IFAC

Cooperation in Networked Prisoner's Dilemma with Individual Learning Feedback

Xiaojie Chen * Feng Fu ** Long Wang *

* Intelligent Control Laboratory, Center for Systems and Control, Department of Mechanics and Space Technologies, College of Engineering, Peking University, Beijing 100871, China (e-mail: xjchen@pku.edu.cn; longwang@pku.edu.cn)
** Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138, USA

Abstract: We introduce a modified learning updating mechanism into the evolutionary Prisoner's Dilemma on Newman-Watts (NW) networks. During the evolutionary process, each individual updates its strategy according to individual deterministic switch in combination with a feedback between its score aspiration and actual score. And individual's score is a linear combination of individual's total payoff and local contribution to its neighbors. We study the cooperation level of the system under this learning feedback mechanism, and find that the cooperation level increases as the relative weight of the local contribution to the score increases. In addition, we focus on the influences of learning rate and intensity of deterministic switch in the strategy updating rule on cooperation. Simulations show that for much low intensity of deterministic switch, cooperation is independent of learning rate to a large extent, and full cooperation can be reached when relative weight is not less than 0.5. Otherwise, cooperation depends on the value of learning rate. Besides, the cooperation level is not sensitive to topological parameters of NW networks. To explain these simulation results, we provide corresponding analytical results of mean-field approximation, and find that simulation results are in good agreement with analytical ones. Our work may shed some light on the maintenance of cooperative behavior in social systems with individual learning feedback.

1. INTRODUCTION

According to the fundamental principles of Darwinian selection, evolution is based on a fierce competition between individuals and should therefore reward only selfish behavior. Yet, cooperative (altruistic) behavior is widespread in natural and social systems (Dugatkin [1997]). How to understand the emergence of cooperation is a fundamental problem. Fortunately, evolutionary game theory has provided a powerful framework to investigate cooperative behavior in systems consisting of competitive individuals (Smith [1982], Hofbauer et al. [1998]). As a common paradigm, the Prisoner's Dilemma game (PDG) has received much attention to study the evolution of cooperation in the literature. In the classical version of PDG, two individuals adopt simultaneously one of the two available strategies, cooperate (C) or defect (D); for mutual cooperation both receive R, and only P for mutual defection, while a cooperator receives S when confronted to a defector, which in turn receives T, such that $T > R > P > S$ and $T + S < 2R$. Under these conditions it is best to defect for rational individuals in a single round of the PDG, regardless of the opponent strategy. However, mutual coop-

eration leads to a higher payoff than mutual defection, but cooperation is irrational. Therefore, the dilemma is caused by the selfishness of individuals. To promote and maintain cooperation, other suitable extensions to the traditional PDG need to be explored accordingly.

During the last decades, based on PDG different mechanisms which favor the emergence of cooperation are summarized in some reviews (Doebeli et al. [2005], Nowak [2006], G. Szabó et al. [2007]). The following mechanisms, e.g., kin selection (Hamiltion [1964]), direct (Axelrod [1984]) and indirect (Nowak et al. [2005]) reciprocity, network (spatial) reciprocity (Nowak et al. [1992], Ohtsuki et al. [2006]), group selection (Traulsen et al. [2006]), chaotic variations to the payoffs (Perc [2006]), inhomogeneous activity of teaching (Szolnoki et al. [2007]), tag-based models (Riolo et al. [2001]), are found to support the emergence of cooperative behavior in biological and ecological systems as well as within human societies (Nowak et al. [2004]). Indeed, the spatial or graph model should be more meaningful in realistic systems because most interactions among individuals are spatially localized. In the evolutionary games with network structure, each individual occupying a node of the network can follow one of the pure strategies (C or D), and collect payoffs by playing the game with its immediate neighbors.

On the other hand, to describe the system to be more realistic, learning theory has application in game theory

to study the evolution of cooperation (Fudenberg et al. [1998], Posch et al. [1999], Borgers et al. [2000]). In learning, the individual's outcome aspirations are used to evaluate whether the individual satisfies its current behavior. If the outcome induced by the individual's behavior exceeds aspirations, the probability increases that the behavior will be repeated rather than searching for a superior alternative. Whereas if the outcome falls below aspirations, the probability decreases that the behavior will be repeated. Based on these above principles in learning theory, Macy et al. [2002] proposed a simple learning model for two-person games, and studied the learning dynamics of cooperation in repeated games. In addition, Individual learning agents have been introduced into networked evolutionary games to study the evolution of cooperation. Recently, Gulyás [2007] investigated the adaptation of cooperating strategies in an iterated PDG with individually learning agents on Watts-Strogatz networks. Szolnoki et al. [2007] studied evolutionary PDG with quenched teaching and learning activity of players on a two-dimensional lattice.

In this present work, to describe the realistic systems, we also introduce learning theory as well as graph theory into the evolutionary PDG, and focus on the evolution of cooperation in evolutionary PDG with individually learning agents on Newman-Watts (NW) networks (Newman et al. [1999]), which are close to realistic social networks. To our knowledge, an alternative way to escape from the dilemma is to consider more sophisticated individuals. For instance, individuals should feel strongly about immediate benefit that affects them directly, but they should also take into account the affairs of others (Pacheco et al. [2006], Nowak [2006]), and different possible measures of success for individuals should be used to assess their performance in games (Fort et al. [2005]). Actually, in realistic systems individuals do not always consider maximizing their immediate benefit as first goal, they also take into account their social responsibility, i.e. contributions to the group individuals belong to. That is to say, in order to go beyond the dilemma, individuals are endowed with heuristic thinking ability: evidently they could not always consider taking advantage of others' help, but sometimes they should help others as donors. This may correspond to the phenomenon in society that, some people would like to give voluntary donations, and do community service. Thus, our starting point is realizing that, in this paper, for long-term interactions with neighbors, individuals not only wish to increase their current total payoffs coming from PDG with their neighbors, but also consider the "local contribution" to their neighbors. We will show that this is an effective way to promote and maintain cooperation at favorable levels. To pursue as much generality as possible without sacrificing the simplicity, this local contribution of a given individual denotes the sum of all the payoffs its neighbors collect against it. Besides, the total payoffs and local contribution form the score through a linear combination. Accordingly, the score aspiration can also be described formally. Individuals are allowed to modify their strategies according to the actual scores and score aspirations. Interestingly, we find that the introduction of the local contribution under the learning feedback mechanism promotes cooperation. We also present theoretical analysis of mean-field approximation, and find that analytical results are in good agreement with numerical simulations.

The remainder of the paper is organized as follows. In the next section, the game model is introduced in detail. In Section 3, simulation results and corresponding theoretical analysis are provided, and finally some important conclusions are made in Section 4.

2. NETWORKED PRISONER'S DILEMMA WITH INDIVIDUAL LEARNING FEEDBACK

We consider the evolutionary PDG with players located on the NW networks, which is a typical small-world model. In the NW network, a parameter p controls the fraction of edges randomly added to the regular ring graph (Newman et al. [1999]). In evolutionary game on graphs, each player who occupies one site of the graph can only follow two simple strategies: C and D, and interacts only with its neighbors in each round. Let us represent the individuals' strategies with two-component vector, taking the value $s = (1,0)^T$ for C-strategist and $s = (0,1)^T$ for D-strategist. For one certain individual x, the total payoff P_x is collected from its neighbors, and the local contribution T_x is the sum of all the payoffs its neighbors collect against it, therefore, they can be respectively written as

$$P_x = \sum_{y \in \Omega_x} s_x^T A s_y \qquad (1)$$

and

$$T_x = \sum_{y \in \Omega_x} s_y^T A s_x, \qquad (2)$$

where the two sums both run over all the neighboring sites of x (this set is indicated by Ω_x), and A is payoff matrix

$$A = \begin{bmatrix} R & S \\ T & P \end{bmatrix}. \qquad (3)$$

The best-studied set of payoff values are $T = 5$, $R = 3$, $P = 1$, $S = 0$ (Axelrod et al. [1981], Axelrod [1984]), which is also adopted in this study. The score U_x of a certain individual x is a weighed mean of the total payoff P_x and the local contribution T_x. Formally,

$$U_x = (1 - h)P_x + hT_x, \qquad (4)$$

where $0 \le h \le 1$ is a parameter characterizing the relative weight between T_x and P_x in the score U_x. With $h = 0$, individuals only consider current total payoffs coming from PDG with their neighbors, that is, the local contribution is ignored. With $h = 1$, individuals only consider the local contribution to the neighbors, that is, the current total payoff is ignored. In learning, considering both its own payoff from the neighbors and the local contribution to the neighbors, each individual expects itself and its neighbors to all play C strategy in steady state, and thus has a score aspiration U_{xa}. Formally,

$$U_{xa} = (1 - h)k_x R + hk_x R = k_x R, \qquad (5)$$

where k_x is the neighbor number of individual x. The aspiration level provides the benchmark which is used to evaluate whether the individual satisfies its current strategy.

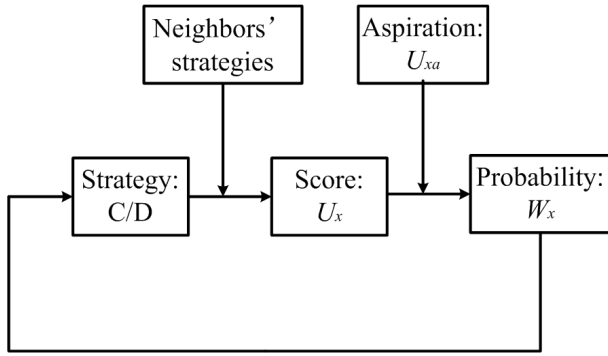Under the mechanism win-stay-lose-shift in two strategy games, the individual will keep its original strategies if

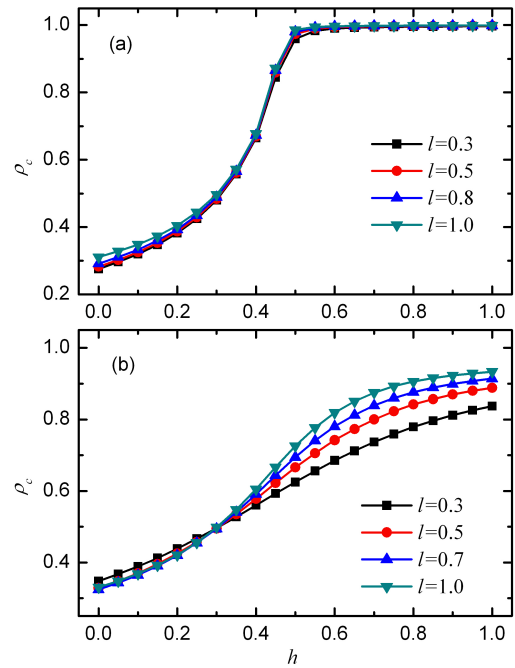Fig. 1. Schematic graph of individual strategy update procedure.

Fig. 2. (Color online) The density of cooperators as a function of h for different learning rates in the case of (a) $\delta = 0.001$ and (b) $\delta = 0.1$. Simulations are carried out on NW networks with $p = 0.5$.

the average payoff of recent rounds is above the aspiration level. Otherwise, it will switch to the opposite strategy. Such switch is deterministic (Nowak et al. [1993], Posch et al. [1999]). And in (Alonso et al. [2006]), individuals adopt the opposite strategy with a deterministic probability which is independent of payoff differences. However, in (Macy et al. [2002]) based on learning models, individuals update their strategies with a probability depending on the difference between the actual payoffs and payoff aspirations. Herein, We consider that individuals update their strategies based on deterministic models, as well as learning models, and Fig. 1 shows the procedure of individual strategy update. During the evolutionary process, each individual updates its strategy as follows: if $U_x \geq U_{xa}$, individual x keeps its original strategy. Otherwise, individual x adopts the opposite strategy with a probability as

$$W = (1 - \delta)l\frac{U_{xa} - U_x}{U_{xa}} + \delta, \qquad (6)$$

where l is the learning rate ($0 < l < 1$), and δ measures the intensity of deterministic switch, that is, the deterministic probability of adopting the opposite strategy. $\delta = 1$ denotes completely deterministic switch, in this case the updating rule is analogous to win-stay-lose-shift mechanism, individuals thus decide to switch to the opposite strategy if they only know $U_x < U_{xa}$. $\delta = 0$ denotes that this switch depends entirely on feedback between actual score and score aspiration. In fact, the probability of adopting the opposite strategy should depend strongly on the difference between actual score and score aspiration, since individuals can accordingly update their strategies with a probability based on their full information, in order to avoid more mistakes in strategy updating process. Therefore, the intensity δ should satisfy the following condition: $0 < \delta \ll 1$, and we reserve the deterministic probability to show that individuals are inclined to adopt the opposite strategy when they are not satisfied. The rationale for this rule is that when the score exceeds the aspiration level, the individual satisfies and keeps its original state. Otherwise, the individual adopts the opposite strategy with a probability W, and W characterizes the exact extent of discontent with its current state, and controls the individual to adopt the opposite strategy in order to reach the aspiration level.

3. SIMULATIONS AND ANALYSIS

Simulations are carried out for a population of $N = 2000$ players occupying the nodes of NW networks. Initially, the two strategies of C and D are randomly distributed among the players with the equal probability 0.5. The above model is simulated with synchronous updating (Hauert et al. [2004]). The key quantity for characterizing the cooperative behavior of the system is the cooperator density ρ_c, which is defined as the fraction of cooperators in the whole population. In all simulations, ρ_c is obtained by averaging over last 1000 generations of the transient time of 11000 generations, and we have confirmed that averaging over larger periods or using different transient times did not qualitatively change the simulation results. Moreover, each date point results from an average over 100 realizations, corresponding to 10 different realizations of a given type of network model with 10 runs of independent initial strategy configurations for each realization.

Fig. 2 shows the simulation results of ρ_c as a function of h for different values of learning rate l. The intensity of deterministic switch δ is set to 0.001 in Fig. 2(a) and 0.1 in Fig. 2(b), respectively. It is found that the density of cooperators monotonously increases with the increasing of h regardless of the values of l and δ. In addition, corresponding to a fixed value of h, there are no differences between the results of ρ_c for different values of learning rate l in Fig. 2(a), that is, ρ_c is independent of the values

of l. Whereas in Fig. 2(b), ρ_c is dependent of the values of l. One can find that when $h < 0.3$, the cooperator density for small learning rate is a little greater than for high learning rate. While for $h > 0.3$, the cooperator density becomes much more favorable as learning rate increases.

In what follows, we carry out theoretical analysis to understand the features shown in Fig. 2. Herein, we can use mean-field method because NW networks prevent the emerge of large degree nodes and can help the extension of the mean-field techniques for small-world structures. Considering the payoff matrix in (3) and assuming cooperators and defectors are distributed uniformly among the networks, the average scores of cooperators and defectors are given as

$$\begin{aligned} U_C &= (1-h)[Rk\rho_c + Sk(1-\rho_c)] \\ &\quad + h[Rk\rho_c + Tk(1-\rho_c)], \\ U_D &= (1-h)[Tk\rho_c + Pk(1-\rho_c)] \\ &\quad + h[Sk\rho_c + Pk(1-\rho_c)], \end{aligned} \qquad (7)$$

and the score aspiration of the population is

$$U_a = (1-h)kR + hkR = kR, \qquad (8)$$

where k is the average degree of NW networks. Subsequently, the transition rates can be written as

$$W_{C \to D} = l(1-\delta)\frac{U_a - U_C}{U_a} + \delta,$$
$$W_{D \to C} = l(1-\delta)\frac{U_a - U_D}{U_a} + \delta,$$

and the dynamical equation of the cooperator density becomes (G. Szabó et al. [2007])

$$\dot{\rho}_c = (1-\rho_c) \cdot W_{D \to C} - \rho_c \cdot W_{C \to D},$$

with $T = 5$, $R = 3$, $P = 1$, and $S = 0$, this yields

$$\begin{aligned} \dot{\rho}_c &= \frac{l}{3}(1-\delta)(1-\rho_c)(2 - 7\rho_c + 10\rho_c h) \\ &\quad + (1 - 2\rho_c)\delta. \end{aligned} \qquad (9)$$

For $\delta \ll 1$, deterministic switch in strategy update process can be ignored, (9) simplifies as

$$\dot{\rho}_c = \frac{l}{3}(1-\rho_c)(2 - 7\rho_c + 10\rho_c h). \qquad (10)$$

Considering (10), if $h < 0.5$, cooperators and defectors coexist in stable equilibrium, and the only stable equilibrium is $\rho_c = 2/(7 - 10h)$. Otherwise, cooperators dominate defectors under this update rule, and the only stable equilibrium is $\rho_c = 1$. Therefore, it indicates that ρ_c monotonously increases as h increases, and it shows that the stable equilibrium is independent of learning rate l from (10). Besides, in the case of learning rate $l = 0$, the dynamical equation of the cooperator density simplifies as $\dot{\rho}_c = 0$. Thus, the stationary solution satisfies $\rho_c = \rho_0 = 0.5$, where ρ_0 denotes the initial density of cooperators, and in this paper ρ_0 is set to 0.5 in all the simulations. The comparison between simulation results and theoretical analysis of ρ_c is shown in Fig. 3. We can find that analytical results are in very good agreement with simulation ones in Fig. 3(a), and there are little differences between analytical and simulation results in Fig. 3(b). Furthermore, we would like to stress that, under small values of δ, the steady state ending up with full cooperation could be reached with $h = 0.5$. That is to say, in
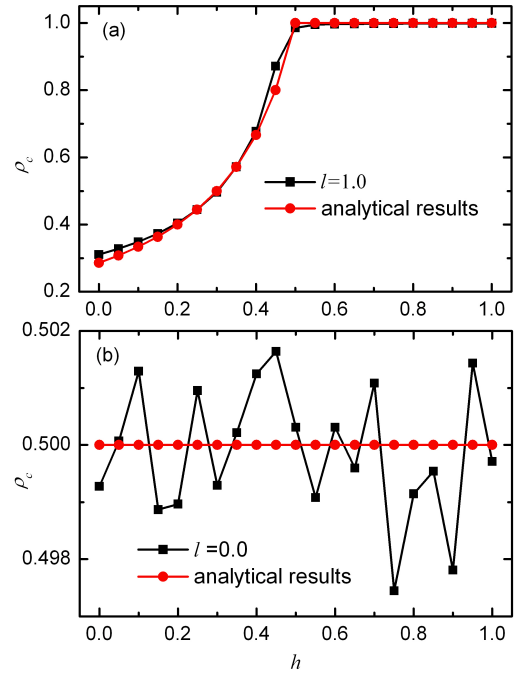

Fig. 3. (Color online) The density of cooperators as a function of h for simulation (closed squares) and analytical (closed circles) results with learning rate (a) $l = 1.0$ and (b) $l = 0.0$. The data points are computed for $\delta = 0.001$ on NW networks with $p = 0.5$.
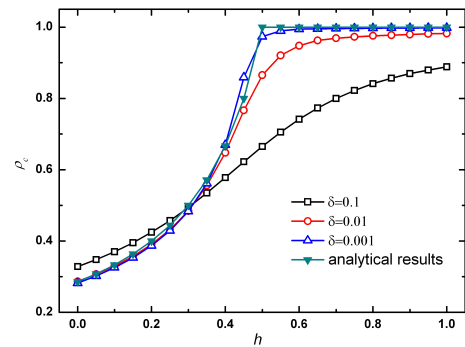

Fig. 4. (Color online) The density of cooperators as a function of h for different values of δ. Open symbols are the simulation results, and closed symbols are the corresponding analytical results for $\delta = 0.001$. We make $p = 0.5$ for NW networks and $l = 0.5$.

order for the emergence of cooperation, individuals do not necessarily consider others' income more than themselves. It is very interesting and fascinating. Of course, if they tend to take into account more contributions than incomes (i.e. $h > 0.5$), full cooperation would be easier to sustain.

Table 1. Simulation results of ρ_c for different values of h with $l = 0.0$ (row 2), and for different values of l with $h = 0.3$ (row 4), respectively.

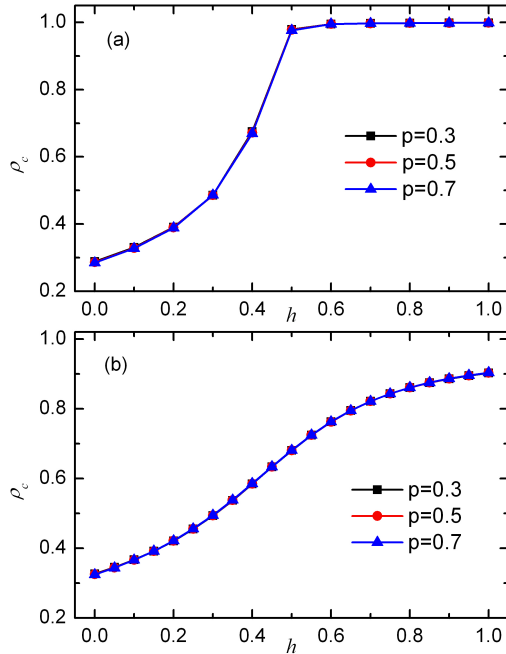| h	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
ρ_c	0.4997	0.5	0.5001	0.4999	0.5001	0.4999	0.4997	0.500	0.5003	0.5002	0.5000
l	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
ρ_c	0.4999	0.4988	0.4973	0.4958	0.4950	0.4941	0.4940	0.4942	0.4952	0.4965	0.49810

Fig. 5. (Color online) The density of cooperators as a function of h for different values of p on NW networks. (a) $l = 0.6$ and $\delta = 0.001$, (b) $l = 0.6$ and $\delta = 0.1$.

On the other hand, when δ does not satisfy the condition $\delta \ll 1$, deterministic switch in strategy update process can not be ignored. It is difficult to obtain the equilibrium points in (9). Nevertheless, especially for $l = 0$, (9) simplifies as $\dot{\rho}_c = (1 - 2\rho_c)\delta$, $\rho_c = 0.5$ is the only stable equilibrium (Alonso et al. [2006]). And interestingly, for $h = 0.3$, (9) simplifies as

$$\dot{\rho}_c = (1 - 2\rho_c)[\frac{2l}{3}(1 - \delta)(1 - \rho_c) + \delta].$$

In this situation, the only stable equilibrium $\rho_c = 0.5$ is independent of learning rate l. Table 1 shows the simulation results of ρ_c for these two situations. We can find that the numerical data all approximate to 0.5, thereby, numerical results are well consistent with theoretical predictions. From these numerical and analytical results, we know that the mean-field approximation can be adopted to study the evolutionary games on NW networks under this learning updating rule. Concretely, the higher the value of h is, the more favorable cooperation becomes. It is because, under this learning update rule, when the relative weight h of the local contribution to the score is high, to reach the

aspiration level, most of individuals would adopt strategy C to make the local contribution to neighbors greater. Otherwise, most of individuals would adopt strategy D to collect much more payoffs from their neighbors. For low δ ($\delta \ll 1$) in strategy update process, the system would eventually reach a steady state with a definite value ρ_c for the fraction of cooperators, and learning rate l would not influence the long run behavior of the system. For other values of δ, when the learning rate l is small, in this case cooperators and defectors tend to be distributed with equal percentage among the players regardless of the value of h. On the other hand, when $h > 0.3$, individuals tend to adopt C strategy with high learning rate. Otherwise, individuals tend to adopt D strategy with high learning rate. Therefore, combining these two factors, cooperation increases as the learning rate increases for $h > 0.3$; whereas cooperation for high learning rate is not more favorable than for small learning rate when $h < 0.3$.

The cooperator density with $l = 0.5$ for different values of intensity of deterministic switch is shown in Fig. 4. We can find that, the density of cooperators for high δ increases much more smoothly than for low δ as h increases. When $h < 0.3$, the cooperator density for low δ is a little greater than for high δ. While for $h > 0.3$, the cooperator density is much more favorable as δ decreases. In fact, when δ is high, the ratio of the deterministic probability to the strategy updating probability tends to be high, hence cooperators and defectors will be distributed with equal percentage among the population (Alonso et al. [2006]). On the other hand, when the ratio of the local contribution to the score is high, individuals tend to adopt C strategy. Therefore, we can understand the simulation results for different values of δ by combining these two factors. Moreover, the analytical results for $\delta = 0.001$ correctly predicts the trends for $\delta = 0.1$, that is, the changes of cooperation for h. However, it is unable to estimate exactly the cooperator density for high δ.

Fig. 5 shows the simulation results of ρ_c as a function of h for different values of probability p on NW networks. It is found that no matter what the value of δ is, there are no differences between cooperation for different values of p, namely, for different values of average degree. Indeed, we can find that the cooperation level is independent of average degree from (9). Moreover, we have confirmed that these simulation results remain valid for different network size N on NW networks. Therefore, it indicates that under this learning updating rule, the cooperation level is not sensitive to the topological parameters. It is shown that cooperation based on this learning feedback mechanism with local contribution is robust, and this mechanism which is analogous to win-stay-lose-shift (Nowak et al. [1993]), can also correct mistakes and exploit unconditional cooperators.

4. CONCLUSION

We have studied the cooperative behavior of the evolutionary PDG under a modified learning updating mechanism on NW networks. Taking into account individual learning in the networked evolutionary PDG, we have defined the score aspiration and actual score for individuals. The score of a given individual is a linear combination of its total payoff and local contribution, which denotes the sum of all the payoffs the neighbors collect against it during the interaction. During the evolutionary process, each individual updates its strategy by comparing its actual score with score aspiration. It is shown that cooperation is promoted when the relative weight h of the local contribution to the score is high. Additionally, when δ in the strategy updating rule is low ($\delta \ll 1$), it is found that the cooperation level is independent of learning rate, and full cooperation can be reached when h is not less than 0.5. Otherwise, cooperation increases as learning rate l increases when $h > 0.3$, whereas cooperation for small learning rate is a little greater than for high learning rate when $h < 0.3$. We have also carried out theoretical analysis of ρ_c using mean-field method, and found that analytical results are well consistent with numerical simulations. Besides, the density of cooperators is qualitatively unchanged for different values of probability p on NW networks.

To conclude, we have illustrated the validity and efficiency of our proposed model in solving the dilemma. With the addition of a simple learning feedback in individual strategy updating process, this model is an extension to the strategy updating rules adopted in (Nowak et al. [1993], Alonso et al. [2006]). The introduced feedback between actual score and score aspiration enables individuals to update their strategies more accurately, and our method might be a plausible stride to explore the role of simple feedback in evolutionary games. The results of this model may shed new light on understanding the cooperative behaviors in society.

REFERENCES

J. Alonso, A. Fernández, and H. Fort. Prisoner's Dilemma cellular automata revisited: evolution of cooperation. *J. Stat. Mech.*, P06013, 2006.

R. Axelrod and W.D. Hamilton. The evolution of cooperation. *Science*, 211:1390–1396, 1981.

R. Axelrod. The Evolution of Cooperation. *Basic Books, New York*, 1984.

T. Borgers and R. Sarin. Naive reinforcement learning with endogenous aspirations. *Int. Econ. Rev.*, 41:921–950, 2000.

M. Doebeli and C. Hauert. Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game. *Ecol. Lett.*, 8:748–766, 2005.

L.A. Dugatkin. *Cooperation among Animals: An Evolutionary Perspective.* Oxford University Press, Oxford, 1997.

H. Fort and N. Pérez. Economic demography in fuzzy spatical dilemmas and power laws. *Eur. Phys. J. B*, 44: 109–113, 2005.

H. Fort and N. Pérez. The Fate of Spatial Dilemmas with Different Fuzzy Measures of Success. *J. Artif. Soc. Soc. Simul.*, 8(3), 2005.

D. Fudenberg and D. Levine. The Theory of Learning in Games. *MIT Press, Cambridge*, 1998.

L. Gulyás. Cooperation in networked populations of selfish adaptive agents: Sensitiviy to network structure. *Physica A*, 378:110-117, 2007.

W.D. Hamilton. The genetical evolution of social behaviour II. *J. Theor. Biol*, 7:17–52, 1964.

C. Hauert and M. Doebeli. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature*, 428:643–646, 2004.

J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics.* Cambridge University Press, Cambridge, 1998.

M.W. Macy and A. Flache. Learning dynamics in social dilemmas. *Proc. Natl. Acad. Sci. U.S.A.*, 99:7229–7236, 2002.

M.E.J. Newman and D.J. Watts. Scaling and percolation in the small-world network model. *Phys. Rev. E*, 60: 7332–7342, 1999.

M.A. Nowak and R.M. May. Evolutionary games and spatial chaos. *Nature*, 359:826–829, 1992.

M.A. Nowak and W.D. Hamilton. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, 364:56–58, 1993.

M.A. Nowak and K. Sigmund. Evolutionary Dynamics of Biological Games. *Science*, 303:793–799, 2004.

M.A. Nowak and K. Sigmund. Evolution of indirect reciprocity. *Nature*, 437:1291–1298, 2005.

M.A. Nowak. Five Rules for the Evolution of Cooperation. *Science*, 314:1560–1563, 2006.

H. Ohtsuki, C. Hauert, E. Lieberman, and M.A. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441:502–505, 2006.

J.M. Pacheco, F.C. Santos, and F.A.C.C. Chalub. Stern-Judging: A Simple, Successful Norm Which Promotes Cooperation under Indirect Reciprocity. *PLoS Comput Biol*, 2(12):1634–1638, 2006.

M. Perc. Chaos promotes cooperation in the spatial prisoner's dilemma game. *Europhys. Lett.*, 75:841–846, 2006.

M. Posch, A. Pichler, and K. Sigmund. The efficiency of adapting aspiration levels. *Proc. R. Soc. London, Ser. B*, 266:1427–1435, 1999.

M. Posch. Win-Stay, Lose-Shift Strategies for Repeated GamesMemory Length, Aspiration Levels and Noise. *J. Theor. Biol.*, 198(2):183–195, 1999.

R.L. Riolo, M. D. Cohen, and R. Axelrod. Evolution of cooperation without reciprocity. *Nature*, 414:441-443, 2001.

J.M. Smith. *Evolution and the Theory of Games.* Cambridge University Press, Cambridge, 1982.

G. Szabó and G. Fáth. Evolutionary games on graphs. *Phys. Rep.*, 446:97–216, 2007.

A. Szolnoki and G. Szabó. Cooperation enhanced by inhomogeneous activity of teaching for evolutionary Prisoner's Dilemma games. *Europhys. Lett.*, 77(30004), 2007.

A. Trulsen and M.A. Nowak. Evolution of cooperation by multilevel selection. *Proc. Natl. Acad. Sci. U.S.A.*, 103: 10952–10955, 2006.