKG), 2012 Eighth International Conference on (pp. 245-248). IEEE.

[35] Yu, H., Kaminsky, M., Gibbons, P. B., & Flaxman, A. D. (2008). Sybilguard: defending against sybil attacks via social networks. IEEE/ACM Transactions on networking, vol. 16, no. 3, pp. 576-589.

[36] Shareh, M. B., Navidi, H., Javadi, H. H. S., & HosseinZadeh, M. (2019). Preventing Sybil attacks in P2P file sharing networks based on the evolutionary game model. Information Sciences, vol. 470, pp. 94-108.

[37] Azzedin, F., & Yahaya, M. (2016). Modeling BitTorrent choking algorithm using game theory. Future Generation Computer Systems, vol. 55, pp. 255-265.

[38] Pavel, L. (2012). Game theory for control of optical networks. Springer Science & Business Media.

[39] Feng, H., Zhang, S., Liu, C., Yan, J., & Zhang, M. (2008). P2P incentive model on evolutionary game theory. In 2008 4th International Conference on Wireless Communications, Networking and Mobile Computing (pp. 1-4). IEEE.

[40] Nisan, N., Roughgarden, T., Tardos, E., & Vazirani, V. V. (Eds.). (2007). Algorithmic game theory (Vol. 1). Cambridge: Cambridge University Press.

[41] Narahari, Y. (2014). Game theory and mechanism design (Vol. 4). World Scientific, IISc Press.

[42] Alexander, J. M. (2002). Evolutionary game theory. Sidney, Stanford Press.

[43] Szabó, G., & Vukov, J. (2004). Cooperation for volunteering and partially random partnerships. Physical Review E, 69(3 Pt 2):036107.

[44] Szabó, G., & Tőke, C. (1998). Evolutionary prisoner's dilemma game on a square lattice. Physical Review E, vol. 58, no. 1, pp. 69-73.

نشریه هوش مصنوعی و داده کاوی

یک سازوکار جدید برای تشخیص و محدودسازی گره‌های سیبل در شبکه‌های اشتراک فایل نظیربه‌نظیر با پهنای باند ناهمگن

مرتضی بابازاده شاره[1]، حمیدرضا نویدی قاضیانی[2,*]، سید حمید حاج سیدجوادی[2] و مهدی حسین زاده[3]

[1] گروه مهندسی کامپیوتر، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران.

[2] گروه علوم کامپیوتر، دانشگاه شاهد، تهران، ایران.

[3] دانشگاه علوم پزشکی ایران، تهران، ایران.

ارسال ۲۰۱۹/۱۱/۱۸؛ بازنگری ۲۰۲۰/۰۴/۰۱؛ پذیرش ۲۰۲۰/۰۶/۲۴

چکیده:

در شبکه‌های اشتراک فایل نظیربه‌نظیر دو نوع گره غیر مجاز با نام رایگان‌سوار و سیبل وجود دارد. رایگان‌سوارها گره‌هایی هستند که قصد دارند بدون هزینه سرویس دریافت کنند. گره‌های سیبل، کاربران واقعی با چندین شناسه تقلبی هستند. تکنیک‌های گوناگونی مانند تیت-فور-تات و سیبل گارد برای مقابله با این کاربرها وجود دارد. روش‌های پیشگیری از رایگان‌سواری به سادگی توسط سیبل‌ها فریب داده می‌شوند و روش‌های فعلی مقابله با سیبل‌ها تنها روی مشخصات شبکه تاکید دارند که هیچ ارتباطی با رایگان‌سواری ندارد. هیچ روشی برای مقابله همزمان با این دو مشکل وجود ندارد. بنابراین، مهمترین هدف این پژوهش طراحی یک روش جامع است که با کمک نظریه‌بازی‌ها با هر دو نوع کاربر مخرب به شکل همزمان مقابله کند. به عبارت دیگر یه سازوکار خلاقانه طراحی شده است که امکان دور زدن آن توسط سیبل‌ها وجود ندارد. مهمترین ایده در این سازوکار، تعریف یک مرکزیت جدید و روابطی برای میزان اعطای پهنای باند است. نتایج شبیه‌سازی روش‌پیشنهادی نشان می‌دهد که با گذشت عمر شبکه گره‌های رایگان-سوار شناسایی شده و سرویس‌های شبکه برای کاربران سیبل کاهش می‌یابد. به کمک سازوکار پشنهادی، شبکه قادر است حداکثر ۴۰٪ گره رایگان‌سوار و ۱۰٪ گره سیبل را تحمل کرده و به درستی به کاربران سرویس ارائه کند.

کلمات کلیدی: شبکه‌های اشتراک فایل، شبکه‌های نظیربه‌نظیر، رایگان‌سوار، حمله سیبل، مکانیزم انگیزه.