Physica A 465 (2017) 153–158

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Role of delay-based reward in the spatial cooperation

Xu-Wen Wang [a], Sen Nie [a], Luo-Luo Jiang [b,*], Bing-Hong Wang [b,c,d], Shi-Ming Chen [a]

[a] School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang, Jiangxi 330013, PR China
[b] College of Physics and Electronic Information Engineering, Wenzhou University, Wenzhou, Zhejiang 325035, PR China
[c] Department of Modern Physics, University of Science and Technology of China, Hefei, Anhui 230026, PR China
[d] School of Science, Southwest University of Science and Technology, Mianyang, Sichuan 621010, PR China

HIGHLIGHTS

- We introduce delay-based reward into spatial cooperation, in which individuals can select various reward amounts on the basis of delay.
- We find that the intermediate differences of reward among kinds of delays are most important contributions to cooperation.
- We show that the individuals who select the fastest and smallest reward in evolving have strongest competition.

ARTICLE INFO

Article history:
Received 26 April 2016
Received in revised form 21 June 2016
Available online 16 August 2016

Keywords:
Spatial cooperation
Prisoner's dilemma game
Delay-based reward

ABSTRACT

Strategy selection in games, a typical decision making, usually brings noticeable reward for players which have discounted value if the delay appears. The discounted value is measure: earning sooner with a small reward or later with a delayed larger reward. Here, we investigate effects of delayed rewards on the cooperation in structured population. It is found that delayed reward supports the spreading of cooperation in square lattice, small-world and random networks. In particular, intermediate reward differences between delays impel the highest cooperation level. Interestingly, cooperative individuals with the same delay time steps form clusters to resist the invasion of defects, and cooperative individuals with lowest delay reward survive because they form the largest clusters in the lattice.

1. Introduction

Game theory that seeks to understand the strategy selection of individuals to maximize the payoffs in specific condition, has attracted more attention [1,2]. Various researches about evolutionary games have focused on strategy selection mechanisms in repeated games [3,4]. Individual within evolutionary games framework, selects the strategy of the next round by referring to the past payoff and opponent's strategy, such as cooperation or defection. For instance, tit-for-tat [5–9], win–stay and lose–shift [10,11] have been explored. Games in structured population, it is found that spatial topology favors cooperation, where individual interacts with its direct neighbors and gains a payoff [12–16]. By means of strategy adoption rules, such as deterministic [17,18] and pairwise-comparison [19–23], individual adjusts its strategy by comparing the performance of neighbors' and itself. Some features, including memory [24], reputation [25,26], diversity [27–30] have

* Corresponding author.
E-mail address: jiangluoluo@gmail.com (L.-L. Jiang).

http://dx.doi.org/10.1016/j.physa.2016.08.014
0378-4371/© 2016 Published by Elsevier B.V.

been incorporated into spatial games. In addition, punishment and reward [31,32], including antisocial pool reward [33], rewarding with links [34] and adaptive rewarding [35] have also been examined from both experimental and theoretical aspects [36,37]. Previous results also show that incorporating of punishment and reward groups into the original game model can produce interesting phase diagram [38].

Because individuals are driven by payoffs during the decision making process, the emerging of delayed and uncertain rewards may affect their strategies selection [39–42]. For example, when an individual is required to choose between a smaller reward available sooner and a larger reward available later, where the decision usually involves reward amount which is so called temporal discounting [43]. That is, individual discounts the value of reward based on the delay. The general formula of discounting in neurobiological and economic science is exponential, while the alternative form in psychology experiments is a hyperbolic $V = 1/(1+Kd)$, where V is the subject value for delay d, and K marks the discounting factor [44].

Players participating in evolutionary games, are also required to make decisions in each round, such as cooperation or defection. Especially, when cooperative players are rewarded on the basis of cooperative rounds. Then, from the perspective of temporal discounting in neurobiological and economic science, incorporating of delay will lead to varied strategy selections [45,46], such as cooperating once to obtain the soonest but smallest reward, or largest but delayed reward by cooperating with several rounds. And how this decision making affects the spatial cooperation, especially when individuals can adopt the strategy and reward selection of neighbors simultaneously. To address this problem, we explore the role of delay-based reward for decision making in spatial cooperation. Here, individuals can select either the larger but available later or smaller and sooner reward based on cooperative rounds. We assume individuals select the reward choices and strategies of neighbors at the same time. We show that, delay-based reward spreads of cooperation on lattice, random and small-world networks. The results prove that, temporal discounting impels the intermediate reward variation for different cooperative rounds to be most important contribution to the cooperation. We find that, cooperators who select the fastest and lowest reward have the strongest competition in the system, rather than who choose delayed and larger reward. This paper is organized as follows: Section 2 presents the model, Section 3 demonstrates numerical results, and the discussion is presented in Section 4.

2. Prisoner's dilemma game with delay-based reward

In the spatial prisoner's dilemma game (PDG), individuals locate on the sites of networks (Newman–Watts small-world networks and square lattice networks with periodic boundary conditions). Initially, each individual is designed either as a cooperator ($s_x = \mathrm{C}$) or a defector ($s_x = \mathrm{D}$) with equal probability, and it plays PDG with its nearest neighbors. For each individual, its payoff is composed of the total payoffs acquired from the neighbors in games, and cooperators can obtain an extra reward payoff r_{E}, which is related to the number of continuous cooperative steps. Following common practice of PDG, the parameter of reward for mutual cooperation $R = 1$, and both the punishment for mutual defection P and the suckers payoff S are set as $P = S = 0$, temptation to defect $T = b$ is the only payoff parameter in PDG. Meanwhile, we assign an integer parameter d to each cooperative individual to ensure they continuously select cooperation in the next d rounds, and d is selected randomly from the interval [1.0, 4.0] initially. Then, the formulation of reward r_{E} for player i with delay d is given by:

$$r_{\mathrm{E}}(i) = \frac{d_i(1 - \mathrm{e}^{-K})}{1 + Kd_i}. \tag{1}$$

Here, on the one hand, nonlinear term $d_i(1 - \mathrm{e}^{-K})$ of Eq. (1) guarantees that the longer the selection of cooperation, i.e. large d_i, the higher the reward. But individual needs to wait longer time to obtain the reward or large delay. On the other hand, denominator of Eq. (1) indicates that the reward $d_i(1 - \mathrm{e}^{-K})$ is also discounted greater by discounting function $V = 1/(1 + Kd_i)$ for larger d_i on the aspects of psychology and neurobiological science. Therefore, reward r_{E} is decided jointly by cooperative rounds and discounting effect. And K represents discounting factor. The larger the K, the more rapid the decay of reward with delay d. A delay $d = 1.0$ represents that the individual chooses to cooperate only once and gets the smallest reward within the shortest time. Given the total payoffs from the previous round, the evolutionary update rule used in this paper is Better-possess-chance strategy [47]. That is, individual x updates strategy by comparing the payoff P_{s_y} of a randomly selected neighbor y and its own payoff P_{s_x}: if $P_{s_x} > P_{s_y}$, individual x holds its strategy unchanged; else, x adopts the strategy of neighbor y's strategy with the probability:

$$W(s_x \leftarrow s_y) = \frac{P_{s_y} - P_{s_x}}{Dk_>}, \tag{2}$$

where $k_>$ is the largest between individual x's degree k_x and individual y's degree k_y, and $D = T - S$ for PDG. Once the strategy of neighbor y is adopted by x, the choice of reward delay d_x of x also converts into d_y. We set the network size $N = 10000$ for all simulations. The individuals update their strategies synchronously.

3. Results

Fig. 1 shows cooperation level of PDG accompanied with delay-based reward in decision making, for square lattice networks. We see that, cooperation level with delay-based reward is greatly improved, compared with the original PDG.