Prisoner's Dilemma Game on Network

Masahiro Ono and Mitsuru Ishizuka

Graduate School of Information Science and Technology,
The University of Tokyo
7–3–1 Hongo, Bunkyo–ku, Tokyo 113–8656, Japan
{mono,ishizuka}@miv.t.u-tokyo.ac.jp

Abstract. We study on the Prisoner's Dilemma game on network to clarify the influence of the network structures on agent strategies, and vice versa. A model is proposed to treat an interaction between the agent strategies and the network formation process. In case of a fixed network, it is observed that the distribution as well as the propagation speed of an agent strategy depends on the network structure. In an experiment combining the agent evolution and the network formation, a novel network that has a few agents connected by all of other agents appears.

Keywords: game theory, prisoner's dilemma, small world network, network formation.

1 Introduction

An emergent mechanism of cooperative strategies has been studied with Prisoner's Dilemma game so far. This game has such a dilemma that a rational person selects "defect" regardless of the fact that the pareto optimum state is achieved if all the players select "cooperate". A question arises here why people in a real world often select a cooperative strategy each other.

To give an solution, a spatial structure was introduced to the game[1][2]. This is a model that players located on a spatial structure, such as a two-dimensional regular lattice, play the game with neighborhoods. The cooperative players appears over generations on the context of the evolutionary game.

In another area, a small world network which appropriately models the social network was recently proposed[3]. Then this network structure was adopted to the Prisoner's Dilemma game, and the dynamics of the game on the network was studied[4][5][6]. It was assumed that a structure of the network was static, though it is generally thought that it affects the players and vice versa.

With regard to the undirected graph with n vertices, there are $2^{n(n-1)/2}$ possible graphs. However, such characteristics as small world[3] or scale free[7] are observed regardless of a large number of possible graphs. A method was proposed to explain the generation of graphs with specific characteristics such as small world or scale free network.

Another approach for this explanation is to use a game theory. In this model, all agents, which correspond to vertices on a graph, are assumed to have perfect

D. Lukose and Z. Shi (Eds.): PRIMA 2005, LNAI 4078, pp. 33–44, 2009.

knowledge of the network and to select agents to be connected rationally for maximizing own gain. It is known that wheel- or star-type network, called Nash network, appeared after the iteration process of the link change[8].

The study combining the characteristics of the agents and the network formation haven't done enough so far. In this paper we propose a model that treats both the agents' strategy and the network formation. Specifically, we focus the Prisoner's Dilemma game, and dynamics of the agents' strategy and the network formation.

In the remainder of this paper, we first introduce the Prisoner's Dilemma and the small world network. We then describe the model we have designed for combined study of agent evolution and network formation. Experimental results are then presented and discussed. Finally we present our conclusions.

2 Backgrounds

In this section, we explain the Prisoner's Dilemma game and the small world network for the preparation of following sections.

2.1 Prisoner's Dilemma

Prisoner's Dilemma is the most popular game in the game theory, because it is an elegant model to express many social phenomena. The name and the typical payoff matrix of this game is due to Albert Tucker in 1950's.

In a symmetric two-player game, the payoff matrix of Prisoner's Dilemma is expressed as Table 1, where R,T,S,P represent Reward, Temptation, Sucker and Punishment respectively. Payoff relations ($T > R > P > S$, $2R > T + S$) exist among them, which raise a dilemma.

Assuming that each player is rational, both players in the game would select the defect strategy. The player 1 considers that he should defect and earn a higher payoff whichever the player 2 cooperates or defects. The player 2 would also defect over the same consideration. After all, each player defects, and (D,D) is an only Nash equilibrium in this game. However, this state is pareto inferior that it's not optimal for both players. This is why this game has a dilemma.

There is a contradiction between the mathematical solution and the real world where we often cooperate each other. This point has been studied for a long time,

Table 1. The payoff matrix of prisoner's dilemma

		Player2	
		C (cooperate)	D (defect)
Player 1	C	R, R	S, T
	D	T, S	P, P

($T > R > P > S$, $2R > T + S$)

and a spatial structure was introduced in the evolutionary game theory. Players are located on the spatial structure, for example, a two-dimensional regular lattice[1][2]. They evolve every generation after they play games between the neighborhoods. In this case, the emergence of a cooperative strategy is observed after some period that the non-cooperative strategy is dominant.

2.2 Small World Network

When we meet someone at first time, we often get to know there is a common friend and say "it's a small world!". The original concept of the small world network was derived from that phenomenon. In 1998, Duncan Watts defined the small world network using two characteristic parameters, i.e., *characteristic path length* and *clustering coefficient* [3].

A graph G consists of vertices and edges. $V = \{1, \ldots, n\}$ is a set of vertices and g is an adjacency matrix of G. $g_{i,j}$ for a pair of vertices $i, j \in V$ indicates an edge between i and j. If $g_{i,j} \neq 0$ then an edge exists. An edge is absent in case of $g_{i,j} = 0$.

Let $d(i, j, g)$ be a function which gives the length of the shortest path between i and j. The *characteristic path length* L is the average [1] of the shortest path length between any two vertices on the network. L is precisely expressed as Eqn. (1).

$$L = \frac{1}{n(n-1)} \sum_i \sum_{j \neq i} d(i, j, g) \tag{1}$$

$N^1(i, g) = \{k \in V | g_{i,k} \neq 0\}$ is a set of vertices adjacent to i. $E(i, g) = \{(j, k) | g_{j,k} \neq 0, k \neq j, j \in N^1(i, g), k \in N^1(i, g)\}$ is a set of the combinations of vertices in $N^1(i, g)$ when an edge exists between them. For simplicity, let $l_i = |N^1(i, g)|$ be the number of edges connected to i. The *clustering coefficient* C is defined as Eqn. (2), and indicates the extent to which vertices adjacent to any vertex v are adjacent to each other as average.

$$C = \frac{1}{n} \sum_i \frac{|E(i, g)|}{{}_{l_i}C_2}, \quad {}_{l_i}C_2 = l_i(l_i - 1)/2 \tag{2}$$

Assuming the numbers of vertices and edges on a graph are fixed, the network structure changes like Fig. 1 as the randomness of network changes. Every vertex is connected to its neighborhoods mutually $p = 0$, where p is a randomness parameter. Edges are changed stochastically as p increases.

In case of $p = 0$, it is a regular network, and both L and C are large. On the other side, it is a random network at $p = 1$, and L and C are small. In the middle between these two extremes, it is called small world network, where L is small and C is large.

The small world networks have been found in many areas so far.

[1] There are two definitions, the average or the median. The median is used in [4].

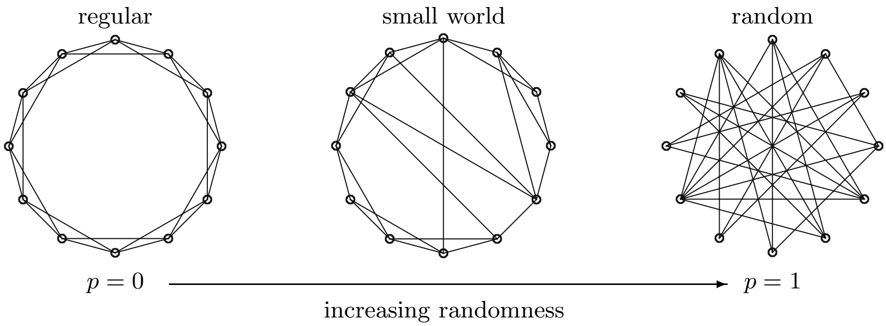

Fig. 1. Small world network

3 Proposed Model

We study here an evolutionary dynamics of the Prisoner's Dilemma game played by agents on the network. It is assumed that a network is changed by each agent which has a meta strategy and a link-change strategy in the evolutionary process. The dynamics of the agents' evolution and the network formation strongly depends on the agent's strategy.

In this section, we propose agent model consisted of two parts. The first part is a basic agent model to play the game. The second part is an advanced model including a network formation function.

As long as it is not noted otherwise the expressions of vertex, agent, node, or player are denoted as agent, and the expressions of edge or link are denoted as link.

We design agent models as follows.

3.1 A Basic Agent Model

Agents are nodes on a networks and players of Iterated Prisoner's Dilemma (IPD) games. They have to select a move C or D at each game with their neighborhoods.

Each agent has a gene and selects a move according to its gene in a deterministic manner. The gene consists of 5-bit; 1-bit determines an initial move and other 4-bit expresses a rule to select a move. Assuming a first-order meta strategy, the information to input is PM (Previous Move) and OPM (Opponent's Previous Move), and each bit of the gene corresponds to possible $4(= 2^2)$ cases.

Table 2 shows several rules of a first-order meta strategy. The agent selects a move regardless of previous moves in case of All C and All D. Tit-for-Tat (TFT) created by Anatol Rapoport is a famous strategy, which repeats opponent's previous move. This strategy is known as a winner of the famous tournament[9]. Pavlov strategy[10] is also known as "Win-Stay, Lose-Shift", which selects the opposite move from previous move when the agent couldn't earn high gain.

This gene-coded model[11] is general in the evolutional game.

Table 2. Meta strategy examples

PM	OPM	strategy examples			
		All C	All D	TFT	Pavlov
C	C	C	D	C	C
C	D	C	D	D	D
D	C	C	D	C	D
D	D	C	D	D	C

PM: Previous Move, OPM: Opponent's Previous Move.

3.2 An Advanced Agent Model with Network Formation Procedure

In addition to the basic agent model, an advanced model with a network formation procedure is introduced mathematically in this subsection.

This is a model that a person in the real society finds new people and changes a link to a new better person from the current worst person.

An assumption that agents have perfect knowledge and information in the simulated world is inappropriate for a multi-agent simulation. Agents with bounded rationality are assumed in this proposed model, where the agents can only know information of their first and second neighborhoods.

The set V includes n agents, and this elements are the players of IPD. Each agent has m links, can attach them to other agents and detach them in their intention. The adjacency matrix g at generation t is denoted by $g(t)$, which is expressed as Eqn. (3) for a pair $i, j{\in}V$.

$$g(t)_{i,j} = \begin{cases} 1 & j \text{ is connected by } i \\ 0 & \text{absence} \\ -1 & i \text{ is connected by } j \end{cases} \tag{3}$$

The value indicates an owner of the edge; that is, j is connected by i in case of $g(t)_{i,j} = 1$, and i is connected by j in case of $g(t)_{i,j} = -1$. This difference doesn't affect an interaction over the link.

The above mentioned $N^1(i, g)$ is also extended to $N^1(i, g(t))$. $N_o^1(i, g(t))$ is the set of agents connected by the link which i owns. It's defined as $N_o^1(i, g(t)) = \{k{\in}V | g(t)_{i,k} = 1\}$ precisely and is the subset of $N^1(i, g(t))$. A set of second neighborhoods of i, which is a set of agents connected by $j{\in}N^1(i, g)$, is denoted by $N^2(i, g(t)) = \{k{\in}V | k{\in}N^1(j, g(t)), j{\in}N^1(i, g(t))\}$.

IPD games are played by the agents for each combination of the agents having a direct connection on the network. Assuming i earns payoff $p_{i,j}$ as a result of the game with j, the average payoff of i is $\bar{p_i} = \sum_{j \in N^1(i,g(t))} p_{i,j}/l_i$.

After all the games have played, each agent performs a link change procedure which is a metaphor of a physical movement. This procedure consists of two steps. The first step is to detach (Eqn. (4)) and the second step is to attach (Eqn. (5)). With regard to i, it is expressed as follows,

$$N_o^1(i, g(t+1)) = N_o^1(i, g(t)) - N_i^{del}(\in N_o^1(i, g(t))) \tag{4}$$
$$N_o^1(i, g(t+1)) = N_o^1(i, g(t)) + N_i^{add}(\in N^2(i, g(t))) \tag{5}$$

N_i^{del} is selected and then detached from i. Also N_i^{add} is selected in $N_o^1(i, g(t))$ and then attached to i. The agents have to evaluate candidates to detach or attach. Assuming j is the adjacent agent and k is the second neighborhood, the indices of the agents for the evaluation with respect to i are follows:

- **for first neighborhoods**
  $U(i, j) = \{l_i, l_j, p_{i,j}, p_{j,i}, \bar{p_i}, \bar{p_j}\}$
- **for second neighborhoods**
  $V(i, j, k) = \{l_i, l_j, l_k, p_{i,j}, p_{j,i}, p_{j,k}, p_{k,j}, \bar{p_i}, \bar{p_j}, \bar{p_k}\}$

An index including a difference is introduced because the above two indices are not enough. An evaluation function is also introduced as follows.

- **generalization(including the difference)**
  $W(U) = \{r - s | r \in U, s \in U \cup \phi\}$
  $W(V) = \{r - s | r \in V, s \in V \cup \phi\}$
- **evaluation function**
  $H = \{\text{argmin}, \text{argmax}, \text{random}\}$

The agents have gene-coded h and w, where $h \in H$ and $w \in W$. The link change procedure depends on the gene as follows.

N_i^{del} is selected by Eqn. (6).

$$N_i^{del} = H_x \left(W(U(i, x)) | x \in N_o^1(i, g(t)) \right) \tag{6}$$

N_i^{add} is selected by Eqn. (7), which evaluates the second neighborhoods set at once, or Eqn. (8), which evaluates the first neighborhoods twice. If Eqn. (8) is used, the agent also has a gene containing $h' \in H$ and $w' \in W$.

$$N_i^{add} = h_x \left(w(V(i, j, x)) | x \in N^1(j, g(t)), j \in N^1(i, g(t)) \right) \tag{7}$$
$$N_i^{add} = h'_x \left(w'(U(j, x)) | x \in N^1(j, g(t)) \right)$$
$$\text{where} \;\; j = h_j \left(w(U(i, j)) | j \in N^1(i, g(t)) \right) \tag{8}$$

If the set of N_i^{del} and N_i^{add} includes several elements or in case of $h = \text{random}$, an element is selected at random from the elements in the set.

4 Experiments and Discussion

In this section, we show two experiments. The first one shows the influence of the network structure on the agent, and the second one shows the network formation combined with the agent evolution.

Basically our experiments are performed as follows. All of the agents are located on a network. They play games. Then they evolve and change links on the network in one generation. An experiment here continued up to 2000 generations.

4.1 Experiment 1

In this experiment, we focus on the influence of the network structure measured by several indices, especially the distribution of the strategies on the network.

A simulation is performed as follows. The small world network g, which random parameter p is given as an initial value and fixed during the simulation, is structured at first. It consists of n agents each of which have m links. These parameters are fixed at $(n, m) = (400, 3)$.

Agents act as the players of the IPD game. They play games with other agents having a direct connection on the network.

They are basic agents and have only a 5-bit gene of playing the game (cf. Table. 2). A unit of a game is an IPD game, which is iterated 100 times. This is not a problem because the agents don't use backward induction due to the fact that they don't know when the IPD game ends in this simulation.

A generation change happens after all the games are finished in every generation. 10% of agents are to die stochastically according to their gain earned in the generation. Then a new agent is to be located there instead of the dead agent; the new agent is generated as a copy from the agent earned most in the first neighborhoods of the dead agent. The mutation of the gene is generated at the probability of 0.02% in copying.

The parameters in the simulation are set as $(T, R, P, S) = (5, 3, 1, 0)$ in Table 1. Noise is introduced to this simulation such that a move selected by each agent is reversed stochastically according to a probability $noise$, 0.02.

Result

An example of the evolutional dynamics of the agent's strategies is shown in Fig. 2. Space occupation (SO) represents the extent how much the strategy occupies the space on the network. SO is defined as Eqn. (9).

$$SO_i(t) = \frac{N_i(t)}{n} \tag{9}$$

where $N_i(t)$ is the number of agents which have the gene of strategy i on the network at the generation t.

It is observed that the cooperative strategies, such as Pavlov strategy in this example, tend to survive. This tendency can be also seen in other cases.

Average gains in case of $p = 0, 1$ are shown in Fig. 3. The average gain falls from the initial value once, then rises and saturates at a point in both cases. At the minimum point, non-cooperative strategies are dominant in the generation. But then they become extinct because they don't earn much with their non-cooperative neighborhoods. The timing that the cooperative strategies increase corresponds to the timing that the average gain rises. At steady state, the cooperative strategies are dominant. It is clear that the speed of convergence to steady state and the saturation point depend on the structure of the network. Lower speed and higher average gain are observed on the network with higher randomness. There are opposite characteristics on the network with lower randomness.

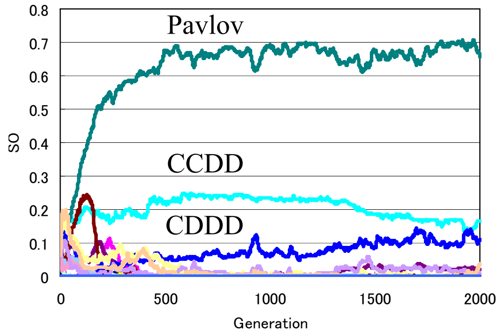

Fig. 2. An example of the generation change($p = 0.003$)

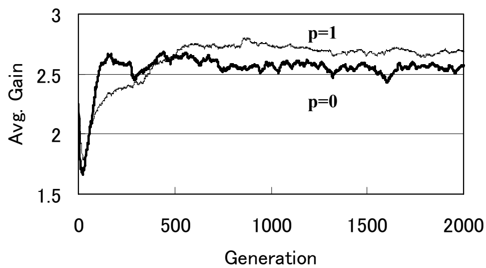

Fig. 3. Average gain($p = 0, 1$)

Next, the Time-Space occupation of strategies is studied. Time-Space occupation (TSO) is defined as Eqn. (10).

$$TSO_i = \frac{1}{T_{max}} \sum_{t}^{T_{max}} SO_i(t) \tag{10}$$

where T_{max} is the maximum number of generations.

In Fig. 4, TSO is on the vertical axis and the rank of strategy , ordered by TSO, is on the horizontal axis. The dominant strategy is located at the left side in the figure. However, dominant strategy here doesn't mean the common "dominant strategy" in the game theory, but which TSO is highest in the simulation. There are 16 strategies on condition that we ignore the initial move of the gene. This TSO follows power-low distribution with a cut-off at about 10th rank, and the inclination varies according to the *clustering coefficient* of the network. In particular, the value at 1st rank varies much.

The TSO of the top strategy in each network is focused in Fig. 5. TSO and the average gain get higher as the *clustering coefficient* C goes lower. Attenuation starts from the range of the small world network to the regular network. The small world network is considered as the network that optimizes the cost to move

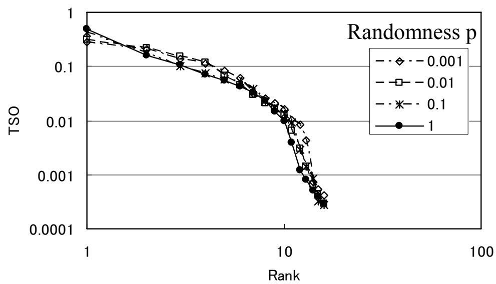

Fig. 4. Time-Space occupation

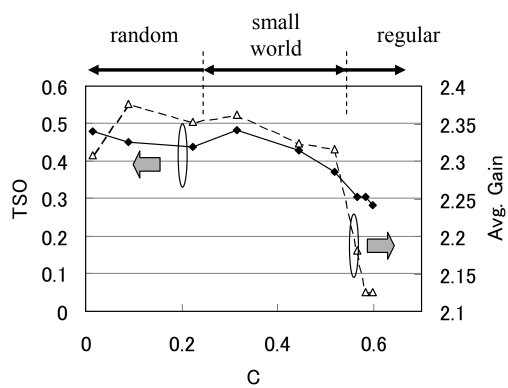

Fig. 5. Time-Space occupation of the dominant strategy and average gain

over the distance in the range where the dominant strategy keeps the biggest occupation.

As a result of the experiment 1, it is clarified that a cooperative strategy such as Pavlov survives regardless of the network structure. The distribution of strategies also depends on the network structure, and the 1st rank strategy is strongly dominant when the *clustering coefficient* C is lower.

4.2 Experiment 2

We focus on the network formation process in this subsection.

All agents here follow the advanced model and have the gene of playing the game and deciding the procedure to change their own links.

They play games with their neighborhoods in the conditions that are the same as experiment 1. After all the games finish, a generation change and a link change occur in every generation.

The initial network is regular ($p = 0$ in the small world network model) when the simulation starts at $t = 0$. The values of *noise* and *mutation* are fixed at 0, because we examine an ideal case at first.

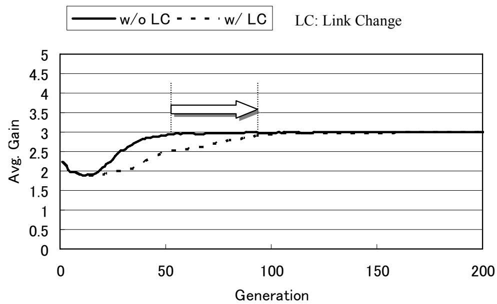

Fig. 6. Average gain

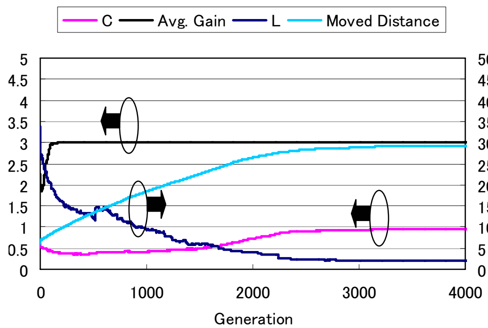

Fig. 7. Evolutional dynamics

Result

The comparison between simulations with and without the link change procedure is shown in Fig. 6.

It is identical in both cases that the dominant strategy changes from non-cooperative one to cooperative one. But the convergence speed in case without the link change procedure is faster than the other. In this process, the agent with the non-cooperative strategy performs the link change procedure to connect the agent with the cooperative strategy to defect them; on the other hand, the agent with the cooperative strategy performs the procedure to make a steady group of their cooperators.

According to the tendency of the connection, the agents with the non-cooperative strategy have only their own links. On the other hand the cooperators have links connected by other cooperators as well as ones they own. The cooperators are survivable at high probability because they earns in proportion to links with other cooperators.

Dynamics of another indices are shown in Fig. 7. There is a monotonic decrease in the *characteristic path length* L. The fact that L increases once in this figure is because an agent with many links died accidentally.

The *clustering coefficient* C decreases once, then gradually increases and approaches to 1. It doesn't seem natural. This network is structured like Fig. 8,

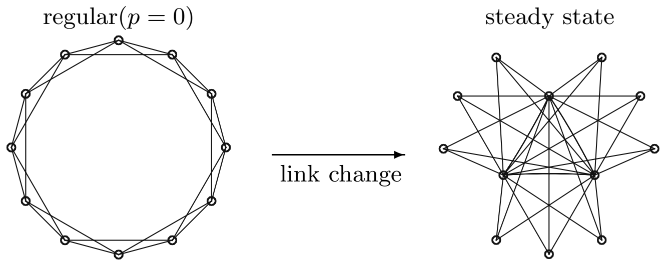

Fig. 8. Network in steady state

which consists of a few *super agents* and many normal agents. A few *super agents* have many links and are connected each other. The other many agents have connections to the *super agents* in this network.

The number of agents that an agent has had the connection before is defined as *moved distance*. The *moved distance* converges to about 30 as the *clustering coefficient* approaches to 1. The *moved distance* converges when many agents have a gene of selecting N^{add} due to the number of links l.

Generally, the agents which select agents to earn more tend to survive. It is natural in early generations. After some generations, agents with many links mean that they tend to cooperate. There are no differences between the gain and the number of links as an index for the evaluation. That's why agents that look for agents with many links also survive easily.

It is unpredictable which type of agents are dominant in this experimental condition. It depends on a dynamic process, and the first dominant type may survive. The gain of neighborhoods is more unstable than the number of links which the neighborhoods have. The *moved distance* increases monotonously in the case that the dominant agent looks for agents with more gain. Otherwise, the *moved distance* is saturated in the steady state.

With regard to the method of selecting N^{del}, the agents that select agents which he couldn't earn much tend to survive.

5 Conclusions

We studied the Prisoner's Dilemma game on network to clarify the influence of the network structures on the agents strategies and vice versa.

Our contribution in this paper is as follows. We here proposed the agent model with strategy evolution and network formation functions. As a result of experiments, it has been shown that the distribution of the strategies depends on the network structure. Also it has become clear that an introduction of movement makes the speed of strategy convergence slow and forms a strange network with a few super agents in the steady state.

Overall, our proposed model and the experiments of this paper give an explanation to the question why people in a real society select a cooperative strategy each other even in the case that a non-cooperative strategy is advantageous for obtaining temporal benefit.

References

1. Nowak, M.A., May, R.M.: Evolutionary games and spatial chaos. Nature 359, 826–829 (1992)
2. Lindgren, K., Nordahl, M.G.: Evolutionary dynamics of spatial games. Physica D 75, 292–309 (1994)
3. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature 393, 440–442 (1998)
4. Watts, D.J.: Small Worlds. Princeton University Press, Princeton (1999)
5. Abramson, G., Kuperman, M.: Social games in a social network. Physical Review E (2001)
6. Masuda, N., Aihara, K.: Spatial prisoner's dilemma optimally played in small-world networks. Physics Letters A 313, 55–61 (2003)
7. Barabasi, A., Albert, R.: Emergence of scaling in random networks. Science 286, 509–512 (1999)
8. Bala, V., Goyal, S.: A noncooperative model of network formation. Econometrica (2000)
9. Axelrod, R.: The Evolution of Cooperation. Basic Books (1985)
10. Nowak, M.A., Sigmund, K.: A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. Nature 364, 56–58 (1993)
11. Lindgren, K.: Evolutionary phenomena in simple dynamics. In: Artificial Life II: Proceedings of the Workshop on Artificial Life held, Santa Fe, New Mexico, February 1990, pp. 295–312 (1992)