A COALITION BASED INCENTIVE MECHANISM FOR P2P CONTENT DISTRIBUTION SYSTEMS

M. V. Belmonte, M. Díaz and A. Reyna
E.T.S.I. Informática, Bulevar Louis Pasteur, N.35, Universidad de Málaga, 29071 Málaga, Spain

Keywords: Group decision making, P2P, Cooperation, Game theory, Coalitions, Economic agent models, Free-riding.

Abstract: P2P systems suffer from free-loaders, peers that consume many more resources or contents (bandwidth) than they contribute. One of the reasons for this is that the mechanisms used for downloading and sharing in the P2P systems, do no take selfish behavior of the peers into account at the design stage. Therefore, it is important to find mechanisms that provide incentives and encourage cooperative behavior among the peers. One possible solution could be to use an economic framework that provides them with incentives. We propose the application of a coalition formation scheme based on game theory to P2P file sharing systems. The main idea for the coalition formation scheme is based on the fact that peers that contribute more get a better quality of service. A peer that participates in a coalition lends "bandwidth" to other peers of the coalition, in exchange for utility and consequently far greater download bandwidth. Simulation results have shown the effectiveness of the mechanism in stopping the free-riding peers and encouraging cooperation, increasing the performance of a P2P network and obtaining an improvement in time download performance.

1 INTRODUCTION

In general, a P2P content distribution system creates a distributed storage medium that allows the publishing, searching and retrieval of files by members of the network (Androutsellis-Theotokis, 2004).

Traditionally, the main problem of the P2P systems is limited to file search. However, the efficient download of content and the fairness in the bandwidth contribution is also an important aim in the design goal of these kind of systems. The early P2P systems (Gnutella, Kazaa,...) lack mechanisms for fairness in bandwidth usage. For this reason, these systems suffer from free-loaders, peers that consume many more resources or contents (bandwidth) than they contribute. In (Sariou et al., 2002) and (Handurukande et al., 2006) empirical studies have observed this behavior in Napster, Gnutella or even eDonkey.

One of the reasons for this problem is that the mechanisms used for downloading and sharing in the P2P systems, do not take the selfish behavior of the peers into account at the design stage. P2P system users act rationally trying to maximize the benefits obtained from using the system's shared resources (Golle et al., 2001). Therefore, it is important to find mechanisms that provide incentives and encourage cooperation. One possible solution could be to use an economic framework that provides incentives. In this sense game theory may be a good tool on which model the interactions between peers in a P2P file sharing system. The idea is to define "the rules of the game" so that the system as a whole exhibits good behavior when self-interested nodes pursue self-interested strategies (mechanism design (Shneidman, 2003)).

Our approach proposes the application of a coalition formation scheme based on game theory to P2P file sharing systems (in (Belmonte et al., 2006b) we presented an early version of this work). The main idea of the coalition formation scheme is the fact that peers which contribute more get a better quality of service. We define a "responsiveness bonus" that reflects the peer's overall contribution to the system, and we use the game theory utility concept to calculate it. It is possible to form a coalition among peers with a re-distribution of the number of bytes to be transferred. A peer that participates in a coalition lends "bandwidth" to other peers of the coalition, in exchange for utility; and this utility will increase its responsiveness bonus. The coalition formation scheme rewards the peers with a higher responsiveness bonus (therefore giving them greater bandwidth to download files), and penalizes the ones that only consume

V. Belmonte M., Díaz M. and Reyna A..
A COALITION BASED INCENTIVE MECHANISM FOR P2P CONTENT DISTRIBUTION SYSTEMS.
DOI: 10.5220/0003140200150024
In *Proceedings of the 3rd International Conference on Agents and Artificial Intelligence* (ICAART-2011), pages 15-24
ISBN: 978-989-8425-41-6

resources, decreasing their responsiveness bonus and consequently their bandwidth.

The proposed incentive mechanism encourages cooperative behavior between the peers preventing the free-riding problem. Using the game theory concept of ”core”, the peers forming the coalition get in return a fair utility in relation to the bandwidth they supply (achieving fairness in bandwidth sharing); And in addition, it allows the formation of coalitions of peers that help each other in downloading files, increasing the performance of the P2P network.

2 DOWNLOADING WITH COALITIONS

In this section, we describe the model of the environment in which the system is deployed and the mechanism of coalition formation among peers. We firstly describe a simplified situation, illustrating the advantages of forming coalitions for P2P downloads and the way of computing and dividing the utility or profit obtained by peer that participates in the coalition. Secondly, and in more general terms, we describe the coalition formation process and how the data and the bandwidth are distributed among the coalition members.

2.1 P2P Network Type

For our work, we have selected a P2P system with a partially centralized architecture and an unstructured network. The first characteristic is related to the degree of centralization of the peer's network, and the second with the fact that the network is created in a non-deterministic way as peers and files are added. When a peer wants to download a file, it directs its request to a supernode and this searches the file in its index (a supernode is a peer that acts as a central index for files shared for a subpart of the peer network). When the file is located, supernode sends to the ”requester” peer an indexed result with the set of nodes that store the requested file. Then, the requester peer opens a direct connection with one or more peers that hold the requested file, and proceeds to download it.

2.2 Coalition Formation Model

Coalition formation is an important mechanism for cooperation in Multi-Agent Systems (MAS). In order to be used by autonomous agents, a coalition formation mechanism must solve the following issues: i) maximize the agents' profit or utility. For every coalition S, coalitional value $V(S)$ must be computed, i.e., the total utility obtained by S as a whole ii) divide the total utility among agents in a fair and stable way, so that the agents in the coalition are not motivated to abandon it. For every coalition S and every agent $i \in S$, payment configuration $x(i)$ must be computed, i.e., the share of $V(S)$ that is assigned to the agent i. iii) do this within a reasonable amount of time and using a reasonable amount of computational efforts. Our coalition formation model allows cooperation to take place among autonomous, rational and self-interested agents in a class of superadditive task oriented domains (Belmonte et al., 2006a). Each agent has the necessary resources to carry out its own task, however it is possible to form a coalition among agents with a new re-distribution of the task that may allow them to obtain benefits. The proposed model guarantees an optimum task allocation and a stable payoff division. Furthermore, computational complexity problems are solved.

In this section the coalition formation scheme is applied with the goal of improving the performance of P2P file exchange systems. In this case, the central idea is based on sharing the task of downloading a file among a set of peers forming a coalition. From the point of view of the peer that wants to download the file there is a clear advantage, since the total download time is reduced. From the point of view of uploading peers, for each one the task of transferring the file is alleviated, since it is divided between the members of the coalition.

Let us consider the simplified situation illustrated in figure 1. In this scenario, p_b, asks p_a for a file Z. This peer p_a forms a coalition S with three other nodes p_h, p_l and p_m to transfer that file. In P2P file sharing systems, every node p_m has an upload b_i^{in} and download b_i^{out} bandwidth dedicated to file sharing. Usually these bandwidths are user defined and indicate maximum values, and b_i^{in} is much lower than b_i^{out}. This simplified scenario can be generalized, we can consider that the downloader peer splits its bandwidth in order to perform simultaneous downloads, determining the b_i^{out} dedicated for each download. The model is valid for both scenarios.

In general, there will be an initial uploading agent p_0 (in the figure p_a) and a set of n additional uploading agents, $p_1, \ldots, p_n$ (in the figure p_h, p_l and p_m), all of which have the file that has to be downloaded and they dedicate their upload capacity b_i^{in} to this transfer. Let us call $size(Z) = T$ the size of the file to download. Let us also assume that $\sum_0^n b_i^{in} \leq b_b^{out}$.

Then an estimation of the time necessary for the transfer of file Z by the coalition S is given by the ratio between the size of the file and the coalition bandwidth(1):

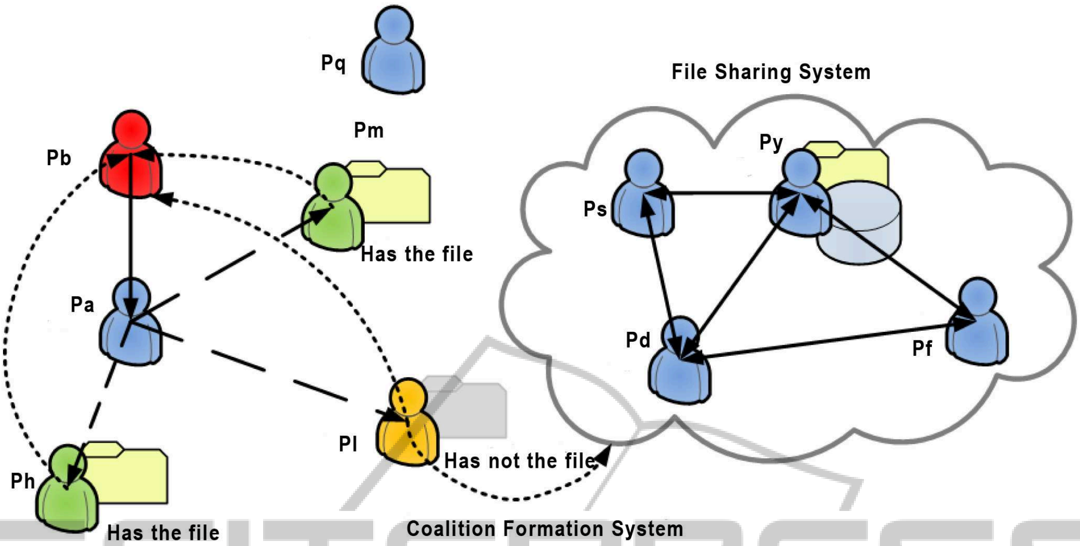


Figure 1: The coalition formation model.

$$t_S = \frac{T}{b_0^{in} + \sum_1^n b_i^{in}} \quad (1)$$

On the other hand, if the coalition is not formed, the time for just an uploading agent p_0 is given by (2):

$$t_0 = \frac{T}{b_0^{in}} \quad (2)$$

Therefore the value obtained by the coalition S can be defined as Δt, the difference between (2) and (1), as shown in (3):

$$\Delta t = \frac{T}{b_0^{in}} - \frac{T}{b_0^{in} + \sum_1^n b_i^{in}} = \frac{T}{b_0^{in}} \frac{\sum_1^n b_i^{in}}{\sum_0^n b_i^{in}} = t_0 \frac{\sum_1^n b_i^{in}}{\sum_0^n b_i^{in}} \quad (3)$$

Of course, if $p_0 \notin S$, $V(S) = 0$. To sum up, the coalitional value for every coalition is given by (4):

$$V(S) = \begin{cases} t_0 \frac{\sum_1^n b_i^{in}}{\sum_0^n b_i^{in}} & \text{if } p_0 \in S \\ 0 & \text{if } p_0 \notin S \end{cases} \quad (4)$$

Now we address the following problem, to define a stable payoff division of $V(S)$ between agents, i. e., given a partition of the set of all agents into different coalitions, to assign an amount $x(i)$ to every agent i. The problem is to distribute the utility in a fair and stable way, so that the agents in the coalition are not motivated to abandon it. Game theory provides different concepts (core, kernel, Shapley value, etc.) for the stability of coalitions (Kahan, 1984). The core is the simplest to define. A payment configuration belongs to the core if there is no other coalition that improve on the payoffs of all of its members.

Formally, let N be the set of all agents, and let us denote $x(S) = \sum_{i \in S} x(i)$. The payoff division lies inside the core iff the following holds: i) for all sets of agents $S \subseteq N$, $V(S) \leq x(S)$ (group rationality); and ii) $V(N) = x(N)$ (global rationality). The existence of the core is not guaranteed in the general case, in a given situation the core may be empty. However, we will show for our case a payoff division that lies inside the core. We will continue the ideas presented in (Belmonte et al., 2006a).

The proposed payoff division scheme is calculated by means of the marginal profit concept. Thus, the payment to each agent is given by the marginal profit according to the resources that describe that agent, and multiplied by the value of the resources. Since the concept of marginal profit is really that of partial derivative, the payment vector x will be computed as follows: for the original uploading agent, p_0, there are 2 parameters (t_0 and b_0^{in}), hence its payment is be given by $t_0 \frac{\partial V}{\partial t_0} + b_0^{in} \frac{\partial V}{\partial b_0^{in}}$. For the remaining agents, there is only one parameter (b_i^{in}), hence their payments are $b_i^{in} \frac{\partial V}{\partial b_i^{in}}$. Finally, we obtain the expressions shown in equations (5):

$$x_i = \begin{cases} t_0 \frac{(\sum_1^n b_i^{in})^2}{(\sum_0^n b_i^{in})^2} & \text{if } i = 0 \\ t_0 \frac{b_0^{out} b_i^{in}}{(\sum_0^n b_i^{in})^2} & \text{if } i \neq 0 \end{cases} \quad (5)$$

In appendix A we prove that equations (5) define a payoff division that lies inside the core.

Finally, we will show that the computations can be done within a reasonable amount of time and using a reasonable amount of computational efforts. It is obvious that the above schema provides a set of explicit formulae that compute payments in time linear with the number of peers.

2.3 Data and Bandwidth Distribution Model

Following on from the example, let us suppose that the peer p_b asks p_0 (p_a in the figure) to download the file Z, and p_0 decides to initiate a coalition to download the file (figure 1). Then, p_0 carries out the following steps:

1. To set the coalition size. If the sum of the upload bandwidth, b_i^{in}, of all the interested peers (let us suppose this value is equal to n) joined with p_0, is lower than the download bandwidth of p_b, b_b^{out}, we are in the trivial case (equation 6). In this case, all the interested peers joined with p_0 will form the coalition. So, $\mid S \mid = N$:

   $$(\sum_{i \in N} b_i^{in}) \leq b_b^{out} \tag{6}$$

   Conversely if the expression (6) is false, we must distribute the bandwidth of p_b, b_b^{out} among all the interested peers[1].

   For this, it is necessary to distribute the b_b^{out} of p_b among all of them. This can be done by ”the progressive filling algorithm” (Ma et al., 2006). Let us suppose w_i is the assigned bandwidth to the interested peers. The algorithm initializes the bandwidth of all the interested peers to 0, $w_i = 0$, $\forall i \in S$. Then, it increases all the bandwidths at the *same rate*, until one or several peers hit their limits, $w_i = b_i^{in}$, $\forall i \in S$. Once the bandwidth assigned to one peer, p_i, reaches its limit, it is taken out of the process. The algorithm will continue to increase the bandwidth of the remaining peers at the *same rate*. The algorithm will finish when all the peers reach their limits, or when the bandwidth of p_b, b_b^{out} is wasted.

   This algorithm provides the *max-min fairness* (Bertsekas, 1992). A bandwidth allocation is max-min fair if and only if an increase of b_x^{in} within its domain of feasible allocation is at the cost of decreasing some other b_y^{in}. So, it gives the peer with the smallest bidding value the largest feasible bandwidth.

2. To split the file size Z, $size(Z) = T$, among the coalition members, $p_i \in S$.

   (a) We estimate the minimum amount of time needed to transfer Z as a function of the known bandwidth limits. Following this, the minimum amount of time can be estimated as follows:

   $$t_i = T \,/ \sum_{i \in S} \min(b_i^{in}, b_i^{out}) \tag{7}$$

   This estimated time is the same initially for all the peers in the coalition S[2].

   (b) Once we know the estimation of the time for each peer, we can carry out a partitioning of the file taking into account the capacities of each peer. Every peer will have to transfer a number of bytes $b_i^{in} * t_i$. The file is divided into blocks of this size that are assigned to the peers.

3. To inform each coalition member of the size of the block to be transferred. p_0 communicates to each peer member of S the number of bytes to be transferred.

3 THE INCENTIVE MECHANISM

As we have already mentioned our incentive mechanism is based on providing a better quality of service to the peers that participate in the coalitions. In order to achieve this, we define a *Responsiveness Bonus*, *Rb*, for every peer. This value reflects the peer's overall contribution to the system (i.e. how much work it has carried out for the other peers in the system). In accordance with the above model, in the proposed payoff division each peer obtains a utility which is proportional to the resources that it supplies. Therefore, the peer p_i that supply a greater bandwidth (uploading peers) will obtain a greater utility, and this utility will increase its *Rb*. Conversely, the value of Rb_i should be reduced when p_i acts as a downloading peer and does not contribute. We consider that an auditing authority is responsible for storing and updating the *Rb*, using proper methods to control the concurrency.

So the value of Rb_i will be calculated as a heuristic function of x_i that can be adjusted with data from the real system behavior or from simulation results. This uses the x_i values obtained by the uploading peers of the coalition as uploading points, Up_i. For the downloading peer of the coalition its downloading points, Dp_i, are calculated as the average of the utility obtained by the uploading peers of the coalition. Each

[1] If there are many interested peers, a maximum size of coalition is established in order to avoid an undue partitioning of b^{out}. This value depends on the b^{out} and the file size.

[2] b_i^{out} will be used to download files from the P2P sharing file system, in the case p_i does not have the file previously and it works for another node in the coalition (it will be 0 if the peer had the file).

peer p_i accumulates Dp_i and Up_i points by adding the points obtained in each coalition formation process in which it participates [3].

$$Up_i = Up_i + x_i \tag{8}$$

$$Dp_i = Dp_i + \sum_{s \in S} x_s / \mid S \mid \tag{9}$$

Rb_i is a value included in the interval $[0..1]$. The correction of the bandwidth is only applied to the download bandwidth $Rb_i b_i^{out}$ (it makes no sense to correct the upload bandwidth, because we would be decreasing the upload capacity of the collaborative peers). Initially the Rb_i of the peers (uploading/downloading) is 1 [4]. A higher responsiveness bonus (Rb_i closer to 1) will mean that p_i will be able to fill all its reserved bandwidth, since it can add more peers to the coalition in order to complete its bandwidth, reducing the download time. Otherwise, an Rb_i closer to 0 will limit the possibility of adding peers to the coalition (in fact, in some cases it will avoid creating any coalition for the download). In this way, our incentive mechanism promotes cooperation taking into account the selfish behavior of the peers.

$$Rb_i(Up_i, Dp_i, Fs_i) =$$
$$\begin{cases} 1 & \text{if } (Up_i - Dp_i) \geq 0 \\ 0 & \text{if } (Up_i - Dp_i) < 0 \wedge Up_i = 0 \wedge Fs_i = 0 \\ 1 & \text{if } (Up_i - Dp_i) < 0 \wedge Up_i = 0 \wedge Fs_i > 0 \\ \frac{Up_i \cdot \gamma}{Dp_i} & \text{if } (Up_i - Dp_i) < 0 \wedge Up_i > 0 \end{cases} \tag{10}$$

The equation 10 computes Rb_i in relation to Dp_i, Up_i, γ and Fs_i (the number of files shared by peer i).

If Up_i-$Dp_i \geq 0$, it means that the peer is contributing to the system more than it is consuming from it, and so $Rb_i = 1$. If, $Up_i = 0$, the peer has not contributed anything to the system and, if, in addition, the number of shared files is 0 obviously the peer is a free-rider and its Rb_i must be 0. Conversely, if the peer has not contributed to the system, but the number of shared files is not 0, it means that the peer wants to contribute to the system but its shared files have still not been downloaded by other peers; So its Rb_i must stay at 1. Finally, if the peer has contributed to the system, but less than what it has been consuming from it, its Rb_i will be proportional to Up_i/Dp_i[5]. The variable γ allows us to regulate this formula in order to increase or decrease the proportional relation between the benefit, Up_i, and the penalty, Dp_i.

4 PERFORMANCE EVALUATION

In this section we describe the simulations we performed and the corresponding results. In order to simulate our coalition formation model for P2P file sharing, we have defined and implemented a generic P2P simulator for service oriented networks. The simulation tool is presented in detail in (Belmonte et al., 2007). Additionally, it should be noted that we are dealing with situations which are different from traditional system simulators, since, we are also trying to model the user behavior. For this reason, and in order to model the user as close to reality as possible, the peers are classified in three categories according to their behavior: free-rider, adaptive and collaborative. We will first describe how the user behavior has been modeled, and then the simulation results.

4.1 Modeling the User Behavior

Free-riding is a consequence of selfish user behavior in file sharing systems. In the case that we want to study actions to take in order to improve cooperation, modifying user's behavior, a key step would be the modeling of users that are going to take part in the simulation. A realistic simulation should include three kinds of users (behaviors):

1. **Free-rider.** Represents the selfish peer which only downloads files and rejects all the incoming file requests.
2. **Collaborative.** Represents collaborative behavior. These peers always try to maximize the system performance, so they offer all their available b_i^{in} and accept all the incoming file requests until their bandwidth is full.
3. **Adaptive.** Represents intelligent behavior, and so, is adapted to the evolution of peer welfare. These users accept download requests as long as they are interested in downloading a file, that is as long as they benefit. When the number of target files is 0 the b_i^{in} will be 0. Otherwise, all the entire available b_i^{in} will be offered so that a high Rb is maintained and all the target files can be downloaded. In case of multiple requests, the b_i^{in} will be divided among all the requests, taking into account the Rb of the requester peers.

[3] Since a new peer that joins the coalition formation system will have its uploading and downloading points set to 0, we allow the peers to download a minimum amount set to a parameter *MinDownload*.

[4] Otherwise their bandwidths would be reduced from the beginning, and the download times of the files would be higher (compared to the scheme without coalitions).

[5] This value is always < 1.

Finally, all of them have a limit of download tries to avoid them repeatedly asking for the same file.

4.2 Experimental Results

We have run simulations of a P2P network of 1000 peers for 2000 units of simulated time (steps). All peers have the same bandwidth capabilities, 1024 kbs for downloads and 512 kbs for uploads. We have defined a collection of files of different sizes, a random number of copies of these files are delivered through the peers at the start of the simulation. The minimum number of copies for a file is 5 and the maximum is the half part of the number of peers that forms the network (500 for our experiments). This means that peers have a random number of initially stored files, between 0 and the whole collection of files. The objective of the simulation is that every peer manages to get the whole file collection, by this we mean, to download the files that are not initially stored. Depending on the peer behavior it will face this objective in different ways. File sizes range from 10000KB to 90000 KB. Table 1 shows the most important simulation parameters.

Table 1: Simulation Parameters.

Numbers of peers	1000
Simulation steps	2000
b_i^{out}	500(min)/1024(max)
b_i^{in}	200(min)/512(max)
File sizes	from 10000Kb to 90000Kb
γ Weight U_p respect to D_p	1 and 2

To measure the impact of the Adaptive users on the system the experiments have been run with two different populations. The first one without Adaptive users, and the same population for Free Riders and Collaborative users (50% Free Riders, 50% Collaborative and 0% Adaptive users), we will hereafter refer to this as Population 1. And the second one with the same population for Adaptive and Collaborative users (40% Free Riders, 30% Collaborative and 30% Adaptive users), hereafter refereed as Population 2. In addition, the simulations have been run with two different incentive policies: No Coalitions (NC), where no incentive mechanism is considered and Coalitions (C), which implements our proposal. After repeating the simulation experiments 100 times we take the average to give the results. To compare how the incentive mechanisms and the user behaviors affect the P2P system, two main metrics have been considered: Downloaded Bytes and Average Time. In addition an analysis of the *Responsiveness Bonus* of the Coalition mechanism is presented.

4.2.1 Number of Downloaded Bytes

Figure 2 shows the evolution of the downloaded bytes distribution for No Coalitions mechanism for experiments run with Population 1 on the left and Population 2 on the right. Similar figures for Coalitions in Figure 3. Downloaded bytes can be interpreted as the benefit obtained from the system. Next, we analyze these results for the different populations.

In Population 1 all work is done by the collaborative users, since free-riders do not collaborate. Figure 2 left shows the evolution of the distribution of the downloaded bytes is round 50 % for both users during the simulation. This means, the collaborative users do all the work, and the benefits are shared equally with free riders. However, when we run the Coalition mechanism, Figure 3 left, free riders are stopped, the percentage of bytes downloaded by free riders drastically decreases after the first 100 steps of simulation. This demonstrates how the coalition formation prevents free-riders from obtaining more bytes as simulation time advances, and so from fully using the system resources.

When Adaptive users are simulated, this is Population 2, distribution of downloaded bytes are affected as shown in figures 2 right and 3 right. With respect to collaborative and adaptive users, both are 30 % of the population, they do all the work and share more or less equally the benefits with free riders. In figure 3 right the evolution of the distribution of the downloaded bytes shows how free riders are again stopped, as in Population 1, and this means that the benefit to be shared between the collaborative and adaptive users, these are those ones that are uploading files. In addition, collaborative users increase the percentage of downloaded bytes during the simulation; However, adaptive users first increase and after decrease the percentage. This is due to the behavior of adaptive users, which are penalized when they are not sharing enough.

Table 2: No Coalition Mechanism Downloaded Bytes.

No Coalition	Population 1	Population 2
Free Rider	157 Gb	123 Gb
Colaborative	156 Gb	92 Gb
Adaptive		91Gb
total	313Gb	306 Gb

Tables 2 and 3 summarize the downloaded bytes per populations and per behavior. In both tables it can be observed that the bytes downloaded by free riders

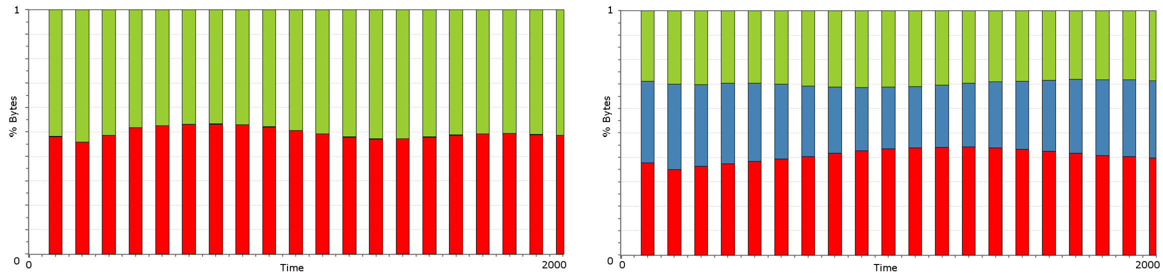


Figure 2: Evolution of the Downloaded Bytes Distribution (% Bytes vs Time) for No Coalitions Mechanism , Population 1 (left) and Population 2 (right). Free Rider in red, Adaptive in blue and green for Collaborative users.

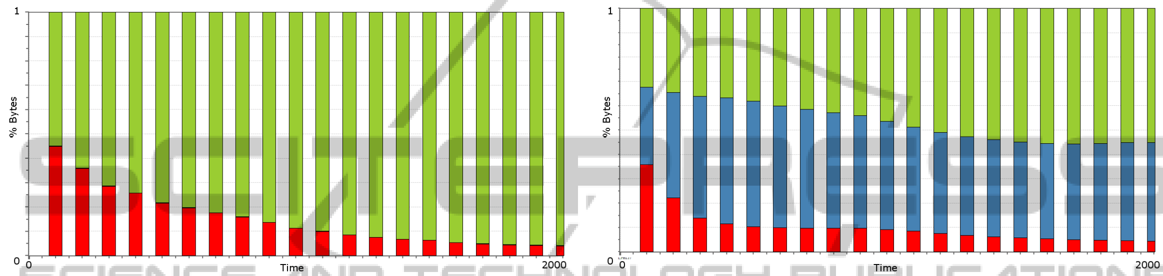


Figure 3: Evolution of the Downloaded Bytes Distribution (% Bytes vs Time) for Coalitions Mechanism, Population 1 (left) and Population 2 (right). Free Rider in red, Adaptive in blue and green for Collaborative users.

Table 3: Coalition Mechanism Downloaded Bytes.

Coalitions	Population 1	Population 2
Free Rider	24 Gb	20 Gb
Colaborative	130 Gb	82 Gb
Adaptive		84 Gb
total	154 Gb	186 Gb

are slightly reduced in Population 2 with respect to Population 1, this is because there are 10% less users in this population, the average bytes per user is very similar. With Population 1 it can also be observed that the coalition mechanism reduces the total bytes downloaded to 50,54%, with respect to No Coalitions, but the 83,61% of this reduction is due to the Free Riders detection. This shows again how the algorithm prevents free-riders from abusing as simulation time advances, and from stressing the system resources; and this leads to a more healthy system.

The Coalition Algorithm (table 3) reacts to the inclusion of Adaptive users by increasing by 20% the total amount of downloaded bytes compared to Population 1, this means that they benefit the whole system. In addition, comparing Coalitions and No Coalitions with Population 2 (when Adaptive users are simulated), the results are better than with the Population 1. In this case, the total amount is reduced by 38,78% where 83,15% is due to the Free Rider's detection.

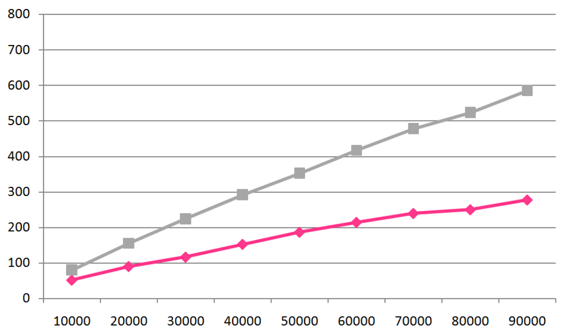
Notice that when using the second population there are fewer Free Riders and Collaborative users in the simulation. This affects the data shared and demanded in the system,and it justifies the smaller amount of downloaded bytes.

4.2.2 Average Time

In addition to the total amount of bytes downloaded during the simulation, the time spent on each download is also a significant measure of the system performance. In Figure 4 the average time for each file for both algorithms is compared (squares for No coalitions and diamonds for Coalitions).

Population 1: When Adaptive users are not considered, Figure 4 left, the best download times are the ones obtained with Coalitions. For the smallest files, the times for No Coalitions and Coalitions are quite similar then, the higher file size is, the greater the time difference. As expected, the benefit of using Coalition is increased as the file sizes grows.

Population 2: When Adaptive users are introduced average download time is increased, this is because the system is more stressed. Adaptive users implement a selfish behavior, but they have to share in order to obtain benefits and they are capable of simultaneous downloads. All of this increases the download time. In Figure 4 right No Coalitions and Coalitions show a very similar slope and smaller values for Coalition mechanism, which also stops Free Riders.



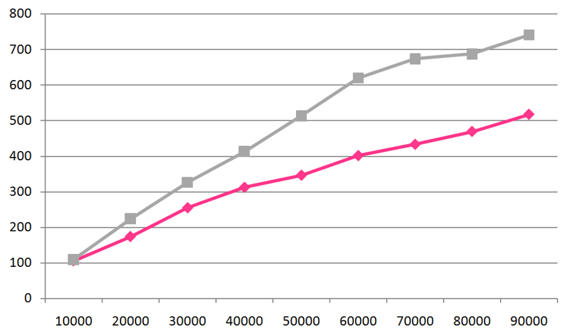


Figure 4: Average Download Time (Time vs Bytes), Population 1 (left) and Population 2 (right).

In this way, the system benefits without penalizing user's downloading times.

4.2.3 Responsiveness Bonus

In Figures 5 left and 5 right we studied the evolution of the *Responsiveness Bonus*, *Rb*. The results obtained confirm the correct behavior of each type of user. The *Rb* of free-rider falls quickly as time goes on. This evolution is correct because they only download and do not upload files, so the *Rb* decreases, and its total bytes downloaded remain static at the beginning of the simulation. In relation to adaptive peers, they maintain a higher *Rb* value, while they are downloading more data, so they are uploading more files for other peers than they are downloading. They share the minimum to keep participating.

Figure 5 left shows the evolution of *Rb* when Coalitions are configured with $\gamma = 1$ and Figure 5 right with $\gamma = 2$. This parameter basically affects Adaptive users, $\gamma = 1$ means that the benefit obtained for each byte provided is 1, this forces Adaptive users to give as much as they want to download. When $\gamma = 2$ the benefit is doubled, this allows Adaptive users to increase and maintain a high reputation with a lower participation. Notice that the *Rb* is a limitation of the download capacity of the peer, so in the first figure, the input bandwidth of Adaptive users is significantly reduced.

Numerically in the first Figure 100% of Free Riders are stopped at step 1100 of the simulation time, and in the second Figure, around step 1700; In fact 50% of Free Riders are detected around step 200 of the simulation time, and in the second figure this is achieved around step 300. The Adaptive users cause a little slowdown on the detection of Free Riders.

5 RELATED WORK

Reviewing the bibliography, several approaches have been proposed to combat the free riding problem. Karakaya (Karakaya et al., 2009) et al. have categorized them into three main types: firstly, incentive mechanisms based on monetary payments: one party offering a service to another is remunerated and inversely, resources consumed must be remunerated or paid for. Secondly, mechanisms based on reputation: it keeps information about the peer reputation, and peers with a good reputation are offered better services. Thirdly, incentive mechanisms based on differential services or reciprocity-based: peers that contribute more get a better quality of service (Cohen, 2003) (Ramaswamy, 2003) (Karakaya et al., 2008) (Garbacki et al., 2006) (Mekouar et al., 2006). Our approach could be included in this category, although its foundation is different and innovative.

Although some of the above approaches (Ramaswamy, 2003) (Karakaya et al., 2008) (Mekouar et al., 2006) are based on differential services, they do not promote a cooperative behavior among peers that improves the download performance in the P2P System. And, in addition they do not achieve fair service differentiation between peers.

Those remaining, more similar to our approach, propose incentive mechanisms that encourage collaboration among peers. For example, 2Fast (Garbacki et al., 2006) is based on creating groups of peers that collaborate in downloading a file. However, compared to our proposal, it does not enforce fairness among the collector and helper peers, and in addition it is not specified how the helper may reclaim its contributed bandwidth in the future. Bit torrent (Cohen, 2003) is also based on collaboration among peers. Its ”tit-for-tat” policy of data sharing works right when the peers show a reciprocal interest in a particular file. However, in bit-torrent, the peers' download band-

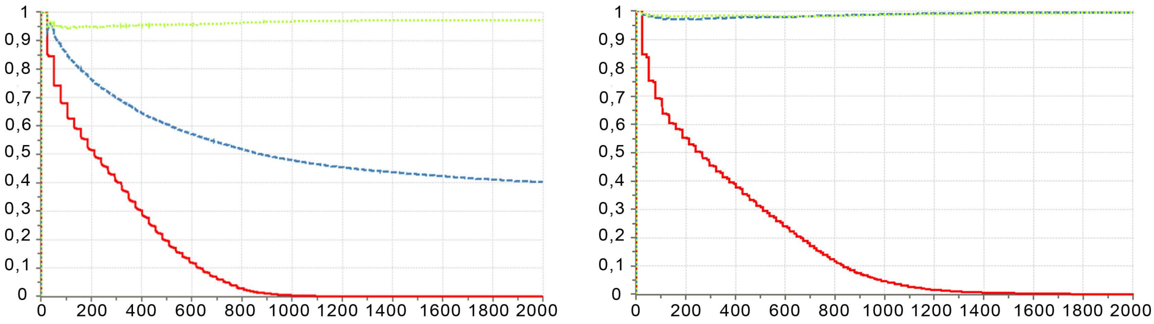


Figure 5: Responsiveness bonus with $\gamma = 1$ (left),$\gamma = 2$ (right). Free Rider in solid line, Adaptive in dashed line and dotted line for Collaborative users.

width is limited to their upload capacity, reducing in this way the achievable download performance. However, in our approach, the system's download capacity is not reduced to its upload capacity; And, using the *Rb* it does not force a "mutual reciprocity" mechanism (like "tit-for-tat"); and thus the bandwidth contributed by a peer can be used in later downloads.

EMule (eMule, 2010) also promotes cooperation among peers. It uses a credit system to reward frequent uploaders and alleviates the free-riding problem. However, credits are exchanged between two specific peers, so content trading can happen only between peers that have mutual interests, and in addition it does not enforce fairness in bandwidth sharing.

Finally, the work of Ma et al. (Ma et al., 2006), also provides service differentiation based on the amount of services that each node has provided to a P2P community, and it uses a game theoretic framework. However, while we use a cooperative approach that proposes coalition formation, they propose a mechanism that makes different requesting users bid for resources, creating a dynamic competitive game.

6 CONCLUSIONS

In this paper we have presented a coalition formation based incentive mechanism for P2P file sharing systems. It is based on game theory and takes into account the rational and self-interested behavior of the peers. In (Belmonte et al., 2006b), the initial idea of applying this model to this problem was presented. Now, we have formally demonstrated the fairness of the model using game theory and, more concretely, the concept of "core". In addition, we have modeled the user behavior and defined the coalition formation model in order to perform simulations, using the simulator presented in (Belmonte et al., 2007). The paper also includes the analysis of the results of these simulations.

Our approach allows any peer with idle bandwidth to participate in a coalition, uploading files for other peers in exchange for utility, and consequently greater download bandwidth; And in addition, it provides, using the "core", a fair utility to the peers forming the coalition in relation to the bandwidth they supply. To achieve this, a *Responsiveness Bonus* that reflects the peer's overall contribution to the system is defined, and the game theory utility concept is used to calculate it.

The simulation results have shown that in relation to downloaded bytes, the coalition mechanism prevents free-riders from obtaining more bytes as simulation time increases. In addition, it reacts to the inclusion of adaptive users increasing by 20% the total amount of downloaded bytes, so they benefit the system. In relation to download time, coalitions are capable of getting the best average download times and stopping free riders at the same time. Finally, the simulation of *Rb* proves that our proposal enables the selfishness but in exchange for sharing data. This helps to keep the systems healthy in despite of self-interested users.

Finally, we are working on the simulation of other approaches in order to be capable of comparing our results with existing proposals. And we plan to generalize the proposed coalition formation algorithm in order to include Quality of Service information. Our idea is to form coalitions in such a way that they are able to provide or guarantee QoS in different aspects depending on the service or application, i.e. real time constraints or fault tolerance.

REFERENCES

Androutsellis-Theotokis, S. Spinellis, D. (2004). A survey of peer-to-peer content distributing technologies. *ACM Computing Surveys*, 36-4:335–371.

Belmonte, M. V., Conejo, R., Díaz, M., and Pérez-de-la Cruz, J. L. (2006a). Coalition formation in p2p file sharing systems. In *Curret Topics in Artificial Intelligence, LNAI 4177*.

Belmonte, M. V., Conejo, R., Pérez-de-la Cruz, J. L., and Triguero, F. (2006b). Coalitions among intelligent agents: A tractable case. *Computational Intelligence. An International Journal*, 22-1:52–68.

Belmonte, M. V., Díaz, M., Pérez-de-la Cruz, J. L., and Reyna, A. (2007). File sharing service over a generic p2p simulator. In *In Proceedings of Semantic, Knowledge and Grid, SKG-07*.

Bertsekas, D. Gallager, R. (1992). *Data Networks*. Prentice-Hall.

Cohen, B. (2003). Incentives build robustness in bit torrent. In *Proceedings of the First Workshop on Economics of Peer-to-Peer Systems*.

eMule (2010). The e-mule web site. http://www.emule.net.

Garbacki, P., Iosuo, A., Epema, D., and Van Steen, M. (2006). 2fast:collaboartive downloads in p2p networks. In *Sixth International Conference on peer-to-peer computing, P2P'06*.

Golle, P., Leyton-Brown, K., Mironov, I., and Lillibridge, M. (2001). Incentives for sharing in peer-to-peer networks. In *ACM Conference on Electronic Commerce*. ACMPress.

Handurukande, S. B., Kermarrec, A., Le Fessant, F., Massouli, L., and Patarin, S. (2006). Peer sharing behaviour in the edonkey network, and implications for the design of server-less file sharing systems. In *1st ACM Sigops/Eurosys European Conference on Computer Systems. EuroSys '06*.

Kahan, J. Rapaport, A. (1984). *Theories of Coalition Formation*. Lawrence Erlbaum Associates Publishers.

Karakaya, M., Korpooglu, I., and Ulusoy, O. (2008). Countrecting free riding in peer-to-peer networks. *Computer Networks*, 52:675–694.

Karakaya, M., Korpooglu, I., and Ulusoy, O. (2009). Free riding in peer-to-peer networks. *In IEEE Internet Computing*, pages 92–98.

Ma, R. T. B., Lee, S. C. M., S., L. J. C., and Yau, D. K. Y. (2006). Incentive and service differentiation in p2p networks: A game theoretic approach. *IEEE/ACM Transactions on Networking*, 14-5:978–991.

Mekouar, L., Iraqi, Y., and Boutaba, R. (2006). Handling free riders in peer-to-peer systems. In *Proceedings of AP2PC 2005. Lecture Notes in Artificial Intelligence 4118*.

Ramaswamy, L. Liu, L. (2003). A new challenge to peer-to-peer file sharing systems. In *6th Hawaii International Conference on System Sciences*.

Sariou, S., Gummadi, P. K., and Gribble, S. D. (2002). A measurement study of peer-to-peer file sharing systems. In *Proceedings of Multimedia Computing and Networking*.

Shneidman, J. Parkes, D. (2003). Rationality and self-interest in peer to peer networks. In *Proceedings of the Second International Workshop on Peer-to-Peer Systems IPTPS'03*.

APPENDIX

We must prove that equations (5) define a payoff division that lies inside the core.

Group Rationality. First note that always $x_i \geq 0$. Moreover, a coalition S such that $p_0 \notin S$ has $V(S) = 0$. So the only thing to prove is that, for every $P \subseteq N - \{p_0\}$, the coalition $S = \{p_0\} \cup P$ has a coalitional value $V(S)$ such that $x(0) + \sum_{i \in P} x(i) \geq V(S)$. Let us define $Q = N - P - \{b_0\}$, $p = \sum_{b_i \in P} b_i^{in}$, $q = \sum_{b_i \in Q} b_i^{in}$. Then

$$\sum_{i \in P} x(i) = t_0 \frac{(p+q)^2 + b_0^{in} p}{(b_0^{in} + p + q)^2} \tag{11}$$

and

$$V(S) = t_0 \frac{p}{b_0^{in} + p} \tag{12}$$

The difference between (11) and (12) is

$$t_0 b_0^{in} q \geq 0 \tag{13}$$

Therefore we have proved group rationality.

Global Rationality. Note that the coalitional value as a function $V = v(t_0, b_0^{in}, b_1^{in}, \ldots, b_n^{in})$ (equation 4) is homogeneous of degree 1. Therefore,

$$\sum_0^n x(i) = t_0 \frac{\partial V}{\partial t_0} + b_0^{in} \frac{\partial V}{\partial b_0^{in}} + \ldots + b_n^{in} \frac{\partial V}{\partial b_n^{in}} = V \tag{14}$$