Fuzzy Strategy Updating in the Prisoner's Dilemma Game

Hongmin Deng, Yong Zhi, Hui Hu
University School of Electronics and Information Engineering
Sichuan University, SCU
Chengdu, China
denghongmin@yahoo.com.cn

***Abstract* - This paper investigates a fuzzy strategy updating mechanism based on a set of fuzzy rules. The proposed scheme is applied to a special Prisoner's Dilemma (PD) game where the player does not know accurate payoffs of its neighbors and even its own (as is often the case in real situations). And the proposed scheme is compared with two popular strategy updating approaches, namely, the Fermi-Dirac function based scheme and the maximal aggregate payoff based scheme (MAX-payoff scheme). The simulation results illustrate that the evolution from cooperation to defection by using fuzzy strategy update converges faster while increasing temptation *T* and also exhibits much sharp step-like characteristic. At the same time, it turns out that self interaction plays more important role in the cooperation of fuzzy strategy update technique. Finally, our model is based on some real situations because fuzzy rules are good at expressing approximate and qualitative knowledge.**

Keywords* - *Fuzzy strategy update, membership functions, fuzzy rule base, approximate payoff, game theory, cooperation

I. INTRODUCTION

The Prisoner's Dilemma (PD) game is one of the most classic games. Numerous examples of the PD game have been found in ecosystems, social sciences, economic activities and so on [1-11].

In the classic PD game, the interactions of the players are usually described by a 2x2 payoff matrix [12]:

$$A = \begin{array}{c c} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array} \tag{1}$$

where C (cooperation) and D (defection) denote two strategies each player can select in each round, and $T > R > P > S$. The entries of the payoff matrix are interpreted as follows: when both players choose cooperation, each player gets reward R ; when they both choose defection, each player deserves punishment P ; when one defects and the other cooperates, the defector obtains payoff T , while the sucker gets payoff S . In this sense, each player should tend to defect because it will get more total payoff irrespective of the strategy of its opponent (for $T > R$ and $P > S$). As a result, each only gets the penalty P . In contrast, when they both choose cooperation ($2R > T + S$), they will get higher total payoffs in the long run. This is the dilemma. Based on the Nowak and collegues' idea, a rescaled form about this matrix where $T > 1, R = 1, P = S = 0$ is usually adopted.

Following their pioneering work, diversified representation schemes of strategy and different neighborhood structures for different games were investigated [13]. Strategy updating techniques for the PD game generally fall into two categories: (1) MAX-payoff strategy updating scheme [12, 14]; (2) Fermi-Dirac function based strategy updating scheme [6, 15].

In the second category, noise is usually added into the Fermi-Dirac function in order to allow for some irrational choices. In the MAX-payoff strategy updating technique, an individual in each site plays against its neighbors including itself. Each individual adopts the strategy of the player who gains the highest total payoff. The total payoff P_i of player i is calculated as follows:

$$P_i = \sum_{\substack{j \\ j \in \Omega_i}} s_i^T A s_j \tag{2}$$

where Ω_i is the neighborhood of player i including itself, A denotes the payoff matrix (1), $(\bullet)^T$ denotes the transpose. And s_i, s_j satisfy the following requirements: If player i chooses defection, $s_i = (0 \quad 1)^T$; If player i chooses cooperation, $s_i = (1 \quad 0)^T$.

In the Fermi-Dirac function based strategy updating technique, each individual plays against the player randomly chosen from its neighbors. And each individual adopts probabilistically its opponent's strategy by the formula of Fermi-Dirac function. For instance, player i calculates its aggregate payoff and compares its score with that of the randomly chosen player j from the neighbors of player i. Eq. (3) is used to determine the probability P_O of player i adopting the strategy of its neighbor j:

$$P_O\left(x_i \leftarrow x_j\right) = \frac{1}{1+\exp[(P_i - P_j)/K]} \qquad (3)$$

where $i, j = 1,2,...N$, and N is the population size, K is the measurement of stochastic uncertainty, called the amplitude of noise or temperature, the inverse ($1/K$) denotes the intensity of selection, P_i and P_j are both determined by Eq.(2).

The above-mentioned two strategy updating techniques well describe most of the evolutionary dynamics. However, sometimes a player doesn't know the exact payoff of its opponents and even its own, how could the player decide its strategy update? Motivated by the fact that fuzzy logic is good at dealing with approximate uncertainty, we adopt a fuzzy rule based strategy updating scheme in this paper. And we obtain a series of reasonable computer simulation results under different spatial structures and membership functions based on a 49-fuzzy-rule base.

There have been a multitude of studies on fuzzy game theories since 1965 [16]. For example, Aubin introduced fuzzy concept in the PD game in early 80s [17]. Campos studied a non-cooperative game with fuzzy payoffs [18] based on fuzzy linear programming. Arfi developed linguistic fuzzy logic which admitted linguistic truth values and predicted its applications in social sciences [19]. Fort and Pérez investigated the economic demography in fuzzy spatial dilemma [20]. More recent works in this area were reported [21-23]. However, most of these studies were focused on the mathematical aspects of fuzzy game theories such as Nash Equilibriums, fuzzy linear programming, etc. There is a paper worthy of mentioning, Borges and Pacheco investigated a fuzzy model to the prisoner's dilemma, where they do not thought the players' decisions as dichotomous choices, but as a variable intensity of cooperation or defection [24]. They introduced three partial decisional factors, in three membership functions and based on three fuzzy expert systems, respectively. In our paper, we construct another fuzzy model to deal with some situations where the players' accurate payoffs are not available. We take the players' payoffs as fuzzy inputs, the winning possibility as fuzzy outputs. Through fuzzy reasoning, we get the possibility of one player forcing its strategy on its opponent in the strategy update.

In this paper, we provide a detailed comparison for the three strategy updating techniques in four different spatial structures: The MAX-payoff scheme, the Fermi-Dirac function based scheme and the proposed fuzzy strategy updating scheme. There are the following findings: (1) The proposed fuzzy approach shows generally faster decreasing from pure cooperation to full defection in various neighborhood structures. (2) The cooperative behavior in our fuzzy approach exhibits more sharp step-like decrease than that of the MAX-payoff strategy updating technique, and the former also has lower thresholds of temptation for the step-like decrease. In contrast, the approach based on Fermi-Dirac function has continuously and monotonously decreasing behavior. Furthermore, for both the MAX-payoff and fuzzy strategy updating techniques, the step-like behavior becomes weak as the number of neighbors grows. (3) It results in more important changes for fuzzy update than other two traditional updates when self interaction is not involved.

The remainder of this paper is organized as follows: Section II proposes the fuzzy strategy updating approach, including the construction of fuzzy sets, fuzzy rules and membership functions, as well as the method of fuzzification and defuzzification for strategy updating. Section III presents the computer simulation results. A comparison analysis of the simulation results is presented in Section IV, followed by conclusion and future outlook in Section V.

II. THE PROPOSED FUZZY STRATEGY UPDATING APPROACH

Fuzzy control primarily utilizes fuzzy logic and inference to imitate human behaviors and make human-like decisions in a real world [25-28]. And it includes three basic processes: Fuzzification, Fuzzy inference and Defuzzification. In this paper we focus on the fuzzy strategy update of the PD game where an individual plays against the randomly selected one from its neighbors.

A . Fuzzification

(1) Fuzzy set

Firstly, we define three variables: two input variables x, y and one output variable z. Each input variable represents the approximate cumulative payoffs including the player itself and all its neighbors. The output variable denotes a possibility of beating the opponent and keeping its own strategy in the next round of game.

Secondly, we denote each of the inputs and output with a fuzzy set { $NB, NM, NS, ZE, PS, PM, PB$ } including six linguistic variables, respectively. For the input variable , each entry of the fuzzy set represents a summation of payoffs of a player. For example, ZE (Zero) means the moderate summation of payoffs, while NB, NM, NS, PS, PM and PB denote Negative Big, Negative Medium, Negative Small, Positive Small, Positive Medium and Positive Big, respectively, with respect to the moderate value ZE. For the output

DOI 10.5013/IJSSST.a.13.3A.02

ISSN: 1473-804x online, 1473-8031 print

variable, each entry of the fuzzy set represents a likelihood of one player keeping its own strategy in the next round of game. Specially, NB denotes a little possibility and PB denotes a great possibility. If a player's accumulative payoff is very small (NB), but that of its opponent is very large (PB), the player will probably revise its strategy and adopt the opponent's strategy, i.e., the chance for the player to beat its opponent is very small (NB), and vice versa. If the two players have approximately equivalent payoffs, the chance of winning by either of them exists in this case.

(2) Fuzzy membership function

We represent the two input variables and one output variable in the same universe of discourse U = {-6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6}. Based on the universe of discourse U we define a membership function in (4):

$$\mu_F(u) \in [0,\ 1] \text{ , where } u \in U \tag{4}$$

A membership function can take different forms, e.g., a triangular function, a Gaussian function, and so on and so forth. In this paper, two types of fuzzy membership functions are applied in the universe of discourse to consider the effects of membership functions with different shapes. The two membership functions are shown in Fig. 1 and Fig. 2, respectively.

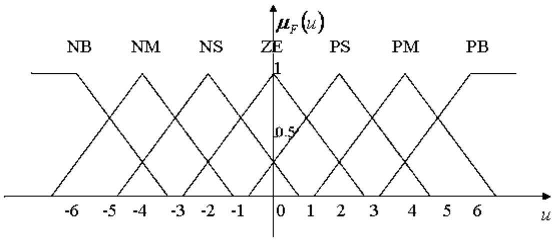

Figure 1. Membership function 1

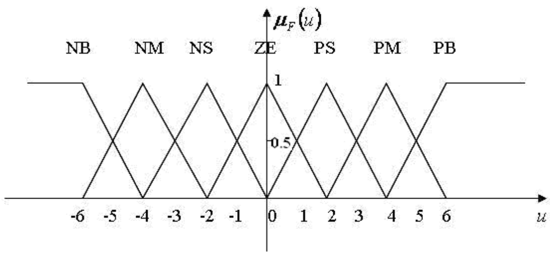

Figure 2. Membership function 2

The membership functions can be represented by mathematical expressions. For instance, a part of membership 1 can be depicted as Eq. (5a) and Eq. (5b), so on and so forth.

$$\begin{cases} \mu_{ZE} = -(7/20)x+1, & if \quad 0 < x < 2.85 \\ \mu_{ZE} = (7/20)x+1, & if \quad -2.85 < x < 0 \\ \mu_{ZE} = 0, & otherwise \end{cases}$$
(5a)

$$\begin{cases} \mu_{PS} = (7/20)x+3/10, & if \quad -6/7 < x < 2 \\ \mu_{PS} = -(7/20)x+17/10, & if \quad 2 < x < 34/7 \\ \mu_{PS} = 0, & otherwise \end{cases}$$
(5b)

B . Fuzzy logic and fuzzy rules

After defining fuzzy sets and fuzzy membership functions in subsection A, we construct empirically a 49-fuzzy-rule, as shown in TABLE I.

The IF-THEN rules in the rule base can be described as: *"If the payoffs of two players in the current round are A, B, respectively, then the possibility of player 1 beating player 2 is C in the next round."*

Let x and y be the payoffs of player 1 and player 2, respectively, and z (the diagonal element in the rule base) be the possibility of player 1 beating player 2 in the next round. Then fuzzy rule inference can be drawn from TABLE I. For example:

R1 : IF x is NB , AND y is NM , THEN z is NS ;

R2 : IF x is NB , AND y is NS , THEN z is NS ;

R3 : IF x is NM , AND y is NM , THEN z is ZE ;

R4 : IF x is NM , AND y is NS , THEN z is NS .

TABLE I THE FUZZY RULE BASE

x \ y	*NB*	*NM*	*NS*	*ZE*	*PS*	*PM*	*PB*
NB	*ZE*	*NS*	*NS*	*NM*	*NM*	*NB*	*NB*
NM	*PS*	*ZE*	*NS*	*NS*	*NM*	*NM*	*NB*
NS	*PM*	*PS*	*ZE*	*ZE*	*NS*	*NS*	*NM*
ZE	*PM*	*PS*	*ZE*	*ZE*	*ZE*	*NS*	*NM*
PS	*PM*	*PS*	*PS*	*ZE*	*ZE*	*NS*	*NM*
PM	*PB*	*PM*	*PM*	*PS*	*PS*	*ZE*	*NS*
PB	*PB*	*PB*	*PM*	*PM*	*PS*	*PS*	*ZE*

C . Defuzzification

Defuzzification is the process of producing a quantifiable result in fuzzy logic, given fuzzy sets and corresponding membership functions. In this paper we adopt COA (Centroid of area) defuzzification method. The COA method represents the final output z^* by computing the centroid of the area of the possibility distribution.

$$z^* = \frac{\sum_{i=1}^{M} \mu(z_i) z_i}{\sum_{i=1}^{M} \mu(z_i)} \tag{6}$$

where z_i is the corresponding support value of the output membership function $\mu(z_i)$ in the ith activated rule, and M is the number of activated rules.

III. Computer Simulations

In our computer simulations, we have the following experimental settings:

Four different neighborhood structures are discussed with the same population size and periodic boundary conditions. Figs. 3(a) and 3(b) denote the Von Neumann neighborhood (n=4) and the Moore neighborhood (n=8) in 2D square lattices. Meanwhile, Figs. 3(c) and 3(d) denote two 3D neighborhood structures where each agent has 6 neighbors (n=6) and 26 neighbors (n=26), respectively. Considering that the topological relationship is abstracted (although the neighborhood structures are denoted by 2D or 3D), it is not defined with the concept of real radius or distance among the neighbors. The simulation results about the evolution of cooperative behaviors in Fig. 4 are calculated after 1000 rounds of iteration and stochastically averaged over 10 trials. The initial proportion of cooperation is independently and randomly selected for all agents as $F_C(0) = 0.5$.

Synchronous updates are adopted where any agent decides its strategy at the same time in Fig. 4.

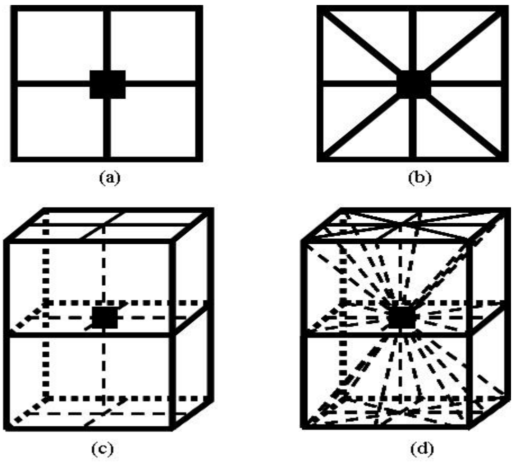

Figure 3. Four kinds of neighborhood structures for the study of an evolutionary PD game is studied (a) Von Neumann neighborhood, 9801(99×99) lattice (b) Moore neighborhood, 9801(99×99) lattice (c) 6 neighbors, 9801(99×33×3) lattice (d) 26 neighbors, 9801(99×33×3) lattice

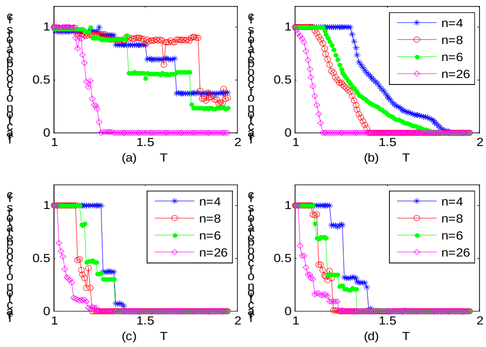

Figure 4. The fraction of cooperators with different strategy updating techniques in periodic boundary conditions and with self interactions. (a), (b), (c) and (d) correspond to the MAX-payoff strategy update, the Fermi-Dirac function strategy update, and the fuzzy strategy update with membership function 1 and membership function 2, respectively

In Fig. 4, it shows the results of the cooperation levels of the three updating strategies as a function of a temptation parameter T in the payoff matrix (1) in the four neighborhood structures. Figs. 4(a) and 4(b) depict the evolution of cooperation by two conventional strategy updates based on the MAX-payoff approach and the Fermi-Dirac function approach, respectively. The proposed fuzzy strategy evolution based on 49 fuzzy rules in TABLE I is examined by two fuzzy membership functions, and the results are shown in Figs. 4(c) and 4(d), respectively.

Comparing Figs. 4(a)-4(d), we observe that the evolution of the cooperative behavior using the fuzzy updating scheme transits faster from pure cooperation to full defection with the increase of temptation T. On the other hand, it exhibits sharp step-like decreases in this fuzzy approach. Meanwhile, in the strategy update based on Fermi-Dirac function, it is continuous and decreases monotonously although it also shows some step-like property with reduce of noise K [15]. The cooperative behavior of the MAX-payoff approach also shows step-like decreases, but in smaller steps in comparison with the fuzzy controller. In the mean time, the results in Figs. 4(c) and 4(d) exhibit the similar changing trends for both membership functions in our study.

To better understand the fuzzy strategy updating technique, we extend our study to calculating the frequency of cooperation with fuzzy strategy updating techniques in dependence on T and R of matrix (1) for different spatial structures. And the contour plots are shown in Fig. 5. It is obvious that when T is comparably

DOI 10.5013/IJSSST.a.13.3A.02 ISSN: 1473-804x online, 1473-8031 print

small and R is relatively large (under the condition $T > R$), that means the temptation is little and the benefits of cooperation are large, so there is good cooperation in this area (dark brown area). With the increase of T or the decrease of R, the cooperation deteriorates gradually till to full defection, where the color from dark brown to dark blue corresponds to cooperation levels from high to low.

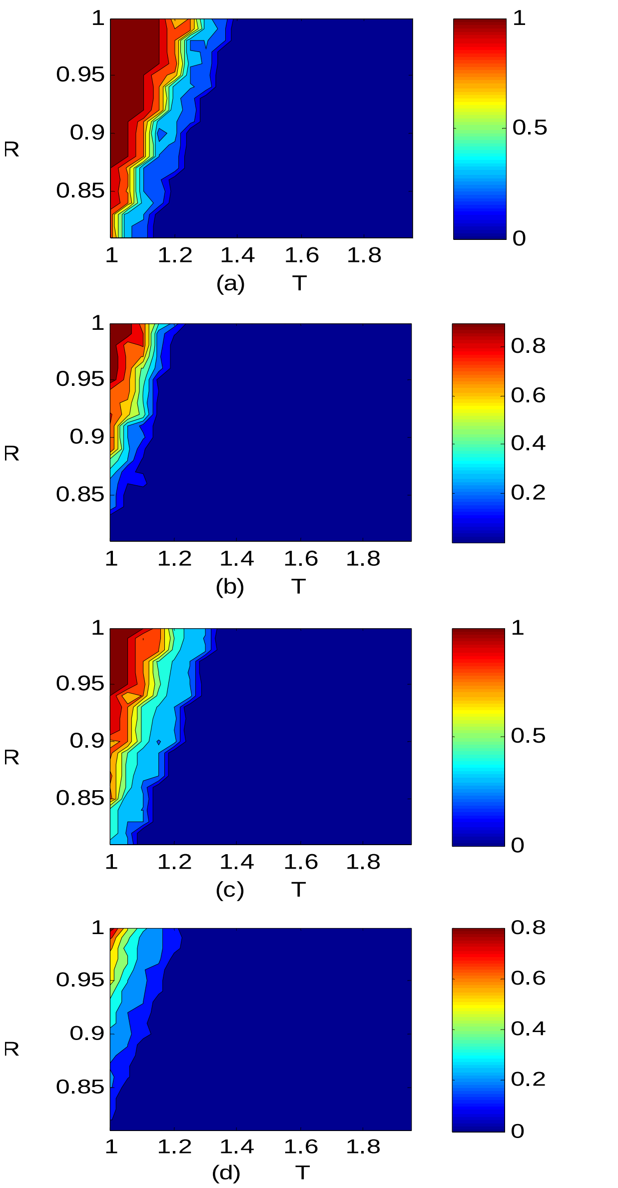

Figure 5. The evolution of cooperators using fuzzy strategy updating techniques in dependence on T and S for different spatial structures under periodic boundary conditions and with self interactions after 100 epochs. (a) Von Neumann neighborhood, 9801(99 × 99) lattice (b) Moore neighborhood, 9801(99×99) lattices (c) 6 neighbors, 9801(99×33 ×3) lattices (d) 26 neighbors, 9801(99×33×3) lattices

IV. ANALYSIS OF THE VARYING TRENDS

In this section we analyze the varying trends in the PD game under different strategy updating techniques and spatial structures.

We examine the results by comparing the fuzzy approach with the other two traditional techniques. And we find the evolution of cooperative behavior in the fuzzy strategy updating technique not only shows the changes in staircase, but also there are sharp transitions from one state to another. While in the MAX-payoff technique, the transition from full cooperation to full defection and from one relatively stationary state to another is slower although it also shows some step-like behavior which is due to the discrete nature of the total payoff [15]. However, the transition from full cooperation to full defection is smoother in the Fermi-Dirac function based approach. The evolution curve of cooperation decreases almost continuously and monotonously although the step-like behavior appeared gradually as the noise parameter K became small enough [15]. That means a high amount of noise causes the levels of the function to become smooth and exhibit a power-law behavior with exponents close to the DP driving point value at both ends of the active region. In other words, randomness leads to its weakening of the step-like feature. Next, through the fuzzy controlling curve in each round, we find the controlling surface based on fuzzy control is unsmooth as shown in Fig. 6. So this also leads to the step-like behavior of fuzzy evolution of cooperation in some sense.

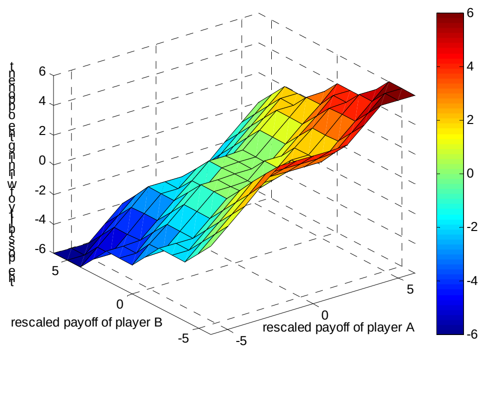

Figure 6. The controlling surface of the fuzzy controller

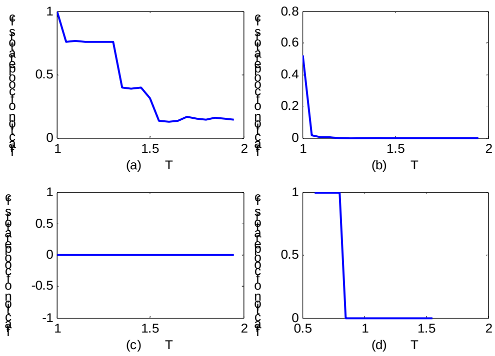

Figure 7. The fraction of cooperators with different strategy updating techniques in periodic boundary conditions and without self interactions (The number of neighbors n=4) (a)the MAX-payoff strategy update (b) the Fermi-Dirac function strategy update (c) the fuzzy strategy update with membership function 1 (T>R) (d) the fuzzy strategy update with membership function 1 (T<=R can be selected) respectively

Finally, we make the simulation in the situation without self-interaction. It shows that the result in fuzzy strategy update steadily converges to extinction (Fig. 7(c)), while the other two strategy updates do not (Fig. 7(a), Fig. 7(b)), although their areas of cooperation shrink. An interesting observation is found that cooperators can be maintained when the parameter T (originally $T > R$, see Section I) is less than 1, but this is not the case in the requirements of our PD game (see Fig. 7(d)). So self-interaction becomes more important for cooperation in the fuzzy updating strategy.

V. Conclusion And Future Work

In this paper we propose a fuzzy strategy updating technique for the PD game where precise payoffs are unavailable, and the evolution from cooperation to defection is comparatively analyzed for three strategy updating techniques (MAX-payoff technique, Fermi-Dirac function based technique and the proposed fuzzy strategy updating technique) respectively. We get the following results.

A common trend for the three strategy updating techniques is a transition from full cooperation to pure defection as temptation parameter T increases. However, there are some significant differences. Firstly, the fuzzy strategy updating technique has generally faster convergence speed than the other two schemes from full cooperation to pure defection. Secondly, there exists more steep step-like decrease in the fuzzy approach. For the MAX-payoff technique, the step-like behavior has much slower attenuation speed. While Fermi-function approach has monotonous decrease and continuous evolution curve. Finally, the different effects on the three strategy updating techniques without considering self interaction show that self interaction plays more important role for fuzzy strategy update.

In this paper, the fuzzy rule base and membership functions designed empirically are heuristic. In future work we will investigate how the optimization of the membership function and fuzzy rule base has influence on the activity of fuzzy strategy updating technique.

Acknowledgment

This work is supported by the National Natural Science Foundation of China under Grant No.61174025. The authors wish to thank all the anonymous referees and editors for their valuable comments and suggestions.

References

[1] A. Yamauchi, J. Tanimoto, A. Hagishima, "An analysis of network reciprocity in Prisoner's Dilemma games using Full Factorial Designs of Experiment," *Biosystems*, vol. 103, pp. 85–92, 2011.

[2] M. Doebeli, C. Hauert, T. Killingback, "The evolutionary origin of cooperators and defectors", *Science*, vol.306, pp.859–862, 2004.

[3] N. Masuda , K. Aihara, "Spatial prisoner's dilemma optimally played in small-world networks," *Physics Letters A*, vol.313, pp.55-61, 2003.

[4] J. M. McNamara, Z. Barta , A. I. Houston, "Variation in behaviour promotes cooperation in the Prisoner's Dilemma game," *Nature*, vol.428, pp.745-748, 2004.

[5] C. Gracia-Lázaro, J. A. Cuesta, A. Sánchez & Yamir Moreno, "Human behavior in Prisoner's Dilemma experiments suppresses network reciprocity," *Scientific Reports*, vol. 2, DOI:10.1038/srep00325, 2012.

[6] M. Perc, "Chaos promotes cooperation in the spatial prisoner's dilemma game," *Europhysics Letters*, vol. 75, pp. 841-846, 2006.

[7] M. Perc , A. Szolnoki, "Social diversity and promotion of cooperation in the spatial prisoner's dilemma game," *Physical Review E* 77, 011904, 2008.

[8] T. Platkowski, "Enhanced cooperation in prisoner's dilemma with aspiration," *Applied Mathematics Letters*, vol. 22, pp. 1161-1165, 2009.

[9] G. Szabó and A. Szolnoki, "Cooperation in spatial prisoner's dilemma with two types of players for increasing number of neighbors," *Physical Review E* 79, 016106, 2009.

[10] S. Számadó, F. Szalai, I. Scheuring, "The effect of dispersal and neighborhood in games of cooperation," *Journal of Theoretical Biology*, vol. 253, pp. 221-227, 2008.

[11] J. Tanimoto, "Promotion of cooperation through co-evolution of networks and strategy in a 2×2 game," *Physica A*, vol. 388, pp. 953-960, 2009.

[12] M. A. Nowak and R. M. May, "Evolutionary games and spatial chaos," *Nature* (London), vol. 359, pp. 826-829, 1992.

[13] H. Ishibuchi, H. Ohyanagi, Y. Nojima, "Evolution of cooperative behavior in a spatial iterated prisoner's dilemma game with different representation schemes of game strategies," *IEEE International Conference on Fuzzy Systems*, pp.1568-1573, 2009.

[14] M. A. Nowak and R. M. May, "The spatial dilemmas of evolution," *International Journal of Bifurcation and Chaos*, vol. 3, pp. 35-78, 1993.

[15] G. Szabó, C. Töke, "Evolutionary prisoner's dilemma game on a square lattice," *Physical Review E*, vol. 58, pp. 69-73, 1998.

[16] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338-353, 1965.

[17] J.P. Aubin, "Cooperative fuzzy games," *Mathematics of Operations Research*, vol. 6, pp. 1-13, 1981.

[18] L. Campos, "Fuzzy linear programming models to solve fuzzy matrix games," *Fuzzy Sets and Systems*, vol. 32, pp. 275-289, 1989.

[19] B. Arfi, "Linguistic fuzzy-logic social game of cooperation, " *Rationality and Society*, vol. 18, pp. 471-537, 2006.

[20] H. Fort, and N. Pérez, "Economic demography in fuzzy spatial dilemmas and power laws," *The European Physical Journal B*, vol. 44, pp. 109-113, 2005.

[21] Y. Chen, M. Larbani, "Two-person zero-sum game approach for fuzzy multiple attribute decision making problem," *Fuzzy Sets and Systems*, vol. 157, pp. 34-51, 2006.

[22] J. Gao, Z. Liu, P. Shen, "On characterization of credibilistic equilibria of fuzzy-payoff two-player zero-sum game", *Soft Computing*, vol. 13, pp. 127-132, 2009.

[23] H. C. Wu, "Core of fuzzy games and their convexity," *Fuzzy Sets and Systems*, vol. 198, pp. 59-69, 2012.

[24] P. S. S. Borges, R. C. S. Pacheco, R. M. Barcia. "A fuzzy approach to the prisoner's dilemma," *BioSystems*, vol. 41, pp.127-137, 1997.

[25] Precup R. E., Hellendoorn, H. "A survey on industrial applications of fuzzy control, " *Computers in Industry*, vol. 62, pp. 213-226, 2011.

[26] F. Farbiz, M. B. Menhaj and S. A. Motamedi, "An Adaptive C-Average Fuzzy Control Filter for Image Enhancement", *Lecture Notes in Computer Science*, vol. 1625, pp. 169-176, 1999.

[27] I. Harmati, K. Skrzypczyk, "Robot team coordination for target tracking using fuzzy logic controller in game theoretic framework," *Robotics and autonomous systems*, vol.57, pp. 75-86, 2009.

[28] J. M. C. Sousa, U. Kaymak, "Fuzzy decision making in modeling and control" (Ed.), *World Scientific Series in Robotics and Intelligent Systems*, vol. 27, 2002.