Incentives Engineering for Structured P2P Systems – a Feasibility Demonstration Using Economic Experiments

Stephan Schosser Klemens Böhm
Universität Karlsruhe (TH)
Department of Computer Science
Am Fasanengarten 5, P.O.Box 6980
76128 Karlsruhe, Germany
Phone +49/721/608-3968
{schosser|boehm@ipd.uka.de}

Rainer Schmidt Bodo Vogt
Otto-von-Guericke-University Magdeburg
Faculty of Economics and Management
Universitätsplatz 2, P.O.Box 4120
39106 Magdeburg, Germany
Phone +49/391/67-18426
{rainer.schmidt|bodo.vogt@ww.uni-magdeburg.de}

ABSTRACT
Structured peer-to-peer systems allow to administer large volumes of data. Several peers collaborate to generate a query result. Analyses of unstructured peer-to-peer systems, namely of those for file-sharing, show that peers tend to shirk collaboration. We anticipate similar behavior in structured peer-to-peer systems. Recently, protocols to counter uncooperative behavior in such systems have been proposed. This article examines the behavior of peers under such protocols, using game theory. A first result of this paper is a set of hypotheses, e.g.: Peers answer queries if more than a certain percentage of their queries is answered. In many situations, free-riding does not lead to a break-down of the system. Trust, reciprocity and reputation building via a feedback mechanism are behavioral motives that increase cooperation. As a second step, we have conducted economic experiments with human participants to validate our predictions. Such experiments are important because we do not need to make any assumptions regarding the behavior of peers. It turns out that the predictions remain valid in these experiments.

Categories and Subject Descriptors
H.1.2 [**Information Systems**]: Models and principles – human factors, human information processing; H.3.4 [**Information Systems**]: Systems and Software – distributed systems, information networks; E.1 [**Data**]: Data structures – distributed data structures

General Terms
Measurement, Design, Economics, Reliability, Experimentation, Human Factors

Keywords
Economic Experiments, Free-Riding, Game Theory, Social Exchange, Structured Peer-to-Peer Networks.

1. INTRODUCTION
Peer-to-peer systems (P2P systems) are an alternative to traditional system architectures. They are distributed and do without any coordinator. Every peer takes part in the work. In return it can consume services, such as content provided by others. The P2P approach is superior to monolithic systems in several respects, such as reliability and scalability, at least in theory.

Experiments with P2P file-sharing systems show [2] that peers controlled by humans tend to free ride, i. e., they use the resources provided by the system while not contributing any resources of their own. We expect similar behavior in structured P2P systems: Queries in structured peer-to-peer networks, similar to other P2P systems, are processed by more than one peer. A query remains unevaluated if one of these peers defects. Identifying such peers is almost impossible – queries cannot be tracked. Hence a peer has to count on the cooperativeness of the other peers. This is also the case in Content-Addressable Networks (CAN), the variant of structured peer-to-peer networks considered here. A peer has information only about a small subset of all nodes, his neighbors. It cannot observe the behavior of other nodes. Hence, a peer does not know how reliable these other peers are.

[4] address free-riding in such structured P2P systems [4], by extending the CAN protocol with an incentives mechanism: Each peer collects information on the interactions with neighboring peers that have satisfied it. They show [4], both analytically and by simulation, that the protocol proposed discriminates well between cooperative and uncooperative peers. However, this result is based on several assumptions that are relatively stiff. The most rigid one is that the only kind of free-riding behavior considered is that peers do not forward or answer queries. Another reason why [4] is not complete is that several questions remain unanswered: Which parameters do result in cooperation in structured peer-to-peer systems? Given this information, can we further improve the design of the system? Do the assumptions hide any situations which could lead to a collapse of the system?

In this article, we analyze the behavior of peers in structured P2P systems from a strategic perspective. This will help us to answer these questions and to deepen our understanding of structured P2P networks. We represent certain aspects of structured P2P systems with idealized game-theoretic models. We perceive the equilibria of the game theoretic models as stable situations of P2P systems. Given these models, we derive our hypotheses: First, we expect cooperation between peers to be correlated with the degree of trust of peers in the system. Next we expect that a peer can identify individual peers, and it will play different strategies with different peers. We also expect that free-riding of a certain share of peers will not let the system break-down. When free-riding occurs, we expect the feedback mechanism to increase co-operation and to reduce the impact of free-riding. Finally we expect the probability of participation in the system to depend on the fraction of queries that the system has answered in the past.

It is well known from experimental economics that theoretic predictions often differ from actual behavior. Consequently, we conduct experiments with human peers in the laboratory. In these experiments, each human participant controls the behavior of one peer of a CAN. To our knowledge, this is the first evaluation of a protocol for peer-to-peer networks with human participants in the laboratory. There is one significant advantage of such an analysis: Simulations or other models rely on assumptions concerning the behavior of peers, e.g., the frequency of queries sent or the criteria for answering queries. Such assumptions depend on the expectations of the designers. With our approach in turn, one does not need to make any assumptions regarding the behavior of peers – it is observed during the experiments!

With these experiments we want to identify the strategies humans might come up with. In addition, we are interested in their effect on robustness, efficiency and network traffic. We admit that the number of participants in our experiments is relatively small, compared to the number of peers in real-world P2P systems. However, the behavior of the participants – in qualitative terms – does not depend on their number in our experiments, as we will explain. On the other hand, we claim that our approach is superior to relying on the assumptions of protocol designers. Finally, our experiments confirm the hypotheses mentioned above.

This article has the following structure: In Section 2 we review related work. Section 3 gives a short introduction to CAN. In Section 4 we formalize structured P2P systems using idealized economic models and derive our hypotheses. Section 5 contains the experimental design. Section 6 presents and discusses the experimental results. Section 7 concludes.

2. RELATED WORK

Empirical studies on the usage of peer-to-peer file-sharing systems like Gnutella [2] or Napster [21] have shown that a majority of users prefers not to share any of their resources with other users. Such users prefer to consume the resources provided by others without contributing in return. Ramasamy and Liu [16] show that this behavior, called free-riding, can even lead to a collapse of these systems or at least may have a severe negative effect. Hence, P2P systems need to be evaluated in this respect before deployment.

In the literature several approaches deal with this problem. Schlosser and Kamvar [22] compare different P2P algorithms using simulations. They model characteristics using observations from unstructured file-sharing networks. On the other hand mathematical models can be used to analyze the performance of peer-to-peer systems. Ge et al. [11] use such a model to compare structured and unstructured peer-to-peer systems. The authors show that structured P2P systems outperform other P2P infrastructures regarding the volume of data transferred. They do not investigate the influence of uncooperativeness on the system, but expect all participants to cooperate.

Strategies to establish relationships within P2P systems are analyzed in [9] and [7]. They provide cost models for internet-like network structures. Each peer is seen as a non-cooperative player, which has a benefit from participating in the system, and it wants to minimize the price of participation. This work regards only the distance between nodes and the degree of connectivity. As an extension [6], the load imposed on each peer is considered in addition. The authors analyze existing P2P structures using game-theoretic concepts like social optima and Nash equilibria.

Golle, Leyton-Brown and Mironov [12] propose a simple game theoretic model to analyze the behavior of peers in a P2P system with central coordinator. To motivate sharing, they introduce several payment mechanisms. Experiments then confirm their theoretical results. A similar approach is described in [17]. Here, peers are modeled as uncooperative players. This work applies reputation-based incentive mechanisms to unstructured P2P systems. The participating peers are expected to behave uncooperatively. It is shown that the mechanisms help to counter free-riding. [10] extends this work by a more detailed analysis. Several reputation-based incentive mechanisms and simulations with different attack strategies are used to demonstrate the usability of the system.

Another approach to eliminate free riders is taken in [5]. Here a differential service-based incentive mechanism is used. It is shown that the strategy of a peer solely depends on the benefit it receives from the system. A peer joins the system if his expected benefit is above a certain threshold.

[23] is a preliminary version of this article that does not consider feedback mechanisms, which we deem one of the most intriguing aspects of this paper. The analysis in this paper also is broader and takes more models into account.

3. CONTENT ADDRESSABLE NETWORK

Structured P2P systems manage (key, value)-pairs. Each peer knows a subset of the peers participating in the system, and it administers a subset of the data. Data is distributed among the peers deterministically: Data objects are assigned to nodes according to their keys. In the last years several such systems, like CHORD [29], Tapestry [30] and Pastry [20], have been proposed. In this paper we focus on Content Addressable Networks (CAN) [18]. In contrast to other approaches, CAN peers only know other peers “close” to it. With CAN one can guarantee that the data object for a given key is retrieved within a certain number of steps, under various model assumptions [4]. We use CAN as a platform since its routing flexibility is relatively high [13]. Participants in the experiments have a choice where to forward the query to.

More specifically, CAN use a hash function to map keys to coordinates in an n-dimensional space. This space is partitioned among the peers. Each peer stores all data objects whose mapped key lies in its partition. In addition, each peer knows its neighbors. Neighbors are the peers which administer a partition of the space that borders the partition of the current peer.

Example. Think of a CAN storing information on movies. The information about the movie "The Crow" might be "Thriller", information stored for "Pride & Prejudice" might be "Drama", and the one of "Chicken Little" might be "Animation". Suppose that the hash function maps "The Crow" to the coordinates (0.3, 0.3). The peer whose partition contains (0.3, 0.3) saves the (key, value)-pair ("The Crow", "Thriller"). Figure 1 illustrates the partition of the coordinate space among the nodes. Each peer of the CAN corresponds to a rectangle, its partition in the coordinate space. The peer administering the data object/meta information corresponding to key "The Crow" is Peer 2. ■

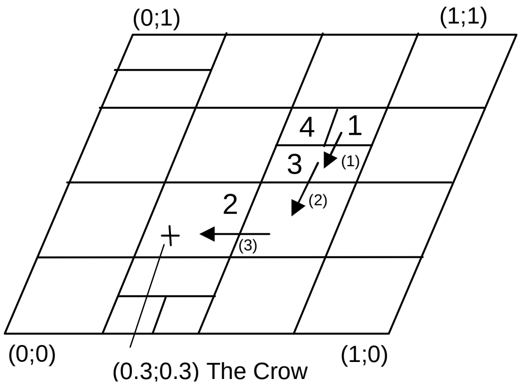

Figure 1. Content Addressable Network

3.1 Basic Routing Algorithm

Each peer obtains positive utility by receiving data objects belonging to a key it is interested in. To evaluate a query, i.e., to find the data object corresponding to a key, the system uses a simple variant of greedy forward routing: A peer sends its query, the key of the object sought, to a neighbor whose partition is closer to the hash value of the key. The recipient checks whether it stores the object. If not, it forwards the query to one of its neighbors. I.e., the peers repeat this step until the query arrives at the peer who has the information desired. Finally, the query result is returned to the issuer of the query.

Example. Figure 1 illustrates the routing algorithm for a query "The Crow" issued by Peer 1. The 2-dimensional hash function maps key "The Crow" to the coordinates (0.3, 0.3). Peer 1 now identifies the neighbors which are closer to the coordinate (0.3, 0.3) than itself and forwards the query to one of them. One possible approach is to use the peer with the smallest Euclidean distance to the mapped key. In our example, this is the Peer 3. The query is now forwarded to this peer (Arrow (1)). This peer does not have the desired query result. Hence, the query is forwarded until it finally reaches Peer 2 (Arrows (2) and (3)). This peer then answers the query by returning the query result to Peer 1. ■

As in other P2P systems, all peers should contribute to the same extent. Hence, the key space forms a torus. There are no edges of the coordinate space. In consequence, the amount of incoming messages should be roughly equal for all participants, assuming that query points are evenly distributed as well.

With the conventional design of CAN described here, it is assumed that all peers forward or reply to all incoming queries.

3.2 Feedback

In the CAN protocol described in [4], a peer may attach information on the reliability of other peers, so called feedback, to outgoing messages. Each feedback object has a value and a subject, the peer described by the feedback. For the moment, let us assume that peers giving feedback behave as intended by the protocol designer. In this case, the feedback value is '+' (positive) if the subject has been reliable and '–' (negative) otherwise. Since a peer cannot explicitly observe the behavior of other peers, it might give false feedback.

Example. Consider again Figure 1. When Peer 1 sends its query to Peer 3 (Arrow 1), Peer 1 attaches information about neighbors of Peer 3, e.g., Peer 4. Suppose that Peer 1 has already interacted successfully with Peer 4 several times. Hence, it would attach positive feedback to the outgoing query. ■

4. MODELS AND HYPOTHESES

In this section we apply several existing game theoretic models to our scenario. Identifying and applying these models is a first contribution of this paper. They help us to understand the strategic options of peers in structured peer-to-peer networks from an economic perspective. Based on these models, we derive several hypotheses on structured peer-to-peer networks. A peer will not only have a choice between cooperating and free-riding in our setup. It is also free in individual decisions, e.g., whether to forward or answer an incoming message, and if it should issue a query or not.

The utility of an actor determines his incentives in an economic modeling. He can obtain utility from money and other factors. Costs result in negative utility, while positive payoffs result in positive utility. Similarly, doing work results in negative utility.

When modeling the CAN, we assign utility to any possible action of a peer as follows: If a query is answered the initial sender receives positive utility. Answering, forwarding or sending a query incurs negative utility. The utility of issuing feedback is zero in our experiments. This is because the effort of generating feedback is negligible for system peers, since they keep track of the behavior of their neighbors anyhow. For simplicity we equate money with utility. This is in line with the fact that the utility of a person increases with the amount of money received. This holds even though the correlation is not necessarily linear.

4.1 Investment Game

We start by comparing a P2P structure to the investment game [3], to better understand cooperation. In this game, two groups of participants (Groups A and B) exist. Participants of Group A obtain a fee of $10. Each of them may then send some of his money to an anonymous counterpart in B, aka. invest. A player in Group B obtains three times the amount of money sent to him, and participants in Group A know about this. The participants in Group B then decide which share of the money to return to their counterpart in A, and how much money to keep for themselves.

The game is comparable to a P2P structure. If a peer sends a query it spends money, i.e., it invests. We interpret the costs for sending a query as allocation for storing data on behalf of others. Another peer can answer the query. This leads to a monetary return that overcompensates the investment. Answering and forwarding queries are investments. This is because answering and forwarding are prerequisites for having queries processed by other peers.

As the game is not repeated, the game-theoretic solution of the investment game is the subgame perfect equilibrium. In equilibrium Player 2 does not return any money to Player 1. This is because he does not have any advantage from doing so. As the game ends after his action, he would only lower his payoff. The efficient solution of this game however is that Player 1 invests all of his money. Based on fairness considerations, Player 2 should return half of the investment of Player 1 times three and keep the other half. In experimental studies of this game, Berg, Dickhaut, McCabe show that the efficient outcome can be observed [3]. This solution can be explained by trust and reciprocity: If Player 1 invests he trusts in Player 2. If Player 2 returns money he shows reciprocal behavior. What is called cooperation in P2P structures can be subdivided into two motives: trust and reciprocity.

From this model we derive the following hypotheses on trust and reciprocity in structured P2P networks:

Hypothesis T (Trust): The query intensity in the system is positively correlated to the average payoff of the peers.

This hypothesis covers the trust argument. The more peers trust in the system, the higher the cooperation. This leads to higher average payoffs of the peers. Trust should be measured as the number of queries sent by a peer, the query intensity. The more queries a peer sends, the higher its trust in the other players.

Hypothesis R (Reciprocity): The strategies of a player are personalized, i.e., they depend on the identity/identification of a peer and not only on the actions of the system as a whole.

Hypothesis R covers the reciprocity argument. Reciprocity should result in strategies that account for the past actions of another peer. If the queries of a peer have been answered by another peer, it will also answer queries of the other peer. Note that this hypothesis is different from observations concerning P2P file-sharing. Our explanation is that behavior in file-sharing systems is based on factors absent from the protocol investigated here. In particular, personalized strategies are difficult to implement there.

4.2 Helping Game

If a peer cannot distinguish between peers, more indirect motives are important. Due to forwarding, it is difficult to attribute query results received to the behavior of one particular peer in structured P2P systems. Hence, indirect interactions are important. A model which considers this is the helping game. There is a population of players who may help each other or not. In each round of the game a pair of players is randomly matched. It is not likely that two players interact with each other frequently. One of the two players is given the chance to donate; the other one will receive the donation. The cost of donating c is lower than the benefit b of the recipient. If one does not donate, both participants receive no reward. Table 1 shows the corresponding payoff matrix. As the donor will not interact with the recipient later, his rational behavior would be to defect. This is because his payoff for defecting is higher than the payoff of donating. Contrary to these theoretical predictions, in experiments with human participants one can observe investment and returning money, but much less than in the case in which partners are known. In these experiments, all peers show cooperative behavior [8].

Another key question in our scenario concerns the power of feedback. Does feedback increase cooperation and help to react to free-riding? The helping game has been extended by a mechanism that allows for reputation building [27]. Here every player in the helping game has a global reputation value. This value increases whenever he donates and decreases if he does not. All participants in the system can see the value. In this way a peer can get an impression of the past actions of another peer. In analogy to the standard helping game no investment should be observed, from the theoretic point of view. Experiments of this game with human participants show that the average payoff increases in the game with reputation, compared to the standard helping game [8], [24]. We derive the following hypothesis from these experiments:

Hypothesis F (Feedback): The average payoff in a peer-to-peer structure with feedback is higher than in one without.

Table 1. Payoff matrix helping game

	Donor	
Recipient	donate	defect
	b / $-c$	0 / 0

This hypothesis deals with the effectiveness of feedback. The experiments mentioned show that feedback increases the average payoff in the system. This can only happen due to more cooperation and less free-riding.

4.3 Service Model by Hens and Vogt

The last model we consider is the service model by Hens and Vogt [14]. Here agents can be in two different states: Either they are in a state where they can deliver a service (State A) or consume a service (State B). Compared to our structured P2P system, delivering a service corresponds to answering or forwarding a query. When delivering a service, an agent looses utility. Consuming a service on the other hand corresponds to receiving a query result in structured P2P systems. Consuming a service yields positive utility. The utility for consuming a service is higher than the one for delivering a service.

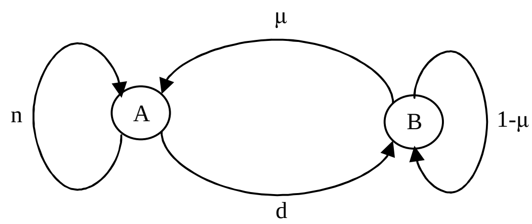

Figure 2. Strategic options in the model of Hens and Vogt

The model is round based. In each round the agents in State A can deliver a service (d) or not (n). An agent in State B does not have any action choices. After each round an agent in State A who offered a service is matched with an agent in State B. If enough agents in State A are willing to deliver a service, every agent in State B can consume one. The number of services offered never exceeds the one of services requested. This is because the number of agents in States A and B are equal, and agents in State B always want to consume the service. Otherwise, a fraction μ of the agents in State B is randomly chosen. They receive the service, while the other agents in State B do not. The agents that have been matched then switch to the other state. Figure 2 illustrates the game. The end of the game is not specified a priori as in many other economic problems. An agent assumes that the number of rounds is infinite.

Between two rounds the utility of each actor is discounted: In consequence, the utility for consuming a service in the future is lower than the utility for consuming a service now. If discounting was not introduced, it would not matter when an agent received a service – the utility would always be the same. As the payoff for consuming a service is higher than the cost of offering a service, it would be rational to always offer a service when in State A. Co-operation would be the dominant strategy. When discounting is introduced, the outcome of the game is less obvious: Think of an agent remaining in State B for several rounds, and the discount factor is high. Then the utility gained from consuming the service and changing to State A is low. It is lower than the utility lost for offering the service, in order to arrive in State B in the first place. Using economic methods one can show when exactly cooperation is rational [14]: This is the case if the expected payoff in the future exceeds the present costs. The standard approach to find equilibria for games with finite horizon is backward induction. Backward induction shows that only non-cooperative behavior is rational. When the horizon is infinite, as in structured P2P systems, a common method to identify strategies is based on stationary equilibria. [14] shows that the key parameter that makes the difference between cooperative and uncooperative behavior of an agent is the fraction of his service requests processed as desired. If the fraction exceeds a certain cut-off value, the agent cooperates.

Another theoretical result is that free-riding of one peer does not lead to a break-down of the system. This is because the other peers have a positive payoff, even when one peer resorts to free-riding.

As each peer may forward or answer a query (provide a service) or receive a query result (receive a service), we see the model of Hens and Vogt as a simplification of the interactions in structured P2P systems. If the number of peers providing a service is below the one of peers receiving a service, this results in peers which do not receive results they have requested. We expect peers in our system to use the equilibrium strategies of the model. This yields the following hypotheses.

Hypothesis S (Cut-off strategies): If the number of queries answered divided by the number of queries issued by a player in the past (past success frequency) exceeds a threshold (cut-off value), the player will cooperate in the P2P system. Otherwise, the player will not answer, forward or send queries.

The model of Hens and Vogt also predicts that free-riding does not lead to a break-down of the system in equilibrium. This leads to Hypothesis FR for structured peer-to-peer systems.

Hypothesis FR (Free-Riding without break-down): Free-riding of $1/k$ of all peers in a structured P2P system will not lead to a break-down of the system.

Here, k is the number of peers in the system. Our experiments show that this holds for k=6. We expect this result to hold for a fraction of 1/6 of free riders in larger P2P systems as well. However, experiments with more participants are extremely expensive both in terms of organization effort and compensation for the participants. This is in the way of a broader validation of this hypothesis. We intend to carry out further experiments with both system peers and human peers in the future, but this is beyond the scope of this article.

5. EXPERIMENT DESIGN

We have implemented a simulation environment for CAN using Java. It allows human participants to control peers. More specifically, a participant in the experiment controls the strategy of one peer. I.e., he decides whether to answer or forward a query, whether to send one, and which feedback to generate. In other words, the participants played against each other. The simulation environment ensures that peers interact properly. It performs the bookkeeping for each peer, as we will explain, so that the human participants can focus on their strategic decisions.

The number of participants in each experiment is based on a claim by Selten [25]. He postulates that “four are few and six are many”. In other words, a small group of more than five persons shows the same behavior as a large group. Section 6.1 will also comment on this issue. Hence, we conducted our experiments with six players. The experiments were played in ten groups, using computer terminals to control the peers. The terminals were separated from each other to prevent communication between participants.

The theoretical models behind our hypotheses are based on utility. Since we cannot directly observe utility, we have paid the participants depending on their behavior in the experiments. During the experiments the participants received points and paid with points. After the game, we computed their monetary reward based on their balance in points. Here, 100 points have resulted in 2 €. The average payoff per participant was 11.05 €.

5.1 Experiment Structure

In the beginning of each experiment the participants were randomly assigned to seats in the laboratory. The experiments lasted approximately 120 minutes. The first 20 minutes were used for orientation and to understand the instructions, given out in written form. After this first phase the participants have been asked to play several rounds to get used to the experiment. Afterwards three games were played. We will refer to these games as treatments.

Each treatment started with 20 rounds which were played without discounting. Afterwards a discounting rate was introduced. In economic experiments, a discounting rate ß corresponds to the probability that the game continues after the current period. It is known from economic theory that the interpretation of ß as a continuation probability or a discounting rate does not affect the results of an economic experiment. We used a discounting rate of 0.1 in our experiments. I.e., we rolled a 10-sided dice, and the treatment ended if the dice showed 1.

After the treatments we conducted a so-called strategy game. The concept of strategy game has been originally introduced in [26]. By then, it has become a common method to identify strategies in game settings, cf. [27]. A strategy game is one where participants are simply asked to write down their strategies. In economic experiments strategy games are typically played after the participants have taken part in several treatments. This allows to extract strategies of experienced players: Participants learn during treatments and refine their strategies, and they tend to have a thorough understanding of the experiment afterwards. In a strategy game

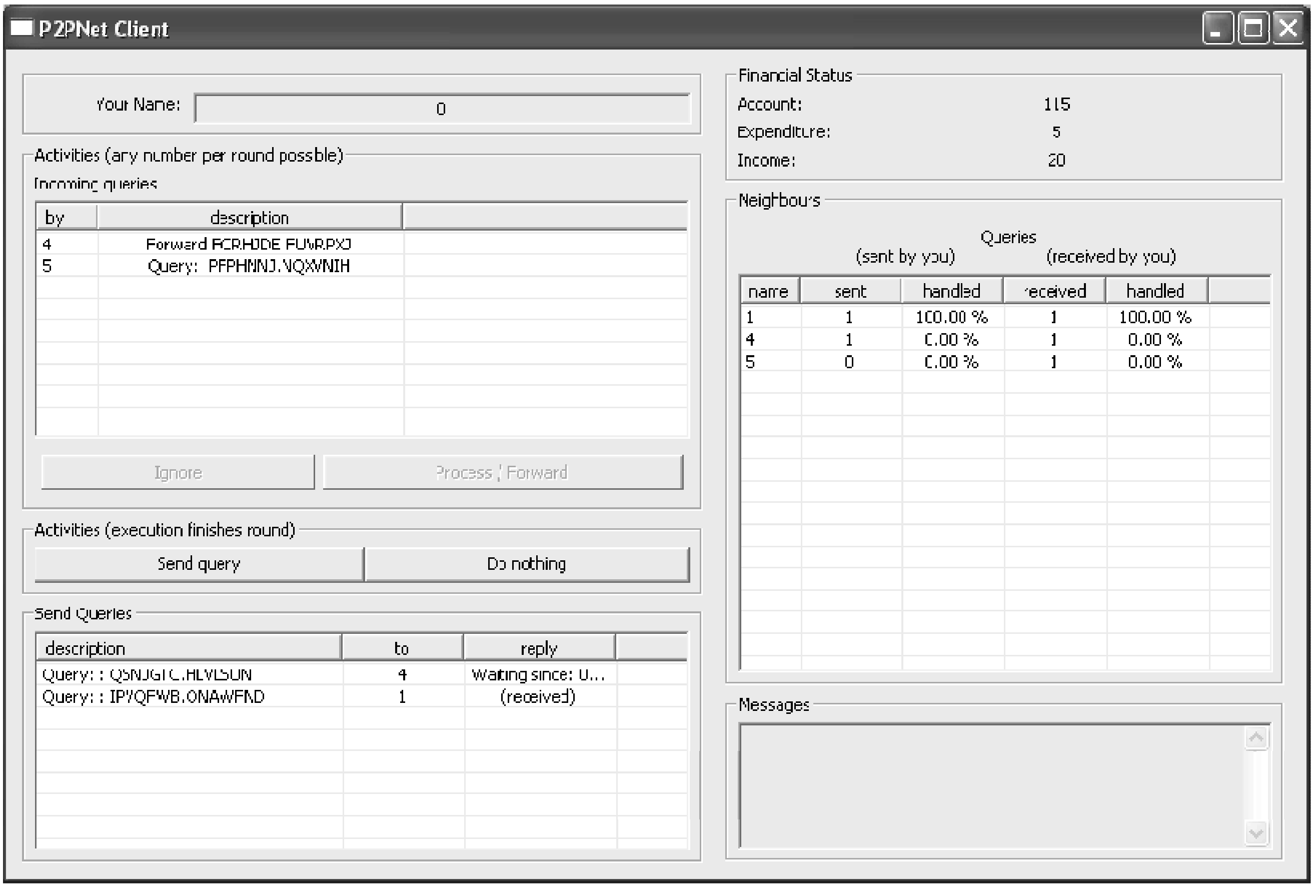

Figure 3. Experiment client

participants describe the strategies they have used during the treatments. Participants base their strategies on the payoff structure. Strategies may also depend on treatment parameters as well as on the history of a treatment. From a theoretic point of view, participants face the same task in the strategy game as in ’real’ plays of the game. The difference is that participants have to decide for all possible situations in a strategy game, compared to one situation in the play of the game. A strategy game reveals a complete strategy from a game-theoretic point of view. This is typically not the case in individual plays of the game.

Finally, each participant was paid depending on his performance.

5.2 Treatment Implementation

In the beginning of each treatment all participants received a list of 200 strings representing keys in a structured P2P network (List A). For every key from one of the lists the system stored one data object. The data objects were distributed such that every participant had to store 200 of them (List B). Even though the system consisted of ‘human’ peers, the human participants did not have to remember the (key, value) pairs themselves; the simulation environment took care of this issue. In every round each participant was allowed to issue a query, i.e., ask for the value corresponding to one of the keys on his List A. More specifically, the simulation environment chose the string. This is because first test experiments had shown that human participants had difficulties to do so. All this reflects the behavior of peers in structured P2P systems: They keep asking other peers for query results. A participant had to decide whether to process an incoming query or not. If he decided to do so, he was informed whether he could answer the query or he had to forward it, i.e., whether the value requested was on the list of the 200 strings he knew. If the query had to be forwarded, the next participant had the same action choices. When sending or forwarding a query, the simulation environment proposed to the participant up to three peers as potential addressees. These peers are neighbors of the current peer. The environment displayed them ordered by their Euclidean distance to the hash value of the query. The participant could then select one of these neighbors. The environment then forwarded the query to this participant. The environment lets the user choose the addressee. This is because the choice of the next participant is a strategic decision, which may not only depend on the probability that the peer had the information, but also on the past behavior of the peer. One reason why we have carried out the experiments is that we wanted to force the participants to develop respective strategies, and to observe them, in different treatments.

The balance of a participant at the beginning of a treatment was 100 points. Answering a query costs 5 points, issuing one costs 2 points, and forwarding costs 1 point. A participant received 20 points for obtaining a data object requested. See Table 2.

Figure 3 is a screenshot of the experiment client. It contains information on a participant with identifier 0. This player has sent one query to Player 1 and one to Player 4. The information in the bottom-left area of the screenshot indicates this. The query sent to Player 1 is already answered, not necessarily by Player 1 himself. The value "(received)" in column 'reply' shows this. The query sent to Player 4 is not answered yet. Player 0 has received one request from Player 5 he can answer himself and one from Player 4 he can only forward. This information is displayed in the upper-left area of the screen. The center-right area is a summary of all neighbors of Player 0. It tells us that Player 0 has sent one query to Player 1 and one to Player 4. (This piece of information is also available in the bottom-left area.) The rate of messages processed is "100.00 %". This indicates that Player 1 has already answered the query. Similarly this area indicates that Player 1 has sent one query to Player 0. The value "100.00 %" in the column labeled 'handled' indicates that this query is already answered. This piece of information is actually redundant, it can also be seen in the lower-left area of the screen. The value "Account" in the upper-right of the screenshot is the payoff of Player 0.

Table 2. Payoff depending on actions

Action	Payoff
Initial balance	100
Answering a query	-5
Sending a query	-2
Forwarding a query	-1
Receiving a query result	20

At the end of each round, the participants were informed about their payoff in the last round, their total payoff and the result of the chance move (whether the game continues). The properties of other participants, such as their score or their action choices, were kept secret. It was impossible for participants to find out the real-world identity of other participants. In addition to the information on their screen, the participants obtained no further information.

We carried out three kinds of treatments. In Treatment 1 we tested for the strategies of human participants in a structured P2P system without feedback. Treatment 2 was a variation of Treatment 1 – we introduced a free rider. In Treatment 3 we let the participants exchange feedback. Leaving this point aside, this treatment has been identical to Treatment 2.

Summing up, each participant had several action choices during the treatments: He could issue a query or not. He could forward/answer an incoming query or not. When forwarding or sending a query, he could choose the addressee. In Treatment 3 each participant could append one feedback object to a query sent. The participant could decide on the nature of the feedback (positive or negative) and freely choose the feedback subject.

We used Treatment 1 and 2 to test all hypotheses except Hypothesis F (the feedback mechanism). Treatment 3 should validate it.

6. RESULTS

This section summarizes the results of our economic experiments with human participants and relates them to our hypotheses. During nine months, we have conducted experiments with 60 students from various disciplines in our experiment laboratories in the University of Magdeburg. The participants were recruited by announcements in the university.

6.1 Hypothesis S (Cut-off Strategies)

To test whether cut-off strategies were used (Hypothesis S), we turned to the strategy game. Experience from other economic experiments shows that players tend to play modified strategies in the beginning and in the end of a treatment. It is difficult to determine the rounds during which this effect occurs. Hence, we resorted to the strategy game to validate the hypothesis. Another argument for the strategy game is that cut-off strategies may only be observed from experienced players. The effects of learning should be excluded as far as possible. Furthermore, we wanted to relate the behavior of the participants to their assessment of other players. This relationship is best detected with strategy games.

The strategy game showed that the strategies of the participants did not differ between Treatment 1 and 2. For the detailed results see [19]. Most strategies consisted of a start phase and a main phase. During the start phase, which ended after a few periods, the participants showed very diverse behavior. As participants have only participated in few interactions in the beginning of the game, their behavior in the start phase could not depend on the past success frequency. Since economic theory only allows analyzing the behavior of peers in equilibrium, and our hypothesis is based on economic theory, we concentrate on the strategies played during the main phase. Table 3 shows the results of our analysis.

The strategies can be classified in the following three classes:

a) action choices depend only on the past success frequency of own queries independent from the sender (with different cut-offs),

b) action choices depend on the past success frequency of own queries and on other factors,

c) the strategies are not in line with Hypothesis S.

Three of the sixty participants in the experiments played strategies which depend on the absolute number of unanswered queries and not on their rate. Five participants used strategies neither depending on the success frequency nor on the absolute number of unanswered queries. The 52 other participants played strategies depending on the past success frequency. 41 of them were in line with Hypothesis S. Eleven participants played slight modifications of a strategy according to Hypothesis S. Three of them tried to anticipate the end of the treatment. Three additionally introduced the condition of doing less than the others (a kind of free-riding). The other five participants played strategies which show differences to cut-off strategies.

To test whether the experiments confirm Hypothesis S, we introduced the Null-hypothesis that the participants select a strategy different from a cut-off strategy. We use a binomial test to test this hypothesis. 41 cut-off strategies (out of 60) contradict the Null hypothesis. The Null hypothesis can be rejected on a significance level of 1%. This test result confirms Hypothesis S.

Consequently, we claim that most players pursue cut-off strategies in structured P2P systems. Only very few participants realized the chance of stopping cooperation in the end of the treatments. Our conclusion is that this calls for countermeasures against defection in the end phase. One possible approach is the introduction of so-called proofs of work [15]. Using such proofs, we could force the peers to donate the system resources whose value exceeds the damage of defection in the end phase.

Table 3. Types of strategies in strategy game

Strategy	Category	# persons
Cut-off strategy only depending on past success frequency	a)	41
Cut-off strategy depending on past success frequency plus end phase	b)	3
Cut-off strategy depending on past success frequency plus the condition to answer less queries than own queries are answered	b)	3
Cut-off strategy depending on past success frequency for queries initially sent, random behavior when answering queries randomly on behalf of others	b)	4
Cut-off strategy depending on past success frequency for queries initially sent, answering queries on behalf of others after several non answered queries were received	b)	1
Cut-off strategy depending on the absolute number of own queries not answered	c)	3
Different types of strategies	c)	5

Clearly, six participants are few, compared to thousand of peers. Nevertheless, we expect the behavior of the human participants in our experiments to be independent from the number of participants, at least in qualitative terms: The outcome of the strategy game is that users will play a cut-off strategy, irrespective of the size of the system. On the other hand, a larger system might be less stable: This is because it is now easier to exceed the threshold value of the cut-off strategies.

6.2 Hypothesis R (Reciprocity)

The results so far show that participants use cut-off strategies. We now analyze whether they are individualized. To do so, we look again at the strategies from the strategy game. We calculated the percentage of strategies which depended directly on the identification of a peer, at least partially. All cut-off strategies observed are individualized (for details see [19]). 55 strategies confirm Hypothesis R, while 5 other strategies do not. Again a binomial test confirms the hypothesis on the 1% level.

On the one hand, this shows that peers tend to cooperate with the players they deem cooperative. On the other hand they will expect cooperation from others if they have been cooperative in the past. Hence it is important to ensure that the system treats cooperative players better than uncooperative ones.

6.3 Hypothesis T (Trust)

We now evaluate the game with regard to the degree of trust. According to Hypothesis T the intensity+in a game quantifies the trust in the system. Intensity is the number of queries that were sent by the peers divided by the maximum number of queries that could have been sent. The number of maximum queries that can be sent is the number of rounds of a treatment multiplied with the number of peers. The average payoff is the sum of payoffs of all participants divided by the number of periods played and the number of peers. Figure 4 shows the correlation between intensity and average payoff for all groups, averaged over all 3 treatments.

A linear regression shows a positive correlation between these factors, see the regression line in Figure 4. The slope of the line relating intensity to average payoff is significantly different (on the 1%-level) from zero. We conclude that Hypothesis T is supported. In other words, we show that trust increases efficiency.

If we look at the strategies in the strategy game every participant started sending queries without knowing anything about the past of other peers (see [19] for details). It also holds that all players react reciprocally and do not free ride. These two facts show that all participants cooperate in a structured P2P network.

Because of this conclusion, we had to introduce free-riding peers to study the impact of free-riding according to Hypothesis FR.

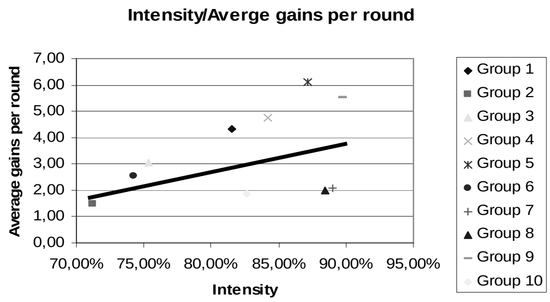

Figure 4. Correlation between average payoff and intensity

This hypothesis has one important implication on P2P data structures: We can expect cooperation between the participating peers.

6.4 Hypothesis FR (Free-Riding without Break-down)

In Treatment 2 a free rider or destructive player was introduced. This participant was instructed to ask for query results and to never process any incoming queries.

Table 4. Average payoff per round

	Treatment 1		Treatment 2		
Group	all peers	destructive peer in Treatment 2	all peers	destructive peer	other peers
1	5.85	7.71	4.11	3.45	4.25
2	2.63	5.90	0.89	1.27	0.82
3	4.40	6.52	1.85	2.35	1.75
4	6.79	10.20	2.81	8.00	1.78
5	7.78	7.36	4.47	3.83	4.59
6	2.31	5.95	0.99	3.67	0.46
7	4.31	-0.50	1.47	2.17	1.33
8	5.30	0.45	0.93	0.50	1.02
9	5.01	7.95	4.61	5.20	4.50
10	2.10	0.09	1.27	3.56	0.81
Mean	4.65	5.16	2.34	3.40	2.13

To test whether free-riding can lead to a break-down of the system (Hypothesis FR), we compare the results of Treatments 1 and 2 in Table 4. In all ten groups the payoff in Treatment 2 with a free rider is smaller than in Treatment 1. A binomial test shows that this result is significant on the 1% level. The efficiency of the system is correlated to the average payoff per round. Hence we conclude that the efficiency decreases. The payoff in Treatment 2 is about half as high as the one in Treatment 1, but it is still positive. If the end of a treatment is known, a P2P system might break down, according to many results in experimental economics, see for example [1], and according to the prediction of the subgame perfect equilibrium. During the first 10 rounds of the treatments cooperation tends to be the strongest. Even if the analysis ignores the first 10 rounds of the treatments, we obtain the same results (see Table 5). This confirms Hypothesis FR.

When comparing the payoff of the destructive player to the average payoff of the remaining peers in Treatment 2, we see that the payoff of the destructive participant is slightly higher than the payoff of the other ones: In 8 out of 10 groups he performs better. However, the average payoff of the destructive player in Treatment 2 is lower than his payoff in Treatment 1. In 7 of 10 groups the destructive participant performs better in Treatment 1. This observation is important: it shows that free-riding does not pay for any participant.

Table 5. Average payoff after round 10

	Treatment 1		Treatment 2		
Group	all peers	destructive peer in Treatment 2	all peers	destructive peer	other peers
1	5.77	8.29	5.07	4.67	5.15
2	3.58	7.80	0.83	0.09	0.97
3	4.61	8.00	0.97	-0.46	1.26
4	7.43	9.73	2.74	5.14	2.26
5	8.40	7.67	4.74	1.86	5.31
6	2.50	1.90	1.07	3.14	0.66
7	4.29	0.21	0.99	-0.57	1.30
8	6.01	2.92	0.68	-2.00	1.21
9	5.71	7.50	4.98	2.00	5.57
10	2.31	-1.23	0.95	4.94	0.15
Mean	5.06	5.28	2.30	1.88	2.39

Participants might not always be interested in obtaining the maximal payoff for themselves. Rather, they just might want to obtain more than others. This is a special form of competition. Hence we calculate the relative payoff. The relative payoff is the quotient of one's own payoff and the average payoff of the remaining participants. If one participant was motivated by his relative payoff, free-riding would be attractive. This is because he earns more than all other participants. For participants who are motivated by their own payoff in absolute terms, free-riding does not pay. This is because the destructive participant receives a lower payoff than in the treatments where he cooperated. Fortunately, in most scenarios, including structured P2P systems, relative payoffs are a notion that is only virtual. In other words, free-riding is not attractive. During our experiments no participant defects, unless he has been instructed to do so. The P2P protocol under investigation protects from free-riding.

6.5 Hypothesis F (Feedback)

If all players were able to differentiate between cooperative players and free riders, free riders should not obtain any payoff. This is the ultimate design objective behind the mechanisms examined in this paper. To ease differentiation between cooperative and uncooperative players we suggest the use of feedback to have more information about other players available.

For a test of Hypothesis F (i.e., whether feedback performs well) we compare the average payoff in Treatment 2 to the average payoff in Treatment 3. Both treatments are with a free-rider. In Treatment 3 feedback was possible. Table 6 shows the result.

As can be seen in Table 6, the payoff of the participants increases on average. This shows the advantage of a feedback mechanism if a free-rider is in the system. This result is in line with our Hypothesis F. If we test Hypothesis F by means of a binomial test we calculate as a result a tendency (significance level 18%) that supports Hypothesis F.

Table 6. Average payoffs in Treatments 2 and 3

	Average payoff in Treatment 2	Average payoff in Treatment 3
Group 1	4.11	3.01
Group 2	0.89	0.90
Group 3	1.85	2.95
Group 4	2.82	4.66
Group 5	4.47	6.14
Group 6	0.99	4.3
Group 7	1.47	0.43
Group 8	0.93	-0.20
Group 9	4.61	6.88
Group 10	1.27	2.17
Average	2.34	3.12

An implication of this result is that a feedback mechanism improves the working of the system. Since our system already works very well, the impact of a feedback mechanism is not as big as it might be in other systems.

7. CONCLUSIONS

Structured P2P networks are useful in a broad range of application scenarios. Protocols for structured P2P systems have to ensure the stability of the system and the cooperation of peers. The system design must rule out free-riding. In this paper we used economic and computer tools to analyze the strategic aspects of structured P2P networks. With the help of game theoretic models we derived several hypotheses on the strategic behavior of peers. The models predict that cooperation between peers is correlated to the trust of peers in the system. They also predict that the strategies of the participants are individualized. Another expectation is that free-riding of some of the peers does not cause a break-down of the whole system. Based on the models we also predict that a feedback mechanism increases cooperation and reduces the impact of free-riding. A final prediction is that the probability of participation in the system depends on the fraction of positive interactions with other peers in the past. We then tested our hypotheses with economic experiments with human peers. The experiments confirm our theoretical predictions.

Economic experiments with humans allow to test which behavior does not occur in structured P2P systems, and which behavior the designers of P2P systems must take into account. As far as we know, we are the first to investigate the behavior of humans who mimick the role of peers in P2P structures. Compared to simulations, this approach can do without any assumptions regarding the behavior of peers. Their behavior is simply observed.

Acknowledgements. We thank the anonymous reviewers, the members of the database group at Universität Karlsruhe (TH), and Erik Buchmann for their valuable comments.

8. REFERENCES

[1] Andreoni, J. Why Free Ride? Strategies in Public Goods Experiments. In *Journal of Public Economics*, 37, 1988, 291-304.

[2] Adar, E. and Huberman, B. *Free Riding on Gnutella*. Technical Report, Xerox PARC, 2000.

[3] Berg, J., Dickhaut, J. and McCabe, K. Trust, Reciprocity and Social History, In *Games and Economic Behavior X*, 1995, 122-142.

[4] Böhm, K. and Buchmann, E. Free Riding-Aware Forwarding in Content-Addressable Networks. In *The VLDB Journal*, Online First, 2006.

[5] Buragohain, C., Agrawal, D. and Suri, S. A Game Theoretic Framework for Incentives in P2P Systems. In *Proceedings of the 3rd International Conference on Peer-to-Peer Computing*, 2003.

[6] Christin, N. and Chuang, J. A Cost-Based Analysis of Overlay Routing Geometries, In *Proceedings of IEEE INFOCOM'05*, 2005.

[7] Chun, B. et al. Characterizing selfishly constructed overlay networks, In *Proceedings of IEEE INFOCOM*, 2004.

[8] Engelmann, D. and Fischbacher, U. *Indirect Reciprocity and Strategic Reputation Building in an Experimental Helping Game*. Working Paper 132, Institute for Empirical Research in Economics, University of Zürich, 2002.

[9] Fabrikant, A. et al. On a network creation game. In *Proceedings of ACM PODC'03*, 2003.

[10] Feldman, M. et al. Robust incentive techniques for peer-to-peer networks. In *Proceedings of the 5th ACM conference on Electronic commerce*, 2004.

[11] Ge, Z. et al. Modeling Peer-Peer File Sharing Systems. In *Proceedings of the IEEE Infocom*, 2003.

[12] Golle, P., Leyton-Brown, K. and Mironov, I. Incentives for sharing in peer-to-peer networks. In *Proceedings of the 3rd ACM conference on Electronic Commerce*, 2001.

[13] Gummadi, K., et al.: The Impact of DHT Routing Geometry on Resilience and Proximity. In *Proceedings of the SGICOMM* 2003, 2003.

[14] Hens, T. and Vogt, B. *Money and Reciprocity*. Working Paper 138, Institute of Empirical Economics, University of Zürich, 2002, in revision.

[15] Jakobsson, M. and Juels, A. Proofs of work and bread pudding protocols. In *Proceedings of the IFIP TC6 and TC11 Joint Working Conference on Communications and Multimedia Security (CMS '99)*, Leuven, Belgium. 1999.

[16] Ramasamy, L. and Liu, L. Free Riding: A New Challenge to Peer-to-Peer File Sharing Systems, In *Proceedings of the 36th International Conference on System Sciences*, 2003.

[17] Ranganathan, K. et al. To share or not to share: An analysis of incentives to contribute in collaborative file sharing environments, In *Proceedings of the International Workshop on Economics of Peer to Peer Systems*, 2003.

[18] Ratnasamy, S. et al. A scalable content-addressable network. In *Proceedings of the 2001 conference on Applications, technologies, architectures, and protocols for computer communication*, 2001, 161-172.

[19] *Results of the Treatments and Strategy Games*: http://www.ipd.uni-karlsruhe.de/~schosser/ec06/

[20] Rowstron, A. and Druschel, P. Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems. In *Proceedings of Middleware 2001 : IFIP/ACM International Conference on Distributed Systems Platforms*, 2001.

[21] Saroiu, S., Gummadi, P. and Gribble, D. A measurement study of peer-to-peer file sharing systems. In *Proceedings of Multimedia Computing and Networking*, 2002.

[22] Schlosser, M. and Kamvar, S. Simulating a P2P File-Sharing Network. In *Proceedings of the First Workshop on Semantics in P2P and Grid Computing*, 2002.

[23] Schosser, S. et al. Strategic Properties of Peer-to-Peer Data Structures and Behaviour of Human Peers in Economic Experiments. In *Proceedings of the Intenational Workshop on Security and Trust in Decentralized / Distributed Data Structures*, 2006.

[24] Seinen, I. and Schram, A. *Social Status and Group Norms: Indirect Reciprocity in a Helping Experiment*. Working Paper, University of Amsterdam, 2001.

[25] Selten, R. A simple model of imperfect competition where four are few and six are many. *International Journal of Game Theory 2*, 141-201, 1973.

[26] Selten, R. Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eine Oligopolexperiments. *Beiträge zur experimentellen Wirtschaftsforschung*. Heinz Sauermann (ed.), 1967, 136-168.

[27] Selten, R., Mitzkewitz, M. and Uhlich, G. Duopoly Strategies Programmed by Experienced Players. In *Econometrica 65*, 1997, 517-555.

[28] Sigmund, M. and Noak, K. Evolution of Indirect Reciprocity by Image Scoring. In *Natur*e 393, 1998, 573-577.

[29] Stoica, I. et al. H. CHORD: A Scalable Peer-to-Peer Lookup Service for Internet Applications, In *Proceedings of the ACM SIGCOMM'01*, 2001.

[30] Zhao, B., Kubiatowicz, J. and Joseph, A. *Tapestry: An Infrastructure for Fault-tolerant Wide-area Location and Routing*. Technical Report, University of California at Berkeley, 2001.