Evolving cooperation in the non-iterated prisoner's dilemma: A distributed, Small World Network inspired approach.

Nicolas Höning (Vrije Universiteit Amsterdam), Tomáš Kozelek (Charles University Prague)

May 9, 2008

Abstract

In complex and dynamic markets, most of the opponents have never interacted before and it is hard to estimate for them wether they can trust each other. We present a distributed framework where agents play the Non-Iterated Prisoner's Dilemma. We propose that agents can ask other agents they already trust if the current opponent can be trusted. These neighbours can in turn ask their neighbours. This process can be iterated until someone knows the opponent, thus a trusted information chain is built. Such a chain has the advantage that no remote messages from completely untrusted sources are used to decide about someones trustability. We show that this framework is computationally feasible, if the agents are connected in a Small World Network. In our framework, such a network evolves and information chains can successfully be built within. We test a strategy using this framework against several default strategies and show that it can very well succeed.

1 Introduction

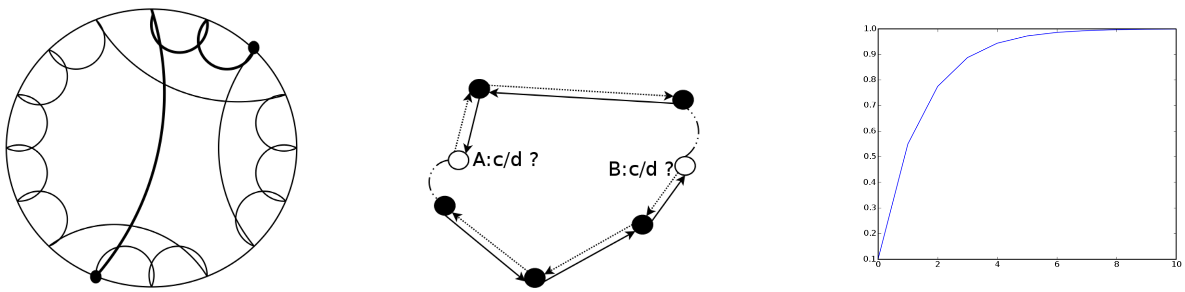
In real interactions between humans, trust is an important notion. We feel best about an upcoming interaction if we can rely on previous interactions with our trade partner. Whenever we didn't have one, we try to find out if someone we do already trust has experienced good interactions with him or knows somebody who has. In complex and dynamic markets, we seldom know our trade partners, so we try to find such chains of trust judgements every day. If we find one, using it to judge wether the interaction can be trusted might protect us from fraud.

We combine two concepts in recent scientific research to model such a world: The Non-Iterated Prisoner's Dilemma and Small World Networks. Through ongoing interactions, a Small World Network is evolved, which makes it feasible to obtain information about opponents. Our goal is to design a fully distributed framework in which cooperative agents would have a potential to cope with lying agents and frauding cartels. This is a design study to explore this research path and the success of such a strategy.

1.1 Concepts

The Prisoner's Dilemma is a non-zero-sum game with two players. It depicts a situation where both opponents need to decide if they cooperate or defect (see section 3 for a payoff matrix). Both would achieve the most if they defect while the other cooperates and perform worst if it is the other way around. The overall payoff is the highest for a situation where both cooperate. The clue is that a short-term rational decision is to defect, while in iterated games overall cooperational behaviour also achieves the best returns for individuals. In the non-iterated version, there is no guarantee that the players will ever see each other again. How can cooperation evolve nonetheless?

Small World Networks describe a phenomenon that seems to be inherent to human societies: The (social) network is connected in such a way that every node is reachable from every other node within a small number of steps (small when compared to the network size - see Figure 1(a) for an example). This property of a network is called the "shortest average path length". Another property is the "clustering index", which is calculated by the average number of neighbours that also know each other.



(a) (b) (c)

Figure 1: **Left**: A Small World Network. Shortcuts between remote agents make short average path lengths possible. **Middle:** Two agents that are about to play and need to decide if to cooperate -*c*- or defect -*d*-. Both manage to find a chain of trust (dotted) to someone who knows their opponent (dashed). Then, the information is passed back along the chain (straight). **Right:** Values in trust table for an opponent starting with 0.1, continued by 10 cooperations.

1.2 Previous Research

Formal research in cooperation of rational individuals was first introduced in Game Theory which was developed by Von Neumann and Morgenstern [9] in the 1940s. The Prisoner's dilemma is a classic Game Theory setup in which cooperation is often studied. The history of computational cooperation research goes back to the early 1980s [3]. Axelrod held tournaments of competing strategies, from which the simple Tit-For-Tat strategy emerged as the most successful. This strategy is 'nice' in general, but will retaliate defection immediately. To prevent all-defecting strategies from winning mixed-population tournaments, the notion of trust became an important issue (e.g. [7, 8, 1]).

Small-World Networks, described by Watts and Strogatz [10], have been very popular since the end of the 1990s, also amongst cooperation researchers, e.g. [2]. A very recent trend is to encode trust in the network structure (e.g. [4]). For example, in Ellis' and Yao's approach, players ask neighbours of the opponent to verify his trustworthiness. They can contact these neighbours through a central authority. These approaches are effective regarding the performance, however they have some unpleasant characteristics (e.g. they require a central authority, they are potentially very vulnerable to cartels of defectors that issue each other high trustability values).

2 Model

In our model, agents play rounds of the Prisoner's Dilemma with random opponents and receive payoff. Periodically, we perform evolutionary replacements on the population to translate payoff success into numerical representation in the population.

Our model is distributed - agents do not need to contact a central authority. Instead, agents with our strategy try to ask other agents if they know the opponent. If they don't, they can ask someone they know (every agent keeps a trust table in which he lists some neighbours and how much he trusts them). Every agent (B) can be asked by some other agent (A) about the trustability of a particular opponent (C). If agent B doesn't have information on agent C in his trust table, then he will in turn ask another agent (D). From the trust table of B, agent D is selected by maximising the formula

$$dist\_bias * dist(D, C) + trust\_bias * trust(B, D)$$

This formula balances a tradeoff between effect and safety. Clearly, the higher the bias on distance, the more probable it is that the opponent will be reached soon. On the other hand, high bias on trust might produce more reliable chains but it is also more likely to fail in finding the opponent at all.
In the tradition of Axelrod's research [3], our agents will default to cooperation if no information on the opponent is accessible. Furthermore, when the interaction is finished, agents update the value for the opponent in their trust table according to

$$\textbf{if } (coop) \textbf{ then } new\_trust = old\_trust + \frac{(1 - old\_trust)}{2}$$

$$\textbf{if } (def) \textbf{ then } new\_trust = old\_trust - \frac{(old\_trust)}{2}$$

This results in a sigmoid curve (Figure 1(c)). If the opponent changes his behaviour, it will have a quick effect on the trust in him, relating to Tit-For-Tat.

Note : We named our strategy *nuts* (**N**icolas **u**nd **T**omas **S**trategy) and its less smart version (which only looks in the own trust table) *nuts_dummy*. Other strategies we employ in different scenarios are *AlwaysDefect* (AD) and *AlwaysCooperate* (AC).

2.1 Relation to the real world

There are many real-world examples of trust-based social networks based on the repeated Non-Iterated Prisoner's dilemma, e.g. peer-to-peer networks, social networks behind online auctioneering systems, ... Our model could resemble an online auctioneering system where users announce what they want to sell and afterwards are contacted by other users willing to buy. The analogies with our model are clear:

- The actions of users represent actions in the Non-Iterated Prisoner's Dilemma. Sending the money for goods is cooperation, whilst not sending money is defection. Similarly, sending goods is cooperation, whilst not sending anything is defection.
- Users tend to know other users with similar interest (which in our model is resembled by a simple metric).
- Users keep information on some particular agents they have interacted with in the past (the trust table in our model), most likely ones that are similar to them.
- Users would like to get information on a person they are about to trade with (the information chains in our model).

Our model introduces one more phenomenon. This is the evolutionary replacement of weak agents by stronger ones. However, it is not impossible to imagine that there is only a limited number of users that can participate and also there are users waiting to enter the system when some of the users in the system are out of money. This could be understood as an analogy of the implemented evolutionary replacement.

2.2 Procedure

Or simulations proceed as follows:

Initialise N individuals
Initialise the world: According to the scenario, assign strategies uniformly to randomly picked agents
for r in $1 \ldots N*4$ rounds **do**
  for i in $1 \ldots N/2$ **do**
    Select an agent i and an opponent j to play against (see also section 2.3 below).
    Agents i and j try to find out if their opponent can be trusted (if they are *nuts* agents).
    Agents i and j decide wether they cooperate or defect.
    Agents i and j receive payoffs and update their trust tables according to their opponents' action.
  end for
  if $(N*4)\%e == 0$ **then**
    Evolutionary update: Select two agents and replace the poorer one with a copy of the other
  end if
end for

2.3 Evolving and using the Small World Network

We would like to evolve the underlying network into a Small World Network. Such a structure can be characterised by low average path lengths L (comparable to L of a random graph of the same size) and a high clustering coefficient C (much higher than C of a random graph). Advantages of such a structure for our framework are obvious: An agent should be able to build a short information chain connecting him to the desired opponent. In particular, we set two goals for a network of size N:

1. The information about neighbours that every single agent maintains about should be small ($O(logN*\sqrt{logN})$)
2. Retrieving information from the network should be fast ($O(logN)$ or constant - given in Small World Networks)

The default way of building a Small World Network structure is to mostly connect neighbouring agents, and sometimes (according to some probability) remote agents [10]. In this model, we follow a more precise approach by Kleinberg [6] to evolve it using a distance metric. This metric works in the sense that agent A should have a higher probability to play against or know agent B than agent C if the distance $d(A,B)$ is lower than $d(A,C)$. Kleinberg calculated that probability with $\frac{1}{d(A,B)^{\alpha}}$.

When agents build information chains (making use

of the evolved network structure), it is also necessary that they have a method of choosing which of their neighbours would be most appropriate for "getting closer" to the desired opponent. They can use the metric to help them decide who is most likely to know him.
The optimal setting for α to evolve a Small World Network is around 2. However, our model needs this metric for two purposes, building the network and information chains. In our first experiments, agents already knew their opponents too often, such that information chains weren't necessary. To resemble a more realistic situation, we split α. We used less bias to locality when building up the network (α, around 1.0) and more bias to locality when agents decide who their closest friends are (β, around 2) - i.e. to decide who to keep in the list of neighbours.

3 Experimental Design

We used a *simple comparative* experimental design. The simulations were run 20 times for each experimental setting.

Independent Variables

.

variable	possible values
N: # of individuals	[150, 300]
K: # of neighbours	$[0.7 * logN * \sqrt{logN}$, $1.1 * logN * \sqrt{logN}]$
α: network	[0.9, 1.2]
β: network	[1.9, 2.4]
strategy scenarios	[($\frac{1}{2}$AD, $\frac{1}{2}$nuts), ($\frac{1}{3}$AC, $\frac{1}{3}$AD, $\frac{1}{3}$nuts), ($\frac{1}{3}$AD, $\frac{1}{3}$nuts, $\frac{1}{3}$nuts_dummy)]
bias to distance when asking	[0.3, 0.7]

In effect, we had $2 * 2 * 2 * 2 * 3 * 2 = 96$ different settings.

Dependent Variables

variable	comment
L	average path length
NP	# of not connected pairs of agents
C	clustering index of network
SN	# of agents of particular strategies
SP	average payoff of the strategy

Fixed Variables

variable	value
limit of steps in asking	6
chain trust threshold	0.3
opponent trust threshold	0.5
# of epochs	4 * N
e: rounds until evol. update	3
bias to trust when asking	1 - dist. bias
steepness of trust update function	2

The payoff matrix is also a fixed variable. It is the standard payoff matrix for a game of Prisoner's dilemma, taken from [3]:

	A	B	payoff for A
traitor benefit (t)	D	C	5
reward (r)	C	C	3
punishment (p)	D	D	1
sucker payoff (s)	C	D	0

4 Results

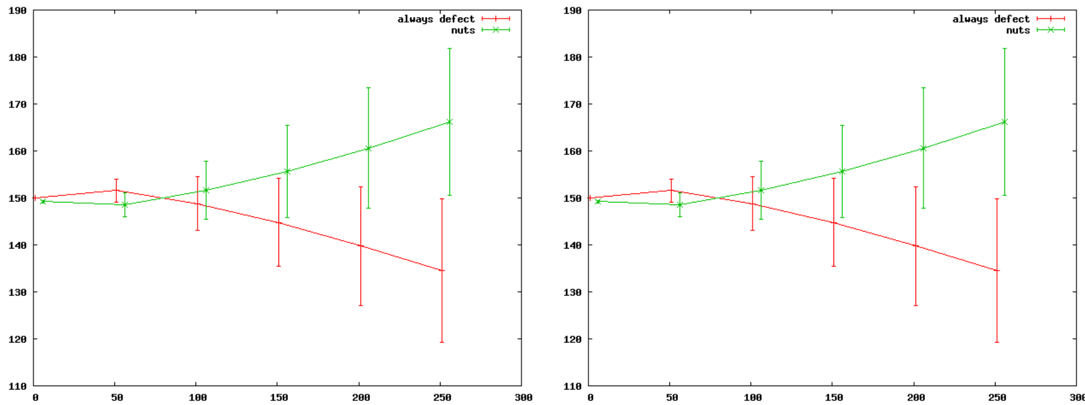
This section reviews our research hypotheses and states in what manner we have achieved them (full set of results at [5]).

- **Hypothesis 1: The underlying structure will evolve to a small-world network.**
  Holds. In every simulation we ran, the network parameters achieved acceptable values. Small World Networks have $L \simeq L_{random}$, but $C >> C_{random}$ (see [10] for details). The table below shows the network parameters at the end, averaged over all our simulations.

	L	L_{random}	C	C_{random}
$N = 150, K = 13$	2.46	2.19	0.30	0.08
$N = 150, K = 21$	2.03	1.89	0.29	0.13
$N = 300, K = 16$	2.62	2.33	0.29	0.05
$N = 300, K = 25$	2.20	2.01	0.26	0.08

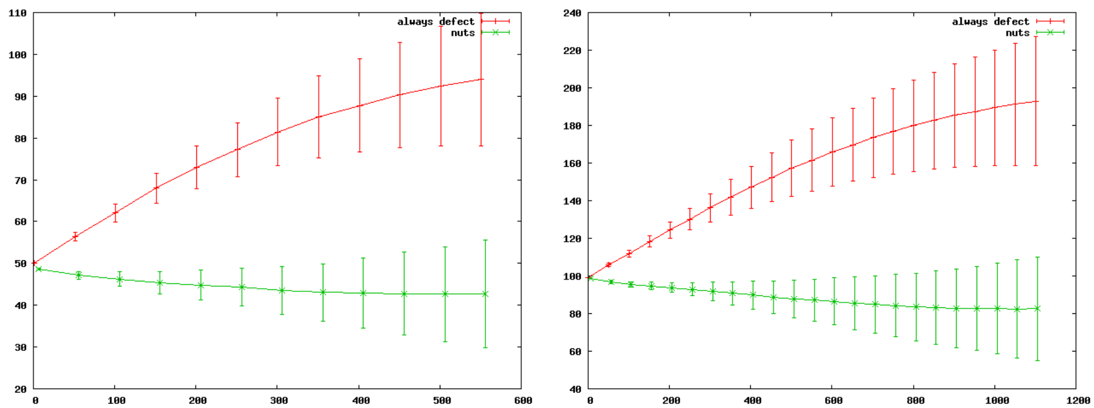
- **Hypothesis 2: Informed strategies will survive in various populations.**
  The simple explanation is that this doesn't necessary hold. Simply averaged over all configurations, our strategy achieves to represent 55.8% of the population at the end of the simulations. That is not a bad value. If we look closer, we see that the success depends on the particular population scenario.

  If we only look at strategy scenario 1, *nuts* versus *AlwaysDefect*, we find that *nuts* wins by far. We conducted a Welsh Two-Sample T-Test



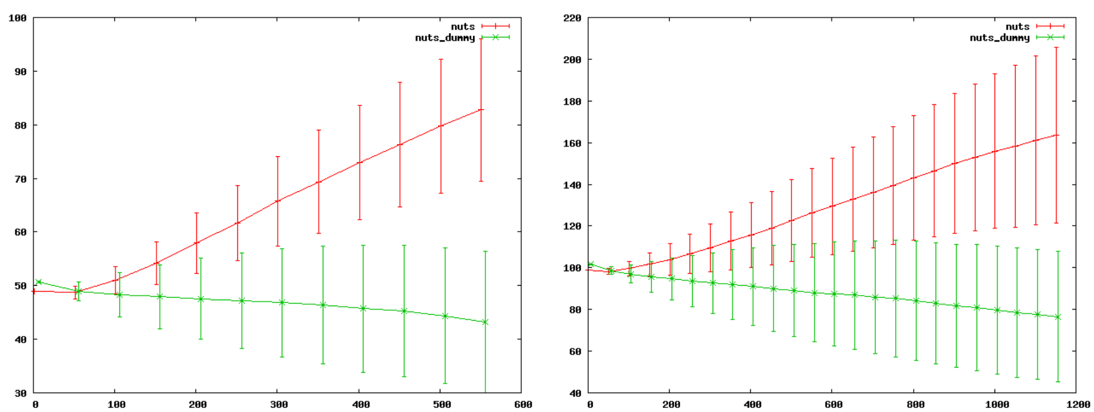
(a) N = 150 (b) N = 300

Figure 2: $AlwaysDefect$ vs $nuts$ in scenario 1 (AD, nuts)



(a) N = 150 (b) N = 300

Figure 3: $AlwaysDefect$ vs $nuts$ in scenario 2 (AC, AD, nuts)



(a) N = 150 (b) N = 300

Figure 4: $nuts$ vs $nuts\_dummy$ in scenario 3 (AD, nuts, nuts_dummy)

($t = 51.1728, df = 712.562, p\text{-}value < 2.2e-16$) to test on the difference of the population representation of both strategies in the end of runs in Scenario 1. These results show that significant differences appeared here (The mean population size of *nuts* was 183.99 and the mean population size of $AlwaysDefect$ was 28.41). Figure 2 shows the differences over the whole runs.

But if we look at these two strategies in the 2nd scenario, where also an always-cooperating strategy was involved, the picture is different. $Nuts$ loses to $AlwaysDefect$. We conducted another Welsh Two-Sample T-Test like the above, only regarding the second scenario ($t = 28.7613, df = 1024.961, p\text{-}value < 2.2e-16$). Here, the $AlwaysDefect$ strategy will significantly outnumber the nuts strategy (the mean population size of *nuts* was 63.47 and the mean population size of $AlwaysDefect$ was 142.78. Figure 3 shows the differences over the whole runs.

We conclude that defecting strategies profit much more from the presence of "dumb", often-cooperating strategies than the *nuts* strategy. In the context of the Prisoners Dilemma, this means that our strategy doesn't exploit cooperators as much as defectors do. Our strategy should recover once pure cooperators die out (and then the setting is the same as in scenario 1).

- **Hypothesis 3: Information chains increase the success of an informed strategy.**
  This hypothesis holds. We conducted a Welsh Two-Sample T-Test ($t = 27.082, df = 1032.923, p\text{-}value < 2.2e-16$) to test on the difference of the population representation of both *nuts* - strategies in the end of runs in Scenario 3 (AD, *nuts*, *nuts_dummy*). The results show with strong significance that the *nuts* strategy that makes use of the information chains will outnumber the "dumb" *nuts_dummy* - strategy. The mean of *nuts* was 125.75 and the mean of *nuts_dummy* was 58.78. Figure 4 shows the differences over the whole runs.

5 Conclusions

We proposed and implemented a distributed reputation system, which evolves a Small World Network. Under various circumstances, we tested the performance of a strategy that is generally forgiving but tries to gather information via the network. We showed some simple cases in which it is successful and which cases can be problematic in the short term. By evolving a Small World Network, we showed that this approach can be computationally feasible.
Of course, further research would deal with more various strategy scenarios and experiment with performance against "lying strategies" and possibly cartels. Regarding the distributed nature of the framework, in which trust information is only passed along local connections, we suspect this to be a fruitful path of research.

References

[1] A. Abdul-Rahman and S. Hailes. Supporting trust in virtual communities. In *HICSS*, 2000.

[2] G. Abramson and M. Kuperman. Social games in a social network. *Phys. Rev. E*, 63(3):030901, Feb 2001.

[3] R. Axelrod. *The Evolution of Cooperation (summary)*. Basic Books, New York, 1984.

[4] T. S. Ellis and X. Yao. Evolving cooperation in the non-iterated prisoner's dilemma: A social network inspired approach. *IEEE Congress on Evolutionary Computation*, September 2007.

[5] N. Höning and T. Kozelek. *http://tomik.krabicka.net/stuff/nt_aso_session_4/table.html*. January 2008.

[6] J. Kleinberg. Complex networks and decentralized search algorithms. 2006.

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

[8] C. O'Riordan. Forgiveness in the iterated prisoner's dilemma. *UK Multi-Agent Systems Workshop*, 2001.

[9] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 1944.

[10] J. D. Watts and H. S. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393, June 1998.