The “self-bad, partner-worse” strategy inhibits cooperation in networked populations

Zhang, Chunyan; Liu, Siyuan; Wang, Zhijie; Weissing, Franz J.; Zhang, Jianlei

Published in:
Information Sciences

DOI:
10.1016/j.ins.2021.11.041

Information Sciences 585 (2022) 58–69

Contents lists available at ScienceDirect

Information Sciences

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

The "self-bad, partner-worse" strategy inhibits cooperation in networked populations

Chunyan Zhang [a,b], Siyuan Liu [a,b], Zhijie Wang [a,b], Franz J. Weissing [c], Jianlei Zhang [a,b,*]

[a] Department of Automation, College of Artificial Intelligence, Nankai University, Tianjin 300071, China
[b] Tianjin Key Laboratory of Intelligent Robotics, Nankai University, Tianjin 300071, China
[c] Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen 9747 AG, The Netherlands

ARTICLE INFO

Article history:
Received 25 June 2021
Received in revised form 30 September 2021
Accepted 14 November 2021
Available online 22 November 2021

Keywords:
Evolutionary game
Cooperation
Strategy evolution
Spatial games
Complex network
Prisoner's Dilemma

ABSTRACT

The emergence and maintenance of cooperation is a popular topic in studies of information sciences and evolutionary game theory. In two-player iterated games, memory in terms of the outcome of previous interactions and the strategy choices of co-players are of great referential significance for subsequent strategy actions. It is generally recognized that there is no single simple and overarching strategy whereby one player X can unilaterally achieve a higher payoff than his opponent Y, irrespective of Y's strategy and response. In this paper, we demonstrate that such strategies do exist in diverse networked populations. More precisely, (i) such strategies can obtain a low payoff for the focal player, however, they also lead to an even lower payoff for that player's partner, in turn lowering benefits of the overall populations; (ii) they are capable of winning with a high probability against opponents with an unknown strategy; and (iii) they have a survival advantage and robust fitness in terms of evolutionary processes. We refer to these as the "self-bad, partner-worse" (SBPW) strategies. Results presented here add to previous studies on strategy evolution in the context of social dilemmas and hint at some insights concerning cooperation promotion mechanisms among networked populations.

1. Introduction

Cooperation is a ubiquitous phenomenon in nature. From the perspective of social development, cooperative behavior is conducive to improving the overall welfare of a society. However, to some extent, cooperation tends to reduce personal benefits and weaken competitive advantages of individuals. Such outcomes lead to a cooperative dilemma[24]. Evolutionary game theory provides an effective mathematical framework for describing and investigating this kind of problem[6,17].

The origins of evolutionary game theory can be traced back to the early 1960s, when Lewontin used game theory to study ecological phenomena as a way of investigating the role cooperation in evolutionary processes[4]. Some paradigmatic scenarios of cooperative dilemmas can be abstracted into various game models, such as the two-player Hawk Dove game and Stag Hunt game, and the n-player Public Good game[18,34]. Among these seminal game theory models, the most famous is the Prisoner's Dilemma (PD) game, which constitutes a powerful metaphor for describing conflicting situations in communities consisting of self-interested individuals[2,42]. In the PD game, if two players X and Y cooperate (C), then each earns R as a reward. If both defect (D), then each gets a scant payment P. While if one cooperates, the naive cooperator obtains T, and

* Corresponding author at: Department of Automation, College of Artificial Intelligence, Nankai University, Tianjin 300071, China.
E-mail address: jianleizhang@nankai.edu.cn (J. Zhang).

https://doi.org/10.1016/j.ins.2021.11.041

the selfish defector gets a larger payment S. To guarantee that the Nash equilibrium is mutual defection, the game must satisfy $T > R > P > S$. In addition, $2R > T$ is necessary to make mutual cooperation the globally best situation. At present, evolutionary game theory has been widely employed in a multitude of disciplines, such as economics, computer science, biology and politics[5,30,41,43]. M. Port et al. presented a socioecological model to explain the transition from solitary life to sociality in biological evolution[27], and F. Huang et al. developed a framework incorporating the adaptive mechanism of reinforcement learning to investigate the evolution of cooperative behaviors in the ever-changing group interaction environment[13].

In real biological systems and social evolutionary processes, interactions between individuals often occur more than once, which highlights the importance of research on iterated games[29]. Compared with a one-round game, interacting players may condition their strategies by taking into account their opponents' previous actions[40]. It is generally assumed that an agent with longer memory may have an advantage over a more forgetful opponent. However, it has been proved that for any strategy of the shorter-memory player X, longer-memory player Y's reward is precisely the same as X. Longer memory gives Y no advantage[28]. Hence, an important class of strategies called memory-one strategies is proposed. Individuals can make their strategy decisions referring to their partners' actions in the previous round. Label the four outcomes of the previous move as $xy \in (CC, CD, DC, DD)$, where C and D represent cooperation and defection respectively. Player X's strategy is $\mathbf{p} = (p_1, p_2, p_3, p_4)$, corresponding to the conditional probabilities playing C under each of the four previous outcomes. From Y's perspective, the strategy is analogously $\mathbf{q} = (q_1, q_2, q_3, q_4)$ in the order of $yx \in (CC, CD, DC, DD)$[12,21].

The "tit-for-tat" (TFT) and "win-stay, lose-shift" (WSLS) variants are two classic examples of memory-one strategies which are generally considered successful because they have strong fitness and can defeat other strategies with a high probability in the self-organized evolution [15,35–37]. TFT and WSLS strategies are denoted as $(1, 0, 1, 0)$ and $(1, 0, 0, 1)$ respectively. The player using the TFT strategy chooses to cooperate at the initial stage of the game, and imitates the action of its partner in the subsequent round[25,14]. It will defect if the co-player has defected, and cooperate if the co-player has cooperated. For the WSLS strategy, the player chooses cooperation in the first round and decides its strategy in subsequent rounds according to the partner's strategy. If both players adopt the same strategy, they will continue to cooperate in the next round. In contrast, if players choose different strategies, they will convert to defect[26,20]. In 2012, Press and Dyson putted forward a hybrid strategy in which a player can unilaterally set its opponent's score and moreover, can enforce an extortionate linear relationship between their rewards. This is referred to as a zero-determinant (ZD) strategy[28]. After that, Stewart et al. identified a set of generous strategies which replace extortionists and dominate in large populations by cooperating with others and forgiving defection[33], and further clarified that a diversity of behavioral choices could cause a population to evolve toward lower levels of cooperation[32]. In addition, the zero-determinant strategy can be applied to cloud computing research. Zhang et al. showed that a trusting party can regulate the actions of betrayers through a zero-determinant strategy, which insures the interests of honest people in the cloud computing environment[44].

In recent years, rapid developments in the study of complex networks have advanced our understanding of evolutionary game theory. Various population structures can be abstracted into network topologies composed of multiple interacting individuals[3,7,11]. Nodes on a complex network are equated with individuals in the real world and connections between nodes are regarded as a certain specific relationships between them[8,16,46]. At present, research on evolutionary game dynamics within complex networks mainly focuses on three trajectories: 1) the influence of network topology on the evolution of game dynamics; 2) the effect of different updating rules on network evolution given various network structures; 3) the co-evolution of game dynamics and network topology[45,23]. After years of development, numerous updating rules have been proposed based on a wide variety of assumptions, such as the "Birth–Death" updating process, the adaptive dynamics [1] based on imitation, as well as the Moran process, etc. In addition to these examples of updating rules, it is also feasible to refine the models discussed above through the introduction of feedback mechanisms[19,31] or other controls in order to improve the fit between systems and desired system states[39].

In diverging from deterministic strategies, such as cooperation and defection, a memory-one strategy $\mathbf{p} = (p_1, p_2, p_3, p_4)$ is more diversified, with each item shifting between 0 to 1, leading to difficulties in terms of analysis and statistics. Machine learning algorithms provide potentially useful approaches for dealing with this puzzle[22]. In this respect, cluster analysis is a method for categorizing data into different groupings based on the statistical similarity between data points[38]. Data belonging to the same cluster are relatively similar, while the data belonging to different clusters have greater statistical divergence. In addition, the dimensionality of the data reduction algorithm can be pre-processed for data with high-dimensional characteristics. It works by removing noise and unimportant information so as to retain the most crucial aspects of statistical patterning, which reduces the complexity of subsequent statistical analysis. This study utilizes principal component analysis (PCA), which is a common variety of factor analysis used for dimension reduction[10]. Given a sample set $D = \{x_1, x_2, \ldots, x_m\}$ with dimension d, we can use PCA to reduce the dimensions of these data to $d\prime$. PCA begins by centralizing all samples: $x\prime_i \leftarrow x_i - \frac{1}{m}\sum_{i=1}^{m} x_i$. Next, we calculate the sample covariance matrix $\mathbf{XX}^T$. Then, the covariance matrix $\mathbf{XX}^T$ is transformed by decomposition into a set of eigenvalues. Finally, we take the eigenvectors $w_1, w_2, \ldots, w_{d\prime}$, which correspond with the maximum eigenvalues for $d\prime$, and we output a projection matrix $\mathbf{W}^* = (w_1, w_2, \ldots, w_{d\prime})$.

In this paper, we explore the evolution of memory-one strategies for the Iterated Prisoner's Dilemma (IPD) game model within complex networks. Clustering and analyzing the dominant strategies evolved from networked populations, surprisingly, we find what we call a "self-bad, partner-worse" (SBPW) strategy. Such strategies work by making the profit of one's partner less than one's self and this also results in the reduction of the payoff for the population as a whole. We explore if

SBPW strategy can get the upper hand encountering a random strategy, examine the differences between SBPW and the various other dominant strategies, and as well compare the evolutionary performances of the TFT, WSLS and SBPW strategies. Our results indicate that SBPW strategies have an obvious evolutionary advantage within the IPD game model.

The rest of this paper is organized as follows: Section 2 presents the game model and network topologies used in this study, as well as the payoff calculation method and strategy updating rule of interactive individuals. In Section 3, we describe the generation process of dominant strategies, based on which the SBPW strategy is found with the aid of AGNES clustering algorithm and PCA dimension reduction method. Further explorations on the SBPW strategy and main results are provided in Section 4. Section 5 ends this paper with discussion and concluding remarks.

2. Model and methods

2.1. Model setting

Assuming that two players, X and Y, are playing the IPD game. The game result of the previous move is denoted as $XY \in (CC, CD, DC, DD)$, where C and D represent cooperation and defection respectively. X's strategy in the subsequent round is then $\mathbf{p} = (p_1, p_2, p_3, p_4)$, in which p_1, p_2, p_3 and p_4 denote the cooperation probability of X in this round given the game outcome in the previous move; and $\mathbf{q} = (q_1, q_2, q_3, q_4)$ denotes Y's strategy in the subsequent round.

The payoff matrix of PD game is set as shown in Table 1. The numbers in the matrix represent payoffs under the various strategy combinations of the two interacting players. For example, the payoff for the focal player is 3 if both individuals choose to cooperate, while if the individual cooperates but the co-player defects, then the payoff for that individual is 0.

2.2. Payoff calculation

According to the principles involved in a memory-one strategy, Player X's strategy can be represented as $\mathbf{p} = (p_1, p_2, p_3, p_4)$ and payoff matrix is $\mathbf{S}_X = (3, 0, 5, 1)$. Similarly, the corresponding terms for Player Y are $\mathbf{q} = (q_1, q_2, q_3, q_4)$ and $\mathbf{S}_Y = (3, 5, 0, 1)$ respectively. After one game interaction, the expected payoffs for Players X and Y can be calculated as Eq. 1.

$$\begin{aligned} S_X &= \frac{\mathbf{v}\cdot\mathbf{S}_X}{\mathbf{v}\cdot\mathbf{1}} = \frac{D(\mathbf{p},\mathbf{q},\mathbf{S}_X)}{D(\mathbf{p},\mathbf{q},\mathbf{1})} \\ S_Y &= \frac{\mathbf{v}\cdot\mathbf{S}_Y}{\mathbf{v}\cdot\mathbf{1}} = \frac{D(\mathbf{p},\mathbf{q},\mathbf{S}_Y)}{D(\mathbf{p},\mathbf{q},\mathbf{1})}, \end{aligned} \tag{1}$$

in which $\mathbf{v}$ is the stationary vector of Markov matrices $\mathbf{p}$ and $\mathbf{q}$. Besides,

$$\begin{aligned} \mathbf{v}\cdot\mathbf{f} &\equiv D(\mathbf{p},\mathbf{q},\mathbf{f}) \\ &= det\begin{bmatrix} -1+p_1q_1 & -1+p_1 & -1+q_1 & f_1 \\ p_2q_3 & -1+p_2 & q_3 & f_2 \\ p_3q_2 & p_3 & -1+q_2 & f_3 \\ p_4q_4 & p_4 & q_4 & f_4 \end{bmatrix}. \end{aligned} \tag{2}$$

In this respect, we define the expected payoff for each node of the networks as the sum of the payoffs obtained by the focal node from the games with all its neighbors divided by its neighbors' number.

2.3. Strategy updating rule

In this study, a Fermi function is used as a strategy updating rule. For any Player A with a neighboring set represented as M, the payoff for A in a given game with its neighbors is s_A. Following that game, Player A randomly chooses a neighbor B from neighboring set M and compares their payoffs s_A and s_B. Player A will then choose the strategy of Player B with probability p_{AB} in the next round and keep its strategy unchanged with probability $1 - p_{AB}$ according to Eq. 3 as follows.

$$p_{AB} = \frac{1}{1 + exp[(s_A - s_B)/k]}, \tag{3}$$

in which, the parameter k satisfying $0 < k < \infty$ represents the rationality of game players and refers to the uncertainty in the strategy transition process. When $k \to \infty$, the player updates its strategy completely at random with the probability of $1/2$ either transferring to its neighbor's strategy or maintaining its own. When $k \to 0$, each agent is absolutely rational, the selection is strong and the strategy which can profit more is more readily adopted. In agreement with previous works and without

Table 1
Payoff matrix of PD game.

	C	D
C	3	0
D	5	1

loss of generality, in this paper, we set $k = 1$. In Eq. 3, there is clearly a higher probability of the focal player adopting its neighbor's strategy if the focal player's payoff in this round is lower than its neighbor's.

2.4. Game environment setting

The randomness and heterogeneity of network structures influences evolutionary processes at the scale of the population. In this case, in order to eliminate the particularity and contingency of game results, we take three typical varieties of complex networks- lattice networks, random regular (RR) networks, and BA scale-free networks- as game environments. We then observe and compare game results in these differently structured populations. A lattice network is characterized by regularity and, when we add some randomness to that, the result is an RR network. The third network type is a BA scale-free network, which embodies both randomness and heterogeneity. Additionally, initial population states also affect the evolution of the system to a great extent. Here, we integrate varying frequencies of the classic strategies (1, 20, and 50 instances of the TFT and WSLS strategies) into each network to simulate different initial population profiles. The dual changes of network structure and initial population state result in the creation of more diverse and realistic social network topologies, so that the evolutionary results of game interactions and the conclusions we draw are more universal.

3. The self-bad, partner-worse (SBPW) strategy

3.1. The generation and evolution of dominant strategies

This study aims to identify the dominant strategy, which succeeds more efficiently during competitions and eventually spreads to entire networked populations as the game environment changes in terms of its network structure and its initial population state. In doing so, we constructed lattice network with 10000 nodes, RR network with 3000 nodes, and BA network with 2500 nodes. Each node on the network represents an individual player with a randomly assigned memory-one strategy. Then variations of 1, 20, and 50 instances of the TFT and WSLS strategies were respectively placed in the networks initially to construct different initial population states for simulating more realistic biological systems and further exploring their impact on the self-organized evolutionary results. Neighboring nodes on these networks interacted with each other repeatedly and the dominant strategy is the strategy with the highest proportion among the population when the system reached steady state. The simulation process used in the determination of such dominant strategies is described in Fig. 1.

We find that, after 600 iterations, the dominant strategy almost takes up the entire networked population, independently of its network structure and initial population state. Fig. 2 shows the spread and evolution of a dominant strategy within the lattice network initialized with 1 instances (20 and 50 instances are provided in Supplementary material) of the TFT and WSLS strategies. We can see that nodes with dominant strategy receive a higher payoff during the game and thus form a dominant-strategy cluster. As the game proceeds, the dominant strategy spreads throughout the whole network and ultimately prevails. The detailed evolutionary processes of the dominant strategies for RR networks and BA scale-free networks can be found in Supplementary material.

3.2. The SBPW strategy discovered using clustering analysis

We derived 1000 dominant strategies for each simulated network. Due to the fact that the randomness of initial strategies and differences in the connections between nodes both influence population evolution to a great extent (even in under conditions of the same network topology), the dominant strategies that eventually spread to the entire population are significantly different from one another. In this respect, the AGNES algorithm is used here to classify and cluster these dominant strategies so as to group similar strategies together and analyze their attributes in common. AGNES is a hierarchical clustering algorithm with bottom-up aggregation mechanism[9]. It first regards each sample in the data set as its own initial cluster and it then determines the closest two clusters to be merged. This process is iterated until the preset number of clusters is reached. Here, the average-distance method is used to measure the distance between two clusters, which is determined by all of the samples included in them. For example, given two strategy clusters $\boldsymbol{C_i}$ and $\boldsymbol{C_j}$, the distance between them can be calculated as Eq. 4.

$$d_{avg}(\boldsymbol{C_i}, \boldsymbol{C_j}) = \frac{1}{|\boldsymbol{C_i}||\boldsymbol{C_j}|} \sum\nolimits_{\boldsymbol{p_x}} \in \boldsymbol{C_i} \sum_{\boldsymbol{p_y}} \in \boldsymbol{C_j} dist(\boldsymbol{p_x}, \boldsymbol{p_y}), \tag{4}$$

where $\boldsymbol{p_x}$ and $\boldsymbol{p_y}$ respectively represent a memory-one strategy characterized by a probabilistic quaternion (p_1, p_2, p_3, p_4). In order to give a more intuitive and visual description of the clustering results, the PCA dimension reduction algorithm is applied to reduce the representation of these dominant strategies from four dimensions to three dimensions, thereby allowing us to present them visually in a 3-dimensional space. Fig. 3 represents the clustering distributions of the three network types initialized with 50 instances of the TFT and WSLS strategies. Fig. 4 shows more detailed information for these clusters, including the cluster centers, the strategy number each center contains, and the expected payoff that the strategy associated with each cluster center would gain in its competition with the other strategy cluster centers. We also provide a Supplementary material with the clustering results for networks initialized with 1 and 20 instances of the TFT and WSLS strategies.

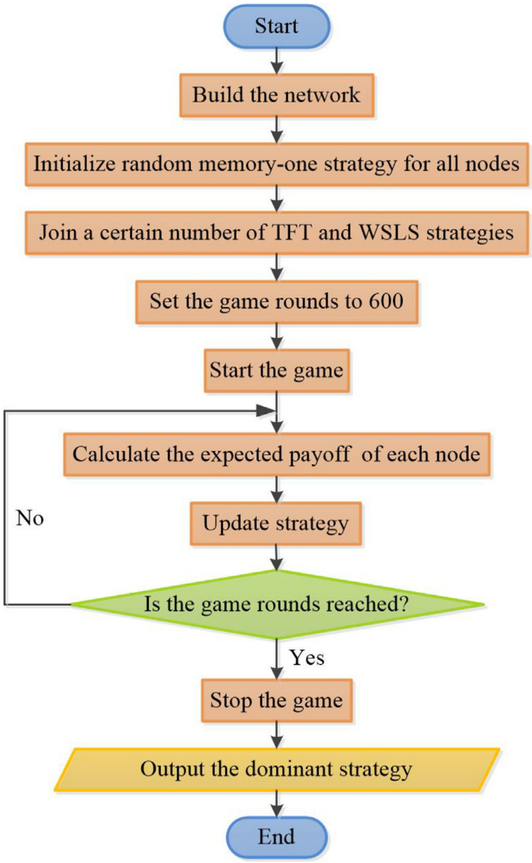


Fig. 1. The flow chart of the simulation process to generate one dominant strategy on complex networks.

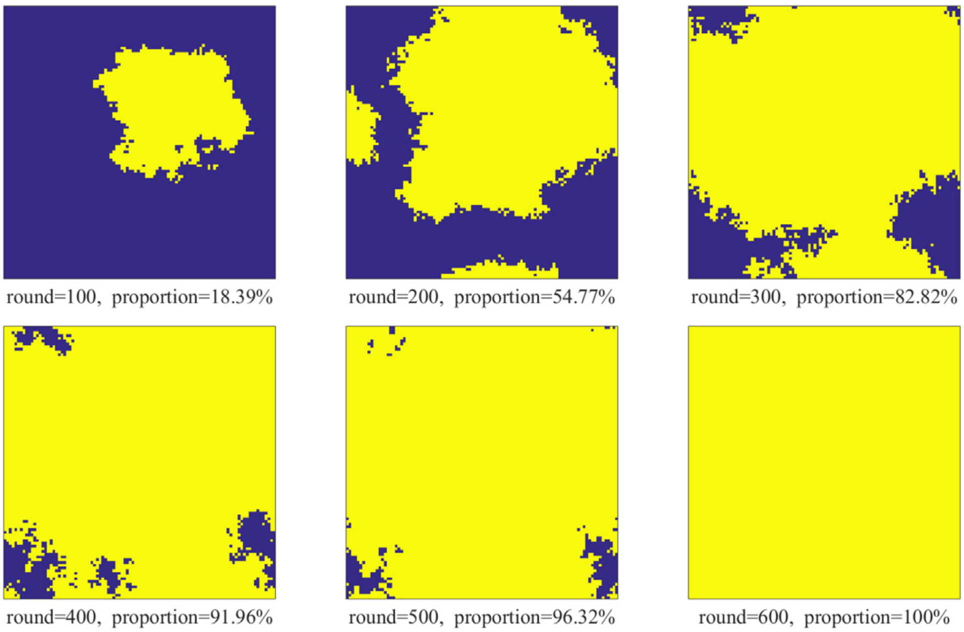


Fig. 2. Snapshots of dominant strategies (yellow) and other strategies (blue) for the lattice networks initialized with 1 instances of the TFT and WSLS strategies. The strategy distributions are shown for the game that have reached 100, 200, 300, 400, 500 and 600 rounds. The 'proportion' of labels refers to the proportion of the dominant strategy among overall number of strategies.

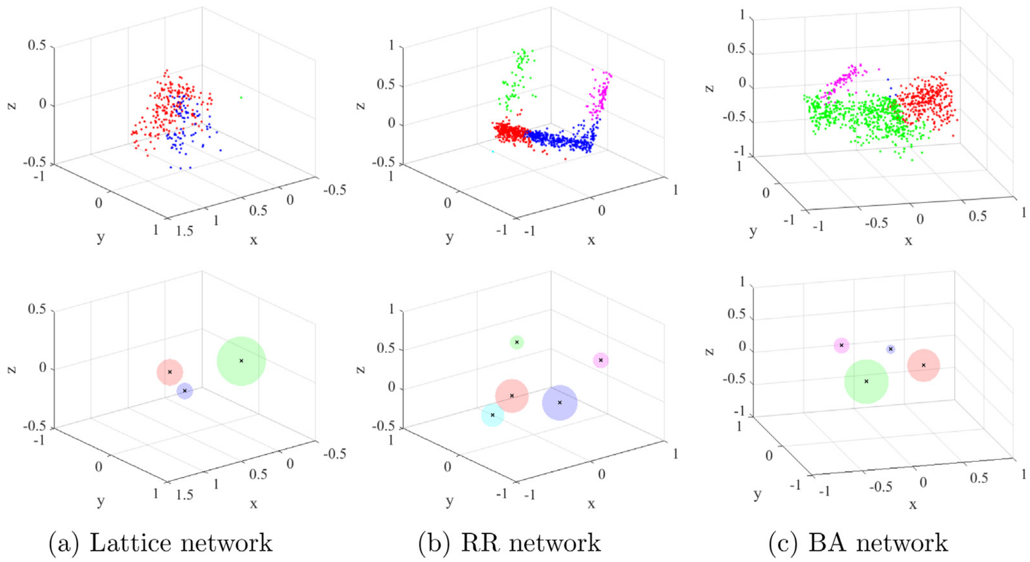


Fig. 3. Dimension-reduced visualization of clustering results for the 1000 dominant strategies on the 4-degree lattice network, the RR network, and the BA scale-free network respectively. Each network is initialized with 50 instances of the TFT and WSLS strategies. Solid dots with different colors on the top panels represent dominant strategies belonging to different clusters evolved from 1000 simulations and the size of colored circles on the bottom panels corresponds to the number of dominant strategies contained by each cluster, with the black '×' showing the center of the respective clusters.

Lattice

Gather into 3 clusters

The center of cluster	The strategy number in the cluster	The expected payoff against with self
C1=[0.9996 0.2700 0.1856 0.1932]	207	2.9730
C2=[1.0000 0.0000 0.0000 1.0000]	711	3
C3=[0.9994 0.1490 0.5999 0.1906]	82	2.9767

The expected payoff battling among clusters

	C1	C2	C3
C1	2.9730	2.9984	2.9779
C2	2.9907	3	2.9888
C3	2.9746	2.9934	2.9767

RR

Gather into 5 clusters

The center of cluster	The strategy number in the cluster	The expected payoff against with self
C1=[0.9923 0.2323 0.1470 0.2390]	333	2.6584
C2=[0.3991 0.3170 0.0554 0.0091]	59	1.0417
C3=[0.9655 0.2333 0.6685 0.1512]	372	2.4834
C4=[0.3612 0.0606 0.9553 0.0950]	75	2.1227
C5=[1.0000 0.0000 0.0000 1.0000]	161	3.0000

The expected payoff battling among clusters

	C1	C2	C3	C4	C5
C1	2.6584	0.8677	2.4668	1.7527	2.9827
C2	1.9309	1.0417	1.6907	1.3952	2.8839
C3	2.4238	0.9110	2.4834	1.9089	2.6394
C4	1.7775	0.9703	2.0960	2.1227	2.0363
C5	2.8611	0.6922	2.5512	2.0630	3.0000

BA

Gather into 4 clusters

The center of cluster	The strategy number in the cluster	The expected payoff against with self
C1=[0.3842 0.2574 0.0813 0.0188]	318	1.0795
C2=[0.9653 0.2501 0.4027 0.1505]	581	2.1935
C3=[0.9973 0.0198 0.0019 0.9795]	25	2.9860
C4=[0.3297 0.0613 0.9421 0.1070]	76	2.0978

The expected payoff battling among clusters

	C1	C2	C3	C4
C1	1.0795	1.6879	2.8484	1.4568
C2	0.9650	2.1935	2.7553	1.7703
C3	0.7648	2.4193	2.9860	2.0452
C4	1.0079	1.7982	2.0670	2.0978

Fig. 4. Results of the cluster analysis including 1000 dominant strategies for the lattice networks, RR networks, and BA scale-free networks initialized with 50 instances of the TFT and WSLS strategy. For each network, there is a SBPW strategy cluster, as marked by red lines, that drives down the network's average payoff but in which the focal player still outperforms its opponents.

In our analysis of these clustering results, we find that irrespective of the structures and initial strategy profiles of their populations, networks will always evolve this particular kind of strategy cluster, which obtains a low payoff for itself but which results in the payoff for its opponent being lower than its own. These results can be seen as C1 for the lattice networks, C2 for RR networks, and C1 for the BA scale-free networks in Fig. 4. Hence, we refer to strategies with this set of characteristics as "self-bad, partner-worse" (SBPW) strategies. We also conducted further supplementary evolutionary simulation experiments on 8-degree lattice networks and ER networks, simulations in which SBPW strategies could similarly evolve and survive with stability in networked populations (see the Supplementary material for more detailed analyses and results).

4. Main results

4.1. The SBPW strategy versus a random strategy

To this point, we have shown that the SBPW strategy drives down the payoff of the whole population but benefits the focal player more than the other dominant-strategy clusters evolved within complex networks. But when the SBPW strategy is confronted by an opponent whose strategy is unknown, can it always maintain its advantage? In this section, we investigate whether the SBPW remains beneficial when encountering an unknown-strategy partner. From a mathematical point of view, we aim to examine whether or with what probability SBPW will maintain higher payoffs when faced with an opponent with a random strategy during IPD game contests.

Assuming that two players, X and Y, play with each other using a memory-one strategy, $\mathbf{p} = (p_1, p_2, p_3, p_4)$ and $\mathbf{q} = (q_1, q_2, q_3, q_4)$ respectively. According to expected payoffs predicted by Eq. 1, the two players' payoff differences can be calculated as shown in Eq. 5. Next, provided that player X uses the SBPW strategy and player Y adopts a random strategy, in which each element of $\mathbf{q}$ (namely $q1, q2, q3, q4$) varies between 0 to 1. We can count the proportion $S_X - S_Y > 0$ in Eq. 5, corresponding with the number of cases in which the SBPW strategy defeats the random strategy with a higher payoff.

$$
\begin{aligned}
S_X - S_Y = 5(&q_4 - p_4 - p_2q_4 + p_4q_2 - p_3q_4 + p_4q_3 + p_1p_4q_1 - p_1p_4q_4 - p_2p_4q_3 \\
&-p_3p_4q_2 + p_2p_4q_4 + p_3p_4q_4 - p_1q_1q_4 + p_2q_3q_4 + p_3q_2q_4 + p_4q_1q_4 - p_4q_2q_4 \\
&-p_4q_3q_4 + p_1p_2q_1q_4 - p_1p_4q_1q_2 + p_1p_3q_1q_4 - p_1p_4q_1q_3 - p_1p_2q_3q_4 \\
&-p_1p_3q_2q_4 + p_2p_4q_1q_3 + p_3p_4q_1q_2 + p_1p_4q_2q_4 - p_2p_4q_1q_4 + p_1p_4q_3q_4 - p_3p_4q_1q_4) \\
&\div(p_4 - p_2 - q_2 + q_4 - p_1q_1 + p_2q_2 + p_2q_3 + p_3q_2 - p_2q_4 - p_3q_3 - p_4q_2 + p_3q_4 \\
&+p_4q_3 - p_4q_4 + p_1p_2q_1 - p_1p_2q_3 - p_1q_4p_1 - p_2p_3q_2 + p_2p_3q_3 + p_1p_4q_4 + p_3p_4q_2 \\
&-p_3p_4q_4 + p_1q_1q_2 - p_1q_1q_4 - p_3q_1q_2 - p_2q_2q_3 + p_3q_2q_3 + p_2q_3q_4 + p_4q_1q_4 \\
&-p_4q_3q_4 - p_1p_2q_1q_2 + p_1p_2q_1q_4 + p_1p_2q_2q_3 + p_1p_3q_1q_3 + p_1p_4q_1q_2 + p_2p_3q_1q_2 \\
&-p_1p_3q_1q_4 - p_1p_3q_2q_3 - p_1p_4q_1q_3 - p_2p_3q_1q_3 - p_1p_2q_3q_4 + p_1p_3q_2q_4 + p_2p_4q_1q_3 \\
&-p_3p_4q_1q_2 - p_1p_4q_2q_4 - p_2p_3q_2q_4 - p_2p_4q_1q_4 - p_2p_4q_2q_3 + p_1p_4q_3q_4 \\
&+p_2p_3q_3q_4 + p_2p_4q_2q_4 + p_3p_4q_1q_4 + p_3p_4q_2q_3 - p_3p_4q_3q_4 + 1).
\end{aligned}
\tag{5}
$$

The results of these games are shown in Fig. 5. For lattice networks, the SBPW strategy defeats a random strategy around 86% of the time and, for RR and BA scale-free networks, SBPW wins nearly 98% of the time. In short, it is clear that the SBPW strategy that evolves within RR and BA scale-free networks are even more dominant when gaming against an unknown-strategy player. Fig. 6 represents the corresponding payoff profiles for the SBPW strategy and a random strategy when they play against each other on various networks initialized with 50 instances of the TFT and WSLS strategies. Payoff comparison for SBPW and random strategies for networks initialized with 1 and 20 instances of these classical strategies are presented in the Supplementary material

4.2. The game within the SBPW cluster

We have now demonstrated that the SBPW strategy has an advantage over other strategies in the IPD game. However, is there potentially a more powerful strategy that might arise from the SBPW cluster, which we have already shown is dominant in comparison with the other strategy clusters? Here, the strategy variants play games among themselves within the SBPW clusters and then we extract the common features of victorious strategies. Fig. 7 indicates that the SBPW strategies with $p_1 = 1$ for the lattice networks and small p_4 for the RR and BA scale-free networks are more advantageous in comparison with other strategy variants. The results shown here are those for networks initialized with 50 instances of the classic strategies (TFT and WSLS). Game results from more diverse situations can be found in Supplementary material.

Inspired that p_1 values of victorious strategies among dominant strategies in SBPW clusters are all 1 for lattice networks, here we assuming that there are two players, X and Y, whose strategies are denoted as $\mathbf{p} = [1, p_2, p_3, p_4]$ and $\mathbf{q} = [1, q_2, q_3, q_4]$, respectively. According to Eq. 1, their expected payoffs after one-round game interaction are the same as predicted by Eq. 6. This tells us that, no matter the strategy value of p_2, p_3 and p_4, payoffs for the two players are the same. Therefore, for the

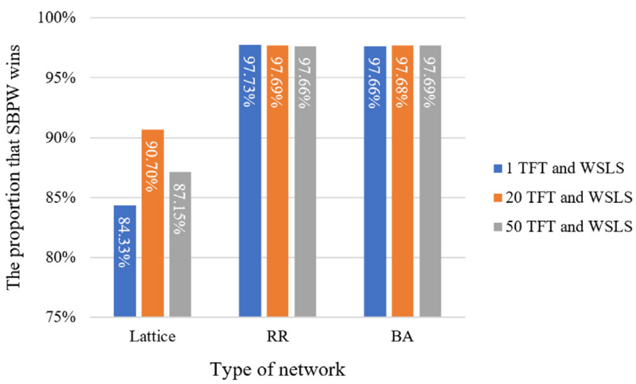


Fig. 5. Percentage of cases in which the SBPW strategy defeats a random strategy. Game environments are lattice networks, RR networks, and BA scale-free networks. Bars with different color correspond to networks initialized with different numbers of TFT and WSLS strategies.

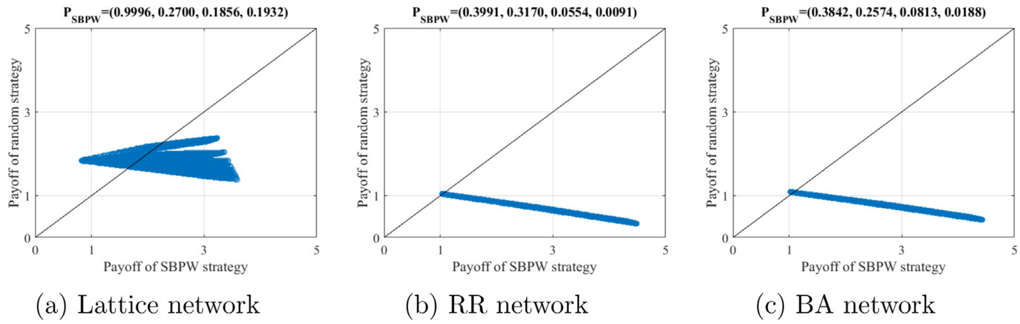


Fig. 6. Payoffs for SBPW and random strategies for IPD games played in complex networks initialized with 50 instances of the TFT and WSLS strategies. In each sub-figure, the strategy **p** of player X is fixed to P_{SBPW}, and the strategy **q** of the partner, player Y (random strategy), can vary, sampling the 4-d cube of ZD strategies (the blue dots correspond to 10^4 different realizations of **q**). 0, 1, 3, 5 on the coordinate correspond with the payoffs for different strategy combinations for two interacting players in the PD game, as shown in Table 1.

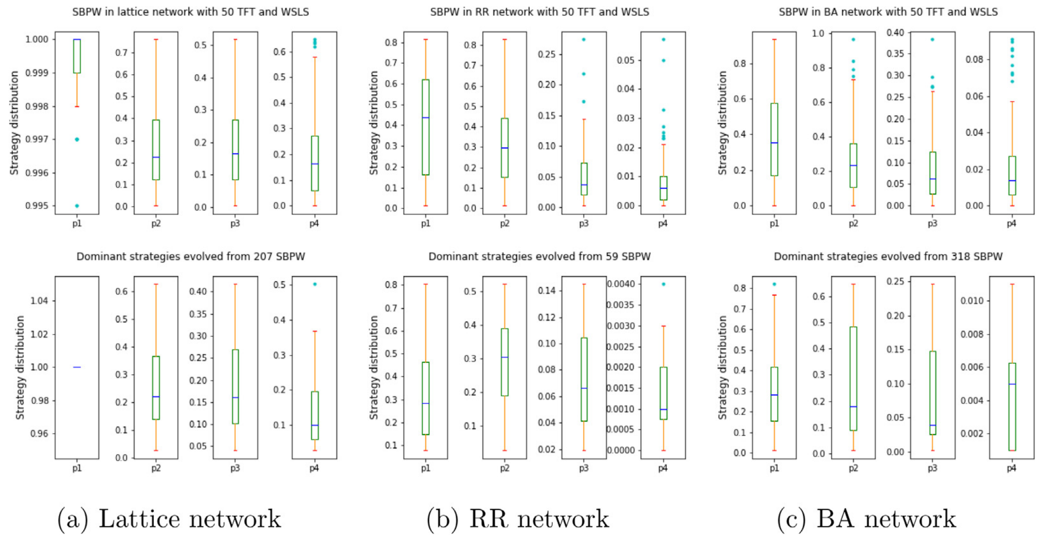


Fig. 7. Boxplots depicting the distribution of 4-d strategy values for memory-one strategies in SBPW clusters and the dominant strategies emerging from these clusters for lattice, RR, and BA scale-free networks. In each sub-figure, the top panel shows the distribution of all strategies for each SBPW cluster and the bottom panel shows the strategy distribution of 50 most dominant strategies winning out from the above SBPW cluster. The three networks are initialized with 50 instances of the classic strategies (TFT and WSLS). For each boxplot, the blue line represents the median, the upper and lower bounds of the green box represent the upper and lower IQR, red lines at the top and bottom represent maximum and minimum values, and cyan dots represent outliers.

lattice network, it is difficult for the system to evolve into a situation in which one strategy occupies the whole network, since two players will obtain the same payoff when their p_1 values are both 1, corresponding to the situation that they update their strategies completely at random.

$$\begin{aligned} S_X = S_Y \quad &= 3(p_4q_4 + p_2p_4q_3 - p_2p_4q_4 + p_3q_2q_4 - p_4q_2q_4 - p_2p_3q_2q_4 \\ &-p_2p_4q_2q_3 + p_2p_3q_3q_4 + p_2p_4q_2q_4 + p_3p_4q_2q_3 - p_3p_4q_3q_4) \\ &\div(p_4q_4 + p_2p_4q_3 - p_2p_4q_4 + p_3q_2q_4 - p_4q_2q_4 - p_2p_3q_2q_4 \\ &-p_2p_4q_2q_3 + p_2p_3q_3q_4 + p_2p_4q_2q_4 + p_3p_4q_2q_3 - p_3p_4q_3q_4). \end{aligned} \tag{6}$$

4.3. The contest between dominant-strategy clusters

Another issue of interest is what would happen if the game was to occur within units representing the various clusters of dominant strategies. In this respect, we identify and label the respective belonging clusters of 1000 dominant strategies that evolved from each network topology. We approach this problem by treating each cluster as one separate population and observing game results between these populations in pairs. Each game between any two populations is repeated 20 times. We then record the cluster to which each dominant strategy belongs.

Results for each of the three kinds of complex networks initialized with instances of 50 TFT and WSLS strategies are shown in Fig. 8, and game results for populations with other initial states can be found in the Supplementary material. Our findings suggest that, for lattice networks, the WSLS cluster can defeat almost any other clusters and the SBPW cluster can beat almost all clusters except the WSLS cluster. For both the RR and BA scale-free networks, the WSLS cluster can defeat almost all clusters except the SBPW cluster and the SBPW cluster will almost always lose to every cluster except the WSLS cluster. It is also worth noting that a common feature of dominant strategies evolved from all three network topologies is that, in addition to the type of each strategy cluster, the size of cluster significantly affects the outcome of the game: the cluster with more strategies has clear advantages in the game. One aspect of the evidence for this conclusion is that the WSLS cluster loses to the C2 cluster, as Fig. 8(c) shows. This owes to the fact that the former group size (25 strategies) is much smaller than the latter (581 strategies). Fig. 9 gives a more intuitive interpretation of the win-loss relationships between the different strategy clusters for lattice, RR, and BA scale-free networks.

4.4. Performance comparison of TFT, WSLS and SBPW

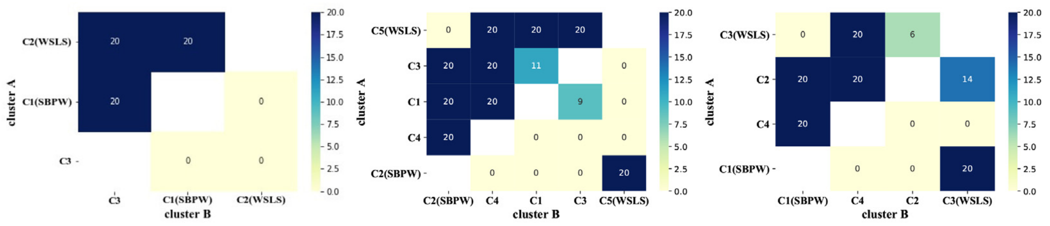
In terms of the distinctive characteristics of the two classic IPD game strategies, TFT always imitates the action of its partner in the previous round and WSLS only cooperates if both players chose the same behavior in the previous round. SBPW is a novel strategy with which a player can always make its payoff higher than its partner's and reduce the welfare of the whole population at the same time. To explore the differences of the performance and survival advantage of the three strategies mentioned above, we count the number of each that evolved among the 1000 dominant strategies for each network structure. These results are shown in Fig. 10.

It is evident that TFT always goes extinct irrespective of the network structure type and the initial number of classic strategies added into the initial populations. For WSLS, within lattice networks, as its initial frequency increases, the probability of it becoming the dominant strategy grows rapidly. This is not true for BA scale-free networks, where larger initial numbers do not convey a survival advantage. For RR networks, the result is intermediate that a high initial frequency of the WSLS strategy among populations only slightly increases its ultimate level of dominance. Thus, we deduce that the randomness and heterogeneity of networks will weaken the survival advantage for the WSLS strategy. As for SBPW, the initial number of classic strategies has little effect on its level of dominance in any network game environment, meaning that the survival dominance of SBPW is stable and robust in different network structures, though its fitness is stronger for the lattice and BA scale-free networks than it is for RR networks.

5. Discussion and conclusion

The strategy that can remember the outcomes of previous rounds of competition has recently attracted considerable attention. It can help game players make the subsequent strategy decisions based on the action of their opponents so as to optimize their behaviors and maximize their own benefits, which is closer to the real-world situation. TFT and WSLS are classic two-player strategies. Based on these foundational models, a novel ZD strategy has been proposed, in which one player can unilaterally determine its opponent's score and create an extortionate linear relation between the payoffs for both players. In this paper, we have investigated a wide range of random memory-one strategies across numerous game scenarios and explored what kind of strategy can be successful and eventually come to dominate the whole population.

In order to realistically recreate real-world conditions, we employed different types of complex networks to simulate various population structures and we integrated different classic strategies into our models in order to simulate a diversity of population states. Clustering the dominant strategies evolved from these various types of complex networks, we discovered the SBPW strategy, which is capable of unilaterally controlling an opponent's payoff lower than that of the focal player. The SBPW strategy also reduces benefits at the scale of the whole population, which results in the cooperative dilemma. When confronted with an unknown-strategy partner, SBPW still has the upper hand with a win probability of about 86% within



(a) Lattice network (b) RR network (c) BA network

Fig. 8. Results of games between the dominant-strategy clusters for lattice, RR, and BA scale-free networks initialized with 50 instances of the TFT and WSLS strategies. Each number in the colored square represents the number of times that Cluster A defeats Cluster B in their 20 simulated games. C1-C5 in each sub-figure corresponds with C1-C5 in Fig. 4. Specifically, the clusters belonging to WSLS and S.BPW are labeled.

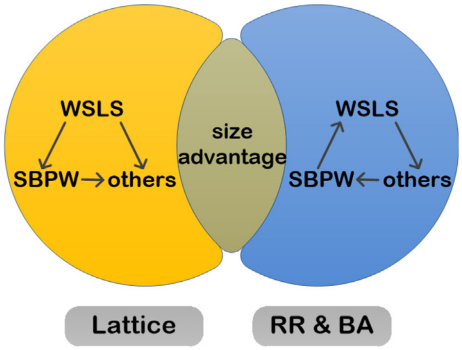


Fig. 9. The win-loss relationships between different clusters for lattice, RR, and BA scale-free networks. Large colored circles represent different network structures. The direction of every arrow in each circle implies a game result within in structured populations. The strategy in which the arrow begins is the winning side, while the strategy where the arrow ends is the losing side. The intersection of two circles represents the common feature of corresponding networks.

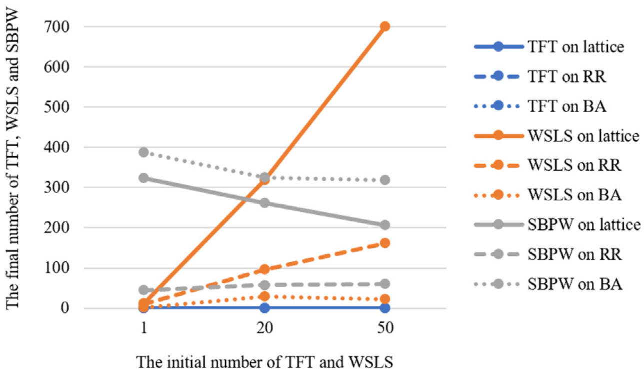


Fig. 10. Comparison of the final dominant strategy counts for the TFT, WSLS and SBPW strategies among the 1000 evolved dominant strategies on lattice networks, RR networks, and BA scale-free networks initialized with 1, 20, and 50 instances of the TFT and WSLS strategies.

lattice network and 98% within RR and BA scale-free networks. Furthermore, we found that SBPW strategies with $p_1 = 1$ on lattice network and with lower p_4 on RR and BA scale-free networks are dominant within IPD games. In addition, we explored the win-loss relationship among SBPW, WSLS and other clusters, after which we found the following: For lattice networks, the most dominant strategy is WSLS, followed by SBPW, and both of these strategies are capable of defeating other strategy clusters. For RR and BA scale-free networks, the three strategies can restrict each other. WSLS has an advantage over the other strategy clusters, which can beat SBPW in most cases and, conversely, the SBPW is stronger than WSLS. Moreover, a size advantage is obvious among the various strategy clusters. The larger the initial population, the more likely it is to win the game. More noteworthily, unlike the TFT strategy, which will ultimately tend to die out within complex networks, and the WSLS strategy for which its survival advantage will be weakened by the randomness and heterogeneity of populations, the SPBW strategy has a significant survival advantage and robust fitness in game evolution over time.

In summary, our findings indicate that complex networks will evolve an SBPW strategy, which leads to the cooperative dilemma for social networks by driving down the profits of opponents and further reducing the benefits at the scale of the population as a whole. Based on this finding, we have made further exploration and concluded that the SBPW strategy has a significant survival advantage and robust fitness in a wide range of game scenarios. Besides, we have characterized some of its unique attributes. Future research might focus on the factors promoting the evolution of the SBPW as a dominant strategy among various network conditions, as well as potential mechanisms that would be capable of regulating the SBPW strategy and thereby improving the broader welfare of social systems.

CRediT authorship contribution statement

Chunyan Zhang: Conceptualization, Methodology, Writing - original draft. **Siyuan Liu:** Methodology, Software, Formal analysis, Writing - original draft. **Zhijie Wang:** Software, Formal analysis. **Franz J. Weissing:** Writing - review & editing. **Jianlei Zhang:** Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We are grateful to the financial support from the National Natural Science Foundation of China (Grant Nos. 62073174, 62073175 and 91848203).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, athttps://doi.org/10.1016/j.ins.2021.11.041.

References

[1] P.M. Altrock, A. Traulsen, M.A. Nowak, Evolutionary games on cycles with strong selection, Phys. Rev. E 95 (2017) 022407.
[2] R. Axelrod, W.D. Hamilton, The evolution of cooperation, Q. Rev. Biol. 79 (1981) 135–160.
[3] S.P. Borgatti, A. Mehra, D.J. Brass, G. Labianca, Network analysis in the social sciences, Science 323 (2009) 892–895.
[4] J.S. Brown, T.L. Vincent, A theory for the evolutionary game, Theor. Popul. Biol. 31 (1987) 140–166.
[5] A. Chacoma, M.N. Kuperman, D.H. Zanette, Payoff nonmonotonic dynamics in an evolutionary game, Adv. Complex Syst. 19 (2016) 1–12.
[6] I.M. Daras, F.J. Weissing, Replication and individual-level emergence are not sufficient for understanding cultural transmission and the evolution of human culture, Phys. Life Rev. 30 (2019) 74–76.
[7] A. Dhand, L. Mccafferty, R. Grashow, I.M. Corbin, A. Barabási, Social network structure and composition in former nfl football players, Sci. Rep. 11 (2021) 1630.
[8] M. Duh, M. Gosak, M. Perc, Public goods games on random hyperbolic graphs with mixing, Chaos Solitons Fractals 144 (2021) 110720.
[9] W.R. Fox, Finding groups in data: an introduction to cluster analysis, J. Roy. Stat. Soc.: Ser. C (Appl. Stat.) 40 (1991) 486–487.
[10] T.R. Gadekallu, N. Khare, S. Bhattacharya, S. Singh, M. Alazab, Early detection of diabetic retinopathy using pca-firefly based deep learning model, Electronics 9 (2020) 274.
[11] J. Gao, B. Barzel, A. Barabási, Universal resilience patterns in complex networks, Nature 238 (2016) 307–312.
[12] N.E. Glynatsi, V.A. Knight, Using a theory of mind to find best responses to memory-one strategies, Sci. Rep. 10 (2020) 132–139.
[13] F. Huang, M. Cao, L. Wang, Learning enables adaptation in cooperation for multi-player stochastic games, J. R. Soc. Interface 17 (2020) 20200639.
[14] G. Ichinose, N. Masuda, Zero-determinant strategies in finitely repeated games, J. Theor. Biol. 438 (2018) 61.
[15] L.A. Imhof, D. Fudenberg, M.A. Nowak, Tit-for-tat or win-stay, lose-shift?, J Theor. Biol. 247 (2007) 574–580.
[16] R. Li, Q. Zhang, T. Chu, On quotients of boolean control networks, Automatica 125 (2021) 109401.
[17] W.J. Li, L.L. Jiang, M. Perc, A limited mobility of minorities facilitates cooperation in social dilemmas, Appl. Math. Comput. 390 (2021) 125705.
[18] L. Liu, X. Chen, M. Perc, Evolutionary dynamics of cooperation in the public goods game with pool exclusion strategies, Nonlinear Dyn. 97 (2019) 1–18.
[19] S. Liu, J. Zhang, The networked division of labor game based on adaptive dynamics, IFAC-PapersOnLine 52 (2019) 156–161.
[20] A. Mamiya, G. Ichinose, Strategies that enforce linear payoff relationships under observation errors in repeated prisoner's dilemma game, J. Theor. Biol. 477 (2019) 63–76.
[21] P. Mathieu, J.P. Delahaye, Experimental criteria to identify efficient probabilistic memory-one strategies for the iterated prisoner's dilemma, Simul. Model. Pract. Theory 97 (2019) 101946.
[22] J.L. Mcclelland, B.L. Mcnaughton, A.K. Lampinen, Integration of new information in memory: new insights from a complementary learning systems perspective, Philos. Trans. R. Soc. B Biol. Sci. 375 (2020) 20190637.
[23] A. Nair, F. Reed-Tsochas, Revisiting the complex adaptive systems paradigm: Leading perspectives for researching operations and supply chain management issues, J. Oper. Manage. 65 (2019) 80–92.
[24] M.A. Nowak, Five rules for the evolution of cooperation, Science 314 (2006) 1560–1563.
[25] M.A. Nowak, K. Sigmund, Tit for tat in heterogeneous populations, Nature 355 (1992) 250–253.
[26] M.A. Nowak, K. Sigmund, A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game, Nature 364 (1993) 56–58.
[27] M. Port, H. Hildenbrandt, I. Pen, F.J. Weissing, The evolution of social philopatry in female primates, Am. J. Phys. Anthropol. 173 (2020) 397–410.
[28] W.H. Press, F.J. Dyson, Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent, Proc. Natl. Acad. Sci. U.S.A. 109 (2012) 10409–10413.
[29] P. Ramazi, M. Cao, Global convergence for replicator dynamics of repeated snowdrift games, IEEE Trans. Autom. Control 66 (2021) 291–298.
[30] R.H. Schonmann, R. Boyd, A simple rule for the evolution of contingent cooperation in large groups, Philos. Trans. R. Soc. f London Ser. B, Biol. Sci. 371 (2016) 20150099.
[31] L. Shi, H. Zhu, S. Zhong, K. Shi, J. Cheng, Cluster synchronization of linearly coupled complex networks via linear and adaptive feedback pinning controls, Nonlinear Dyn. 88 (2017) 859–870.
[32] A.J. Stewart, T.L. Parsons, J.B. Plotkin, Evolutionary consequences of behavioral diversity, Proc. Natl. Acad. Sci. U.S.A. 113 (2016) E7003–E7009.
[33] A.J. Stewart, J.B. Plotkin, From extortion to generosity, evolution in the iterated prisoner dilemma, Proc. Nat. Acad. Sci. 110 (2013) 15348–15353.
[34] G. Szabó, G. Fáth, Evolutionary games on graphs, Phys. Rep. 446 (2007) 97–216.
[35] M.A. Taha, A. Ghoneim, Zero-determinant strategies in repeated asymmetric games, Appl. Math. Comput. 369 (2019) 124862.
[36] C. Tang, C. Li, X. Yu, Z. Zheng, Z. Chen, Cooperative mining in blockchain networks with zero-determinant strategies, IEEE Trans. Cybern. 50 (2020) 4544–4549.
[37] C. Tang, X. Li, M. Cao, Z. Zhang, X. Yu, Incentive mechanism for macrotasking crowdsourcing: a zero-determinant strategy approach, IEEE Internet Things J. 6 (2019) 8589–8601.
[38] J. Wang, Y. Gao, W. Liu, W. Wu, S.J. Lim, An asynchronous clustering and mobile data gathering schema based on timer mechanism in wireless sensor networks, Comput., Mater. Continua 58 (2019) 711–725.
[39] T. Wen, Y. Deng, Identification of influencers in complex networks by local information dimensionality, Inf. Sci. 512 (2020) 549–562.
[40] X. Xu, Z. Rong, Extortion boosts cooperation through redistributing strategies in assortative networked systems, IFAC-PapersOnLine 52 (2019) 267–271.
[41] Y. Xue, W. Luan, Y. Yang, H. Wang, Evolutionary game for the stakeholders in livestock pollution control based on circular economy, J. Clean. Prod. 282 (2021) 125403.

[42] H. Yamamoto, I. Okada, T. Taguchi, M. Muto, Effect of voluntary participation on an alternating and a simultaneous prisoner's dilemma, Phys. Rev. E 100 (2019) 032304.
[43] Z. Yang, C. Yu, J. Kim, Z. Li, L. Wang, Evolution of cooperation in synergistically evolving dynamic interdependent networks: Fundamental advantages of coordinated network evolution, New J. Phys. 21 (2019) 073057.
[44] D. Zhang, Y. Tian, C. Yue, M. Fan, Cooperate delegation of computation for rational party using zero-determinant strategy approach, IEEE Access 8 (2020) 27734–27743.
[45] Z. Zhou, T. Ren, H. Xiao, L. Xing, Y. Cui, X. Zhang, Z. Wang, The iterative scheme and the convergence analysis of unique solution for a singular fractional differential equation from the eco-economic complex system's co-evolution process, Complexity 2019 (2019) 1–15.
[46] Y. Zhu, J. Zhang, Q. Sun, Z. Chen, Evolutionary dynamics of strategies for threshold snowdrift games on complex networks, Knowl.-Based Syst. 130 (2017) 51–61.