Evolving Cooperation in the Non-Iterated Prisoner's Dilemma: A Social Network Inspired Approach

T. S. Ellis, X. Yao

Abstract— **Online service provision is becoming increasingly decentralized as system designers pursue the benefits gained from utilizing nodes at the periphery of the network. However, distributing control means relying on the cooperation of participating agents, and it is a significant challenge to design mechanisms that incentivise optimal global behavior in a population of selfish, rational agents. This is particularly evident in peer-to-peer file-sharing, where a high incidence of selfish behavior in the form of downloading without uploading, leads to the network losing the benefits of a decentralized network. In this paper a notion of reputation based on simple social network analysis is used to significantly improve cooperation rates in the one-shot game of prisoner's dilemma, where without such a technique the dominant strategy would be for all agents to defect.**

I. Introduction

The question of how cooperation arises between selfish agents is a problem of significant importance to biologists, for it appears to contradict evolutionary principles; in particular, the survival of the fittest. Fortunately, mechanisms such as reciprocal altruism [23] and kin-selection [8] have rescued the situation. The problem is also beginning to arise in computer systems. While, traditionally, systems have favored the top-down, hierarchical approach, the success of distributed systems in nature (flocking, schooling, herding, ant colonies, termite mounds, evolution), and in human society (the internet, economic markets), is generating interest in using similar techniques in the design of computer systems. One of the most significant areas where such a transformation is taking place is that of online content distribution. As the number of internet users and the bandwidth available per user increases, the client/server paradigm becomes less effective due to its lack of scalability. Costs rise, and as single nodes become increasingly important, so their vulnerability to malicious attack and unreliability also increases.

The rest of the paper is split into the following sections. A brief review of previous work and concepts central to the paper is presented in Section II. Section III introduces the conceptual model of a socially inspired reputation system, and results are presented in Section IV. This is followed by the conclusion (Section V) and some ideas for further work (Section VI).

II. Previous Work

A. Tragedy of the Commons

Recent decentralized file-sharing applications such as Napster and Gnutella [19] have shown that decentralized file-sharing systems can work, but they do suffer from certain problems. The most fundamental is that of free-riding - a phenomenon that Hardin [9] termed 'the tragedy of the commons'. This occurs when agents are autonomous, selfish, rational and strategic, and when the optimal action for the individual does not produce the best outcome for the population as a whole. This effect leads to the startling observation by Adar and Huberman [5] that nearly 50% of all query responses on the Gnutella network are made by 1% of peers. This means that 1% of hosts are effectively acting as centralised hosts for the network, and thus, all the benefits associated with a decentralised network are lost.

T. S. Ellis and X. Yao are with the School of Computer Science, University of Birmingham, UK (email: t.s.ellis@cs.bham.ac.uk, x.yao@cs.bham.ac.uk).

B. Prisoner's Dilemma

There are many existing game-theoretic strategic games that represent these complex social dilemmas in terms of simple payoffs and a limited number of actions. The most common is the prisoner's dilemma [6], where players make simultaneous decisions whether to cooperate or defect, and their subsequent payoff depends on their action and the action of their opponent. The payoff matrix for the Prisoner's Dilemma is shown below, where b is benefit and c is cost.

	C	**D**
C	b-c (r)	-c (s)
D	b (t)	0 (p)

The absolute values in the payoff matrix are actually unimportant - it is the ordinance that matters [17]. It is usual to describe the outcome P_{DC} as temptation, P_{CC} as reward, P_{DD} as punishment and P_{CD} as the sucker payoff - thus, the ordering in the prisoner's dilemma is

$$t > r > p > s$$

where $2 * r > t + s$. When this game is played once (the one-shot PD), the strictly dominant strategy, and thus the single equilibrium, is for both players to defect. However, if the game is repeated (Iterated PD[2]), there are several equilibria. One equilibrium is still for both players to always defect but the threat of punishment can make cooperation a Nash equilibrium.

Tit-for-tat: In 1984, Robert Axelrod ran a competition to find the best strategies for playing the Iterated Prisoner's Dilemma [2]. The winner was submitted by Anatol Rapoport, and was called tit-for-tat. Its strategy was to copy the action of its opponent: if its opponent defected, it would defect (punish), but if its opponent decided to cooperate, it would also cooperate, thus rewarding cooperative behavior. This is not a Nash equilibrium strategy, but it is, on average,

1-4244-1340-0/07/$25.00

superior to most other strategies. By analyzing the top ranking strategies, Axelrod found that they generally had the following characteristics: they should be nice (do not defect before your opponent, retaliating (defect in response to defection), forgiving (do not hold grudges), and non-envious (try to maximize your own score, not beat your opponent). In fact, superior strategies have since been found which also have these qualities, such as generous TFT (cooperates despite opponent defection 1/3 of the time [14]) and win-stay, lose-shift, which continues the current action as long as it receives R or T, but changes strategy if it receives the sucker's payoff. However, despite these variations, the central idea is that by taking into account future punishment, cooperation can become a rational behavior.

Unfortunately these strategies only work for the IPD; repeated interactions with the same agent. In a real-life scenario such as a large peer-to-peer network however, repeated interactions with the same agent are unlikely [12]. Therefore, a more accurate model of a large network might be a series of one-shot games of prisoner's dilemma with different agents. In this situation, strategies such as TFT do not work, as an agent's history is not known.

C. Reputation Systems

Reputation systems are a common method of gaining information about an agent's history. The following section provides a brief taxonomy of current approaches towards reputation systems (for a review of individual systems, see [1],

1) Centralized Reputation Systems: For these systems, a central authority collects ratings from members of the community who have interacted with another agent [1]. Examples of such a system can be found at Ebay[1] and Amazon[2]. These systems require the central authority to collate and process the interaction information sent to them, and to communicate this information back to potential participants. Because the reputation information, once processed, is the same for everybody, this is a form of global reputation, and can help to foster indirect reciprocity. Indirect reciprocity can be expressed in the phrase, 'you scratch his back, I'll scratch yours', and is also observed in the animal kingdom [18]. Mui et al [13] describe it as observed reputation, and model it by assigning a subset of agents in the population to be observed by each agent - in the case of a centralized system, this subset is the entire population of agents for every agent. The agent then uses these observed histories to guide future interactions. While centralized systems are the simplest reputation systems conceptually, the centralized authority provides a single point of failure which lowers robustness, and can be the object of malicious attack or subversion.

2) Decentralized Reputation Systems: Decentralized reputation systems pose a far harder problem to the system designer than does a centralized system. For example, while we would normally assume the centralized authority to be a a-priori trusted entity, we can make no such assumption about the autonomous, selfish agents that make up a distributed system. The simplest method of decentralized reputation is a form of direct reciprocity, which refers to the phenomenon of 'you scratch my back, I'll scratch yours' - a behavior noted in vampire bat colonies [25]. Mui et al [13] refer to this form of reputation as 'encounter-derived', and model it as each agent keeping a history of their own interactions with other agents in the system. This is shown to be the least effective method of reputation, particularly as the size of the population increases [12][13]. A more effective method is to propagate reputation scores via inter-agent communication, such as 'word-of-mouth' in human society. Mui et al [13] show it to be the most efficient reputation mechanism tested. However, communicating reputation messages around a network can be costly in terms of bandwidth, as can storing reputation scores for every agent in the population. There is also no guarantee you will find another agent that has interacted previously with your chosen opponent. A further type of reciprocity that can lead to maintenance of cooperation in the spatial PD game, is 'network reciprocity', where agents are constrained to interact only with their neighbors. The graph structure can be defined as a one-dimensional lattice [11], a two-dimensional lattice [15][16], as a random, regular graph [4], or as a small-world network [7]. Although this is very interesting from a biological perspective, particularly its interpretation as allowing for kin selection [8], it is not applicable to most peer-to-peer applications, as these could not function with a neighbor-only interaction rule.

D. Peer-to-peer Networks

Peer-to-peer networks are a good example of a system that requires cooperation between nodes (that is, contributing to the public good by the sharing of files), but where the rational action is not to contribute. Reputation systems have been considered as a solution to this problem, and indeed Kazaa[3] does include a 'share ratio' that enables some degree of reciprocation. However, there is a general problem of how to store these reputation values in a secure decentralized manner that would be resistant to tampering with the client, but not dependent on a vulnerable centralized server. Structured P2P systems have been devised that use observers [10]; that is, for each agent, a sub-set of the remaining agents are assigned to observe this agent's behavior. Because the observing agents can easily be found by hashing the key of the observed agent, it is very easy to find the observing agents for a particular opponent agent. However, many situations require a non-structured network such as highly dynamic environments; structured networks have a high cost for adding and removing data and users and are thus inappropriate for certain tasks.

E. Reputation in Multiple Choice Prisoner's Dilemma

Yao and Darwen [26] have previously looked into trying to evolve cooperation in a population of agents playing

[1]www.ebay.com
[2]www.amazon.co.uk
[3]http://www.kazaa.com

prisoner's dilemma when the possible actions for each agent are not binary (cooperate or defect), but instead range over varying degrees of cooperation [26]. In order to generate the new payoff function they used the following formula:

$$p_A = 2.5 - -0.5c_A + 2c_B, (-1 \leq c_A, c_B \leq 1)$$

They also reduced the length of IPD games which makes cooperation even more difficult to evolve. They found that while adding varying degrees of cooperation made the game more complex, and thus harder to co-evolve cooperative strategies in a population, adding a reputation variable that indicated whether an opponent agent had exploited (scored higher) than its previous opponents significantly improved the chances of cooperation evolving. Their major result was that reputation appeared to mitigate the disruptive effect of adding more levels of cooperation - while adding extra levels would normally massively reduce the likelihood of cooperation, with a reputation value this was no longer the case. This work differs from the current work in that we look at one-shot PD rather than short games of IPD, and we do not assume that the reputation scores are stored centrally. Yao and Chong [20] also investigate levels of cooperation in the multi-choice PD, and find a similar result to the above. They explain the decrease in cooperation as being due to the difficulty in resolving the intention of an opponent's intermediate choices, and also find that reputation increases cooperation, which they put down the ability of cooperative agents to start mutually cooperating right from the start of the interaction.

F. Social Networks

Social networks are a tool for representing the relations between, traditionally, people, although interpreted more broadly we can replace 'people' with 'agents'. Within the field of social networks, the most fundamental measure of an agent's importance, or popularity [21] can be inferred from the position of the agent within the network, or its 'centrality'. It is necessary here to define exactly what we mean by centrality. An agent in a social network has both a 'local centrality', which is simply the degree (the number of neighbors), and a 'global centrality', which measures how many hops it is from that agent to every other agent in the network. There is a further definition of centrality which measures importance by the degree to which an agent connects two parts of the network, but this need not concern us here. In his 1987 paper, 'Power and centrality: a family of measures', Bonacich [3] devised the following generalized formula for measuring centrality.

$$c_i = \sum_j r_{ij}(\alpha + \beta c_j) \qquad (1)$$

In this equation c_i refers to the centrality of node i. This is calculated by summing over j the strengths of all links connected from i to j, multiplied by α which is an 'arbitrary standardizing constant' [21], plus β multiplied by the centrality of node j. If β is set to 0, this is a definition of local centrality (in other words, it does not take indirect links into account), and very similar to the one we shall be using in this paper.

G. How the current work differs

The model presented in this work is designed as a method for storing a reputation measure in a distributed fashion that encapsulates the notion of indirect reciprocity. It does this without requiring individuals to store either their own image/reputation score, or to store the values of other agents. Instead an agent's image score is embedded in the network and can be inferred by measuring the degree of an agent. It is impossible to subvert ones own image, as links are established by mutual consent. The social network itself is not stored in any one agent; rather it emerges as a result of each agent storing a link to its own neighbors. Thus, for an agent to observe a part of the network, it must ask an agent to specify its links. It is also important to note that as we are only measuring *local* centrality, the time it takes to compute is constant with respect to the size of the network.

III. The Model

We imagine the problem of service provision by a multi-agent system. This service could be file transfer, or the loaning of network infrastructure to a communications company. For a service to be provided, a contract is agreed by both sides, but of course, both agents must decide whether to adhere to this contract. For the purposes of this model, we assume no centralized enforcement agency, as in real life these are often rather slow and ineffective, and are also probably not the prime factor in encouraging agents to keep to contracts. This is also a very important assumption for P2P file sharing, where a centralized authority represents a vulnerability. Therefore, when this problem is abstracted to the prisoner's dilemma, cooperation represents adherence and defection represents reneging. The game then proceeds as follows: A population ($\mathbf{P}$) of one hundred players is maintained. Players are chosen randomly from the population to take part in pairwise encounters wherein a single game of the prisoner's dilemma is played. Players are chosen without replacement; that is, once they have played a game they cannot be chosen to play a further game during that round. Note that at all times the game is single-shot rather than iterated, and therefore the cooperation equilibrium selected in the IPD is not available. This is because we are modeling indirect reciprocity, and IPD in analogous to direct reciprocity. A single round consists of fifty games (each agent plays one game). At the conclusion of each round the evolutionary process is carried out, as explained in Section III-A. Players are removed from the population and replaced with randomly generated individuals according to the '*churn*' parameter. This reflects the open nature of most networks and business environments, where agents with various strategies may enter the system at any time, thus upsetting any non-Nash equilibrium that may have been established.

The social network element of the model occurs from the rule that a successfully completed transaction results in a

```
method calculateCentrality(node_j)
{
  count = 0;
  B = getReferals(node_j);
  for(each node_k in B)
  {
    if(B.contains(node_j)) count = count + 1;
  }
  return count;
}
```

Fig. 1. The algorithm used to calculate local centrality. Node_j (n_j) is the object node, while $B \subset N$ is the set of nodes that have a mutual link with n_j. The conditional clause checks that $n_j \in B$ - if it is false, n_j was lying about that particular link, and it does not contribute to the centrality score.

link formed between those agents, and repeated cooperation between the same agents increases the strength of this connection. Therefore, if we have agents $a_i \in \mathbf{P}, a_j \in \mathbf{P}$, and r_{ij} represents the strength between them:

$$r_{ij} = \begin{cases} r_{ij} + 1, & if \text{ coop(i, j)} \\ 0, & if \text{ defect(i, j)} \end{cases}$$

A '*max_links*' parameter is set for the system, so that the following condition must hold:

$$\sum_j r_{ij} \leq \textit{max\_links}$$

Therefore, the *max_links* parameter refers to the maximum allowed sum of the strengths of each connection. If it has a maximal number of links, and another cooperation occurs, the oldest existing link is replaced with the new one. A crucial point to note here is that links can only be formed by mutual consent from both agents concerned. This means that an agent cannot lie about his own degree, as each link can be investigated; that is, the neighbor questioned as to whether the link is genuine. A failed transaction (when either agent defects) breaks the link. The algorithm which would be used in order to calculate the local measure of centrality is shown in Figure 1. Node_j (n_j) is the object node, while $B \subset N$ is the set of nodes that have a mutual link with n_j. The conditional clause checks that $n_j \in B$ - if it is false, n_j was lying about that particular link, and it does not contribute to the centrality score.

The agents are not spatially organized, so agents can chose to interact with any other agent in the population (the population is well-mixed). This reflects the fact that in the virtual world (and indeed post-industrial human society), there is no spatial separation, unlike the biological world where agents can necessarily only interact with nearby neighbors. Therefore, it is important to note that the links that are established as part of the reputation system do not represent viable interactions between agents - agents can interact with any other agent in the network. This is demonstrated in Figure 2. The network can be considered as two layers. The first is fully connected, which means any agent can 'reach' any other, while the second represents which agents have cooperated with each other recently. Two types of agent strategy encodings and evolutionary methods were considered in order to evaluate the effectiveness of the model. These are described in detail below.

A. Discriminators Versus Pure Strategies (Configuration One)

In this set-up the world is populated with three types of agent, two of which are pure strategies while one is a mixed strategy: pure cooperators, pure defectors, and discriminators. The discriminators in the first set of experiments base their decision on the centrality of their opponent node, using the simplified version of Equation 1:

$$c_i = \sum_j r_{ij} \tag{2}$$

We then use evolutionary game theoretic methods [24] to select an evolutionarily stable equilibrium [22]. An evolutionarily stable strategy is defined as one that cannot be invaded by a mutant, and is similar to a Nash equilibrium [17] but for the subtle difference that rationality is not assumed. In this model discriminators cooperate probabilistically, as determined by the following function which uses the previously defined value for c_i (Equation 2).

$$P(j) = \frac{c_i^\alpha + \beta}{c_i^\alpha + \beta + 1} \tag{3}$$

As can be seen in Figure 3, changing α makes the step steeper - the step function in Figure 3 is produced with $\alpha = 10$, while β sets the likelihood of cooperation when the opponent has no links. This can be considered a 'stranger policy', or the degree of forgiveness analogous to the 'generous IPD' strategy mentioned in Section II-B. All agents have the same values for α and β, and Section IV-A shows how the behavior of the model varies with different values of β. The idea behind this configuration is to see which strategy becomes dominant in the population - thus some method is required of causing successful strategies to spread throughout the population, while weaker strategies

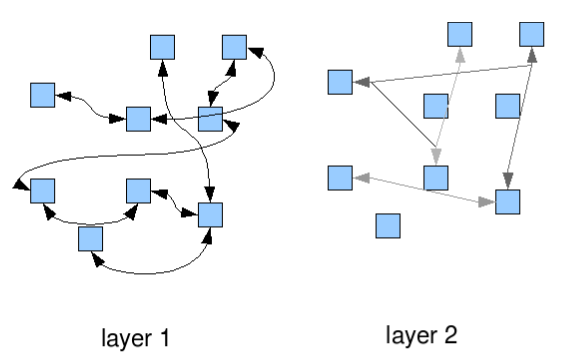

Fig. 2. For the purposes of the model, the network can be considered as two layers. The first is fully connected, which means any agent can 'reach' any other, while the second represents which agents have cooperated with each other recently.

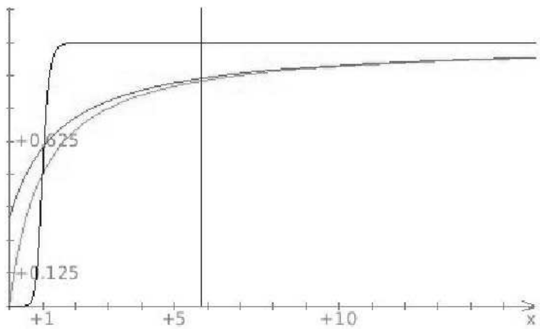

Fig. 3. Varying parameters α and β in Equation 3

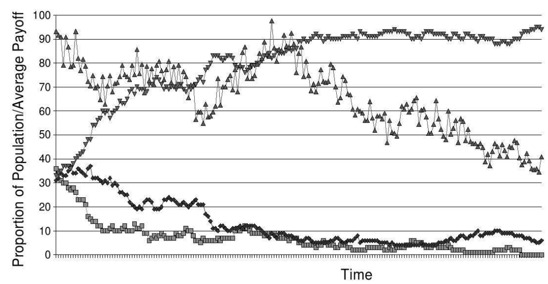

Fig. 4. Proportion of cooperators, defectors and discriminators, and average payoff to each agent over time when β = 0

die out and are replaced. The simulation was then run for two thousand runs, each run consisting of one hundred rounds, and each round consisting of one hundred games. The population size of agents is set to 100 in this experiment. At the end of each round, two agents are selected from the population at random, and the strategy of the stronger agent replaces that of the weaker. This is only performed once per round in order to slow down the process by which successful strategies can spread.

B. Evolutionary Dynamics (Configuration two)

In the second set of experiments, a different type of strategy encoding is used. Rather than use replicator dynamics to determine the equilibrium strategy, which merely shows that discriminators are more successful than either all-cooperators or all-defectors, agent strategies are represented instead as a continuum, from all-cooperate to all-defect, and all possible thresholds and levels of forgiveness in-between. In other words, while configuration one demonstrates the ability of the mechanism to allow a population of agents that use the extra network information to become dominant in the population that includes pure strategies, this configuration shows that it is possibly to evolve agents even in a population without pure strategies - where each opponent agent is also evolving at the same time. This form of evolution is known as co-evolutionary learning, and differs from evolutionary computation in that the evaluating function is a moving target formed by the rest of the population [26], rather than a static evaluation function. This problem is therefore much harder than that in configuration one, and also allows us to see exactly what type of discriminating function is optimal and reaches an equilibrium (if, indeed, an equilibrium is reached at all). Agent strategies are represented by two values: α and β. α is the probability an agent will interact with another agent that has a lower social standing (centrality) than a given threshold, and β is that threshold. If the centrality of the agent is above the threshold, the probability of cooperation is 100%. These experiments aim to show whether a collection of homogeneous (in terms of strategy space) agents will select an equilibrium strategy that maintains a levels of cooperation higher than would be expected without the dynamically constructed network. In these experiments, rather than simple dynamics in configuration one, where weak strategies are replaced with stronger ones as the system is run, mutation was also used in order to evolve entirely new strategies. After each round of the game, two agents were chosen at random. The weaker of the two was removed from the population, and replaced with a slightly mutated (according to a mutation rate) version of the fitter parent. Mutation was performed by randomly adding or subtracting a small amount from the α value, or adding/subtracting an integer from the threshold value (with far lower probability). The population size varies depending on the experiment, but the default size is 100.

IV. Results

A. Configuration One

Figures 4, 5, and 6, show the population dynamics during a typical simulation run for three values of beta: $\beta = 0$, $\beta = 0.5$, and $\beta = 0.1$ along with the average payoff values. Note that the payoff values are divided by three in order to keep them within the same scale. In Figure 4, where $\beta = 0$, cooperators and defectors become practically extinct very rapidly. Cooperators tend to die out very quickly because they do not differentiate between a fellow cooperator and a defector - thus they are often exploited by defectors who immediately subsume their population. However, once the cooperators have died out, defectors no longer have anybody to exploit, and so they are soon replaced by discriminators which becomes the equilibrium strategy. Once the discriminators have become the dominant strategy however, the average payoff begins to decrease. This is because every time an agent leaves the network and a new one joins, the new agent has no links and so is unable to find a cooperative partner. We see from this that a purely discriminatory strategy is not enough to maintain cooperation.

Figure 5 shows the population dynamics when β is set to 0.5. Here, discriminating agents have a probability of 0.5 of cooperating with an agent without any links. This is shown to be too generous; defectors are able to exploit these discriminators, and thus discriminators soon become the dominant strategy in the population. Payoff levels are minimal (a score of 20 indicates every member of the population receives the punishment payoff - total payoffs are divided by 4 for scaling reasons).

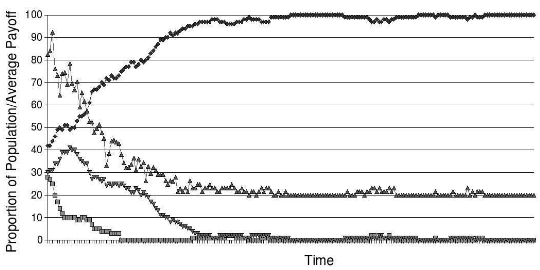

Fig. 5. Proportion of cooperators, defectors and discriminators, and average payoff to each agent over time when β = 0.5

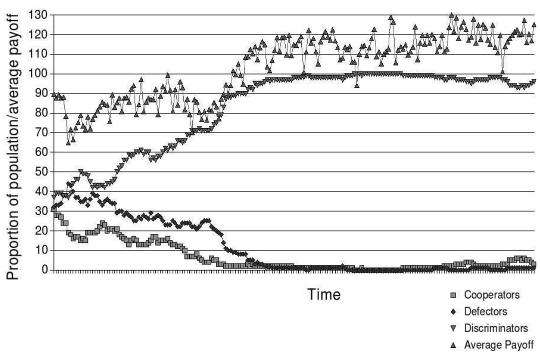

Fig. 6. Proportion of cooperators, defectors and discriminators, and average payoff to each agent over time when β = 0.1

However, in Figure 6, when $\beta = 0.1$, discriminators become the equilibrium strategy, and cooperation is maintained. Therefore, two properties are demonstrated. The fact that discriminators take over the population demonstrates that by using the extra network information, they can tell the difference between an uncooperative agent and a fellow discriminator or cooperator. We can also see that a population of discriminators cannot maintain levels of cooperation unless they are willing to interact with low-image individuals, and this has certain implications for the uses to which this system could be put. For example, in a high-risk scenario, it may be unwise to trust such individuals at all. In this case, it may be that the system can maintain a certain level of risk-averse agents, as long as it is not the entire population. This will be investigated in further work. Finally, in order to see exactly how the the value of beta affected the population, the population was run for 10000 runs with increasing values for beta. The resulting population proportions can be seen in Figure 7. As can be seen, with low values of β, discriminators become very dominant, but overall payoff in the system remains low. For higher values of β defectors become dominant and cooperation still remains low. At around β = 0.1 however, there is a high degree of cooperation, despite around 10% defection.

B. Configuration Two

When the simulation was run using the more open-ended strategy representations, some interesting results were ob-

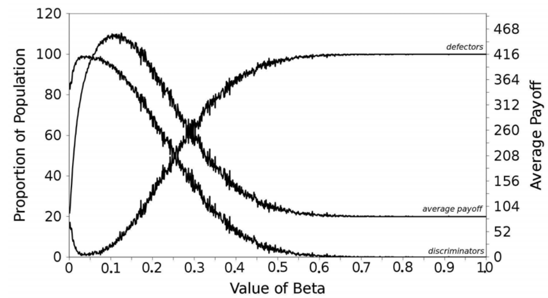

Fig. 7. Proportion of cooperators, defectors and discriminators, and average payoff per agent as β is increased

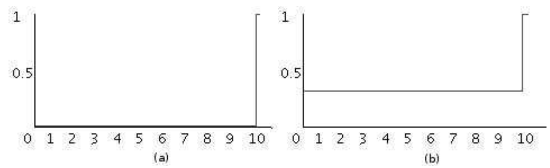

Fig. 8. Average Probability of Cooperation Function

served. First of all, we show the behavior of the system when the network links were not taken into account - instead random numbers were generated so that there was no correlation between an agent's actual number of links. As expected, the agents soon fall into a 'all-defection' attractor. The average evolved function is shown in Figure 8 part (a). It can be seen that the threshold has risen all the way to the maximum and there is no degree of generosity towards agents with low-degree. This is typically what one would expect when defection is the dominant strategy. The values of α, β and the average payoff are shown in Figure 9. Note that in these graphs all values are normalized and multiplied by 100 in order to be viewable on the same graph. As can be seen, the α value which represents the generosity of the agent falls quickly to zero, while the threshold value rises quickly to the maximum. This leads average payoffs for the system to also fall to the punishment payoff, which is clearly inefficient.

Figure 8 part (b) shows the average evolved function for the system when agents are allowed to observe the centrality of the opponent agent. While the threshold is still very high in relation to part (a), indicating that the agents are very discriminatory, the generosity value is around 0.3. This tallies with the work by Nowak and Sigmund [14], who found the 'generous TFT' strategy in the IPD very successful, which would forgive an agent for defections 1/3rd of the time. Figure 10 shows the evolutionary dynamics of the average α and β parameters of the population when the agents are allowed to use degree as a measure of trustworthiness. It can be seen that the value of beta seems to be highly correlated to the over-all payoff of the system. The oscillating nature of the beta value seems to suggest a higher beta value increases the fitness of an agent, and of the overall population (it

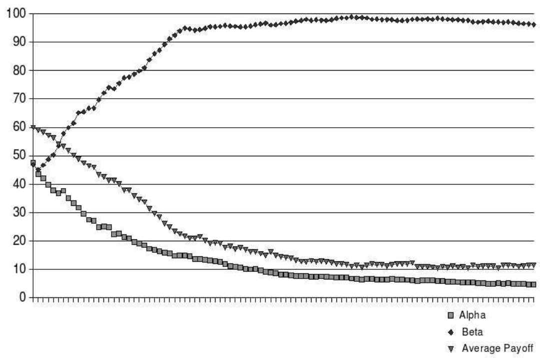

Fig. 9. Values of α, β and Average payoff over time

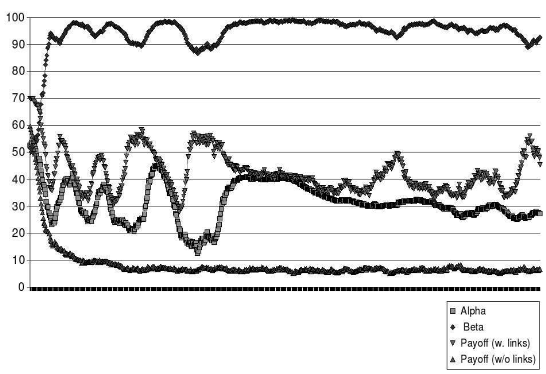

Fig. 10. Values of α, β and Average payoff with and without links

increases the overall payoff), but only up to a certain point after which it leads to exploitation. At this point, the fitness of high alpha value agents decreases, and alpha values in the population begin to fall. However, too low a value and an agent's number of links decreases, thus causing other agents to stop cooperating with it. It is interesting that the threshold for unconditional cooperation remains at between 9 and 10. This threshold is vital however,and cannot be removed in order to only evolve an alpha value - it is the threshold which provides the selection pressure for agents to maximize their number of links - without this there would be no incentive to have a non-zero alpha value, and cooperation would never be maintained. Therefore, we can see from this result that the possibility of generating an extra link motivates an agent to cooperate not only with agents it perceives as being of a cooperative type, but even with unknown or possibly deviant agents, although only with a certain probability.

A very important task is to determine the relationship between the required maximum number of links allowed per agent and the population size. Therefore, an experiment was conducted whereby the cooperation level of the system was measured while the maximum number of links was altered from 0-20 for 4 different population sizes: 50, 100, 500, and 1000. The experiment was run for a number of rounds equal to the size of the population multiplied by 100. As can

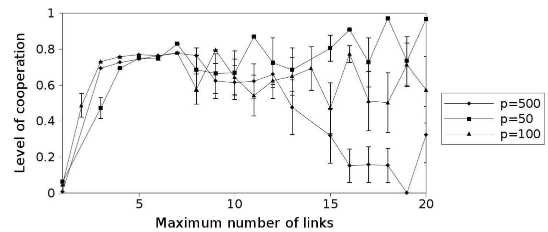

Fig. 11. The level of cooperation with various 'maximum links' settings, over three different population sizes

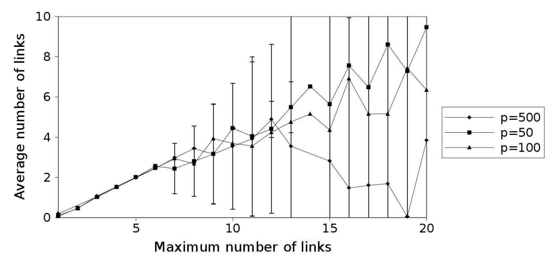

Fig. 12. The average number of links per agent in relation to the maximum number of links allowed, for three different population sizes

be seen in Figure 11, the most effective maximum number of links for each population size is actually between 5 and 10. Any more and performance begins to either drop off entirely, or at least become quite unpredictable. As this is a complex system, a greater number of links appears to make the system more sensitive to differences in starting conditions (the initial random population) and to thus fall into unpredictable basins of attraction. Figure 12 sheds some light onto this, by showing that while the average number of links increases linearly with respect to the maximum number of allowed links up until 6, beyond this the relationship is far less distinct, perhaps implying that that inherent randomness in the opponent that you meet means that past a certain maximum number of links, an agent's centrality actually becomes a poor predictor of how cooperative the agent is.

V. Conclusions

In this work we demonstrate that it is possible to maintain reasonably high levels of cooperation in the one-shot prisoner's dilemma by using a notion of reputation derived from social network analysis. Rather than use a pre-established social network, the network is dynamically created at runtime. Through several sets of experiments, we showed that discriminators become the stable equilibrium strategy when their opponent types are the pure strategies of always-defect and always-cooperate, and we found an optimal value of beta, which is analogous to a stranger policy. We also demonstrated using co-evolutionary learning that when agents are allowed to evolve both a strategy for dealing with low-image opponents, and a threshold for whom they would always cooperate with, levels of cooperation can be maintained at a fairly high rate and agents are happy to cooperate with un-trusted agents roughly 30-40 percent of the time. This is due to the benefit of gaining a new link making taking the risk of receiving the sucker's payoff worth it. Thus, even this simple extra network information appears to be sufficient to enable long-term cooperation to evolve in a population of selfish reward-seeking agents. This model is particularly suitable for applications such as peer-to-peer file sharing due to the simplicity in establishing links in such a system, the constant time necessary to retrieve reputation scores, and the inherent difficulty in subverting ones own reputation score.

VI. Further Work

The nature of the prisoner's dilemma means that the sociograph that is dynamically created is undirected. It also implies that both agents in a transaction have some duty to perform (such as with Ebay). However, in a peer-to-peer networking environment, there will often be no such mutual duty to cooperate. Therefore, it would useful to see if a version of the public goods game where transactions were one-way, thus creating a directed sociograph, would perform equally as well. We also assume that the formation of links is enforced - in a real system this is too strong an assumption, and so it would be necessary to see how the system performs under conditions of voluntary feedback. A final question that requires further investigation is exactly how well the system may perform under the effects of collusion. Any reputation system is potentially vulnerable to collusion, as agents leave unjustified positive feedback in order to boost each others scores. Further work will investigate exactly how various sizes of collusion sub-network affect the evolution of strategies.

Acknowledgment

This work was supported in part by the Engineering and Physical Sciences Research Council as part of the Market Based Control of Complex Computational Systems project, grant number GR/T10671/01. Many thanks also to the anonymous reviewers for their invaluable comments and suggestions.

References

[1] R. Ismail A. Josang and C. Boyd. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43:618–644, 2007.

[2] R. Axelrod. *The evolution of cooperation*. Basic Books, New York, 1984.

[3] P. Bonacich. Power and centrality: a family of measures. *American Sociological Review*, 52, 1987.

[4] G. Szabo C. Hauert. Prisoner's dilemma and public goods games in different geometries: compulsory versus voluntary interactions. *Complexity*, 8:31–38, 2005.

[5] B.A Huberman E. Adar. Free-riding on gnutella. *First Monday*, 5(10), 2000.

[6] M. M. Flood. Some experimental games. *Research Memorandum RM-789, RAND Corporation, Santa Monica CA*, 1952.

[7] M. Kuperman G. Abramson. Social games in a social network. *Physical Review E 63:030901*, 2001.

[8] W.D. Hamilton. The genetical evolution of social behaviour ii. *Journal of Theoretical Biology*, 1964.

[9] G. Hardin. The tragedy of the commons. *Science*, 162(3859):1243–1248, 1968.

[10] So Young Lee O-Hoon Kwon Jong Kim Sung Je Hong. A reputation management system in structured peer-to-peer networks. *Enabling Technologies: Infrastructure for Collaborative Enterprise*, 2005.

[11] A. Shaked I. Eshel, E. Sansone. The emergence of kinship behaviour in structured populations of unrelated individuals. *International Journal of Game Theory*, 28:447–463, 1999.

[12] I. Stoica K. Lai, M. Feldman and J. Chuang. Incentives for cooperation in peer-to-peer networks. *Workshop on Economics of Peer-to-Peer Systems*, 2003.

[13] A. Halberstadt L. Mui, M. Mohtashemi. Notions of reputation in multi-agent systems: a review. *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 1*, pages 280–287, 2002.

[14] K. Sigmund M. A. Nowak. Tit for tat in heterogenous populations. *Nature*, 355:250–253, 1992.

[15] R.M. May M.A. Nowak. Evolutionary games and spatial chaos. *Nature*, 359:826–829, 1992.

[16] M. Oliphant. Evolving cooperation in the non-iterated prisoner's dilemma: the importance of spatial organisation. *Proceedings of Artificial Life IV*, 1994.

[17] M. J. Osborne. *An Introduction to Game Theory*. Oxford University Press, 2004.

[18] A.S. Grutter R. Bshary. Image scoring and cooperation in a cleaner fish mutualism. *Nature*, 441:975–978, 2006.

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

[20] X. Yao S. Y. Chong. Multiple choices and reputation in multiagent interactions. *IEEE Transactions on Evolutionary Computation*, 2006.

[21] J. Scott. *Social Network Analysis, Second Edition*. Sage, 2006.

[22] J. M. Smith. *Evolution and the theory of games*. Cambridge University Press, 1982.

[23] R. L. Trivers. The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 1971.

[24] J. W. Weibull. *Evolutionary game theory*. MIT Press, 1995.

[25] G.S. Wilkinson. Reciprocal altruism in bats and other mammals. *Ethology and Sociobiology*, pages 85–100, 1988.

[26] P. J. Darwen X. Yao. How important is your reputation in a multi-agent environment. *Proc. of the 1999 IEEE Conference on Systems, Man, and Cybernetics, IEEE Press, Piscataway, NJ, USA*, pages 575–580, 1999.