Indirect Partner Interaction in Peer-to-Peer Networks – Stimulating Cooperation by Means of Structure

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}

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
{bodo.vogt@ww.uni-magdeburg.de}

ABSTRACT
Structured P2P systems are a prominent representative of a class of systems where participants communicate via a given structure. Peers in such systems must show cooperative behavior, to avoid performance degradation of the system. Economic literature has proposed various mechanisms to stimulate cooperation. Their effectiveness strongly depends on the interaction scenario. The two important scenarios are the *partner scenario*, where participants interact repeatedly, and the *stranger scenario*, where participants tend to interact only once. The use of shared histories is beneficial in the stranger scenario, but it is not necessary in the partner scenario. This paper is based on the observation that the systems investigated here do not match either of these scenarios. Thus, we propose *indirect partner interaction* as a new interaction scenario, i.e., peers interact indirectly via a sequence of peers. To study peer behavior in this new scenario without any assumptions, we have carried out economic experiments. They give way to the following results: Participants interacting on behalf of strangers show roughly the same degree of cooperative behavior as with the other mechanisms examined, like partner design or punishment. While participants tend to rely on the shared history if no other information is available, they use the network structure as basis for their strategic decisions whenever possible. The presence of a shared history does not lead to an increase of the payoff earned in such a system. We conclude that the settings investigated here do not need shared histories to stimulate cooperation.

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, Feedback, Game Theory, Social Exchange, Structured Peer-to-Peer Networks.

EC'07, June 11–15, 2007, San Diego, California, USA.

1. INTRODUCTION
Peer-to-peer systems (P2P systems) are coordinator-free systems administering large amounts of data. A peer gains positive utility from obtaining results for queries it itself has issued, while its utility when processing queries is negative. The difference of the positive and the negative utility of a peer is its *payoff*.

Structured P2P systems are a useful variant of P2P systems: A distinctive feature is that each data object is assigned to a certain peer. A peer exchanges messages with a few other peers only, its *contact peers*. In other words, each peer only *interacts directly* with a certain fraction of all peers. However, message exchange often is on behalf of a distant peer, which does not have contact with the peers exchanging the message. In other words, *indirect interaction via partners* occurs. There is a broad range of systems featuring this kind of interaction, and structured P2P systems are just a prominent representative (see Section 7). In what follows, we may use the term 'structured P2P systems' to refer to this entire class of systems.

In P2P systems that are operational, free riding is a problem [1][21]: In the absence of incentives, in order to maximize their payoff, peers tend to consume resources provided by others while not contributing to the system. In consequence, the system encounters a significant decrease in performance or even collapses [18]. We expect such free-riding behavior in structured P2P systems as well, and it even is more critical: There typically are several peers forwarding a query message. If only one of them defects, the query is lost.

To counter free riding, the behavior of a peer in the past can serve as a criterion for how its contacts should treat it in the future [7][14]: In a nutshell, only peers which have been cooperative, i.e., have contributed to the system, may benefit from it. The starting point for this work is the following observation: in P2P systems, two sources of behavior information exist: (a) A peer monitors the behavior of its contacts, i.e., own observations. (b) It can rely on observations made by other peers, aka. *feedback* or *shared history*. Both own observations and shared history have significant advantages and disadvantages: While own observations are always truthful and are not subject to manipulation, they only consider a small share of the interactions in the system, namely direct interactions. Shared histories in turn can observe a larger fraction of the interactions, i.e., both direct and indirect ones. But guaranteeing the truthfulness of feedback is difficult.

The objective of this paper is to assess the usefulness of shared histories in structured P2P systems. In this context, the distinction between *global shared histories* and *local shared histories* [3] is important: With local shared histories, only some peers, for instance a peer and its contacts, exchange feedback. Global shared histories in turn are open to any peer. In this article, we focus on local shared histories. This is because local shared histories are more "trustworthy" from the perspective of the peers, in contrast to global ones containing feedback from just any peer (we argue).

While research on the benefit of feedback in P2P systems (albeit not structured ones) exists, e.g., [7], such work is typically based on assumptions. For instance, [7] relies on the assumption that feedback is truthful and readily accessible. We in turn envision results that are 'assumption-free' regarding peer behavior. This is the case if results are based on the actual behavior of participants in such systems.[1] Thus, our method of choice is economic experiments. With such experiments, we simply observe the behavior of human participants in structured P2P systems. In the course of the evolution of mankind, humans have kept refining their strategies. This is why many researchers, be they economists, be they biologists, deem human strategies mature and sophisticated [11]. We conclude that the strategies observed can serve as a basis for stable and efficient protocols for structured P2P systems.

From an economic perspective, two interaction scenarios exist: In a *partner scenario*, participants interact primarily with the same partners. In such a setting, one typically relies on own observations [15]. The second scenario is the *stranger scenario*. A stranger is a participant one is not likely to interact with again in the future. In this scenario, one typically uses the shared history [4]. One important insight of ours is that structured P2P systems have characteristics of both scenarios: It is not clear how to weigh the two mechanisms (own observations, shared history), and neither of those scenarios in isolation maps to structured P2P systems. Thus, we introduce the notion of *indirect partner interaction*, a new economic scenario combining aspects of the stranger and the partner scenario. More specifically, indirect partner interaction has two distinctive properties:

- A participant provides a service (forwards or answers messages) on behalf of a stranger.
- Direct interaction always occurs between partners. (A partner forwards a message to another one.)

An important finding from our experiments is that the degree of cooperation between participants in indirect partner interaction is about as high as in a partner scenario or in a stranger scenario with feedback. In more detail, we are able to demonstrate the following points regarding the behavior of participants in structured P2P systems: Participants do use feedback when interacting directly with others they have not interacted with and are not linked to by any network structure. But as soon as participants can interact repeatedly with others, they do not use feedback any more, even when interacting on behalf of strangers. In a system in steady state, participants even do not rely on feedback on newcomers, but rather wait for own observations. The fact that participants do not rely on feedback is interesting. This is because participants do not have any information on strangers, except that they are connected to them via contacts that deem each other cooperative. We conclude that the use of feedback does not increase the payoff of the participants in indirect partner interaction scenarios. We find this surprising, since we had expected local shared histories to be more beneficial. Finally, we discuss why we expect our results to hold for any system where participants interact via distant communication channels using a certain network structure.

This paper is organized as follows: Section 2 reviews structured P2P systems. Section 3 summarizes related work. Section 4 derives our hypotheses concerning the usage of feedback, which we try to verify in Section 6. Section 5 describes the design of our experiments. Section 7 features a discussion, Section 8 concludes.

2. STRUCTURED P2P SYSTEMS

This section is a short introduction to structured P2P systems and Content-Addressable Networks (CAN). Although we deem these structures important – our experiments were based on a CAN –, we expect our findings to hold for any system featuring indirect partner interaction.

Structured P2P systems such as CAN [19], Chord [27] or Pastry [20] administer large sets of (key, value)-pairs. Each (key, value)-pair is mapped to a certain location in the key space. Therefore, each key is represented by a coordinate, the query point. Each peer is responsible for a certain fraction of the key space. We call this fraction of the key space the *peer zone*. Each peer manages all (key, value) pairs whose key lies in its zone. In addition to its zone, each peer also knows some other peers, its *contacts*, and the coordinates of their zones. A peer processes a query (which simply is a key) as follows: It first transforms the key into the corresponding query point. If the query point falls into its zone, it returns the value. Otherwise, it forwards the query to the contact which is closest to the query key. This procedure recurs until the peer knowing the (key, value)-pair is reached. The query result is then returned to the initial sender.

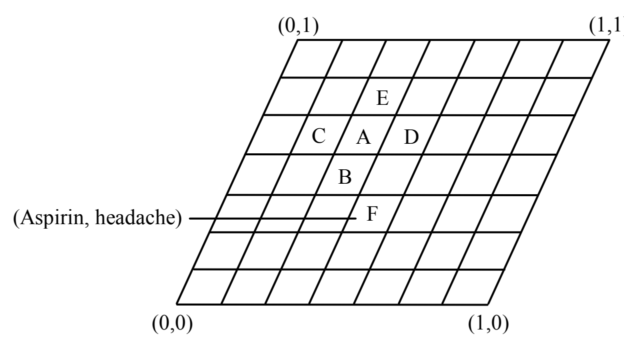


Figure 1. Content-Addressable Network

Example: Figure 1 illustrates this for a Content-Addressable Network (CAN). All keys have been transformed to two-dimensional coordinates using a hash function. For instance, the key of the (key, value)-pair ("Aspirin", "headache") is mapped to the coordinates (0.45, 0.3). Each square is the zone of a peer. Hence, Peer F stores the value corresponding to "Aspirin". In CAN, the contacts of a peer are exactly the peers with adjacent zones. Thus, Peers B, C, D, and E are the contacts of Peer A. If Peer A issues the query "Aspirin", it first maps "Aspirin" to the coordinates (0.45, 0.3) using the public hash function. It then forwards the query to the contact that is closest to this point, Peer B. Peer B does not know the (key, value)-pair sought and forwards the query. This procedure recurs until the query arrives at Peer F. Finally, Peer A obtains the result value. ▪

The example reveals two important properties of structured P2P systems: First, each peer directly interacts with its contacts only. Second, several peers have to cooperate for a query to be answered.

[1] In what follows, we use the terms 'participant' or 'player' to refer to the human controlling the behavior of the peer; we use the term 'peer' when referring to the node that is part of the network and interacts with others.

If only 5% of the peers do not cooperate, 40% of the queries are not answered, under realistic assumptions [3].

Peers can monitor the behavior of their contacts. Each peer can distinguish between cooperative and uncooperative contacts, and it can decide autonomously which contacts it deems cooperative/uncooperative. Further, a peer decides itself if it processes a query obtained from a contact, and which contact to forward a query to, should this be necessary. Our expectation had been that a peer does not necessarily forward queries obtained from contacts it deems uncooperative (reciprocity). Experiments with human participants have confirmed this [22].

A mechanism to discriminate between cooperative and uncooperative contacts is effective if it classifies few cooperative participants uncooperative and vice versa. In structured P2P systems, a peer can distinguish between peers based on two kinds of information: First, *own observations* of behavior, and, second, *feedback* issued by other peers which (claim to) have already interacted with the peer in question. A peer can attach feedback to any message it sends to another peer [6]. A peer may collect incoming feedback in its local shared history, be it positive, be it negative. A peer can then see how many negative and positive feedback items it has received from other peers regarding a certain peer.

This gives way to one fundamental question: Under which circumstances do participants rely on feedback, as opposed to own observations? The remainder of this paper will answer this question for structured P2P systems.

3. RELATED WORK

Computer scientists have used feedback to establish cooperative behavior in online communities. Most of these approaches are applicable to structured P2P systems. We discuss such approaches in this section.

Some approaches use feedback to predict the cooperativeness of peers: [16] does so by collecting observations in a global repository and evaluating them. They show that their mechanism can decrease the negative impact of uncooperative participants. This approach relies on assumptions; it assumes that feedback is truthful.

To make feedback more reliable, i.e., to decrease the impact of untruthful feedback, several authors propose peers to weight feedback based on own observations [26][6]. These mechanisms have the following drawback: While peers might behave cooperatively when processing queries, they might issue feedback that is not truthful. For structured P2P systems, it has not been analyzed so far whether such feedback mechanisms are beneficial, given the eventuality that peers might issue feedback strategically.

Another idea is the design of incentives rewarding honest feedback [13][17]. Both approaches need a central, trusted authority which observes the quality of feedback and punishes dishonest peers. The approaches are not readily applicable to structured P2P systems where any central node does not exist.

[12] proposes to decentralize the mechanism described in [13], using several peers to evaluate the outcome of interactions. The peers then store their observations. Peers interested in the cooperativeness of a peer query all of these peers for their observations. This leads to a new problem: How to ensure that the observing peers return honest feedback? [14] uses reputation agents to save reputation information. The problem here is similar: Participants have to trust the reputation agents.

Thus, the design of feedback mechanisms to establish cooperation is a central issue in the literature. But the questions whether other mechanism might be better, and if they avoid the problems related to feedback has not been investigated in our current context. From a different perspective, we are not aware of any mechanism ensuring reliable feedback which is readily applicable to structured P2P systems.

While the work described here uses the same methodology as [22], the problem investigated there is a different one. [22] does not investigate the usefulness of feedback in structured P2P systems in detail.

4. HYPOTHESES

Much work in experimental economics has addressed cooperation, trust and reciprocity, see for example [2][4][10]. It shows that trust and reciprocity determine the behavior of humans in many situations where they interact. Many papers deal with the problem of establishing cooperation in social dilemma situations, e.g., [4][9]. These are situations where opportunistic behavior is rational, but does not lead to an efficient outcome. In economic experiments, humans show cooperative behavior depending on the design of the interaction. The degree of cooperation crucially depends on subtle variations of the experiment design.

The structure of structured P2P systems is complicated, compared to the one of the vast majority of economic games described in the literature. We are not aware of any economic game that can serve as a model of a structured P2P network. This calls for new models, together with experiments to analyze their characteristics and respective strategies. Designing such experiments and formulating hypotheses to be verified in such experiments is not obvious. In the remainder of this section, we therefore discuss the behavior of individuals observed in economic experiments on several idealized games and the relevance of this behavior for structured P2P networks. We focus on design variations that typically lead to cooperation. In the last subsection of this section, we derive hypotheses on the behavior in structured P2P systems.

4.1 Punishment

Punishment has turned out be effective in order to establish cooperation: Fehr and Gächter [9] have demonstrated this with the following public goods experiment: A group of four players receives an endowment of 20 money units each. Each participant can contribute a self-determined fraction of this money to a group project, and he keeps the remaining money. All participants then receive 0.4 money units for each money unit invested in the group project by any participant: While the benefit of the investor for one money unit invested is lower than his investment (0.4 money units), the payoff of the group as a whole is higher (1.6 money units). Hence, from a game-theoretic perspective, it is in the best interest of a participant to keep his money – irrespective of the amount of money invested by the other participants. But this is not efficient: While each player earns 20 money units when not investing, he would earn 32 money units if all players invested.

Fehr and Gächter carried out two versions of this experiment: One without punishment option, which corresponds to the experiment described, and one with such an option. The punishment option introduces the following modifications: The participants were allowed to punish each other after they had been informed about the investments of the other players. The implementation of punishment has been as follows: A player had to assign a certain number of points (between 0 and 10) to the player he wanted to punish. Each point assigned resulted in a deduction from the balance of the player to be punished (3 money units), but also from the one of the punisher (1 unit).

Several groups of participants played this game in parallel. The game was conducted repeatedly. After each round new groups where formed at random, but in a way that no participants played with each other more than once. This rules out effects such as direct reciprocity or reputation.

It would be rational for a player to neither punish nor cooperate – punishment and cooperation are costly and yield no additional payoff. Future players however may benefit from punishment in current rounds: A punished player might increase his investment in the periods that follow. In the presence of punishers even selfish players would have a reason to cooperate in the punishment version of the experiment. In the experiment in the laboratory, a large fraction of the players did punish defectors.

P2P systems allow for punishments as well: if a peer defects, i.e., does not process any queries, it is straightforward to punish by not processing its queries in return. This punishment is possible at no cost. This is even 'better' than in the Fehr and Gächter design where punishment incurs costs. Thus, we hypothesize that punishment fosters cooperation in our setting as well (Hypothesis "Indirect Punishment" formulated in Subsection 4.4). However, in contrast to Fehr and Gächter, it typically is not possible to punish a peer directly which does not cooperate. This is because interaction is indirect. It is only possible to punish a neighbor peer by not forwarding a query obtained from it or the whole group of peers by not forwarding any queries. The experiments of Fehr and Gächter do not tell us how punishment looks like in our setting, in particular when a peer wants to punish a distant peer.

4.2 Partner Design vs. Stranger Design

Another feature that enhances cooperation is to let individuals directly interact repeatedly. This is called a *partner design*, in contrast to a *stranger design* [15]. One can expect different behavior of humans in these two scenarios, due to the fact that humans tend to cooperate in small groups (as a family or among friends), but they tend to defect in anonymous surroundings.

Keser and van Winden [15] examine this behavior in public good experiments and find clear indications of partner-partner interaction, compared to interaction between strangers: They conducted public goods experiments similar to the experiments described above with one slight modification: Keser and von Winden played two different treatments of the game: A stranger treatment corresponding to the no punishment option used in the Fehr and Gächter experiments and a partner treatment where the groups were the same in all rounds. After each round each player was informed about the total investment of the group in the last round. The experiments show that the payoff in the partner treatment is significantly higher than the one of strangers. In other words, when humans directly interact with individuals they know, they tend to behave more cooperatively, compared to interactions with strangers.

As stated before, structured P2P systems are not partner-partner interactions. This is because most of the time a peer is processing queries of peers it does not know. This is different from partner-partner interaction: It is not the partner but another peer which benefits from query processing. Nevertheless a peer keeps directly interacting with the same contacts: Peers process messages by indirect interaction while being organized in a certain given structure, e.g., a grid for CAN. In what follows, we refer to interaction via a certain structure as *indirect partner interaction*. To sum up, structured P2P systems are different from systems of partners or strangers. Mechanisms to motivate cooperation in such a scenario cannot be easily derived from existing models.

4.3 Feedback

Reciprocity induces cooperative behavior: In a partner-partner design, a participant can reciprocate cooperative behavior of his interaction partner. Such behavior is referred to as *direct reciprocity*. When participants do not interact repeatedly, only *indirect reciprocity* takes place: Here, a participant reciprocates if he has information on the outcome of past interactions.

In the literature on interaction on the internet, there is a consensus that such reciprocity is indirect, and that reputation mechanisms like feedback will increase cooperation significantly (see for example [4]). A helping experiment by Seinen and Schram analyzes the effect of feedback [23]: Interaction was between pairs of participants. They were randomly assigned the roles of donor and recipient. The donor had the choice of 'helping' the recipient. If he did, the recipient received a payoff higher than the cost of helping. Otherwise, both participants received a payoff of 0. Afterwards the recipient was informed about the decision of the donor. Two treatments of this experiment were conducted repeatedly: one no-information treatment, which was the same as the game described before, and one feedback treatment. In the second treatment, the donor was informed about the 6 last action choices of the recipient when he was a donor.

The experiments showed that by giving the participants information on the behavior of other participants in previous rounds, the degree of cooperation increases significantly. We want to find out how important this effect is in structured P2P networks, compared to the influence of the other mechanisms.

4.4 Behavior in Structured P2P Systems

Given the results of the experiments mentioned so far as reference points, we formulate several hypotheses.

According to Section 4.1, participants punish uncooperative participants even if it is costly for them. As punishing is not expensive in our scenario – a participant can drop a query of another participant at no cost –, we expect participants in structured P2P systems to punish as well. We further expect the participants to use the information available to them when deciding on punishments: If there is only feedback, we expect participants to punish on the basis of feedback. If they only have own observations, we expect them to use these as a basis for their decisions. We formulate the following hypothesis:

Hypotheses "Indirect Punishment": Individuals will indirectly punish depending on the information available.

a) If only feedback is available they will use the feedback.
b) If only a network structure is available, they will rely on own observations regarding their neighbors.

Our next two hypotheses relate to the issues from Section 4.2 and Section 4.3: It is unclear whether behavior observed in indirect partner interaction scenarios is comparable to one in the partner or in the stranger design. Thus, two mechanisms to increase cooperation might be applicable: If the behavior was comparable to the one in the partner design, participants would punish based on their own observations. In the other case, participants would rely on feedback. Both alternatives, i.e., the use of own observations or of feedback, as well as combinations might occur. [5] observes the effects of different mechanisms to stimulate cooperation. They compare stranger and partner scenarios with and without feedback mechanisms in a simpler setting without structure. The authors conclude that feedback increases cooperation compared to a stranger scenario, but that the degree of cooperation increases further in a partner scenario. Hence, we expect the characteristics of the partner scenario to be most influential and hypothesize as follows:

Hypothesis "Feedback free Strategies": In structured P2P systems (i.e., the network structure is given) where information about the contacts is available (indirect partner interaction), participants use strategies without feedback.

Hypothesis "Feedback Has no Impact on Payoff": A feedback mechanism in structured P2P systems does not increase the payoff.

Finally, the analysis of these hypotheses is important: If we confirm these hypotheses, we can conclude that implementing a feedback mechanism is not necessary in structured P2P systems – participants already show a high degree of cooperation based on the interaction structure in the system.

5. EXPERIMENT DESIGN

We have carried out economic experiments with human participants. To do so, we developed an experiment environment which allows participants to easily assume the part of peers. Each participant controls the strategy of one peer[2]: He decides if the peer under his control forwards/answers a query (for each query individually), and which contact to forward a query to.

The experiment takes place in rounds, and each peer may issue one query per round. The player decides if his peer does so (for each round individually). At the end of a round, all participants were informed about their current total payoff, the one of the last round, and whether the game continued. Properties of other participants were kept secret.

In the beginning of each treatment a participant had a balance of 100 points. Depending on his strategy choices, this balance was either increased or decreased: For obtaining the result of a query he had issued, he received 20 points, forwarding cost 1 point, answering a query cost 5 points, and issuing a query cost 2 points. These values reflect the costs with a structured P2P system: Forwarding or issuing a query is relatively cheap, as only one message is sent. Answering consists of sending a message and searching for information. While the exact utility of a query result clearly is domain specific, it is relatively high in many cases. Further, a setting where the benefit of receiving a query result is low does not make for interesting experiments – participating in such a system would not be beneficial, and this would be obvious.

The experiment environment hides certain details from the participants: For instance, it manages the data saved in the peer zones, calculates the distance to query points etc. Further, when a participant had to forward a query, the experiment environment showed the participant a list of his neighbors sorted by the probability that they had the desired query result, i.e., sorted by the distance to the point in the key space representing the query.

[2] A website containing screenshots, the exact description of the game (the written instructions for the players) and the detailed experiment results are available at: http://www.ipd.uni-karlsruhe.de/~schosser/ec07/

We conducted the experiments with 6 participants each, according to Selten [24]. He showed that even small groups of more than five human participants show the same behavior as large groups. We will discuss this aspect in more detail in Section 7.

Each experiment was played by twelve groups. The terminals of the participants were separated from each other to prevent any communication between participants. In addition, players could not identify the participants controlling the contacts of their peers.

At the beginning of an experiment, all participants were randomly seated in the laboratory. The experiments lasted about 80 minutes each. During the first 20 minutes we introduced the participants to the game: The instructions to the game were in written form, and several test rounds were played. We then carried out several treatments. Each treatment consisted of twenty rounds without discounting, and discounting with a discounting rate of 0.1 was introduced after twenty rounds. In other words, a ten-sided dice with one 1 was rolled. If it showed 1 the treatment ended, otherwise another round was played. Rolling a dice is a common trick used in economic experiments so that participants do not show 'end-game behavior': They do not anticipate the end of the game and behave as if the game went on forever.

For each treatment all peers were assigned a zone of the same size. All peers had the same number of neighbors and could therefore expect the same number of queries each. The experiment environment itself generated the queries randomly.

We conducted two treatments, a *trust treatment* and a *feedback treatment*. In the trust treatment, a participant knew which share of queries issued by it the other peers had answered, i.e., the experiment environment displayed the value of α to him.

$$\alpha = \frac{number\ of\ query\ results\ received}{number\ of\ own\ queries\ issued}$$

Hence, while a participant knew which fraction of his queries was answered, the experiment environment did not reveal which participant he had actually forwarded a particular query to. I.e., α is the fraction of query results received from <u>any</u> participant in the system from the perspective of the participant under consideration. In other words, there is one value of α, not different ones for the various peer the current peer interacts with. While this setup is artificial to some degree, we investigated it in order to analyze the effect of feedback in isolation, i.e., no own observations. Our second treatment, i.e., the *feedback treatment*, was similar to the trust treatment, except that a participant had to attach one feedback object whenever issuing, forwarding or answering a query. The participant chose the participant the feedback referred to and decided if the feedback object was positive or negative. A participant saw an aggregate of all feedback objects it had received in prior rounds, for every contact. – In order to rule out order effects, we changed the order of the two treatments from group to group.

Another characteristic of the setup was that there always was one uncooperative player: He only issued queries and never answered or forwarded any query on behalf of others. Having such a participant is necessary to analyze the impact of punishment: In preliminary experiments, we did not have such a free rider, and all participants cooperated. The effects of sanctioning would not have been observable.

After the treatments we conducted a strategy game [29]. Strategy games are common to identify strategies of human players in experiments. In a strategy game participants are asked to describe in natural language the strategies they used in the game itself. Strategy games are typically played after the participants have taken part in several treatments. While participants learn during the treatments and refine their strategies, they tend to have a thorough understanding afterwards.

In strategy games the participants formulate their strategies based on the payoff structure. The strategies may also depend on treatment parameters and on the history of the treatment. The participants face the same tasks in the strategy game as during the treatments, but now at an abstract level. In other words, the difference is that the participants have to specify their decision for all situations they might encounter, while they only react to one situation at a time in the 'regular' game. From a game-theoretic point of view, the strategy game reveals a complete strategy.

After the treatments and the strategy game, all participants were paid depending on their success in the treatment: Their monetary reward corresponded to the points earned during the treatment. 100 points corresponded to 2.00 €, the average payoff per participant was € 13.18.

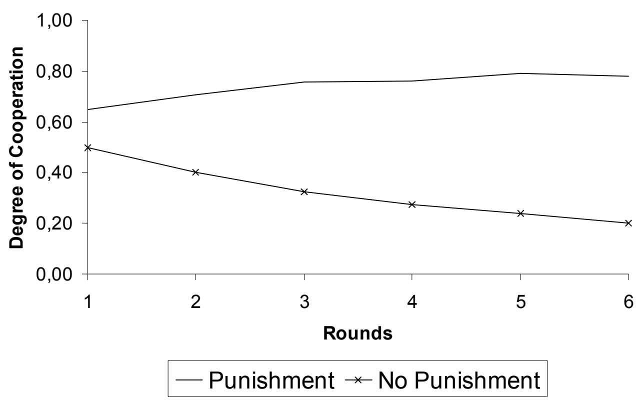
6. RESULTS

Figure 2 compares different mechanisms to increase the degree of cooperation, including ours. We look at the evolvement of cooperation over time. For our experiments, we analyzed the behavior of all participants except for the free rider. He had to follow our specific instructions; his behavior therefore is not interesting. We then calculated the degree of cooperation by dividing the number of queries processed per round by the number of queries received per round.

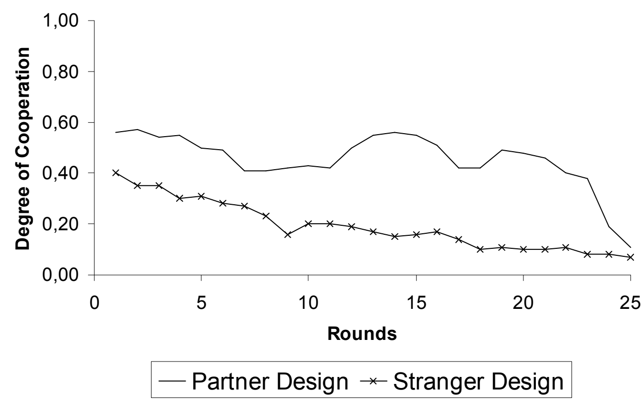
Figure 2a shows the results of the experiments conducted by Fehr and Gächter [9]. The degree of cooperation is the money invested divided by the maximum amount that could be invested. The degree of cooperation decreases over time in the absence of a punishment mechanism. If punishment is possible, the degree of cooperation is higher and does not decrease in later rounds. The results of Keser and van Winden [15] in Figure 2b are similar: The degree of cooperation in treatments of strangers decreases over time. The degree of cooperation is higher in a treatment where only partners interact. Further, it does not decrease over time. Similar results are obtained by Seinen and Schram [23] (see Figure 2c): They compare the behavior of participants who could exchange feedback or not. Having feedback increases the degree of cooperation.

When comparing these results to ours (see Figure 2d), two observations are interesting: First, having feedback does not influence the degree of cooperation. Second, the degree of cooperation for the feedback as well as for the trust treatment is roughly as high as the degree of cooperation with any of the other mechanisms in Figure 2.

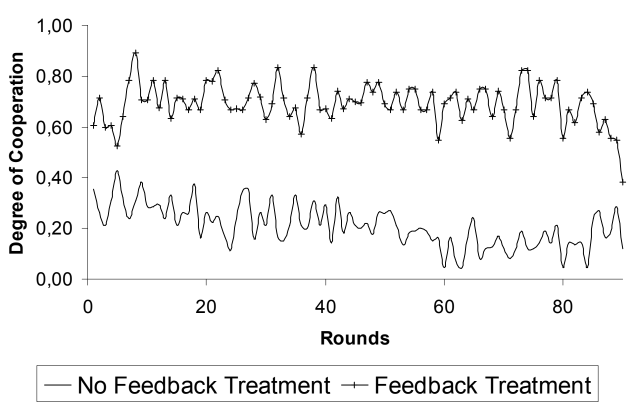
There even is the same level of cooperation in the presence of a free rider as with the other mechanisms without a free rider. This shows that the level of cooperation in structured P2P systems is even higher as in the partner scenario or in a stranger scenario with feedback, and such systems are robust against free-riders (about one out



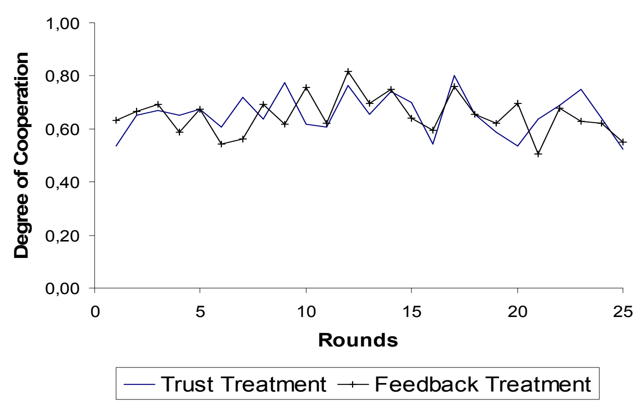
a) Treatments conducted by Fehr and Gächter [9]



b) Treatments conducted by Keser and van Winden [15]



c) Treatments conducted by Seinen and Schram [23]



d) Our Treatments

Figure 2. Evolvement of degree of cooperation for different experiments

of six peers). The high level of cooperation results from the underlying structure: The participants have the choice with whom to interact directly. A participant can always interact with another participant which it deems cooperative. As soon as participants started to behave selfishly, they were excluded from the system.

Except for the impression that our results are similar to the other results shown, any further analyses based on the results of the treatments are difficult. Namely, various explanations for the different phenomena, such as the degree of cooperation or the development over time, are conceivable, and the sheer numbers do not tell us which ones apply. Further, our experiments are not identical to the other experiments: Our scenario is different, as is the notion of rounds. Thus, in what follows, we try to rely on the strategy game for our analyses whenever possible when trying to verify our hypotheses.

6.1 Hypothesis "Indirect Punishment"

We expect that participants forward messages in a structured P2P system based on the feedback collected if no other information is available (Hypothesis "Indirect Punishment", (a)). To verify the hypothesis, we analyzed the strategy game: We showed the participants patterns of negative and positive feedback objects on a peer. Then we asked the participants for their strategy, i.e., if they would forward a query on behalf of this peer, given that they had no other information on it.

Table 1 shows the results of the strategy game regarding feedback. We have assigned the strategies to one of two groups:

- Strategies ignoring feedback (Group a))
- Strategies using feedback (Group b))

To test whether the experiments confirm the hypothesis that participants will use feedback if only feedback is available, we introduced the Null hypothesis, i.e., participants do not punish based on feedback. Of the 70 persons participating in the strategy game, 63 participants (all participants of Group b)) used feedback. Hence, we reject the Null hypothesis on a significance level of 1%. This confirms Hypothesis "Indirect Punishment".

Table 1. Feedback strategies observed in strategy game

Strategy	**Category**	**# persons**
Always cooperate	a)	3
Always defect	a)	2
Random strategies	a)	2
Sum category a)		*7*
Cooperation if more positive than negative feedback is received	b)	43
Cooperation if at least one positive feedback item existed	b)	4
Cooperation if less than n negative feedback items existed	b)	16
Sum category b)		*63*
No answer given		2

If no feedback is available, we also expect punishment to depend on the number of queries answered (Hypothesis "Indirect Punishment", (b)). We expect so-called cut-off behavior, similar to [22]. When using cut-off strategies, participants forward or answer queries if they received them from participants who answered more than a certain fraction of their queries in turn. In [22], we observed 91,7% cut-off strategies. Given this result from [22], a test analogous to the feedback treatment revealed significance on the 1% level. I.e., we confirmed that players use cut-off strategies for feedback. In other words, we expect indirect punishment to be used in systems of indirect partner interaction.

6.2 Hypothesis "Feedback-free Strategies"

To analyze whether our experiment is in line with the hypothesis that participants use feedback, we continued analyzing the strategy game. We asked the participants which strategies they played during the experiment, with regard to three aspects:

- How they chose the recipient when issuing queries.
- For which peers they answered queries.
- For which peers they forwarded queries.

Table 2 contains the respective numbers.

Table 2. Description of game

Strategy	# persons
Sending queries:	
Feedback independent strategies	45
Feedback dependent strategies	14
Answering queries:	
Feedback independent strategies	39
Feedback dependent strategies	20
Forwarding queries:	
Feedback independent strategies	41
Feedback dependent strategies	18

To confirm our hypothesis, we formulated three different Null hypotheses: 1. Participants choose the addressee based on feedback when sending queries. 2. Participants answer queries for other peers based on feedback. 3. Participants forward queries based on feedback. We used a binomial test to validate that the strategies used do not depend on feedback. Based on this test, we can reject all three hypotheses on a significance level of 1%. This confirms Hypothesis "Feedback-free Strategies".

Hence, while feedback is beneficial in the absence of any other information, participants do not use it as soon as other information, such as own observations, becomes available.

6.3 Hypothesis "Feedback Has No Impact on Payoff"

Finally we want to verify the hypothesis that feedback does not lead to a significant increase of the payoff of the participants. Therefore we analyzed the payoff sum of cooperative participants from the fifth to the twenty-fifth round. We did not take the other rounds into account to avoid startup and end-game effects [25]. Table 3 shows the points the participants earned in one round on average.

We derived the Null hypothesis: The payoff of the feedback treatment is higher than the one of the trust treatment. For 8 out of 12 of the groups, the payoff of the feedback treatment was lower, compared to the trust treatment. Thus, we reject the Null hypothesis at a significance level of 1% using a binomial test. Hence, the hypothesis that feedback has no influence on the payoff is confirmed.

The payoff sum over all treatments is 30.54 in the trust treatment compared to 28.36 in the feedback treatment. This indicates that performance in the trust treatment without feedback is better.

Table 3. Payoff during our treatments

Group	Trust Treatment	Feedback Treatment
Group 1	3,32	3,99
Group 2	1,21	3,21
Group 3	-0,06	-0,45
Group 4	2,34	1,42
Group 5	3,74	3,53
Group 6	4,53	1,51
Group 7	-0,06	-0,66
Group 8	-0,55	-0,81
Group 9	4,70	3,06
Group 10	0,24	1,82
Group 11	1,39	1,46
Group 12	3,81	2,22

7. DISCUSSION

In what follows, we discuss three interesting points, namely whitewashing, the design of a feedback repository, and the size of the structured P2P system. We then discuss the generality of our results.

Whitewashing. According to van Winden, participants who perceive their counterpart in an interaction as a partner are more likely to cooperate, in particular in the beginning of the interaction [15]. This would be problematic when it comes to whitewashing [8], i.e., a participant enters the system, issues queries until he is revealed as an attacker, quits the system, rejoins under another identity etc. Using feedback might mitigate this effect: By issuing honest feedback such negative behavior might be identified much faster. However, feedback is not necessarily honest.

We do not expect whitewashing behavior to occur in our setting: [22] has shown that participants in structured P2P systems play cutoff strategies: Participant A only processes queries on behalf of Participant B if B in turn has handled a certain fraction of A's queries. A participant who does not process queries on behalf of others is identified after a few rounds. His queries are not forwarded anymore. In this way, peers joining the system first have to prove their willingness to contribute by processing messages: Our experiments show that 'established' peers are not willing to handle any of their queries otherwise. This means that whitewashing is unattractive. From a slightly different perspective, even in the presence of new peers joining the system, feedback is not needed. This insight has another positive consequence: In a system without feedback, strategic manipulations of feedback are not possible.

Even if a whitewashing peer could successfully issue one query after joining the system and would earn some money during the first few rounds, we can ignore this phenomenon, for two reasons: First, if 1/6 of the participants is uncooperative the system does not collapse [22]. In other words, a structured P2P system would even tolerate a relatively large fraction of uncooperative peers. Second, the payoff of a peer decreases if it switches from cooperative to uncooperative behavior [22]. In other words, a rational participant behaves cooperatively.

Global vs. local shared history. In our experiments, we showed that participants do not rely on local shared histories of feedback in most situations. From the perspective of a participant, such feedback should be more trustworthy than global shared histories: The participant only collects feedback information issued by his contacts, i.e., participants he knows relatively well. As the participants do not even rely on this kind of feedback, we do not expect them to rely on global shared histories in the first place where feedback is likely to be truthful to a lower degree.

System size. The size of the structured P2P system used in the experiments does not have any influence on our results: In structured P2P systems of any size, a peer interacts with two kinds of peers, its neighbors and distant peers. Since there is no central node within such a system, each participant can only rely on the information (own observations and feedback) it collects itself. If a participant realized that he was part of a system of 6 peers, this would not influence his behavior: He still would not have any other information on the contacts, and he could only trust the forwarders of the message. Of course, having a bigger network might result in different absolute numbers, but the strategy game has not yielded any indication that this would affect the behavior of participants.

Another issue is that other structured P2P systems tend to need fewer hops to route a query through the network than CAN: For instance, the number of hops with Chord is logarithmic in the number of participants [27], compared to a two-dimensional CAN where it grows with the root of the number of participants [19]. We used CAN as a basis for our experiments because an implementation of local shared histories is available for it [3]. We expect our findings to hold for any other structured P2P system as well: The participants in the experiments could not even see which kind of structured P2P system had been in use! Hence, even if the number of the peers had an influence on peer behavior in qualitative terms (which we do not have any indication for), this would not be a problem: From a practical perspective, the number of hops is more or less bounded, irrespective of the number of peers.

Generality of Results. Our findings carry over to other systems as well, namely any system where autonomous nodes interact using distant communication channels, such as certain variants of the sensor networks or of the internet. Such scenarios could be organized similarly to ours: The nodes could interact with few other nodes repeatedly, and there is a global scheme specifying these relationships. This would give way to a system of indirect partner interaction, motivating participants to show a higher degree of cooperation compared to systems without such a structure. Our results also have some relationship to social networks (even though further specific experiments will be necessary). Humans often are part of structures, and they cannot freely choose their position in these structures. Think of your relatives, neighbors, colleagues and the teachers of your children which you typically cannot choose yourself. (Clearly, social networks are not a perfect example of systems of indirect partner interaction – one can choose friends and business partners, to give other examples. Further, not all social networks have a given underlying structure, e.g., the set of all ebay participants and the relationships among them. But we think that assuming that any social network does not have any structure a priori is not realistic. A thorough investigation which social-network phenomena indirect partner interaction can explain is beyond the scope of this article.)

8. CONCLUSIONS

Structured P2P systems are interesting in the sense that there are two ways of establishing cooperation among peers, and it is unclear how important they are. First, the fact that neighboring peers interact repeatedly may foster cooperation. Second, indirect reciprocity may occur between distant peers which are unlikely to interact again in the future. Feedback can establish this. In this paper we analyzed whether indirect reciprocity is beneficial if the peers can also establish cooperation in indirect partner interaction. In other words, we analyzed whether participants establish cooperative behavior in systems with interaction via a sequence of partner relations. We showed that feedback is an accepted mechanism in systems of human interaction if there is only indirect reciprocity. Participants do not rely on feedback any more if the participants interact with unknown peers via a sequence of partners. Under such circumstances, feedback does not have a positive effect on the behavior of the participants.

We found this result surprising: Based on the literature such as [16], we had expected structured P2P systems as systems characterized by the stranger design and indirect reciprocity. In such a scenario, feedback would be beneficial. But we have shown that designing the structure of the interaction may make feedback mechanisms unnecessary. We also obtain higher levels of cooperation than in a stranger scenario with feedback and without structure, and the system is robust against free riders.

We observe behavior of participants which is guided by trust and reciprocity. The success of such behavior is not influenced by leaving and joining peers: The strategies observed support cooperative peers and isolate uncooperative ones. Since this behavior relies on own observations, participants cannot issue feedback strategically to mislead other peers.

9. ACKNOWLEDGEMENTS

We thank the anonymous reviewers, the members of the database group at Universität Karlsruhe (TH) and the members of the empirical economics group at Otto-von-Guericke Universität Magdeburg for their valuable comments. In particular, we thank Rainer Schmidt for much help with the experiments.

10. REFERENCES

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

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

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

[4] Bolten, G.; Katok, E. and Ockenfels, A. Trust among Internet Traders: A Behavioral Economics Approach. In: *Analyse und Kritik*. 26 (1), 2004, 185-202.

[5] Bolton, G.; Katok, E. & Ockenfels, *A. How Effective are Online Reputation Mechanisms? - An Experimental Investigation.* Working Paper, 2002.

[6] Buchmann, E.; Böhm, K. and von der Weth, C.: Towards Truthful Feedback in P2P Data Structures. In: *Proceedings of the 14th International Conference on Cooperative Information Systems*. 2006.

[7] Feldman, M.; Lai, K.; Stoica, I. and Chuang, J. Robust Incentive Techniques for Peer-to-Peer Networks. In: *Proceedings of the 5th ACM Conference on Electronic Commerce*. 2004, 102-111.

[8] Feldman, M.; Papadimitriou, C.; Chuang, J. and Stoica, I.: Free-Riding and Whitewashing in Peer-to-Peer Systems. In: *Proceedings of the ACM SIGCOMM Workshop on Practice and Theory of Incentives in Networked Systems*. 2004, 228-236.

[9] Fehr, E. and Gächter, S. Altruistic punishment in humans. In: *Nature*. 415, 2002, 137-140.

[10] Fehr, E. and Gächter, S. *Do Incentive Contracts Crowd out Voluntary Cooperation?* Working Paper, 2001.

[11] Hagen, E. H. and Hammerstein, P.: Game theory and human evolution: A critique of some recent interpretations of experimental games. In: *Theoretical Population Biology*. 69, 2006, 339–348.

[12] Han, J. and Liu, Y.: Dubious feedback: fair or not? In: *Proceedings of the 1st International Conference on Scalable information Systems.* 2006.

[13] Jurca, R. and Faltings, B.: "CONFESS". Eliciting Honest Feedback without Independent Verification Authorities. In: *Proceedings of the Agent-Mediated Electronic Commerce*. 2004.

[14] Jurca, R. and Faltings, B.: An Incentive Compatible Reputation Mechanism. In: *Proceedings of the IEEE International Conference on E-Commerce*. 2003, 285-292.

[15] Keser, C. and van Winden, F.: Conditional Cooperation and Voluntary Contributions to Public Goods. In: *Scandinavian Journal of Economics*. 102, 2000, 23–39.

[16] Marti, S. and Garcia-Molina, H.: Limited Reputation Sharing in P2P systems. In: *Proceedings of the 5th ACM Conference on Electronic Commerce*. 2004.

[17] Miller, N.; Resnick, P. and Zeckhauser, R.: Eliciting Honest Feedback: The Peer Prediction Method. In: *Management Science*. 51(9), 2005, 1359-1373.

[18] 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.

[19] Ratnasamy, S.; Francis, P.; Handley, M.; Karp, R. and Schenker, S.: A Scalable Content-Addressable Network. In: *Proceedings of the 2001 Conference on Applications, Technologies, Architectures and Protocols for Computer Communications*. 2001, 161-172.

[20] Rowstron, A. and Druschel, P.: Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems. In: *Proceedings of the IFIP/ACM International Conference on Distributed Systems Platforms (Middleware)*. 2001, 329-350.

[21] Saroiu, S.; Krishna Gummadi, P. and Gribble, S. D. A Measurement Study of Peer-to-Peer File Sharing Systems. In: *Proceedings of Multimedia Computing and Networking*. 2002.

[22] Schosser, S.; Böhm, K.; Schmidt, R. and Vogt, B.: Incentives Engineering for Structured P2P Systems - a Feasibility Demonstration Using Economic Experiments. In: *Proceedings of the ACM Conference on Electronic Commerce*. 2006.

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

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

[25] Selten, R. and Stöcker, R.: End Behavior in Sequences of Finite Prisoner's Dilemma Supergames. In: *Journal of Economic Behavior and Organization*. 7 (1), 1986, 47-70.

[26] Srivatsa, M.; Xiong, L. and Liu, L.: TrustGuard: countering vulnerabilities in reputation management for decentralized overlay networks. In: *Proceedings of the 14th International Conference on World Wide Web*. 2005, 422-431.

[27] Stoica, I.; Morris, R.; Karger, D.; Kaashoek, M. F. and Balakrishnan, H.: Chord: a scalable peer-to-peer lookup protocol for internet applications. In: *IEEE/ACM Transactions on Networking*. 11, 2003, 17-32.