# ITERATED PRISONER'S DILEMMA ON AN ADAPTIVE NETWORK WITH CONTINUOUS LINKS

MENGTING LEI

Abstract. Previous mathematical studies of Prisoner's Dilemma have been focused on simulating this game in networks with discrete links, where an agent plays equally against all its opponents in one time step. Such models could be more realistic if we consider the limited resources of each individual. Therefore in this research, we examine the game by allowing each agent to divide its resources into different portions for different opponents. The long-time behavior of the agents and the distribution of their connections to individual opponents are studied.

## 1. Introduction

Prisoners' Dilemma (PD) is a classical example of game theory. In the original setting that was framed by Merrill Flood and Melvin Dresher, the police are convicting two suspects of a crime but do not have adequate evidence. As a result, the two players are encouraged to testify their confederate for committing the crime (defect), although they are also given the option to remain silent (cooperate). There are three assumptions for the game: (1) Both suspects make rational choices. In other words, they will always seek to maximize their benefit and minimize their penalty. (2) The two suspects are not loyal to each other. Therefore they only consider about their own outcomes. In addition, altruism is not possible. (3) There is no communication between the two suspects so an individual does not know which strategy his or her opponent will choose.

The final result in terms of the duration of imprisonment depends on the decisions of both players. The four possible scenarios and the corresponding results in the original setting are presented in *Table 1.* For example, if one suspect plays cooperate (we refer it as C in the following text) while the other defects (we refer it as D), the D player will be set free and the C player will be kept in jail for 3 years.

| | Player 1 cooperates | Player 1 defects |
|---|---|---|
| Player 2 cooperates | -1, -1 | 0, -3 |
| Player 2 defects | -3, 0 | -2, -2 |

*Table 1*

In the iterated version of PD, an individual plays against the same opponent repeatedly. In this way a player is able to adjust his or her strategy according to previous outcomes. While there is no best strategy for the one-time PD game, empirical results show that there are some strategies that lead to better average payoffs in the long run. The tit-for-tat strategy, for example, has long been deemed as one of the simplest winning strategies (Axelrod, 1980).

In 2005, Zimmermann and Eguíluz simulated iterated PD on an adaptive network, where the players are represented as nodes and their interactions as links. In the model, nodes are allowed to play against multiple opponents with the same strategy simultaneously. As a result, they introduced the concept of "best neighbor", which is the neighbor with the highest aggregate payoff. To simulate these rules in the computer, each time step is divided into the following three stages.

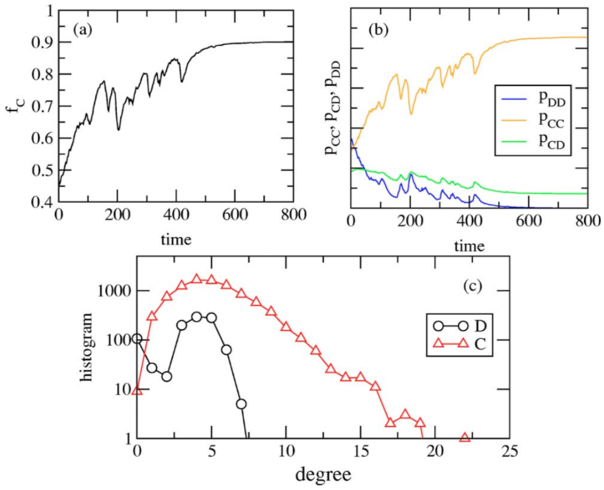
Stage I: Each node plays the same strategy to all its neighbors and collects payoffs from all its neighbors according to the designated payoff matrix in *Table 2*. The $\beta$ value is a variable that satisfies $1 < \beta < 2$.

| | C | D |
|---|---|---|
| C | 1, 1 | 0, $\beta$ |
| D | $\beta, 0$ | 0.1, 0.1 |

*Table 2*

Stage II: Each node imitates the strategy of its best neighbor only if the aggregate payoff of the node is less than that of its best neighbor. If multiple neighbors of a node have the same highest payoffs, then the best neighbor is randomly selected from them.

When the population size $n = 10,000$, the percentages of C nodes and C-C links always increase until an equilibrium around 90% occurs. In contrast, both C-D and D-D links decrease until the respective equilibrium values are reached. Ultimately, D-D links are going to become extinct.




*Graph 1*

In real life, it is difficult to give an example of an individual with unlimited resource, which is seen in this model as a node can technically play against 9,999 opponents at the same time. The resource can be in the form time, attention, and energy. Therefore in our research, we try to modify the original model to include considerations of the limited resource of individuals.

## 2. Our Model

In our model, we first generate a network with designated numbers of nodes and links. The initial percentage of C nodes is $c$%. Link weights are randomly assigned such that the sum of link weights of one node does not exceed one. Different from Zimmermann and Eguíluz's model, all nodes perform the following three stages in each time step.

Stage I: Each node still collects payoff from all its neighbors according to the payoff matrix in $Table 2$. However the total payoff $\mu_i$ of a node $i$ at time $t$ is calculated as the sum of the product of payoff $s_j$ node $i$ receives from neighbor $j$ and the respective link weights $x_ij$. $V(i)$ is the set of neighbors of node $i$.

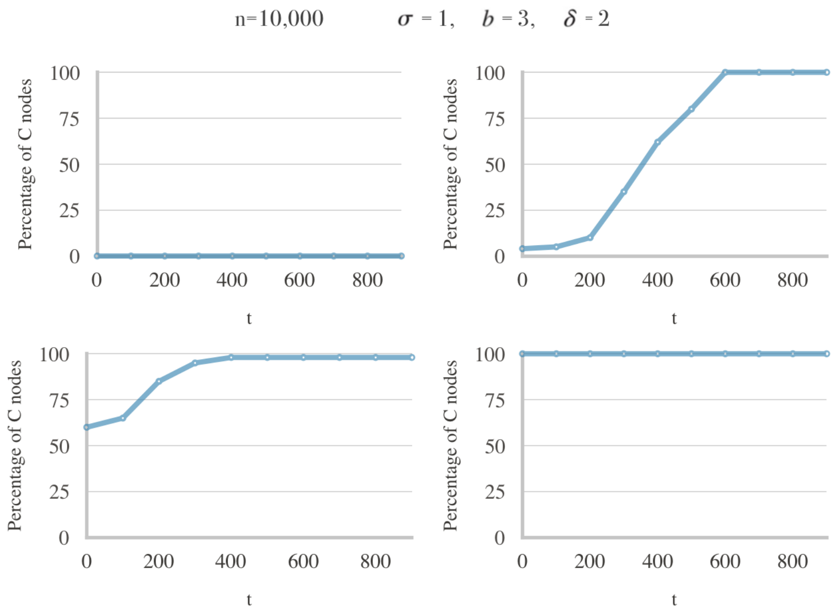
$$\mu_i = \sum_{j \in V(i)} s_j * x_{ij} \tag{2.1}$$

Stage II: Each node imitates the strategy of its best neighbor only if the aggregate payoff of the node is less than that of its best neighbor. If multiple neighbors of a node have the same highest payoffs, then the best neighbor is randomly selected from them.

Stage III: Increase all C-C links by multiplying their link weights by $\epsilon = 1.1$. Decrease all D-D links by dividing their link weights by $\epsilon = 1.1$. Our goal is to remain the sum of link weights of one node not exceed 1. However we have not figured out a meaningful method to realize this rule. Therefore in the current simulation the sum of link weights of some nodes will eventually exceed 1. While we certainly want to improve this stage in the next step, our current code should produce similar results to those in the original paper, thus testing if our code has any bugs.

## 3. Results

After running the simulation under the same conditions as the 2005 paper, we observe the following phenomena. The percentages of C nodes and C-C links are both increasing monotonically to 100%. On the contrary, the percentages of D nodes, C-D and D-D links are all decreasing monotonically to 0%. Such results were surprising and we are still in the process of making sense of them.




*Graph 2*

## 4. Discussion of the results and the presentaion

To understand the global extinction of D nodes, first we may have to re-check the accuracy of our code. Is there any grammar mistake? Does the code fully execute our designated rules? Therefore in the next step, we want to simplify the model to the original one and make sure they yield the same results.

Regarding the presentation, I have to admit that I did not do a good job in conveying the reasoning behind the adjustments made upon the original model. Nonetheless, I was pleased to see that our classmates were interested in this subject and offered many constructive feed-backs. For example, many have responded that the demonstration of our rules could be clearer if we first explain it in a simpler model, say a network with population $n = 3$. Professor Li also approached me after class and asked me to consider about the cases when the parameters in the payoff matrix are assigned different values, such as the ones in $Table$ 3.

| | C | D |
|---|---|---|
| C | 1, 1 | 0, 10 |
| D | 10, 0 | 0.1, 0.1 |

*Table 3*

Surprisingly, after applying these conditions to the model, we saw a completely opposite result: the global extinction of C nodes. Analogically, we can possibly make other parameters in the game, such as $p$ and $\epsilon$, variables with certain restrictions and then see whether the results would be different. However at this time, we should first try to make sense of the differences between our results and the original ones since we are using the exactly same parameter values.

In the future, there are many possible advancements to our three stages. To name one, we can also make the node weights variables. We can impose a certain probability distribution, such as the Poisson distribution, to the initial node weights to model the wealth distribution in human society.

## References


Axelrod, R. (1980). *Effective choice in the prisonerâs dilemma.* Journal of Conflict Resolution, 24(1): 3-25.

Axelrod, R. (1980). *More effective choice in the prisonerâs dilemma.* Journal of Conflict Resolution, 24 (3), 379-403

Axelrod, R., Hamilton, W. D. (1981). *The evolution of cooperation.* Science, 211(4489), 1390-1396.

Kaznatcheev, Artem (2015). *Theory, Evolution, and Games Group.* Short history of iterated prisoner's dilemma tournaments.

Zimmermann, M. G. EguiĚluz, V. M. (2005) *Cooperation, social networks, and the emergence of leadership in a prisonerâs dilemma with adaptive local interactions* Science, 211(4489), 1390-1396.