

# Topological evolution of trust game theory networks

J. Mario Siqueiros-García

January 8, 2015

## 1 Introduction

A problem of game theory and cooperation on networks is to find how cooperation can emerge from selfish agents [1]. Our model is an implementation of game theory on complex networks. As the majority of models of game theory on networks do, we compute each agent's and the whole network utility or fitness (according to a pay-off matrix), but we do different in computing as well the trust among agents (represented in the weight of the link  $w_{i,j}$ ) and the global trust for all the network.

The purpose of the model is to have a better understanding of the behavior of a network of selfish agents when trust among them is also an important variable. Trust can increase or decrease as it depends on the agent's behavior and the state of its neighbors.

Though our model is inspired in Hopfield networks as we are interested in optimizing global utilities (fitness and trust), there are many differences to consider. 1) Hopfield networks are coordination, Ising-based models (agents or neurons states can be 1 or  $-1$ ). Our model is based on game theory (the states of the agents are *defect* or *cooperate*). 2) Because of point 1, in a Hopfield network optimality can be found when all agents are in state 1 or in state  $-1$ . This is not the case in our network because all being *defective* does not lead to an optimal state, only all *cooperative* will do the job, actually all defective will lead to the lowest possible average fitness, while all cooperative will produce the highest [1]. For this reason, we are interested in reaching a configuration in which global fitness and trust remain positive.

## 2 Model

### 2.1 Structure

Being a game theory based model, agents in the network can be *cooperative* or *defective*. We implement the model using different topologies: full-connected network (as in a Hopfield network), erdos-renyi random network, Watts-Strogatz small-world and Barabasi-Albert scale-free networks. Distribution of *cooperative* and *defective* agents in the network is random.

The connections have the following restrictions:

- $w_{ii} = 0 \forall i$  (No node is connected to itself).
- $w_{ij} = w_{ji}, \forall i, j$  (Connections are symmetric).

### 2.2 Updating

Agents have a fitness or utility which can increase or decrease according to their state and the state of their neighbors, following a pay-off matrix:

	Cooperate	Defect
Cooperate	1, 1	-2, 2
Defect	2, -2	0, 0

Every node has a fitness record. Fitness for node  $i$  is the sum of all the outcome values of the pay-off matrix when  $i$  playing with its neighbors  $j$ .

$$f_i = \sum pm_{ij}$$

Where  $pm$  stands for payoff matrix.

Every pair of connected nodes has a weight that we have conceptualized as *trust* and its value change following a payoff matrix:

	Cooperate	Defect
Cooperate	2	-1
Defect	-1	-2

Each node has a trustability  $\tau$  record that is the sum of the weight of all its links with its neighbors  $w_{ij}$ .

$$\tau_i = \sum w_{ij}$$

Agents update their state according to the following rules:

$$S_i = \begin{cases} S_{i(t)} & \text{if } f_i/\tau_i > \theta, \\ \neg S_{i(t)} & \text{otherwise.} \end{cases}$$

In this case, the agent is sensitive to the changes in fitness (so to keep the rule in line with selfish behavior of game theory). If the value of the fitness of  $i$  is lower than its trustability, then the  $i$  will change to a different state.

In the following case the agent is sensitive to the changes in trustability (different to the rule of selfish behavior of game theory, but perhaps the agent is sensitive to its own reputation). If the value of the trustability of  $i$  is lower than its fitness, then the  $i$  will change to a different state.

$$S_i = \begin{cases} S_{i(t)} & \text{if } \tau_i/f_i > \theta, \\ \neg S_{i(t)} & \text{otherwise.} \end{cases}$$

### 3 Global values

We keep track of the whole network fitness and trust <sup>1</sup>. Though fitness cannot be optimized as in a Hopfield network, we want to find a way in which trust can remain positive.

Global fitness is defined as:

$$F = \sum f_i/N$$

Where  $N$  is the number of agents in the network. Global fitness is seen as fitness average (nowak2006)

Global trust:

$$T = 1/2 \sum w_{ij}$$

---

<sup>1</sup>Remember that our model, being close to the concept of Evolutionary Game Theory, it cannot be seen as an optimizing strategy for fitness.

The sum of the weight of all links is multiplied by .5 so to avoid counting the value of each link twice.

## References

- [1] Nowak, Martin A. (2006) Five rules for the evolution of cooperation. *Science*, 314:1560-1563.