Payoff-based accumulative effect promotes cooperation in spatial prisoner's dilemma*

Liu Yong-Kui(刘永奎)[a)†], Li Zhi(李 智)[a)], Chen Xiao-Jie(陈小杰)[b)], and Wang Long(王 龙)[a)b)]

[a)] *Center for Complex Systems, Department of Automatic Control Engineering, Xidian University, Xi'an* 710071, *China*

[b)] *State Key Laboratory for Turbulence and Complex Systems, Center for Systems and Control, College of Engineering, Peking University, Beijing* 100871, *China*

(Received 13 October 2009; revised manuscript received 31 January 2010)

We study the effect of accumulative payoff on the evolution of cooperation in the evolutionary prisoner's dilemma on a square lattice. We introduce a decaying factor for the accumulative payoff, which characterizes the extent that the historical payoff is accumulated. It is shown that for fixed values of the temptation to defect, the density of cooperators increases with the value of the decaying factor. This indicates that the more the historical payoff is involved, the more favourable cooperators become. In the critical region where the cooperator density converges to zero, cooperators vanish according to a power-law-like behaviour. The associated exponents agree approximately with the two-dimensional directed percolation and depend weakly on the value of the decaying factor.

Keywords: payoff, accumulative effect, prisoner's dilemma, cooperation

PACC: 0250, 0175, 0270

1. Introduction

Cooperation among selfish individuals plays an essential role in modern human societies. Revelation of mechanisms supporting the evolution of cooperation is therefore of key interest within many branches of social and natural sciences.[1] A common mathematical framework for addressing this issue within groups of selfish individuals is evolutionary game theory,[2−4] and in particular, the prisoner's dilemma[5,6] has been widely adopted for this purpose. In this simple game, two individuals may choose simultaneously between cooperation (C) or defection (D). They obtain some payoffs depending on their choices. For mutual cooperation, they both receive the reward R, but only the punishment P for mutual defection. If one individual cooperates while the other defects, then the cooperator gets the payoff S, while the defector receives the payoff T. These payoffs satisfy the characteristic payoff ranking of the prisoner's dilemma: $T > R > P > S$. It is additionally required that $2R > T + S$ for repeated interactions. Under these conditions, mutual cooperation warrants the highest collective payoff that is equally shared between the two individuals. Mutual cooperation is, however, challenged by the defecting strategy that promises the defector a higher income at the expense of the neighbouring cooperator. Thus, the prisoner's dilemma has inspired numerous investigations of suitable extensions that enable cooperative behaviour to emerge and sustain.

To date, several mechanisms, e.g., kin selection,[7] direct reciprocity,[1] indirect reciprocity, network reciprocity, and voluntary participation[8,9] (see Ref. [10] and references therein) that facilitate the emergence and persistence among selfish individuals have been uncovered. Especially, the spatial version of the evolutionary prisoner's dilemma[6] can facilitate the emergence of cooperation in the absence of strategic complexity. As direct extensions to spatial games, evolutionary games on complex networks[11−19] have been intensively studied, which were comprehensively reviewed in Ref. [20]. However, spatial structure is sometimes inefficient in the promotion of cooperation, and some other ingredients, e.g., memory effects,[18−23] dynamical preferential selection,[21] and different teaching activities[22] are needed to promote cooperation to the desired level. It has been shown that historical memory[18,23] plays a key role in the evolution of cooperation. Furthermore, in real-world

*Project supported by the National Natural Science Foundation of China (Grant Nos. 70671079, 60674050, 60736022 and 60528007), the National Basic Research Program of China (Grant No. 2002CB312200), the National High Technology Research and Development Program of China (Grant No. 2006AA04Z258) and 11-5 Project (Grant No. A2120061303).

†Corresponding author. E-mail: yongkuiliu@163.com

© 2010 Chinese Physics Society and IOP Publishing Ltd

situations, it is often observed that influential individuals, who either are more successful in collecting payoffs[24] or have a larger teaching capability[22] are more likely to enforce their strategies on the opponents.

In the traditional case, strategy updating of individuals is based on the evolutionary outcome in the current time step, and historical records are completely irrelevant. However, in real social systems, past records are inevitably taken into account by individuals and affect their decision-making.[18] Based on this consideration, we present a model of the evolutionary prisoner's dilemma on a square lattice with von Neumann neighbourhood, i.e., each individual has four neighbours, where individuals update by simply imitating their neighbours' current strategy in each time step based on historical accumulative payoffs. Thus, the strategy of the individual who is more successful in accumulating payoff is more likely to be adopted. Interestingly, we find that the introduction of accumulative payoff can remarkably promote the cooperative behaviour in the prisoner's dilemma.

2. Model

We consider the evolutionary prisoner's dilemma with players located on a square lattice with period boundary conditions. In accordance with previous studies,[6,12] the prisoner's dilemma is characterized by $T = b > 1$, $R = 1$, and $P = S = 0$, where b represents the advantage of defectors over cooperators, being typically constrained to the interval $1 < b \leq 2$.

The game is iterated in accordance with the simulation procedure comprising the following elementary steps. First, every individual engages simultaneously in pairwise interactions within its von Neumann neighbourhood to obtain payoffs, which is the sum over all interactions with its neighbours. Next, all individuals are allowed to modify their strategies simultaneously by imitating a randomly chosen neighbour's strategy with a probability depending on the accumulative payoff difference, i.e., at time t ($t \geq 1$), the probability for individual i to adopt a (randomly chosen) neighbour j's strategy is

$$W(i \leftarrow j) = \frac{1}{1 + \exp\{[P(i,t) - P(j,t)]/\kappa\}}, \quad (1)$$

where κ determines the amplitude of noise, which is simply set to be 0.1 for all simulations. $P(i,t)$ and $P(j,t)$ are the sums of the payoff U_t at time t and the payoff accumulated up to time $t-1$, that is, for individual i

$$P(i,t) = P(i,t-1)\tau + U_t. \quad (2)$$

Starting with the initial condition $P(i,0) = 0$, we can rewrite Eq. (2) in terms of U_r ($1 \leq r \leq t$) as

$$P(i,t) = U_1\tau^{t-1} + U_2\tau^{t-2} + \ldots + U_{t-1}\tau + U_t, \quad (3)$$

where $0 \leq \tau < 1$ is a decaying factor for the accumulative payoff. One can find that, for $0 < \tau < 1$, as time passes, incomes in earlier time steps contribute less to the overall accumulative payoff. In particular, for $\tau = 0$, we recover the traditional model where the historical payoffs, i.e., U_r ($r < t$), are irrelevant in the strategy updating process. In the limit $\tau \to 1$, all the historical payoffs are exactly accumulated. For intermediate values of τ, there exists a compromise between the above two limiting cases. Repeat the above process and then the system evolves. In our model, τ controls to what extent the historical payoffs are concerned and thus will be the focus of our studies (note that we restrict our discussions to the interval $0 \leq \tau < 1$ because the case of $\tau = 1.0$ demands perfect rationality of individuals and thus may be unrealistic).[18]

3. Results

Simulations were carried out for a population of $N = 100 \times 100$ individuals (we find that simulations with 200×200 or even larger system sizes do not qualitatively change the results). Initially, an equal percentage of cooperators and defectors is randomly distributed among the population. The key quantity for characterizing the macroscopic behaviour of the system is the density of cooperators $\rho_{\rm C}$ in the population. After the transient state, the system reaches a dynamical equilibrium state where the amplitude of the fluctuations of $\rho_{\rm C}$ is considerably smaller than the corresponding average value.

Figure 1 shows the simulation results of $\rho_{\rm C}$ as a function of b for several different values of τ. We observe that in comparison with the traditional case where no historical payoff is involved, i.e., $\tau = 0.0$, cooperation is remarkably promoted for each positive value of τ. Especially, when τ is larger than a certain value, there are two thresholds of the payoff parameter b for each value of τ. When $b < b_{\rm c1}$, a homogeneous cooperator state emerges. The density of cooperators

monotonically decreases with increasing values of b for $b > b_{c1}$. When $b > b_{c2}$, the system evolves into the absorbing state of all defectors. The critical value of b_{c1} clearly increases with τ (see the inset of Fig. 1), which is in contrast with the result reported in Ref. [23] by Qin *et al.* who found that b_{c1} reaches its maximum value near $\tau = 0.72$ and when $\tau > 0.72$, cooperation is inhibited for large b (e.g., $1.75 < b < 1.8$). While in our model, the cooperation level always increases with increasing τ. This indicates that our model is essentially different from theirs.

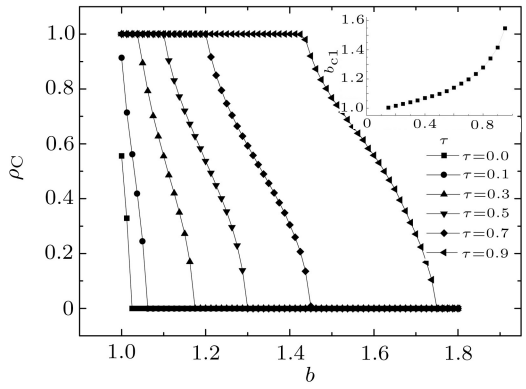


Fig. 1. Density of cooperators ρ_C as a function of the payoff parameter b for several different values of τ. The inset shows the critical value of b_{c1} versus τ. Each data point is obtained by averaging over 20 runs of simulations with different initial conditions, and for each evolutionary process, we average the last 5000 generations of the total 25000 generations.

It is well known that spatial structure enables cooperators to form clusters and thus cooperators along the boundary can outweigh their losses against defectors by gains from interactions within the cluster.[25] Here we also investigate the spatial distributions of cooperators and defectors on the square lattice. To gain a complete picture on how cooperators (or defectors) organize during the evolutionary process, we investigate the cluster formation process at different time steps for $b = 1.15$ and $\tau = 0.7$, as shown in Fig. 2. One can find that, initially cooperators and defectors are randomly distributed on the square lattice (see Fig. 2(a)). However, after the initial moment, individuals with the same strategy join together forming complex patterns which continuously move and change shapes. More precisely, cooperator density decreases to the minimum value $\rho_C = 0.214$ at $t = 4$ (see Fig. 2(b)) where cooperators have been maximally invaded by defectors and cooperators can only form small gangs. Note that at this moment because there are not so many cooperators can be exploited by defectors, their advantage is greatly reduced. Thus, with the increment of t, cooperators gradually extend their territories and at $t = 30$ (see Fig. 2(c)), cooperators unite to form a single connected cluster whose pattern is highly irregular. When t is further increased, cooperators eventually wipe out defectors from the population.

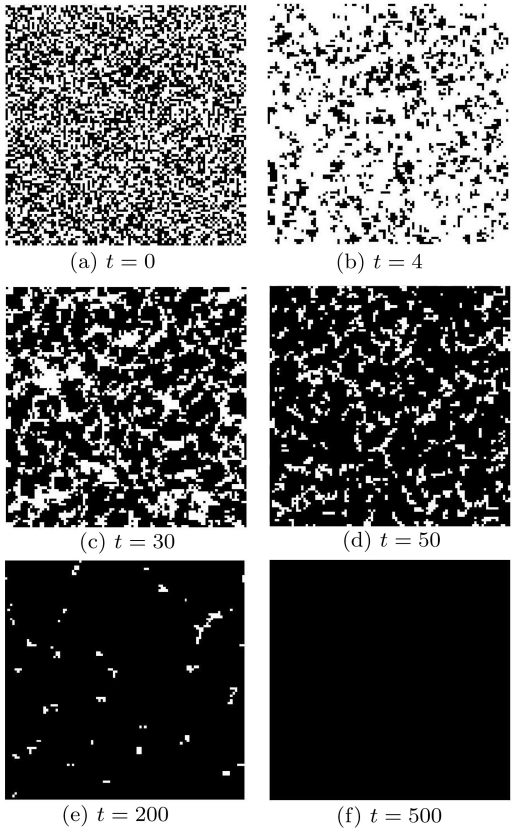


Fig. 2. Snapshots of typical spatial patterns illustrating the distributions of cooperators (black) and defectors (white) on a square lattice for $b = 1.15$ and $\tau = 0.7$ at different time steps t. (a) $t = 0$ [$\rho_C(0) = 0.50$], (b) $t = 4$ [$\rho_C(4) = 0.217$], (c) $t = 30$ [$\rho_C(30) = 0.694$], (d) $t = 50$ [$\rho_C(50) = 0.8366$], (e) $t = 200$ [$\rho_C(200) = 0.9829$], (f) $t = 500$ [$\rho_C(500) = 1.0$].

In the following, we investigate the mean accumulative payoffs of cooperators and defectors along the boundary during the evolutionary process, which gives us a complete picture on how the competition between cooperators and defectors leads to the observed results shown in Fig. 1. Boundary individuals are defined as those who have at least a neighbour adopting a different strategy. The results are shown in Figs. 3 and 4 for $b = 1.15$ and several values of τ (note that for the cases where the absorbing states of all cooperators or all defectors can be reached, we only show the results for the transient process, because there is no

interface between cooperators and defectors in the stationary state for such cases once the absorbing state is reached). The inset of Fig. 3 shows the time evolution of the corresponding density of cooperators. We

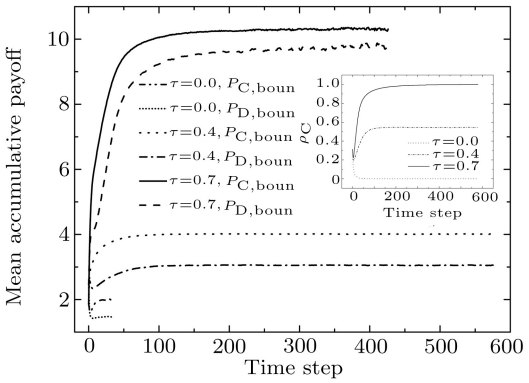


Fig. 3. Time evolution of the mean accumulative payoffs of cooperators and defectors along the boundary ($P_{\mathrm{C,boun}}$ and $P_{\mathrm{D,boun}}$) for $b = 1.15$ and $\tau = 0.0, 0.4, 0.7$, respectively. The inset shows the corresponding time evolution of the density of cooperators. Results are obtained by averaging over 100 realizations with different initial states.

observe that for $\tau = 0.0$ and $\tau = 0.7$, the system evolves to the absorbing states of all defectors and all cooperators very quickly, while for $\tau = 0.4$, after the transient process, cooperators and defectors coexist in the stationary state. We can understand the result by observing the competition between cooperators and defectors in the evolutionary process. From Fig. 3, for all values of τ, boundary cooperators always have a higher mean accumulative payoff than boundary defectors, indicating that cooperators overall have advantage over defectors in the competition process. To gain more insights, we show the accumulative payoff difference (that is $P_{\mathrm{C,boun}} - P_{\mathrm{D,boun}}$) between cooperators and defectors in Fig. 4. Clearly, larger values of τ make cooperators gain more advantage than defectors, i.e., for $\tau = 0.7$, in the first 100 time steps, cooperators can have a much higher mean payoff than defectors. Thus, cooperators can wipe out most defectors in the population in this time range (see the inset of Fig. 3, for $\tau = 0.7$, cooperators dominate the population at $t = 100$). On the contrary, for $\tau = 0.0$, although cooperators also have advantage over defectors on average, but the amplitude of the advantage is so small that it cannot prevent the invasion by defectors, and thus cooperators vanish ultimately. For $\tau = 0.4$, cooperators and defectors are to some extent evenly matched and thus can coexist in the stationary state.

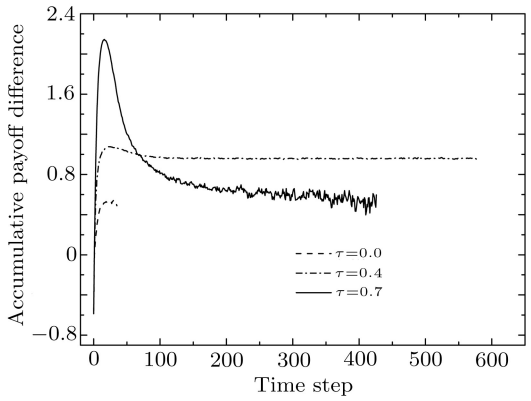


Fig. 4. Accumulative payoff difference between boundary cooperators and boundary defectors. Parameters are the same as in Fig. 3.

Now it is clear that in terms of accumulative payoff, cooperators can gain advantage over defectors. Moreover, the stronger the effect of accumulative payoff, the more advantageous cooperators become. The key aspect for such phenomenon is that cooperators are advantageous in accumulating payoff. Since the invasion process occurs only along the boundary, we focus on the boundary individuals. Cooperators along the boundary usually turn their defecting neighbours into cooperators, which continuously increases their incomes; on the contrary, successful defectors breed defectors in their neighbourhood, and thus introduce a negative feedback that reduces their incomes. Thus, in a single time step, defectors may obtain a higher payoff than cooperators, but in the long run, their accumulative payoff is usually lower than that for cooperators, that is, cooperators can have accumulative advantage in collecting payoff over defectors (see Figs. 3 and 4).

In Refs. [25] and [24], the authors discussed the critical exponents of b_{c2} and showed that near the extinction threshold b_{c2}, the average density of cooperators vanishes as $\rho_{\mathrm{C}} \sim (b_{\mathrm{c2}} - b)^{\beta}$, and the value of β agrees with the directed percolation (DP) exponent. While in Ref. [23], the authors found that the critical exponent is not universal but depends on the so-called memory factor (see the reference for details). Inspired by these works, we have also investigated the divergence behaviour of cooperators in the vicinity of the extinction thresholds for several values of τ, as shown in Fig. 5. According to our simulations, near the extinction thresholds, cooperators vanish according to a power-law $\rho_{\mathrm{C}} \sim (b_{\mathrm{c2},\tau} - b)^{\beta_\tau}$, where β_τ are a set of exponents associated with the value of τ (see the plots). These values of β_τ are approximately consis-

tent with the critical values $\beta \approx 0.584$ of the two-dimensional DP and depend weakly on τ, which is in contrast with the results reported in Ref. [23]. The estimated exponents $\beta_{0.1}$, $\beta_{0.3}$, $\beta_{0.5}$, $\beta_{0.7}$, $\beta_{0.9}$ are 0.624, 0.632, 0.623, 0.660 and 0.675, respectively. The large errors are caused by the limits of the computational conditions.[24]

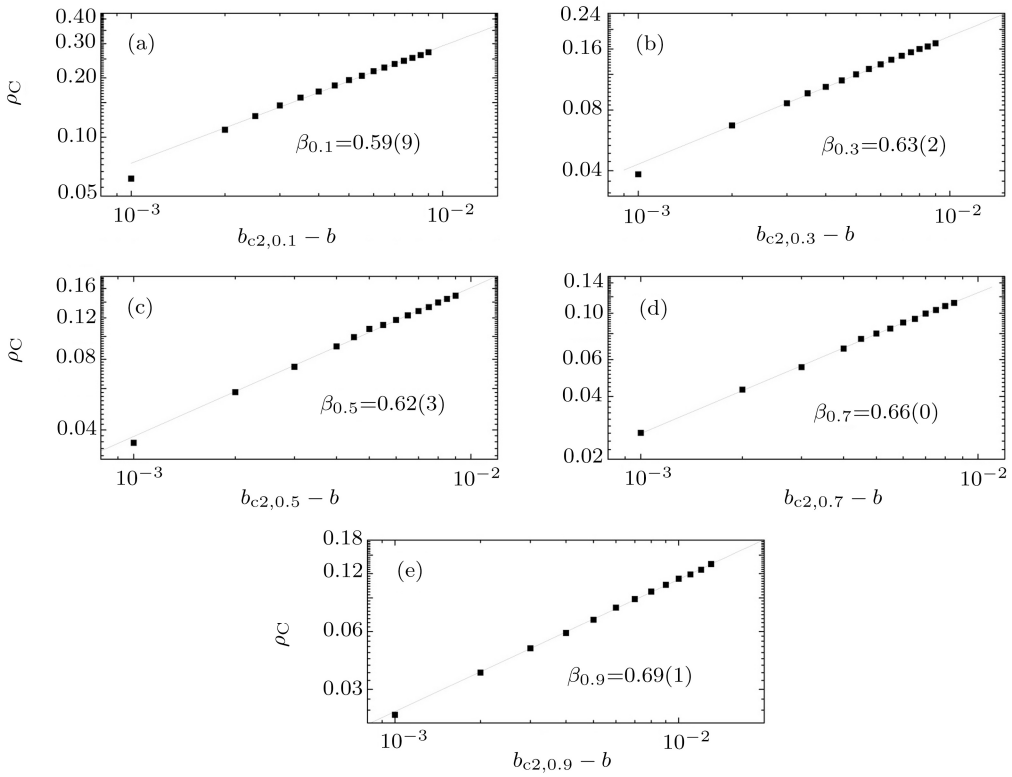


Fig. 5. Log–log plots of the average density of cooperators ρ_C as a function of the distance to the extinction threshold $b_{c2,\tau} - b$ for several values of τ: (a) $\tau = 0.1$, $b_{c2,0.1} \approx 1.057$, (b) $\tau = 0.3$, $b_{c2,0.3} \approx 1.175$, (c) $\tau = 0.5$, $b_{c2,0.5} \approx 1.296$, (d) $\tau = 0.7$, $b_{c2,0.7} \approx 1.448$, (e) $\tau = 0.9$, $b_{c2,0.9} \approx 1.744$. The solid lines are the power laws $\sim (b_{c2,\tau} - b)^{\beta_\tau}$ which fit the data with the exponents β_τ (the detailed values are shown in the plots, where the figures between the parentheses represent the statistical uncertainties of the last digit). The system size is 800×800 in our simulations.

4. Discussion and conclusion

In this section, let us first discuss the difference of our model from previous related ones. In Ref. [18], Wang *et al.* studied the SG[26] where individuals update strategies according to the finite memory of the optimal strategy which is determined by introspection. While in our model, individuals update by learning from their neighbours. Another model that is close to the present one is proposed by Qin *et al.*[23] In their model, they distinguish the accumulative payoffs obtained with C and D, respectively and individuals are more likely to adopt the strategy (C or D) leading to a higher accumulative payoff. While in our model, we assume that individuals are more likely to adopt the strategy of the successful individuals who have a higher accumulative payoff, no matter what strategy is used to obtain the payoff. More importantly, in comparison with previous memory-related models, the present one is more simpler, and requires less cognitive capability of the participating players.

Interestingly, we also find that cooperation in the snowdrift game[26] can also be significantly enhanced under the current updating rule for $\tau > 0$ in compar-

ison with the case of $\tau = 0.0$. In addition, regarding the role of κ, we find that generally small value of κ facilitates the emergence of cooperation because in our case cooperators are advantageous with respect to defectors, and thus small amount of noise is overall favourable to cooperators. It is noteworthy that in Ref. [27], the authors investigated the accumulated wealth distribution by adopting evolutionary games taking place on scale-free networks. They found that the total wealth value is closely related to the density of cooperators: the system's total wealth is high when the density of cooperators is high, and a low total wealth shows up when the density of cooperators is low. In their model, the payoff for a single time step is used for strategy updating, which is different from our model where the historical accumulated payoff is used when an individual is updated. We have shown that in our case cooperation can be promoted with increasing values of τ, which means that the total wealth is higher for a larger value of τ for the same value of b.

In summary, we have studied the prisoner's dilemma on a square lattice where strategy updating is based on the accumulative gains. It is shown that cooperation is enhanced to a higher degree with increasing the effect of accumulative payoff. Investigation of the spatial patterns indicates that the updating rule favors the formation of cooperator clusters and thus promotes cooperation. We find that for a larger value of the decaying factor, cooperators can have more advantage over defectors, and this is because the former are advantageous in accumulating payoff. In addition, we find that cooperators vanish according to a power-law-like behaviour near the extinction threshold for all values of the decaying factor τ. Our work shows that in terms of accumulative payoff cooperators are to some extent favoured by natural selection and thus may provide an escape for the social dilemma.

References

[1] Axelrod R 1984 *The Evolution of Cooperation* (New York: Basic books) p145

[2] Hofbauer J and Sigmund K 1998 *Evolutionary Games and Population Dunamics* (UK: Cambridge University Press) p101

[3] Nowak M A 2006 *Evolutionary Dynamics: Exploring the Equations of Life* (Cambridge: Harvard University Press) p45

[4] Smith J M 1982 *Evolution and the Theory of Games* (UK: Cambridge University Press)

[5] Axelrod R and Hamliton W D 1981 *Science* **211** 1390

[6] Nowak M A and May R M 1992 *Nature* **359** 826

[7] Hamiliton W D 1964 *J. Theoret. Biol.* **7** 17

[8] Hauert C, de Monte S, Hofbauer J and Sigmund K 2002 *Science* **296** 1129

[9] Szabó G and Vukov J 2004 *Phys. Rev.* E **69** 036107

[10] Nowak M A 2006 *Science* **314** 1560

[11] Ebel H and Bornholdt S 2002 *Phys. Rev.* E **66** 056118

[12] Santos F C and Pacheco J M 2005 *Phys. Rev. Lett.* **95** 098104

[13] Ohtsuki H, Hauert C, Lieberman E and Nowak M A 2006 *Nature* **441** 502

[14] Yang H X, Gao K, Han X P and Wang B H 2008 *Chin. Phys.* B **17** 2759

[15] Gómez-Gardeñes J, Campillo M, Floría L M and Moreno Y 2007 *Phys. Rev. Lett.* **98** 108103

[16] Lin H and Wu C X 2007 *Acta Phys. Sin.* **56** 4313 (in Chinese)

[17] Santos F C, Rodrigues J F and Pacheco J M 2006 *Proc. R. Soc.* B **273** 51

[18] Wang W X, Ren J, Chen G R and Wang B H 2006 *Phys. Rev.* E **74** 056113

[19] Liu Y K, Li Z, Chen X J and Wang L 2009 *Chin. Phys.* B **18** 2623

[20] Szabó G and Fáth G 2007 *Phys. Rep.* **446** 97

[21] Wu Z X, Xu X J, Huang Z G, Wang S J and Wang Y H 2006 *Phys. Rev.* E **74** 021107

[22] Szolnoki A and Szabó G 2007 *Europhys. Lett.* **77** 30004

[23] Qin S M, Chen Y, Zhao X Y and Shi J 2008 *Phys. Rev.* E **78** 041129

[24] Guan J Y, Wu Z X, Huang Z G, Xu X J and Wang Y H 2006 *Europhys. Lett.* **76** 1214

[25] Hauert C and Szabó G 2005 *Am. J. Phys.* **73** 405

[26] Hauert C and Doebeli M 2004 *Nature* **428** 643

[27] Hu M B, Wang W X, Jiang R, Wu Q S, Wang B H and Wu Y H 2006 *Euro. Phys. J.* B **53** 273