Online vs. Offline Behavior: How to Design Strategic Agents for Distributed Coordinator-Free Environments

Stephan Schosser
Universität Karlsruhe (TH)
schosser@gmx.de

Klemens Böhm
Universität Karlsruhe (TH)
boehm@ipd.uka.de

Bodo Vogt
University of Magdeburg
bodo.vogt@ww.uni-magdeburg.de

Abstract

Highly distributed coordinator-free systems with many agents, such as peer-to-peer systems, are receiving much interest in research and practice. In such systems, free riding continues to be a severe and difficult problem. To reach a high degree of cooperation, researchers have proposed numerous incentive mechanisms against free riding. Our concern in turn is a more profound question: Under which circumstances do human individuals indeed resort to effective strategies in those settings? This question is important as humans control the agents. In economics, mimicking the behavior of humans is an accepted method for strategy design. In particular, this holds for complex settings where formal analyses must rely on simplifying assumptions that do not hold in reality. By means of extensive human experiments in one specific setup, namely structured P2P systems, this paper provides evidence that strategies in the settings under investigation often are the result of a non-strategic perspective. This perspective lets participants overlook obvious strategies that are effective. Further, our experiments reveal the following: Online players, i.e., individuals taking part in the system directly, intuitively tend to find better strategies than offline players, i.e., individuals who just implement agents. Offline players have difficulties predicting the strategies of others and overestimate the quality of their strategies. We conclude that a combination of 'online' and 'offline' strategy design is a cost-efficient and effective solution.

1. Introduction

In research and practice, distributed coordinator-free systems with many agents are receiving much interest, be they economic systems such as social networks, be they technical systems such as peer-to-peer systems (P2P systems). So far, such systems tend to rely on voluntary cooperation. But free riding still is a severe and difficult problem. Incentive mechanisms against it must be designed properly to foster cooperation. Game theory is often used to this end. It sees the participants as utility maximizers who are rational. It identifies the strategies one can choose, the strategy space, and calculates their utility, given the strategies of others, and the equilibria. Mechanism design has a slightly different objective, namely modifying the participant utilities so that rational participants use strategies leading to an efficient outcome. However, it is not enough that the system designer knows the good strategies – participants must also know and actually use them. For instance, [26] has shown that various incentive mechanisms in P2P file sharing systems do not foster cooperation as expected. Free riding is almost as common as without any mechanism. This is symptomatic of one central problem when designing such mechanisms. Having designed these, do participants indeed resort to the strategies expected? One result from behavioral economics [5] is that humans, although capable of finding good strategies in certain situations they are familiar with, have difficulties to determine strategies that are optimal under rationality. The authors observe that strategies depend on how they are designed.

In open distributed systems where different organizations and individuals (referred to as participants in this paper) control the agents, each participant can choose the code that defines the behavior of his agents. Hence, it is important to study if there are any patterns/regu-larities in the perception of the participants that are in the way of choosing strategies that are optimal under rationality, and, if so, how they can be avoided. This paper takes a first stab at this difficult problem.

We do so by choosing one specific object of investigation, namely Content-Addressable Networks (CAN) [17], a prominent variant of structured P2P systems [2], and by analyzing different approaches to strategy design. In CAN, strategies exist which lead to an efficient equilibrium – a first contribution of this paper is to show that this is the case using game theory and to identify these strategies. Using behavioral experiments, we then analyze whether participants find and use these strategies for 'their' agents or not. To do so, using that game-theoretic optimum as a reference point, we compare two approaches: (a) Individuals who have not been exposed to the problem domain before come up with their strategies by playing (online strategy design). (b) Participants first spend several weeks implementing various variants of CAN agents. Having gained this experience, we ask them to implement a strategy that successfully cooperates with other agents (offline strategy design). – This analysis is complete (in the economic sense): We use all existing methods, i.e., game theory, online strategy design and offline strategy design, to analyze the problem.

Carrying out such a comparison is not straightforward. Both approaches are expensive and time-consuming. Further, to ensure comparability, the general conditions need to be as equal as possible. This is extremely difficult: Small changes in the design can affect the results observed significantly, as we will explain.

We show that both groups of participants resort to strategies with features that are known to curb free riding, namely reciprocity and cut-off strategies [12]. (Here, reciprocity means that participants cooperate with others who have cooperated in the past and defect otherwise. With cut-off strategies, Participant A distinguishes between cooperativeness and uncooperativeness of Participant B by observing B's share of cooperative actions in the past. If it is above a certain threshold, A cooperates with B, otherwise not.) However, both groups differ in one central aspect. Online players intuitively find strategies leading to systems that value cooperation and therefore reach the efficient equilibrium. Offline players in turn tend to be good in designing strategies which react successfully to known attacks. But they have difficulties foreseeing the consequences of their strategies and predicting the strategies of others. In turn, their strategies are less cooperative, and the resulting systems are less stable. We found this result surprising. Offline players have spent several weeks studying the system, and we had expected them to have a deeper insight than online players. Further, we show that this result is due to the perspective of developers on strategic problems. Their mechanism development focuses on certain attacks and countermeasures against these attacks. The economic advantages or disadvantages of the proposed strategies often receive less attention. We conclude that successful strategies for distributed coordinator-free system are reached by a combination of 'online' and 'offline' strategy design in a cost-efficient way. Our experiences should be of interest to anyone intending to design agent strategies in realistic setups.

Paper outline: Section 2 reviews related work. In Section 3 we introduce the basics of structured P2P systems and derive the equilibria in such systems. We describe our experiments in Section 4 and the evaluation in Section 5. We discuss our results in Section 6. Section 7 concludes.

2. Related Work

[14] introduced the principle of mechanism design. The system designer creates incentives so that the optimal outcome is reached if all participants behave rationally. [25] proposed to apply mechanism design to computer-science problems. As agents follow a strategy given by a designer, their behavior is always rational. Thus, computer-science problems promise to be a good application for mechanism design. [7] identifies two tasks when designing multi-agent systems: First, the designer has to specify a protocol which covers the range of actions of the participants and defines how the action choice translates into the outcome of the system. Second, each strategy of an agent has to be specified. The authors propose mechanism design to analyze such systems. As deriving the optimal strategy is computationally complex, they propose computational mechanism design as a solution. Here, a central instance derives the optimal strategy choices, given the preferences of the participants, and chooses the corresponding strategy allocation. While this works well in systems with a central coordinator, in open, distributed systems the agents themselves have to accomplish this task.

[6] proposes to let the participants define their optimal strategy. A central authority only decides which allocation to choose, given the strategies chosen. [16] introduces the *price of anarchy*. It defines the difference between the efficiency reached if the central chooses the strategy leading to a system-wide optimum and the efficiency if the participants choose the best strategy.

[10] proposes distributed algorithmic mechanism design where the participants calculate the efficient equilibrium. They do not analyze whether agents have an incentive to use strategies leading to this equilibrium. To overcome this problem, [15] proposes three principles: First, computation needs to be carefully distributed across agents. Second, agents should only perform computations necessary to reveal their information. Third, several agents must conduct the computation, and deviations have to be punished. To follow these principles, they suggest incentives for the participants. Nevertheless, they still need a central instance to enforce the principles. Whether the participants can enforce the optimal strategies and follow their optimal strategy remains an open question.

Empirical analyses of P2P systems give an insight into this problem. Existing P2P systems are analyzed regarding the behavior of the participants. Analyses of *unstructured* P2P file sharing systems using this method show that a majority of the agents tends to free ride if there is no incentive mechanism [1][19]. [26] analyzes such systems with incentive mechanisms. The authors show that while the mechanisms result in an increased participation of the agents, their contributions are only as high as needed to benefit from the system. However, as there currently is no structured P2P system that is operational in the real world, such empirical analyses are not feasible.

Another approach is behavioral experiments. Human participants assume the role of agents. Using such experiments, we have shown that inexperienced participants play cut-off strategies [20] and do not use feedback if they can rely on information they have gathered themselves [21]. While this work analyses the behavior of participants in structured P2P systems, it does not address the question how humans in such settings actually arrive at their strategies. It does not study the offline approach and its implications on the strategies.

[5] analyzes the ability of humans to determine behavior that is optimal under rationality. They show that, while humans find optimal solutions for specific interaction problems, they find it difficult to find rationally optimal strategies in unfamiliar scenarios. Given that the strategy of each agent is designed by a human, this raises the question whether humans use those optimal strategies, given a well-designed mechanism.

Behavioral economists have observed that the strategies of individuals in cooperation situations depend on *how* the individuals have designed their strategies. Economists distinguish between strategies the participants played against others in behavioral experiments (hot game) and strategies specified without actually playing against other participants (cold game). [13] reviews various behavioral experiments in this context. They observe that strategies differ between hot and cold games and justify these differences by differences in the perception of the behavior of the counterpart. While in the hot game cooperation of the counterpart is seen as a signal for cooperativeness, such signals may not be perceived as a cooperative act in the cold game. [3] shows that participants tend to show fair behavior more often if they have to decide on each interaction, compared to specifying a general strategy. According to [4], those differences are the consequence of two effects: The positive self-image effect is that people view themselves favorably, e.g., they see themselves as generous, fair and cooperative. The consensus effect is that participants resort to behavior they would expect from themselves when predicting the behavior of others. Offline strategy design is similar to cold games to some degree. In both cases the participant has to describe a strategy and anticipate the behavior of others. Nevertheless, participants using offline strategy design are much more experienced than participants in cold games. They spend several weeks becoming familiar with the problem domain and the system. Further, those comparisons of hot and cold games have used simpler setups than ours. All this means that we cannot infer that offline strategy design has the same effects as cold games.

[22] is a preliminary shorter version of this paper, which introduces the notions of online and offline behavior. In contrast to this current paper, it does not examine which strategies result with the two approaches, and how they differ.

3. Structured P2P Systems

In this section, we give a short introduction to structured P2P systems. We show formally that mechanisms in these systems exist which lead to efficient equilibria.

3.1. CAN Fundamentals

Structured P2P systems [17][18] manage large sets of (key, value)-pairs. We focus on Content-Addressable-Networks (CAN) [17] as an example for such systems. A function which assigns every key a location in the key space is given. We assume that all peers know this function. The key space is partitioned into zones. Each agent manages one zone, i.e., all data objects whose key maps to the zone. In addition, it has a list of all agents with adjacent zones, its neighbors, and their zones. The system links each agent to his neighbors. A query is a request for the value belonging to a key. When processing a query, the key is transformed to its location in the key space, the query point. If the query point lies in the zone of the current agent, it returns the (key, value)-pair sought. Otherwise, it forwards the query via one link to the neighbor closest to the query point. This step is repeated until the agent having the query result is reached.

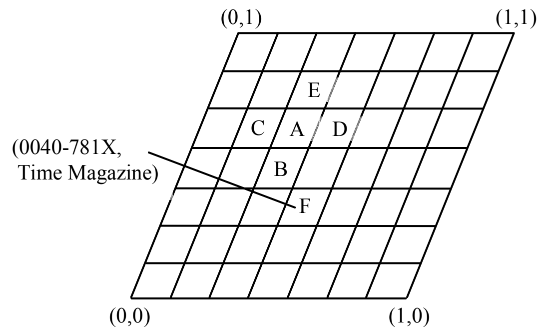

Figure 1. Content-Addressable Network

Example: Figure 1 shows a CAN. In this example, a hash function maps all keys to two-dimensional query points. For example, it maps the key of the pair (0040-781X, "Time Magazine") to (0.45, 0.3). Each rectangle is a zone, i.e., Agent F administers the information belonging to key ISSN "0040-781X". If another agent, e.g., Agent A, is interested in information on ISSN "0040-781X", it forwards this query to a neighbor, i.e., Agents B, C, D or E. To do so

, the agent first calculates the query point corresponding to the ISSN, i.e., (0.45, 0.3). It then forwards the query to the neighbor closest to the point, i.e., B. Note that A does not know F, only the position of the query point. As B does not know the query result, it forwards the query to one of its neighbors, until the query reaches F. F returns the query result to Agent A. ▪

We assume that receiving query results is beneficial, while work results in negative utility. Issuing a query for instance costs 2 points in our experiments, forwarding 1 point and answering 5 points. If a query result arrives at the initial issuer, it receives 20 points. (At the beginning of each experiment, a participant has 100 points.) For the issuer, receiving query results is a big benefit, while issuing and forwarding incurs small costs. The costs for answering are higher than the ones of forwarding. This is because answering does not only mean sending a message, but also searching for the query result. The average sum of the costs of forwarding, answering and issuing a query needs to be smaller than the benefit of receiving a query result. Otherwise, an agent would not have any interest to participate.

3.2. Formal Model

When modeling distributed systems from an economic perspective, we are interested in the utility of agents. Depending on it, a participant chooses the strategy. To reflect this, we assign a value to every action of an agent according to the criteria from Section 3.1, depending on the cost and benefit structure. Given this, we derive a value which is the sum of all costs and benefits of a participant when choosing a certain strategy. Using a monotonically increasing function, we can map this value to a utility value.

We refer to the set of agents in our system as $I = \{1,...,n\}$. The action space of each agent consists of three possible actions: issuing a query (Action S), forwarding a query (Action F), and answering a query (Action A). From the perspective of Agent i, sending a query has value $v_i(S)$, answering a query has value $v_i(-A)$, and forwarding a query has value $v_i(F)$, with $v_i(A) < v_i(S) < v_i(F) < 0$. (The values of Actions A, S and F are always negative.) To facilitate the analysis, we make the following assumptions:

- *Actions take place in rounds*. A participant decides once per round how to process the incoming queries.
- The value of a query result received is $v_i(R) > 0$.
- *We do not distinguish between forwarding and answering queries.* Instead, participants process queries (Action P). Further, we do not distinguish between queries, i.e., participants process all incoming queries or none.
- *We also neglect the stochastic aspects of the P2P* structure. Every agent must process the same number of queries in each round. $v_i(P) < 0$ is the value of processing all queries received per round.
- *An agent processes a query in the round after it received it (if it chooses to process it)*. Each agent has a belief μ_i. It is the probability, according to the expectation of the agent, that a query it has issued is answered eventually. μ_i does not change in equilibrium.
- Further, obtaining the query result later reduces the utility by the discounting factor ß per round. E.g., if receiving the query result in this round has value w, receiving it one round later has value w·ß.

Agent i can now choose from one of four strategies in each round: First, it can leave the system (s^L) by neither issuing nor processing queries. Second, it can behave altruistically (s^A) by only processing and not issuing queries. Third, it can free-ride (s^F) by only issuing and not processing queries. Fourth, it can fully cooperate (s^C) by processing and issuing queries. For the time being, we limit the analysis to pure strategies. The combination of all four strategies defines the strategy space of agent $S_i = \{ s^L, s^A, s^F, s^C \}$. The value of a strategy for an agent is as follows:

- Strategy Leave: $v_i(s^L) = 0$
- Strategy Altruism: $v_i(s^A) = v_i(P)$
- Strategy Free-Ride: $v_i(s^F) = \mu_i ß\, v_i(R)+v_i(S)$
- Strategy Cooperate: $v_i(s^C) = \mu_i ß\, v_i(R)+v_i(S)+v_i(P)$

Query results are only discounted once. This is because our analysis focuses on infinitely repeated games (see Subsection 3.2.1): In equilibrium, participants play the same strategy every round. Hence, we can expect a query result to arrive every round. $v_i(R)$ is not the benefit of receiving a query result for the query issued in this round, but of the query result received this round.

Obviously, the other pure strategies dominate Strategy Altruism. Its utility is negative. Therefore, we do not consider it any further. Given the remaining strategy space, each rational agent chooses the strategy which maximizes its value:

$$\max\{v_i(s^L), v_i(s^F), v_i(s^C)\}$$
$$= \max\{0, \mu_i\beta\, v_i(R)+v_i(S), \mu_i\beta\, v_i(R)+v_i(S)+v_i(P)\}$$

This value represents the value of one round. Given this formula, we derive equation v_i^t of the value in t:

$$v_i^t = \max\{0, \mu_i\beta\, v_i(R)+v_i(S), \mu_i\beta\, v_i(R)+v_i(S)+v_i(P)\} + \beta\, v_i^t$$

As agents can only process queries they received in the preceding rounds, and query results cannot be obtained in the round they are issued, one should resort to Strategy Leave in the last round. Using backward induction, it can be shown that this holds for all previous rounds. However, economists have observed that participants behave as if a game would last forever if it does not end in the next few rounds. We take this into account by analyzing a game lasting for infinitely many rounds $T = \{1, ...\}$.

3.2.1. Calculating the Equilibria

Based on equation v_i^t, we derive the stationary equilibria and characterize them [12]. A stationary equilibrium is a Nash Equilibrium with time-invariant choices of the agents [11]. An equilibrium is reached if each agent plays the best-response strategy, given the strategies of others. Each rational agent chooses the action which maximizes v_i^t, given β and μ_i. A stationary equilibrium consists of beliefs μ_i^* and strategy choice $s_i^* \in S_i$ (for $i = 1, ..., n$) such that:

- s_i^* is optimal given μ_i^*, and s_i^* fulfills v_i^t
- μ_i^* is consistent with all choices s_j^* for $j \neq i$

's_i^* fulfills v_i^t' means that s_i^* yields the highest payoff for Participant i. The statement that μ_i^* is consistent with all choices s_j^* means that agents form their beliefs based on the strategies observed. Thus, if all other participants play Leave, the belief of an agent will reflect this. In this case its belief will be close to 0.

Note that one equilibrium is reached if all agents follow Strategy Leave. Then the benefit of each participant and the system as a whole is 0. I.e., the equilibrium where nobody interacts is one possible outcome, for all possible parameter values. We determine the other equilibria based on the following considerations.

Each rational Agent i chooses the strategy which maximizes its utility v_i^t. At first sight, Strategy Free-Ride seems to dominate Strategy Cooperate since $v_i(P)$ is negative. But this is not true. It is not always possible to find μ_i^* consistent with Strategy Free-Ride. E.g., if all agents choose to free-ride, $\mu_i^* = 0$ holds for all $j \neq i$. In this case Strategy Leave dominates Free-Ride.

However, given $\mu_i^* = 1$ for all i, an equilibrium exists if all agents choose Strategy Cooperate. In equilibrium, two further inequalities must hold: First, each agent must have positive payoffs – otherwise it would leave the system. Second, deviation from the equilibrium strategy by one agent reduces its payoffs. The following two inequalities reflect this:

I: $\mu_i\, ß\, v_i(R) + v_i(S) + v_i(P) > 0$

II: $(1 - 1/n)\, \mu_i\, ß\, v_i(R) + v_i(S) + v_i(P) < 0$

Inequality II ensures that deviation by only one agent lets the system break down. If one participant resorts to free riding, this removes at least 1/n of the payoff of each cooperative participant on average: He will not answer queries any more. If the resulting payoff now is below 0, cooperative participants have an incentive to leave the system. It breaks down. Without the second equation, deviating from Strategy Cooperate would be beneficial. Free-riding would increase the payoff without increasing the costs.

Other constellations are that Agents $j = 1, \ldots, l$ choose to free-ride and Agents $k = l + 1, \ldots, n$ choose to cooperate. These are equilibria if the inequalities hold:

III: $\mu_j\, ß\, v_j(R)+v_j(S) > 0$ for $j = 1, \ldots, l$

IV: $\mu_k\, ß\, v_k(R)+v_k(S)+v_k(P) > 0$ for $k = l + 1, \ldots, n$

Further, all resulting μ_j and μ_k need to be consistent with the observation of the system. In equilibrium, agents do not have an incentive to deviate from their strategy: First, Agent j does not have an incentive to deviate from its strategy, as it would decrease its utility. Second, if Agent k deviated, μ_j and μ_k would not be consistent with the system any more. This reflects that, if one Agent k chooses to free-ride, others have an incentive to switch to Leave, as they would not expect to benefit from the system any more. Given Inequalitites III and IV, we derive conditions for the beliefs of the agents in the system. Agent *j* participates if

$\mu_j > v_j(S) / (ß\, v_j(R))$

holds. Agent k participates if the following holds:

$\mu_k > (v_k(S) + v_k(P)) / (ß\, v_k(R))$

As equilibrium is reached if all agents leave, participants either follow Strategy Leave, or Strategy Free-Ride and Strategy Cooperate respectively, if the corresponding inequality is met. Such strategies are called cut-off strategies: An agent participates if its belief is above its cut-off value, otherwise it leaves the system.

3.2.2. Equilibria for Indirect Partner Interaction

If agents only interact with a fraction of the other agents they are linked to and can identify them, they can play different strategies with them. Agents derive beliefs concerning each other Agent μ_i^j with $j \neq i$ instead of beliefs μ_i concerning the system as a whole. Agent *i* can then choose one different strategy for every Agent *j*. Agent *i* then plays one game with each of its neighbors. This game is referred to as binary game since only two agents interact. The equilibria derived in Section 3.2.1 remain the same, but μ_i is now replaced with μ_i^j. The equilibria for every binary relation remain the same. Agents use cut-off strategies in every binary relation, i.e., they behave reciprocally. The cut-off values of the two agents do not need to be equal, as their utility functions are different. Differentiating between agents has one advantage: Free-riding is not possible any more. In equilibrium the belief of the other agent has to be consistent with the behavior it perceives. If an agent believes that the other agent is not processing requests, it will choose Strategy Leave towards this agent. Then both agents will end up in Strategy Leave.

3.2.3. Efficiency

The number of free-riders determines the efficiency of the system. Many different combinations of free-riders and cooperative agents are possible, and different efficiencies between 0% and 100% can be obtained. As soon as agents can identify other agents (see Section 3.2.2) high efficiency is guaranteed: Free-riding is not beneficial any more, as agents can use reciprocal cut-off strategies to ensure cooperation. This situation does not change if one agent earns more than another one. Due to reciprocal cut-off strategies, no participant should deviate from cooperation. Structured P2P systems have the properties of indirect partner interaction. Hence, it is rational for all agents to play cut-off strategies as long as they can benefit from the system.

4. Experiment Design

Section 3.2 has shown formally that reciprocal cut-off strategies lead to efficient equilibria in structured P2P systems. We now analyze under which conditions humans (who control the strategies of the agents) find these strategies and actually use them. To this end, we conduct two kinds of experiments: *Online experiments* correspond to the online design process. In such experiments, each participant has immediate control over one agent, which interacts with agents controlled by other participants. Offline experiments in turn match the offline design process. Here, each participant is asked to implement the strategy of one agent. Implementations by different participants then interact with each other in a test bed that we have built ourselves.

In all experiments each participant specifies the strategy of one agent, either by implementing it a priori or by controlling it 'online' using a GUI. Here, 'strategy' means whether to forward, answer or issue a given query, and, if forwarded, which agent receives the query. The experiment environment on the other hand manages the data objects and the neighbors of each agent and controls the query processing (except for the strategic decisions just mentioned). E.g., when it comes to forwarding, it calculates the distance of each possible recipient to the query point. The experiment environment also assigns zones to participants. Besides that, it generates all queries. It does so randomly and distributes them equally over the key space.[1]

In all experiments, all agent zones have the same size. All agents have the same number of contacts and can expect to receive the same number of queries. All experiments are conducted in rounds. Each agent is allowed to issue one query per round, while it can answer or forward all messages it has received.

For each query the agent wants to issue or forward, the experiment environment shows the participant a list of potential other agents to receive the query and their distance to the query point. This distance defines the order of the list. I.e., the agent with the lowest distance is the first one in the list, the second closest the second one, etc. Hence, the order of the list represents the probability that one agent knows the query result. In this way, the agent/participant can weigh between the reputation and the distance of the other agents.

We do not consider any additional aspects, such as the possibility to add new links to additional contacts, as we do not expect such aspects to influence the strategies used. They would only distract the participants from the essential properties of the system.

Based on its behavior, the agent receives points. Contingent on these points, the participant is paid in the end of the experiment. Here, similar to real structured P2P systems, the agents only know which of their queries have been answered and whom they have sent them to initially. The experiment environment does not reveal whether the first agent in the forwarding chain or another one has dropped it. Each agent always knows which of its queries were answered in preceding rounds, and how much it has earned in preceding rounds. It does not know any other system properties.

After both online and offline experiments, we conduct so-called strategy games [24]. In such a game, participants are asked to describe their strategies in own words. More specifically, we confront them with various system states, on an abstract level. The participants have to specify the behavior of their agent in the given situation. The combination of several such situations lets us observe complete strategies. Although the strategies implemented in the offline experiment are already complete, such complete strategies are not observable in online experiments. The strategy game helps us when comparing both experiments. As the participants filled out the same form, the differences were easy to grasp.

We play one treatment using both methods (offline experiments and online experiments). In this treatment, one participant, the *distinguished agent*, receives 5 times the payoff the other agents earn, and all participants are informed about the ID of the distinguished agent. This helps to analyze the impact of bounded rationality on the development process: According to game theory the same behavior, namely full cooperation, is rational in both treatments (see Section 3.2). According to behavioral economics, humans tend to favor equality among all participants [8]. Hence, they might behave differently in this treatment to 'punish' the agent earning more.

4.1. Online Experiment

In the online experiments six participants are part of the system. A higher number of participants does not have any *qualitative* effect on the behavior, and humans tend to be unable to distinguish between more than five participants, according to Selten [24]. With our setup, no participant can figure out the assignment of agents to other participants. All terminals are separated from each other to prevent communication among participants. Participants are randomly seated in the beginning of the experiment. Each treatment lasts 20 rounds. Then a six-sided dice is rolled after each round. If it shows one, the game ends, otherwise it continues. By rolling the dice we try to reduce end-game behavior. All participants are paid contingent on the success of their strategy. They receive € 2.00 per 100 points. This payment corresponds to the utility a participant takes from participating in the P2P system in the real world (v_i^t in the formal model). This should lead to roughly the same payoffs for the participants in the online experiments as for the ones in the offline experiments.

4.2. Offline Experiment

We conduct the offline experiments with a group of computer-science students. These students participate in a laboratory course lasting one semester. Initially, fourteen students participated in the course, throughout the semester three of them left. Only eleven students took the course for the whole semester. In the beginning of the course, we introduced the participants to the theoretical foundations of structured P2P systems. To 'warm up', we asked them to implement a first strategy, namely an agent which receives a higher payoff than others when in a system with other student implementations. The timeframe for this assignment was two weeks. We then informed all participants about the success of each implementation. We determined 'success' by running a structured P2P system consisting of all student implementations. Each agent was randomly assigned to a zone. The points an agent earned in this run defined its success.

We repeated this throughout the course (5 times) and asked the participants to keep refining their strategies. In the first two iterations, we gave the participants all strategies other participants had implemented in the preceding iteration and one additional strategy to compare their new implementation to. In the first iteration, the additional strategy was one which always cooperated. In the second

[1] The description of the experiment software together with screenshots, log files, implementations of the participants, as well as the software itself can be downloaded from http://www.ipd.uni-karlsruhe.de/~schosser/iat08

one, the additional agent always showed free-riding behavior. We asked the participants to implement strategies which are at least as good as the additional strategy. The assignments of the next two iterations were similar in spirit, so that the participants could familiarize themselves further with the problem domain. The fifth iteration then was the 'interesting one' when we paid the participants contingent on performance, as described below. We provided the students with a simulation environment where different implementations can be plugged into. We also used this environment for our evaluation. It calculated the points each participant earned with his implementation. – According to a questionnaire at the end of the semester, a participant spent 6.67 hours on average per week working on his implementation.

To keep costs down, we conducted only one iteration with monetary incentives, and we only evaluated this iteration. We instructed the participants that the payoff is € 1.00 per 1000 points they would earn with their strategy. We also informed the participants that we would evaluate their strategy in a system of agents of other participants and strategies developed by us. We informed them that each treatment would last at least 1000 rounds, afterwards a six-sided dice would be thrown, similarly to the Online Experiment. If the dice showed 1 the experiment ended, otherwise it continued for another 50 rounds.

5. Experimental Results

We now address the similarities in online and offline strategy design before describing the differences.

5.1. Similar Strategy Properties

As described in Subsection 3.2 we expect the participants to use reciprocal cut-off strategies in both experiments. First, to analyze whether participants play cut-off strategies, we compare the strategy games of both experiments. We classified the strategies observed into three categories: cut-off strategies depending on the past success frequency of own queries (A), cut-off strategies with modifications (B), strategies that are not cut-off strategies (C). The values in Table 1 confirm that participants play cut-off strategies on a significance level of 1% for the online experiments and 15% for the offline experiments using a binomial test. 44 of 60 (eight of eleven) participants in the online (offline) experiments play reciprocal strategies. This confirms that participants behave reciprocally on a significance level of 1%, using a binomial test. As predicted by the formal model, the participants realized the potential of reciprocal cut-off strategies and resorted to them.

5.2. Differences between the Strategies

We now analyze the degree of cooperativeness shown in the experiments. To make online and offline experiments comparable, we simulated systems of six agents for twenty rounds using the implementations resulting from the offline experiments. We conducted simulations with every possible combination of the implementations we had received and repeated those 5 times using different positions for each agent. The average degree of cooperation in the online experiments is 82% (51% in the offline experiments). With other mechanisms, such as punishment (64%) [9] or feedback (69%) [23], similar degrees of cooperation are reached on average as in the online experiments. Systems without any cooperation-enhancing mechanisms lead to degrees of cooperation (40% - 50%) [23], comparable to the offline experiments. We conclude from this comparison that participants in offline experiments cooperate more than in settings with mechanisms known from literature, while they cooperate less in offline experiments.

Table 1. Strategies in Online vs. Offline Experiment

Strategy	*Cat.*	*# pers. (online)*	*# per. (offline)*
Cut-off strategy depending on past success frequency (CSPSF) only	A	35	5
CSPSF plus end phase or start phase	B	9	1
CSPSF plus limit for answering queries of others per round	B	0	1
Cut-Off strategy depending on the absolute number of own queries not answered	C	0	1
Unconditional cooperation	C	11	0
Unconditional cooperation plus condition that account is high	C	3	0
Unconditional cooperation plus limit for number of anwers	C	1	0
Different free-riding strategies	C	0	2
Different types of strategies	C	1	1
Σ		60	11

To further analyze the degree of cooperation, we carry out simulations using the strategies implemented by the offline participants. We conduct simulations lasting 1000 rounds with 100 agents. We first run two reference experiments, one with 100 cut-off strategies and one with 100 fully cooperative strategies. In both cases, the payoff per agent is around 8,500 points. Hence 8,500 is the number of points when all participants cooperate. Next, we simulate a system consisting of one strategy defined by a participant and 99 fully cooperative strategies. Figure 2 (white bars) shows the success of the participant strategies. They all do well. For instance, the Strategy of Participant 3, when playing against 99 fully cooperative agents, yields 17,699 points. Strategies 3 to 7 do not handle many messages. Hence, their costs in this setting are small, and their payoff is higher than 8,500 points. We then conduct simulations with 99 cut-off strategies and one participant strategy. For instance, the Strategy of Participant 3, when playing against 99 agents with cut-off behavior, yields -1,190 points. Here, five out of eleven participant strategies do not make any profit, for various reasons: Strategy 3 only handles 20% of the messages. It is a rather obvious free-riding strategy. Strategies 4 and 5 are cut-off strategies, but limit the number of queries handled per round to a certain value. Strategies 6 and 7 are other, rather short-sighted, free-riding variants. Next, we conduct experiments with systems consisting only of strategies of one student. Five out of eleven systems collapse, i.e., no agent processes message any more after about 200 rounds (Strategies 3 to 7). We find this result surprising. We have never observed any collapse in online experiments. Further, we had expected the participants to implement strategies which are at least robust against cut-off strategies and against themselves.

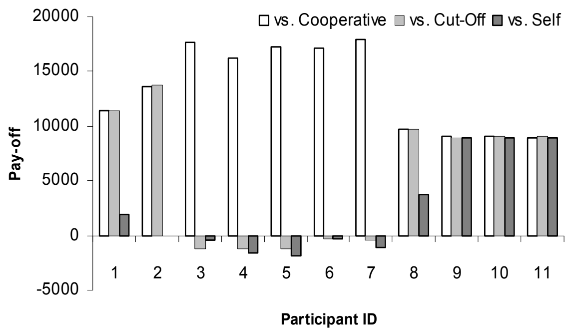

Figure 2. Simulation results using offline strategies

Next, to analyze the influence of heterogeneity, we look at the outcome of the strategy games again. Does the behavior change towards participants earning more? We first analyze the results of the online experiments. According to Table 2, 34 (37) participants answer (forward) queries of the distinguished agent as if they were received from any other agent. Four (five) participants even give preferential treatment to them. Only twelve (eight) participants are less cooperative towards the distinguished agent. The situation changes with the offline experiments. Here, six (four) participants out of eleven change their strategy towards the distinguished agent. No participant is more cooperative towards the distinguished agent. In this case, we cannot draw any conclusion concerning the influence of heterogeneity on the behavior. While participants in the online experiments intuitively realize that behavior towards the distinguished agent targeting at fair outcomes is not beneficial, offline participants do not see it. Two participants even justified free riding against the distinguished agent by claiming that it would always cooperate.

Table 2. Strategy Changes under Heterogeneity

	Economic		*Simulation*	
	Ans.	*For.*	*Ans.*	*For.*
Same strategy towards all	34	37	5	7
Queries of distinguished agent are handled preferred to others	4	5	0	0
(Some) queries of distinguished agent are not handled	12	8	6	4
Σ (without distinguished agent)	50	50	11	11

6. Discussion

We observed that developers tend to implement great reactive strategies based on the strategies they expect to observe. I.e., they develop strategies that can deal very well with attacks or other strategies – as long as they are expected. These strategies tend to be rational and mature. Nevertheless, some developers expect very naïve and cooperative strategies to react to. They also seem to have problems seeing the impact of their behavior on the strategies of others. In consequence, offline players frequently do not reach the efficient equilibrium of structured P2P systems, in contrast to online players. This is the main insight of this paper. Since the motives behind the behavior of online and offline players are the same, this difference is indeed due to the different setups only. Similar misinterpretations are observable in existing P2P file-sharing systems. Some of them enforce a minimum degree of cooperation to benefit from the system, such as bittorrent [26]. The system will then end up with many participants playing at this minimal level. These systems are an enhancement of systems without any reputation mechanism, such as kazaa [19] and gnutella [1], which converged to a state where a majority of participants was free riding. In our opinion, this is the result of a non-strategic perception of such systems. Participants do not find the strategies most beneficial for them.

Currently, developers of software for distributed systems develop a first version of their software and make it accessible to users. After the software is deployed, some users misuse it. The software is then patched against the misuse, and the procedure starts again. Clearly, by improving software over several iterations, successful strategies might eventually emerge. Simulating these iterations in a laboratory using online strategy design reduces development costs of nodes in distributed coordinator-free systems: Misuse can be detected at an earlier stage. Software for distributed coordinator-free systems that has been developed using this method should be more mature at deployment time.

7. Conclusions

Economic efficiency of systems with many agents depends on the agent strategies. To create them, two approaches exist: With offline strategy design, the designers implement them. Online strategy design in turn is based on strategies that humans use when mimicking system nodes. We have chosen structured P2P systems, a prominent kind of distributed coordinator-free system, as our object of investigation. Our first result has been to identify strategies in these systems that lead to an efficient equilibrium, using game theory. They feature reciprocity and cut-off behavior. We then compared online and offline strategy design, to analyze whether participants reach that efficient outcome with both approaches. This comparison is important: It helps to understand how well-designed incentive mechanisms can find their way into real systems. Both online and offline participants have recognized the usefulness of reciprocity and cut-off behavior. Online participants intuitively refrain from free riding and reach high payoffs. Offline players in turn are more uncooperative and try harder to achieve fair outcomes. By doing so, they do not reach that efficient equilibrium, in contrast to online players. Offline strategies are the result of focusing on reactive behavior and underestimating the strategies of others. Behavioral experiments with humans as part of the strategy design appear to be promising to circumvent these problems. Having shown that strategies strongly depend on how they have been designed, we propose to use a combination of online and offline strategy design when creating strategies for distributed coordinator-free environments.

8. References

[1] Adar, E. & Huberman, B.: *Free Riding on Gnutella*, Technical Report (Xerox PARC), 2000

[2] Androutsellis-Theotokis, S. & Spinellis, D.: A survey of peer-to-peer content distribution technologies, In: *ACM Computing Surveys*, 2004.

[3] Bazerman, M. et al.: Explaining how preferences change across joint versus separate evaluation, In: *Journal of Economic Behavior and Organization*, 1999.

[4] Brosig, J. et al.: The hot versus cold effect in a simple bargaining experiment, In: *Experimental Economics*, 2003.

[5] Cosmides, L. & Tooby, J.: Better than Rational: Evolutionary Psychology and the Invisible Hand, In: *The American Economic Review*, 84 (2), 1994.

[6] Conitzer V. & Sandholm T.: Self-interested Automated Mechanism Design and Implications for Optimal Combinatorial Auctions, In: *ACM Conference on Electronic Commerce*, 2004.

[7] Dash, R. et al.: Computational Mechanism Design: A Call to Arms, In: *IEEE Intelligent Systems*, 2003

[8] Fehr, E. & Schmidt, K.: A Theory of Fairness, Competition, and Cooperation, In: *The Quarterly Journal of Economics*, 1999.

[9] Fehr, E. & Gächter, S.: Altruistic punishment in humans, In: *Nature*, 2002.

[10] Feigenbaum, J. & Shenker, S.: Distributed Algorithmic Mechanism Design: Recent Results and Future Directions, In: *International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*, 2002.

[11] Fudenberg, D. & Tirole, J.: *Game Theory*, MIT Press, 1991.

[12] Hens, T. & Vogt, B.: *Money and Reciprocity*. Working Paper, University of Zürich, 2002, in revision

[13] Hoffman, E. et al.: Behavioral foundations of reciprocity: Experimental economics and evolutionary psychology, In: *Economic Inquiry*, 1998.

[14] Hurwicz, L.: The design of mechanisms for resource allocation, In: *American Economic Review Papers and Proceedings*, 1973.

[15] Parkes, D. & Shneidman, J.: Distributed Implementations of Vickrey-Clarke-Groves Mechanisms, In: *International Conference on Autonomous Agents and Multiagent Systems*, 2004.

[16] Papadimitriou, C.: Algorithms, Games, and the Internet, In: *ACM Symposium on Theory of Computing*, 2001

[17] Ratnasamy, S. et al.: A Scalable Content-Addressable Network, In: *Conference on Applications, Technologies, Architectures and Protocols for Computer Communications*, 2001.

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

[19] Saroiu, S. et al.: A Measurement Study of Peer-to-Peer File Sharing Systems, In: *Multimedia Computing and Networking,* 2002.

[20] Schosser, S. et al.: Incentives Engineering for Structured P2P Systems - a Feasibility Demonstration Using Economic Experiments, *ACM Conference on Electronic Commerce*, 2006.

[21] Schosser, S. et al.: Indirect Partner Interaction in P2P Networks - Stimulating Cooperation by Means of Structure, In: *ACM Conference on Electronic Commerce*, 2007.

[22] Schosser, S. et al.: Do Humans Identify Efficient Strategies in Structured Peer-to-Peer Systems? (Short Paper), In: *International Conference on Autonomous Agents and Multiagent Systems*, 2007.

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

[24] Selten, R.: Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eine Oligopolexperiments, In. *Beiträge zur experimentellen Wirtschaftsforschung*. 1967.

[25] Varian, H.: Economic Mechanism Design for Computerized Agents, In: *USENIX Workshop on Electronic Commerce*, 1995.

[26] Zghaibeh, M. & Anagnostakis, K.: On the Impact of P2P Incentive Mechanisms on User Behavior, *Joint Workshop on The Economics of Networked Systems and Incentive-Based Computing,* 2007.