

PSO Algorithm for IPD Game

Xiaoyang Wang^{1*}, Yibin Lin²

¹Business School, Sun-Yat Sen University, Guangzhou, Guangdong, China, ²School of Software, Sun-Yat Sen University, Guangzhou, Guangdong, China

Abstract

Mechanisms promoting the evolution of cooperation in two-player, two-strategy evolutionary games have been discussed in great detail over the past decades. Understanding the effects of repeated interactions in multi-player with multi-choice is a formidable challenge. This paper presents and investigates the application of co-evolutionary training techniques based on particle swarm optimization (PSO) to evolve cooperation for the iterated prisoner's dilemma (IPD) game with multiple choices in noisy environment. Several issues will be addressed, which include the evolution of cooperation and the evolutionary stability in the presence of multiple choices and noise. First is using PSO approach to evolve cooperation. The second is the impact of noise on the evolution of cooperation is examined.

Keywords:

co-evolution, iterated prisoner's dilemma (IPD), particle swarm optimization (PSO), social component, noise

1. Introduction

In the modern society, individuals are often given a large degree of freedom and anonymity. This in turn allows some individuals to take advantage of the efforts of the others, which is called free riders. Achieving a satisfactory understanding of how cooperation can be promoted and maintained in the presence of free riders remains an open and challenging question. IPD provides a way to simulate and analysis the behaviours of players in some dynamic and complex environment, and it has been widely studied in many areas [1-5], such as sociology, economics, biology, and other interactions to occur within a group of selfish individuals [29]. In this well-known nonzero-sum game [2], two isolated players repeatedly choose one of the two actions: cooperate and defect. The dilemma for the players is that they will eventually receive low payoff from mutual defection whereas higher payoff can be obtained from mutual cooperation [6]. In the IPD defection is not always the best choice. Groups of co-operators tend to outperform groups of defectors, but defectors in any mixed group would have the "upper hand" over co-operators [30]. Generally speaking, cooperation levels and social cohesion are correlated with the size of groups and magnitude of the payoff values for a given social dilemma [31].

Over the past decades, a large number of studies have examined mechanisms promoting cooperative behaviour in social dilemmas. Axelrod [7] pointed out the importance of the IPD game and proved that a specific form of cooperative play can make the strategies perform well by using tournament competitions. Many researchers also confirmed that by using an evolutionary

algorithm [8-12], even though the evolution is not stable, the cooperative play can also be learned from a set of random strategies [13]. The classical IPD model is extended to bridge the gap between the model and real-world dilemmas, especially about social and economic ones. One of them is the consideration of various situations such as a multiplayer version [6], a spatial version [14-16], stochastic strategies [17], random pairing [18], multiple objectives [19], noisy games [20], Q-Learning [21] and the trust based model [22,28-31].

As with many other heuristic researches, PSO has also been used in the context of prisoner's dilemma game. Franken and Engelbrecht [23, 24] investigated two different approaches using PSO to evolve strategies, one is Binary PSO and the other is neural networks. Chio [25] integrated strategies from the prisoner's dilemma into the PSO algorithm. They use these strategies to represent different methods to evaluate each particle's next position. Those studies mainly focused on the IPD game played by two players and with two choices, cooperation and defection. Use PSO algorithm to evolve cooperation in multiple choices IPD game with multiple players were not adequately considered.

Motivated by the findings above, this paper presents a detailed study of using PSO approach to evolve cooperation in the IPD game with multiple choices and noise in a co-evolutionary environment. Two major research issues are examined. First, the behaviour performances of players with different environmental factors are examined, such as the number of choices they encounter, the size of population et al. Second, the influence of noise is tested on the evolution of cooperation in the IPD game with multiple choices. Finally, implications of experimental results to real-world dilemmas are discussed.

The rest of the article is organized as follows. Section 2 briefly introduces the IPD game. The PSO approach and the application to IPD game are described in detail in section 3. Section 4 presents and analysis the experimental results in noisy environment, followed by Section 5 which concludes this paper and does some remarks for future work.

2. IPD Game

2.1. Classical IPD Game

A player would receive payoff as the payoff matrix set when his opponent makes his choice in IPD game [2]. Referring to the payoff matrix given by Table1, the payoff matrix satisfies following three conditions:

- 1) $T > R$ and $P > S$ (defection always receives less);
- 2) $R > P$ (mutual cooperation wins mutual defection); and
- 3) $R > (S+T)/2$ (alternating does not pay)

There are many possible values for T (temptation), R (reward), P (punishment) and S (sucker) that satisfy the above conditions (we use $T=5$, $R=4$, $P=1$, $S=0$ in all our experiments unless specific stated).

Table 1. Payoff matrix of prisoner’s dilemma game

		B	
		Cooperate	Defect
A	Cooperate	R	T
	Defect	S	P

2.2. IPD with Multiple Levels of Cooperation

The IPD game with multiple levels of cooperation follows previous work in [10, 11, 2-5]. The multiple choices IPD game is defined by the payoffs obtained through the following linear interpolation in this paper:

$$P_A = 2.5 - 0.5C_A + 2C_B, \quad -1 \leq C_A, C_B < 1$$

Where p_A is the payoff of player A, given that c_A and c_B are the cooperation levels of the choices that agents A and B make respectively. Note that the payoff matrix must satisfy the following conditions in a multiple choices IPD game:

- 1) for $c_A < c'_A$ and constant c_B : $p_A(c_A, c_B) > p_A(c'_A, c_B)$;
- 2) for $c_A = c'_A$ and $c_B < c'_B$: $p_A(c_A, c_B) < p_A(c'_A, c'_B)$;
- 3) for $c_A < c'_A$ and $c_B < c'_B$: $p_A(c'_A, c'_B) > \left(\frac{1}{2}\right) (p_A(c_A, c'_B) + p_A(c'_A, c_B))$.

The above three conditions are similar to those for the classic IPD game. The four corners are the same payoffs for the 2-choice IPD and that any 2×2 sub-matrix of the $n \times n$ matrix is itself a 2-choice IPD.

Given the payoff equation and the three conditions above, an n -choice IPD can be formulated [2, 4]. However, only the number of choices is fixed where the line between full cooperation and full defection is symmetrical distribution of the zero. For example, consider the case of an IPD game of 4 choices as Table 2 shows.

Table 2. Payoff matrix of four-choice IPD game

		B			
		+1	$+\frac{1}{3}$	$-\frac{1}{3}$	-1
A	+1	4	$2\frac{2}{3}$	$1\frac{1}{3}$	0
	$+\frac{1}{3}$	$4\frac{1}{3}$	3	$1\frac{2}{3}$	$\frac{1}{3}$
	$-\frac{1}{3}$	$4\frac{2}{3}$	$3\frac{1}{3}$	2	$\frac{2}{3}$
	-1	5	$3\frac{2}{3}$	$2\frac{1}{3}$	1

3. PSO Approach

The PSO technique was introduced by Kennedy and Eberhart [26]. Inspired by the flocking behaviour of birds, PSO has been applied successfully to function the optimization, game learning, data clustering, and image analysis and neural networks training [24, 25]. PSO involves “flying” a swarm (or population) of n -dimensional particles, and through a problem space, each possible solution to the optimization problem need to search a single optimum or multiple optima. Each particle has its own velocity, a memory of the best position it has obtained thus far (referred to as its personal best position), and knowledge of the best solution found by other particles (referred to as the global best solution).

In the PSO algorithm, each particle adjusts its position in a direction toward its own personal best position in a direction toward its own personal best position and the global best position. The velocity of the particle is calculated using:

$$V_{id}^{k+1} = \omega V_{id}^k + c_1 \text{rand}(0,1)(P_{id}^k - X_{id}^k) + c_2 \text{rand}(0,1)(P_{gd}^k - X_{id}^k) \quad (1)$$

where for particle i , the position vector can be represented by $X_i = (X_{i1}, X_{i2}, \dots, X_{id})$, V_{id} presents the velocity of the i th particle on the specific d -dimension. ω is the inertia weight, c_1 and c_2 are acceleration coefficients. P_{id}^k is the best position of particle i on dimension d at iteration k , and P_{gd}^k is the d th dimensional best position in the whole particle swarm at iteration k .

For particle i , its position x_{id} is changed according to:

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1} \quad (2)$$

Combining with the prisoner’s dilemma framework, each particle in the swarm is associated a behaviour (or “strategy”), and the purpose of the behaviours is to maximize the payoff of individual or the group. The strategy is used to compute its position in the next step. An interpretation of multiple levels of choice between cooperation and defections is employed to the PSO approach.

4. Experimental Studies

In this section, the experimental setup of the paper will be described in more detail. The experiments in the first part are conducted in PSO approach to show the behavioural performances by varying levels of choice and the number of population without noise. The experiments in the second part are conducted with varying degrees of noise levels which are given by the noise probability of occurrence.

4.1. Evolution of cooperation by using PSO approach

The experiments of this part test the co-evolutionary process of behaviour performances in multiple choices IPD game. In order to keep the uniform distribution of choices during the initialization process, each choice is set to players with the same probability. For example, if $PNum=8$ and $L=4$, then each player has 25% probability to get one of the four choices. Each iteration two players are chosen for the game, and the average payoff is calculated between two players.

Figure 2 shows the average frequencies for the choices that are played at each generation of the PSO approach. Because of the limitation of the figure, if the number of choices is higher than four, some specific choices are shown in Figure 2. From Figure 2, it can be found that the frequency of full cooperation (+1) is significantly higher than other choices, indicating that there are more full cooperative plays by the evolving strategies. If the $PNum$ is fixed, comparing Figure 2(a) and (b), (c) and (d), (e) and (f), as the number of choices increases, players choose several levels of choices instead of full cooperation (+1). This is quite different from the classical IPD game with two choices [27], where the mutual cooperation runs quite consistently and quickly. It illustrates that more choices have made cooperation more difficult to evolve.

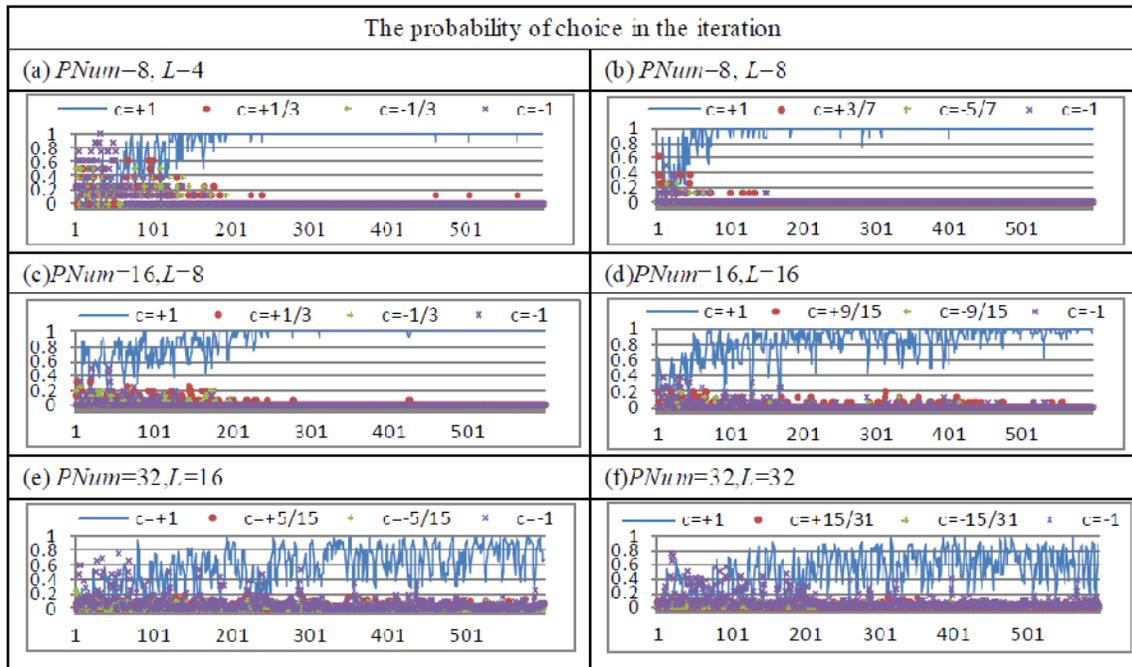


Figure 2. The probability of each choice is calculated during the iteration. The figure can't show each choice in detail if $L > 4$, however, some specific points are show in this figure.

Figure 3 shows the plots of average payoff of each generation. The average payoffs in Figure 3 reflect the changes of the behaviour performances. As the increase of population and choice number, more middle levels of choice are chosen. This experimental finding is rather similar to some phenomena in our human society, first is full cooperation or full defection is easier to emerge in a group of people with small population size than in a larger one; the second is more choices for people leads to more complex choosing diversity.

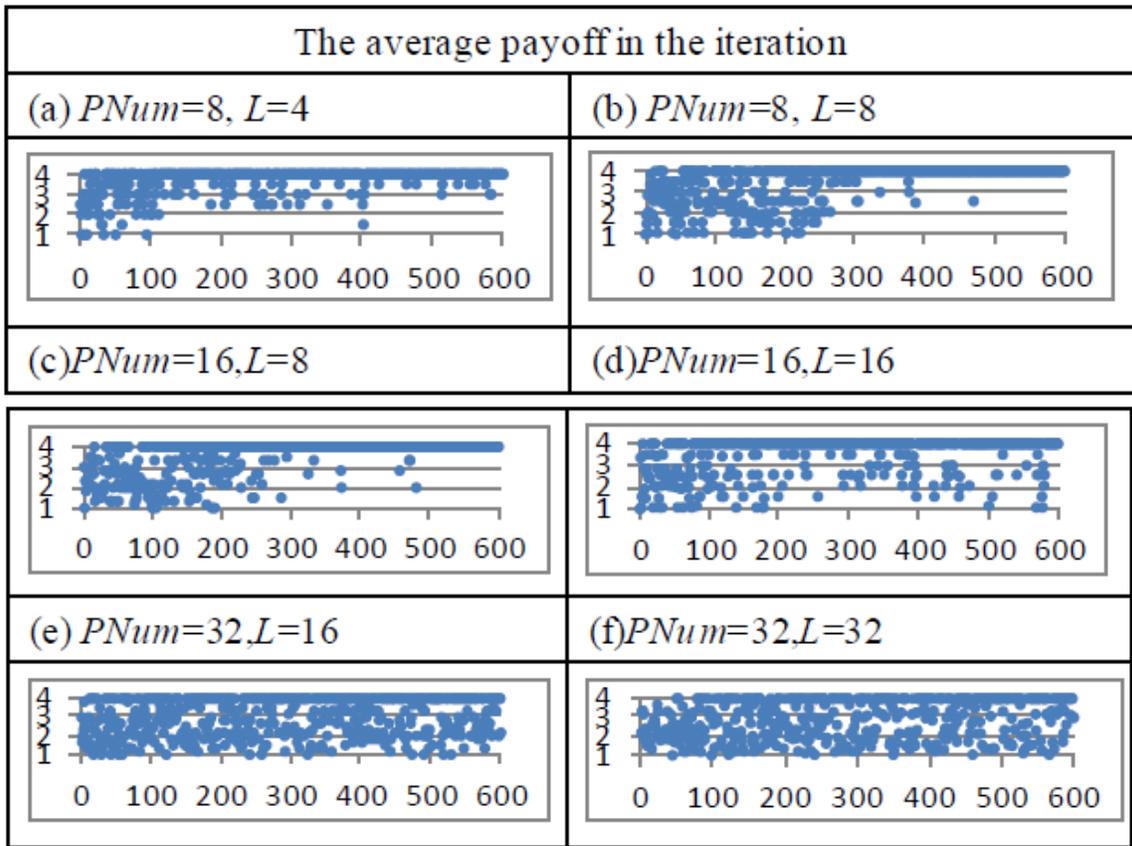


Figure 3. The average payoff of each player in the multiple choices IPD game.

4.2. Evolution in Noisy Environment

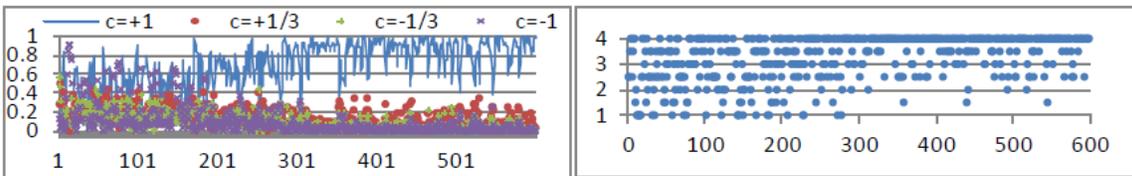


Figure 4. (a) The evolution of frequencies of each choice in 4-choice IPD and without noise. (b) The average payoff of each player in the 4-choice IPD game and without noise.

Figure 4 shows simulation results of a co-evolving population of PSO approach playing the IPD game with 4 choices and without noise, the average payoff is calculated between two players in each generation. It's noted that the mutual cooperation is the most popular choice among the PSO approach.

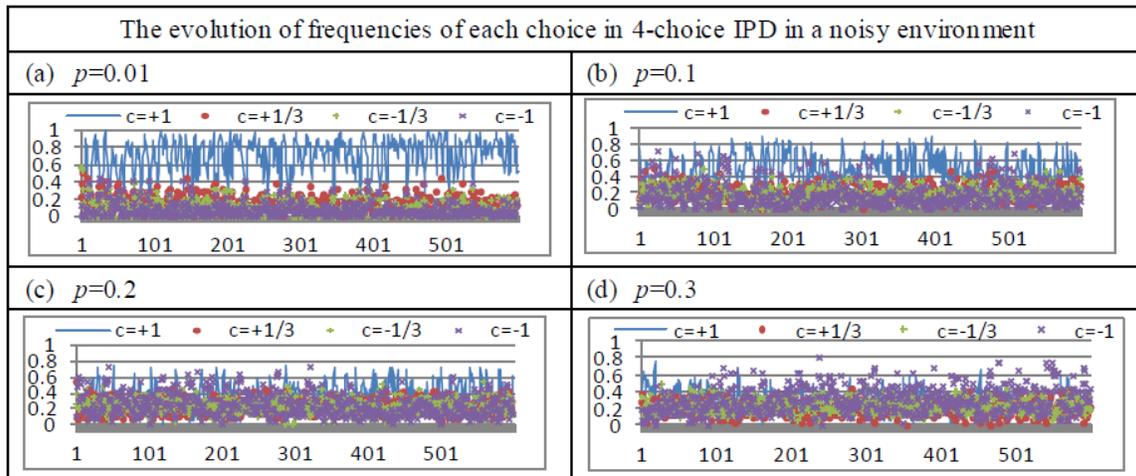


Figure 5. The evolution of frequencies of each choice in 4-choice IPD with noise.

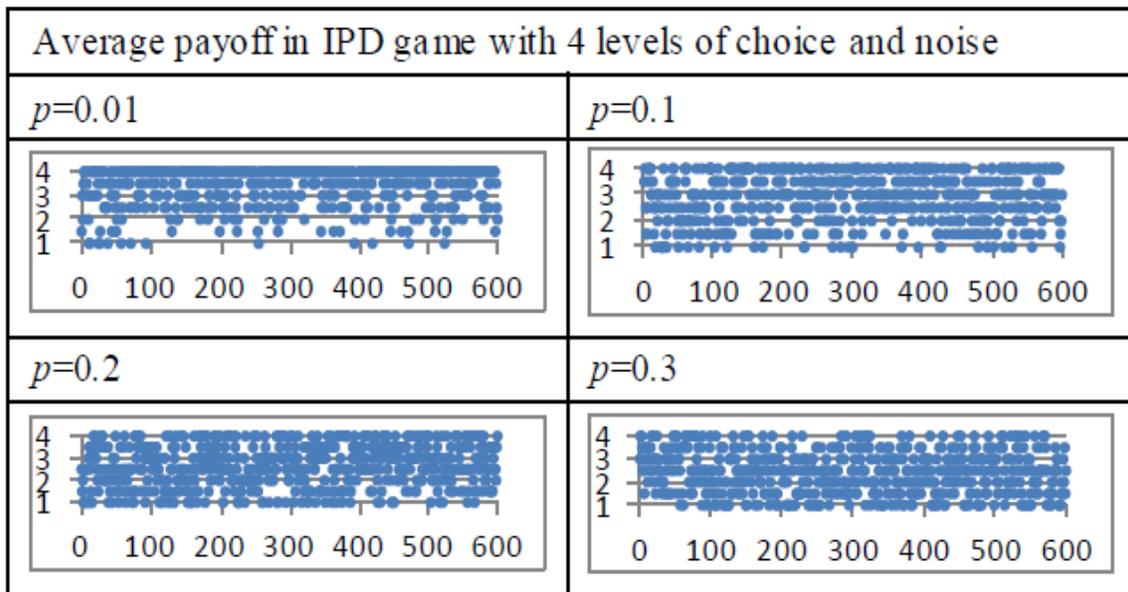


Figure 6. The average payoff of each player in the 4-choice IPD game with noise.

From Figure 5 we find that with increasing noise in the IPD game, cooperation becomes less likely as generosity invites more exploitation. There is an increasing tendency for runs to evolve toward mutual defection as p increased. Observation on individual runs showed a consistent trend toward defection.

Figure 6 shows the average payoff of agents in noisy environment. The distribution of plots indicates that the choosing diversity is more complicated as the noise probability increases.

5. Conclusion

In this paper, a new interpretation for the multiple choices IPD game based on PSO approach was proposed. Two main issues were presented.

The first used PSO to co-evolve the cooperation as evaluators, chosen one or the other strategy according to the fitness from the previous iteration in the IPD game. Two performances were addressed, the evolution of strategies and the stability of the evolution. The second exploited the influence of noise to the PSO algorithm based IPD game.

To simulate the real-world, the impact of noise is tested. Increasingly the levels of noise discourage cooperation because of increasing misunderstandings in the interactions. With high levels of noise introduced in the game, strategies were evolved to more defection.

Future work will include comparisons against existing game theory and IPD approaches, an investigation of the effects of using different payoff matrices to influence cooperative behaviour, competing against neighbours rather than competing against the complete the competition pool. As the theoretical study of the PSO techniques is extremely necessary and important, more detailed studies of the combination of evolved strategies and better PSO algorithms in the co-evolutionary training technique should be carried out.

6. References

- [1] S.Y. Chong, X. Yao. Behavioral Diversity, Choices and Noise in the Iterated Prisoner's Dilemma. *IEEE Transactions on evolutionary computation*, vol.9, no.6, 2005, pp.540-551.
- [2] S.Y. Chong, X. Yao. Multiple Choices and Reputation in Multiagent Transactions. *IEEE Transactions on evolutionary computation*, vol.11, no.6, 2007, pp.689-711.
- [3] S.Y.Chong, P.Tiño and X.Yao. Measuring Generalization Performance in Coevolutionary Learning. *IEEE Transactions on evolutionary computation*, vol. 12, no. 4, Aug 2008, pp.479-505.
- [4] S.Y.Chong, P.Tiño and X.Yao. Relationship Between Generalization and Diversity in Coevolutionary Learning. *IEEE Transactions on computational intelligence and AI in games*, vol.1, no.3, Sep 2009, pp.214-232.
- [5] S.Y.Chong, P.Tiño, D. C. Ku, and X. Yao, Improving Generalization Performance in Co-Evolutionary Learning. *IEEE Transactions on evolutionary computation*, vol.16, no.1, Feb 2012, pp.70-85.
- [6] H. Ishibuchi, K. Takahashi, K. Hoshino, J. Maeda and Y. Nojima, Effects of configuration of agents with different strategy representations on the evolution of cooperative behaviour in a spatial IPD game, *IEEE Conference on Computational Intelligence and Games*, 2011.
- [7] R. Axelrod, *The evolution of cooperation*. Basic Books, New York, 1984.
- [8] R. Axelrod and L. Davis, *The evolution of strategies in the iterated prisoner's dilemma*, *Artificial Life II*, Ed. SAn Mateo, CA: Morgan Kaufmann, 1987, pp. 32-41.
- [9] B.F. David, On the relationship between the duration of an encounter and the evolution of cooperation in the iterated prisoner's dilemma, *Evolution of computation*, vol. 3, no. 3, 1996, pp. 349–363.
- [10] P.J. Darwen and X. Yao, Coevolution in iterated prisoner's dilemma with intermediate levels of cooperative: Application to missile defense, *International Journal of Computational Intelligence and Applications*, col.2, no. 1, 2002, pp. 83-107.
- [11] P.J. Darwen and X. Yao, Why more choices cause less cooperation in iterated prisoner's dilemma, *Proceedings of the Congress on Evolutionary Computation*, Seoul, Korea, 2001, IEEE Press, pp. 987-994.
- [12] P.J. Darwen and X. Yao, Co-evolution in iterated prisoner's dilemma with intermediate levels of cooperative: Application to missile defense, *International Journal of Computational Intelligence and*

- Applications, vol.2, no. 1, 2002, pp. 83-107.
- [13] H. Ishibuchi and N. Namikawa, Evolution of iterated prisoner's dilemma game strategies in structured demes under random pairing in game playing, *IEEE Transactions on evolutionary computation*, vol.9, no.6, 2005, pp.552-561.
 - [14] H. Ishibuchi, H.Ohyanagi and Y.Nojima, Evolution of Strategies With Different Representation Schemes in a Spatial Iterated Prisoner's Dilemma Game, *IEEE Transactions on Computational Intelligence and AI in Games*, vol.3, no.1, 2011,pp.67-82.
 - [15] H. Ishibuchi, H.Ohyanagi and 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*, 2009, pp. 1568 -1573.
 - [16] K. Brauchili, T. Killingback and M. Doebeli, Evolution of cooperation in a spatially structured populations, *Journal of Theoretical Biology*, vol.200, no.4, 1999, pp.405-417.
 - [17] M.Bonaventura,A. Kaye; B.Rachel,G. Rebecca, H. Kate and W. Natasha, Human friendship favours cooperation in the Iterated Prisoner's Dilemma, *Behaviour*, vol.143, no.11, 2006, pp. 1383-1395.
 - [18] H.Ishibuchi and N.Namikawa, Evolution of iterated prisoner' dilemma game strategies in structured demes under random pairing in game playing, *IEEE transactions on evolutionary computation*, Vol.9,no.6, 2005, pp.552-561.
 - [19] L.J.Eshelman and J.D.Schaffer, Real-coded genetic algorithms and interval-schemata, in *Foundations of genetic algorithms 2*. San Mateo, CA: Mogan Kaufman, 1993, pp.187-202.
 - [20] R.Chiong and M.Kirley, Iterated n-player games on small-world networks, *Proceedings of the 13th annual conference on Genetic and evolutionary computation*, NY, USA, 2011.
 - [21] K.Moriyama, Utility based Q-learning to facilitate cooperation in Prisoner's Dilemma games, *Web Intelligence and Agent Systems: An International Journal*, IOS Press, vol. 7, 2009, pp. 233-242.
 - [22] B. Chen, B. Zhang and W.D.Zhu, Combined trust model based on evidence theory in iterated prisoner's dilemma game, *International Journal of Systems Science*, vol. 42, no. 1, Jan. 2011, pp.63-80.
 - [23] N.Franken, A.P.Engelbrecht, Comparing PSO structures to learn the game of checkers from zero knowledge, *The 2003 Congress on evolutionary computation*, 2003, pp.234-241.
 - [24] N.Franken, A.P.Engelbrecht, Particle swarm optimization approaches to coevolve strategies for the iterated prisoner's dilemma, *IEEE Transactions on evolutionary computation*, vol.9, no.6, Dec.2005, pp.562-579.
 - [25] C.D.Chio, P.D.Chio and M.Giacobini, An evolutionary game-theoretical approach to particle swarm optimization, *Proceedings of the 2008 conference on Applications of evolutionary computing*, pp. 575-584.
 - [26] J.Kennedy and R.C.Eberhart, Particle swarm optimization, in *Proc. IEEE International conference of neural network*, vol. 4, 1995, pp.1942-1948.
 - [27] Fogel D.B, The evolution of intelligent decision making in gaming, *Cybernetics and Systems: An International Journal*,vol.22,1991, pp.223-236.
 - [28] S.Suzuki,E.Akiyama, Evolution of indirect reciprocity in groups of various sizes and comparison with direct reciprocity, *Journal of Theoretical Biology*, vol.245(3), 2007,pp.539-552.
 - [29] S.Suzuki,E.Skiyama,Three-person game facilitates indirect reciprocity under image scoring, *Journal of Theoretical Biology*, vol.249, no.1, Nov. 2007, pp. 93-100.
 - [30] X.Yao and P.Darwen, An evolutionary approach to the N-player iterated prisoner's dilemma, *Evolutionary Computation: Theory and Applications*, World Scientific, 1999, pp.331-351.
 - [31] R.Chiong and M.Kirley, Effects of Iterated Interactions in Multi-player Spatial Evolutionary Games, *IEEE Transactions on evolutionary computation*, DOI:10.1109/TEVC.2011.2167682.