A Mechanism of Sharing Aspiration to Promote Cooperative Behavior in a Group

Manahan Siallagan*1, Hiroshi Deguchi[2], Manabu Ichikawa[3]

[1, 2, 3]Computational Intelligent and Systems Science, Tokyo Institute of Technology, Japan

*1manahan09@cs.dis.titech.ac.jp; [2]deguchi@dis.titech.ac.jp; [3]ichikawa@dis.titech.ac.jp

Abstract- **Many experimental studies have been built to investigate social dilemmas. There are some clues in these studies that can be used to create mechanisms to overcome social dilemmas. This research deals with a simulation model that uses a mechanism of sharing aspiration based on some of such clues, i.e., individuals' expectations, information seeking and communication, obtained from previous experimental studies on social dilemmas. A mechanism of sharing aspiration is combined with a learning process to promote cooperative behavior in a group. Simulation results show that the mechanism of sharing aspiration shaped by learning can promote cooperative behavior in a group.**

Keywords- Aspiration Levels; Learning; Public Goods Game; Social Dilemmas; Sharing Aspiration

I. INTRODUCTION

The tension between individual interests and collective interests is a primary concern in social dilemmas. Individuals in social dilemmas are better off by contributing nothing or very little of one's own resources to the public goods. However, if all group members act in this way, public goods will not be provided. Also, self-interested individuals can take advantage of cooperative individuals (the individuals who contribute to the provision of public goods) [1].

Laboratory experimentation has been conducted to investigate this issue. On one hand, there are some experimental studies related to the role of a leader in a group. The results suggest that the effectiveness of the leader's role to improve the cooperative behavior in a group not only depends on the leader's characteristics but also depends on individuals' expectations [2, 3]. On the other hand, information seeking and communication are also related to solving social dilemmas. Accessibility of information refers to the individuals' strategies, that is, people have clear preferences of particular pieces of information and that information preferences vary systematically across individuals as a function of their contribution strategies [4]. When the information about aggregated contribution to the public is provided, the cooperative behavior of individuals would vary from strong co-operators to strong free riders [5].

The interaction of individual opinions in group discussion can improve the cooperative behavior [6]. This interaction is also known as communication effect on cooperation [7]. As long as communication persists, cooperation rates are high and stable. Conversely, without communication, cooperation rates gradually decline. The communication is achieved in subsequent periods, which allowed the individuals to share information and reach a better understanding of their tasks [8].

As we can see in the above findings of experimental studies on social dilemmas, individuals' expectations, information, and communication have an important role in order to spur cooperative behavior in social dilemmas. However, this role also depends on the way that individuals identify, discuss, and commit to make cooperation possible [8]. The identification is related to the way that individuals perceive others in terms of similarities and closeness. The more similar and closer the individuals are, the easier the communication happens. The discussion is about the cues on the character and motives of individuals that are exchanged during the communication process. If the cues of the character and motives are about "the dilemmas", the cooperative behavior could be improved. Conversely, the cooperative behavior could be decreased if the cues of the character and motives are not about "the dilemmas". The commitment is a promise to keep cooperative behavior during the interaction. All these factors would determine the level of cooperation in the social dilemmas.

In line with laboratory experimentation to investigate the social dilemmas, game theory has formalized the issue as cooperation problems. The problems are represented as a mixed-motive two-person game, i.e., prisoner's dilemma game or n-person game, i.e., public goods game. In recent studies, the analysis of the games has been shifted from high game theory to low game theory. In high game theory the players are modeled as hyper rational. In low game theory the players are modeled as simple adaptive learners [9]. The interpretation of the player as a learner has resulted in a number of learning models such as Bush-Mosteller Stochastic learning model [10]. This model is also known as reinforcement learning model, which is designed to capture the "Law of Effect" [11]. Positive reinforcement increases the tendency to play an action, while negative reinforcement decreases it. Positive or negative reinforcement is judged by a cognitive factor to stimulate their action. The standard cognitive factor is aspiration level. The difference between payoff and aspiration level will generate a stimulus. This aspiration level is not static but evolves slowly as a player gains experience.

The variant of this model, i.e., payoff matching model, has successfully described human behavior in experimental studies of social dilemmas [9, 12, 13]. This model predicts that players will learn to cooperate depending on the payoff structure. In the theoretical analysis and simulation approach, a large number of researches have been examined to solve the prisoner's dilemma game. The results showed that cooperative behavior could emerge and survive in the long run. However, the emergence of cooperative behavior depends on certain payoff conditions [14], sufficiently slow speed of updating the aspiration level [15], or a combination of these two factors [16].

Based on the above information, we claim that there are three crucial factors to overcome social dilemmas, i.e., individuals' expectations, information and communication. We use these three factors to develop a mechanism of sharing information, i.e., sharing the level of aspiration. In a group, people may have different aspirations toward their relationship. We can consider this aspiration as a goal or an expectation which is what the people are willing to achieve. We assume the members of the group can interact (communicate) with each other to share their aspiration level (goal or expectation). The information that one person would use depends on the closeness of this person to another person. We adopt a social comparison theory [17] to build a mechanism for sharing information. The information that has been received by one person is used to update his/her aspiration level by comparing with his/her current aspiration level. This concept reflects a process that involves identification, discussion and commitment within the interaction.

We combine the sharing aspiration process with the learning process. The discrepancy between the current payoff and the aspiration level will generate a stimulus, which would be used to update the probabilities of the actions. We use three models of learning, i.e., Roth-Erev, Borgers-Sarin, and Satisfying, which are based on stimulus-response mechanism shaped by a learning force. This situation reflects the fact that people may learn with different models of learning. Within this framework, we want to promote the cooperative behavior in a group. The rest of the paper is organized as follows: in Section 2, we describe the public goods game that represents the social dilemmas; in Section 3, we discuss three learning models that use a stimulus-response mechanism shaped by learning forces; in Section 4, we build the sharing aspiration mechanism; in Section 5, the results of simulation of several scenarios are shown and Section 6 concludes with a summary and discussion of the results.

II. PUBLIC GOODS GAME

A conflict situation in a group can be modeled as a Prisoner's Dilemma (PD) game. In contrast to the more familiar PD game in which the strategy space of each player is binary (cooperate/defect), the strategy space in in-group conflict that we consider here is discrete. Consider a group consisting of N players. Assume that each player has the same endowment denoted by e. At each time t, every player is faced with a decision of allocating a units of his/her endowment. We assume the strategy space is any discrete number that does not exceed the endowment. This assumption can be made more general if the strategy space is continuous, i.e. any fraction does not exceed the endowment [18].

Let a_i be the amount contributed by player i, where $a_i \in A = \{0,1,2 \ldots e\}$, and let $X = \sum_{i=1}^{N} a_i$be the total contribution of N players. The payoff of player i at time t is given by,

$$R_i(t) = (e - a_i) + g\frac{X}{Ne} \tag{1}$$

where g is public good that all of N players generate if each player contributes his/her entire endowment e. The second term in Eq. (1) is equal to zero if each player contributes nothing, equal to g if each contributes his/her entire endowment e, and equal to some intermediate value between 0 and g if $a_i \in A\backslash\{0, e\}$. The game has the PD property if $0 < g < Ne$, the equilibrium solutions for player i is to contribute nothing, i.e., $a_i = 0$ [19].

III. ASPIRATION-BASED LEARNING MODELS

In this section we describe three models of learning based on aspiration. In these models, players' behaviors are based on two properties. First, they have a stimulus, i.e., the discrepancy between payoff and aspiration level which divides outcome into two subsets, i.e., satisfactory and unsatisfactory. Second, players learn via trial-and-error, and become more inclined to try actions that satisfy their aspiration level and less likely to try those that do not. Aspiration level is endogenous, i.e., they adjust to a player's experience (payoff).

Figure 1 shows a scheme of the learning models [16]. A player takes an action based on the probability of the action. Aspiration and the payoff will generate a stimulus of the action. Positive outcomes increase the probability that the associated action will be repeated, while negative outcomes reduce it. In this sense, the stimulus of a player will encourage him/her to find a "good" action (in this game is to contribute to public goods). In real life situations, an individual will receive a stimulus from a discrepancy between achievement (outcome) of his/her works and what he/she expects from the work (aspiration). However, the real life situations involve a random exposure to environmental variables and humans are sometimes inertial, i.e., they do not invariably adapt or learn as well. To accommodate these circumstances, we can introduce randomness to extend the models. Thus with probability ε_1, a player does not adjust his/her probability in a current period. With probability ε_2, a player may not adjust his/her aspiration level and after having decided which action to undertake, he/she may select the wrong action with probability ε_3. However, the current research does not consider this extended model and it will be considered in future research.

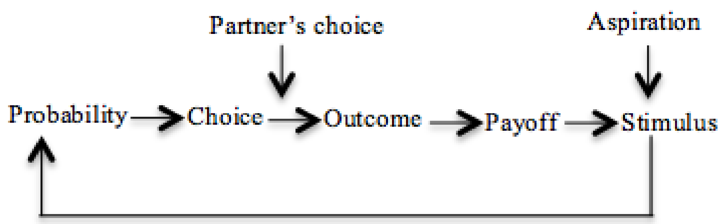

Fig. 1 Scheme of Aspiration-Based Learning Model

Let $\rho_i(t)$ be the aspiration level for player i at time t, let $pr_i(a,t)$ be the probability of an action $a_i \in A$ for player i at time t, and $R_i(t)$ be a payoff for player i at time t. A stimulus associated with payoff $R_i(t)$ and aspiration level $\rho_i(t)$ for taken an action $a \in A$ for player i is $S_i(t) = R_i(t) - \rho_i(t)$. We described the three learning models and differences among them as follows.

A. *Roth-Erev Learning Model*

This model uses the propensity to evaluate the probabilities of an action [9, 12, 13]. The actions that have been played and have succeeded tend to be played with greater frequency than those that have been less successful over time. Let $q_i(a,t)$ be the propensity of an action $a \in A$ at time t. The propensity to take an action $a \in A$ for player i is updated by setting:

If action a was chosen at t,

$$q_i(a,t+1) = \max\{\nu, (1-\phi)q_i(a,t) + (1-\epsilon)S_i(t)\} \tag{2}$$

Otherwise,

$$q_i(a,t+1) = \max\{\nu, (1-\phi)q_i(a,t) + \epsilon S_i(t)\} \tag{3}$$

Parameter $\nu > 0$ is a technical parameter to ensure that propensities remain positive. Parameter ϕ slowly reduces the importance of past experience and parameter ϵ prevents the probability of choosing any action from going to zero. The probability of choosing action $a \in A$ at time t for player i is proportional to past average propensities,

$$pr_i(a,t) = \frac{q_i(a,t)}{\sum_{a \in A} q_i(a,t)} \ \forall\, a \in A \tag{4}$$

The aspiration level is linearly adjusted in the direction of outcome experienced,

If $R_i(t) < \rho_i(t)$

$$\rho_i(t+1) = (1-\omega^-)\rho_i(t) + \omega^- R_i(t) \tag{5}$$

If $R_i(t) \geq \rho_i(t)$

$$\rho_i(t+1) = (1-\omega^+)\rho_i(t) + \omega^+ R_i(t) \tag{6}$$

The parameters ω^- and ω^+ control the adjustment of the aspiration level following negative and positive rewards.

B. *Borgers-Sarin Learning Model*

This model directly uses stimulus to update the probability of an action [20]. The actions that are successful today are more likely to be used tomorrow. Unsuccessful actions are less likely. The stimulus, i.e., the discrepancy between payoff and aspiration level, is used as a threshold that divides all possible current payoffs into successful and unsuccessful ones, hence indicating which actions are coded as success and which as failures. A player's aspiration level itself adjusts to experience, and reflects prior payoffs.

The probability of choosing action $a \in A$ at time t for player i is given as follows,

If $S_i(t) \geq 0$ and a was chosen at t,

$$pr_i(a,t+1) = \big(1-\lambda S_i(t)\big)pr_i(a,t) + \lambda S_i(t) \tag{7}$$

Otherwise,

$$pr_i(a,t+1) = \big(1-\lambda S_i(t)\big)pr_i(a,t) \tag{8}$$

If $S_i(t) < 0$ and a was chosen at t,

$$pr_i(a,t+1) = \big(1+\lambda S_i(t)\big)pr_i(a,t) \tag{9}$$

Otherwise,

$$pr_i(a, t+1) = \big(1 + \lambda S_i(t)\big)pr_i(a, t) - \lambda S_i(t) \tag{10}$$

The parameter λ controls the effect of rewards in $pr_i(a, t+1)$. We normalize the value of $S_i(t)$ by dividing it with $Z = \max[R_i(t) - \rho_i(t), \rho_i(t)]$. This value guarantees that the absolute value of $\lambda S_i(t)$ always lies between zero and one. The aspiration level is also linearly adjusted in the direction of outcome experienced,

$$\rho_i(t+1) = (1-\beta)\rho_i(t) + \beta R_i(t) \tag{11}$$

where $0 \leq \beta \leq 1$.

C. Satisfying Model

This model is inspired by the theory of satisfying behavior stating that players tend to get satisfied rather than optimize [21]. Instead of choosing a strategy giving the highest payoff, a player sets a reference point representing the payoff level he/she wants. This reference point is called aspiration level. The player searches a strategy giving his/her payoff higher than the aspiration level. Once he/she finds such a strategy, he/she keeps the strategy at the next step. Otherwise, he/she drops the strategy and chooses an alternative randomly at the next step.

Each player i will make decision based on following criteria [22]:

If $S_i(t) \geq 0$,

$$a_i(t+1) = a_i(t)$$

Otherwise,

$$select\ an\ action\ a \in A\ randomly$$

The aspiration level is also linearly adjusted in the direction of outcome experienced via learning rate $0 \leq \delta \leq 1$,

$$\rho_i(t+1) = (1-\delta)\rho_i(t) + \delta R_i(t) \tag{12}$$

D. Differences of the Three Learning Models

All three learning models use a stimulus, i.e., the discrepancy between payoff and aspiration levels, to determine an action to be chosen. Moreover, the three models also use the same formulation to update the aspiration level, i.e., linearly adjusted in the direction of outcome experienced via learning rate. The main difference is the way that a stimulus is used to determine the action that will be chosen. The Roth-Erev model uses the stimulus to calculate the propensity of the action and uses this propensity to update the probability. In the Borgers-Sarin model, the stimulus is used directly to update the probability of an action and also as the condition to update the probability. In the satisfying model, the stimulus is only used as a condition to change the action. This model is similar to satisfying theories that state that decision makers search for new alternatives if and only if today’s action is unsatisfactory, i.e., yields a payoff that falls below the decision maker’s aspiration level [23]. On the other hand, Roth-Erev and Borgers-Sarin models are reinforcement learning models that use a probability of an action. If such action produces a satisfactory payoff in the current period, then the player will not decrease his/her probability of that action. Therefore, the state space of reinforcement learning models is more complex in terms of the probability over all available actions.

IV. MECHANISM OF SHARING ASPIRATION

The theory of social comparison states that a person tends to make a self-evaluation based on comparison with other persons. In this situation, the information of others would determine the behavior of the person in the future. Therefore, the competitive environment may be occurring in this process and there is a pressure toward uniformity [17].

We assume each player updates his/her aspiration level by sharing the information of other players’ aspiration level and then compares this information with his/her current aspiration level. The sharing process is based on the interaction scheme that is given in the beginning of the simulation. Within the interaction, a player will obtain the information about the other players' aspiration level depending on the closeness of the player in the given interaction scheme. The closeness is represented by weights.

Let $\alpha_i(t)$ be a level of information that will be received by a player i at time t, and let T_i be a set of players that interact with player i. Let n_i be a number of players who interact with player i, and let n_k be a number of players who interact with player k. We calculate $\alpha_i(t)$ as follows:

$$\alpha_i(t) = w_i\rho_i(t) + \textstyle\sum_{k\in T} w_k\rho_k(t) \tag{13}$$

where $w_k = \frac{1}{1+\max\{n_i,n_k\}}\ \forall\ k \in T_i$, and $w_i = 1 - \sum_{k\in T_i} w_k$, w_k is a set of the weights that represents the closeness of a player i in the given scheme. Eq. (13) looks similar with distributed algorithm for distributed averaging problem [24], but differs in term of what information will be communicated. We explain the model as follows:

Suppose the scheme of interaction is given in the beginning of simulation as shown in Figure 2.

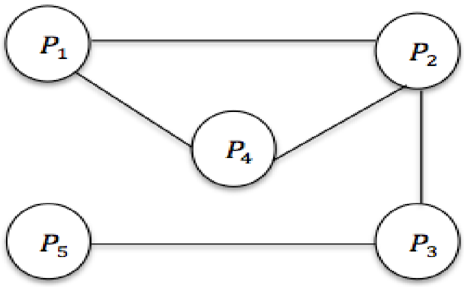

Fig. 2 Scheme of interaction

We can describe this interaction in form of matrix as shown in Table 1. The value of an entity $P_{ij} = 1$ if there is a connection between player i and player j, otherwise $P_{ij} = 0$. From Table 1 we can count $n_i \; \forall \; i \; \in \; \{1,2,3,4,5\}$, i.e., $n_1 = 2$, $n_2 = 3$, $n_3 = 2$, $n_4 = 2$, and $n_5 = 1$. Level of aspiration information for player P_1 can be calculated as follows: P_1 interacts with P_2 and P_4, $w_2 = \frac{1}{1+\max\{2,3\}} = \frac{1}{4}$ and $w_4 = \frac{1}{1+\max\{2,2\}} = \frac{1}{3}$, so we get $w_1 = 1 - (w_2 + w_4) = 1 - \left(\frac{1}{4} + \frac{1}{3}\right) = \frac{5}{12}$. After that we can calculate,

$$\alpha_1(t) = w_1\rho_1(t) + w_2\rho_2(t) + w_4\rho_4(t) = \frac{5}{12}\rho_1(t) + \frac{1}{4}\rho_2(t) + \frac{1}{3}\rho_4(t)$$

TABLE 1 MATRIX REPRESENTATION OF INTERACTION

	P_1	P_2	P_3	P_4	P_5
P_1	0	1	0	1	0
P_2	1	0	1	1	0
P_3	0	1	0	0	1
P_4	1	1	0	0	0
P_5	0	0	1	0	0

As we can see player P_1 will get the aspiration information from P_2 and P_4 with different weights.

We update aspiration level of a player i based on the three learning models above as follows:

Roth-Erev:

If $\alpha_i(t) \geq \rho_i(t)$,

$$\rho_i(t+1) = (1-\omega^+)\alpha_i(t) + \omega^+ R_i(t) \tag{14}$$

If $\alpha_i(t) < \rho_i(t)$,

$$\rho_i(t+1) = (1-\omega^-)\rho_i(t) + \omega^- R_i(t) \tag{15}$$

Borgers-Sarin and Satisfying:

If $\alpha_i(t) \geq \rho_i(t)$,

$$\rho_i(t+1) = (1-h)\alpha_i(t) + hR_i(t) \tag{16}$$

If $\alpha_i(t) < \rho_i(t)$,

$$\rho_i(t+1) = (1-h)\rho_i(t) + hR_i(t) \tag{17}$$

Borgers-Sarin model and Satisfying model have the same formulation to update the aspiration level. We replace parameter β and δ with h that represents the learning rate. In these formulations, a player adjusts his/her level of aspiration by comparing his/her level of aspiration to the level of information that he/she gets from interaction, and uses it if the level is higher or equal to his/her aspiration level. It means that a player raises his/her aspiration level to the group's aspiration level.

V. SIMULATION AND RESULTS

In this simulation a set of players play the public goods game and repeat the game for a number of times. We define a set of players as a group. We do not consider a group as a dynamic group that involves the process of group formation, group membership, and group cohesion. We assume each player in a group follows a learning model that is embedded to him/her without the capacity to interpret what one learns. In this sense, the players are only aware of what actions are successful (give a satisfaction) through learning.

Under this circumstance, we consider two scenarios. The main purpose of the first scenario is to compare a group in which players interact with a group in which players do not interact. We define the interact group as a group in which the players share information about their aspiration level through the mechanism of sharing aspiration. Through this mechanism we expect that the players share information about their aspiration level, which in turn influences their aspiration toward cooperative behavior. Therefore, we can assert that the mechanism of sharing aspiration spurs cooperative behavior in interact group under such circumstance. Conversely, non- interact group is a group in which the players do not share information about their aspiration level. The scheme of interaction for this setting is given in Figure 2.

The main purpose of the second scenario is to investigate the effectiveness of sharing aspiration mechanism in relations with the number of players in a group and the strength of connectivity. We vary the number of players in a group, i.e., 6, 8, and 10 and two strengths of connectivity, i.e., strong connectivity and weak connectivity. The scheme of interaction for strong connectivity and the number of players can be seen in Figure 3. Figure 4 shows the weak connectivity and the number of players.

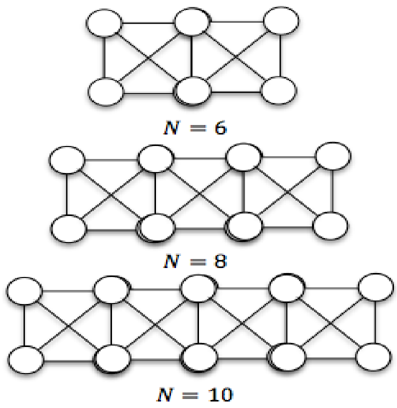

Fig. 3 Scheme of interaction for strong connectivity

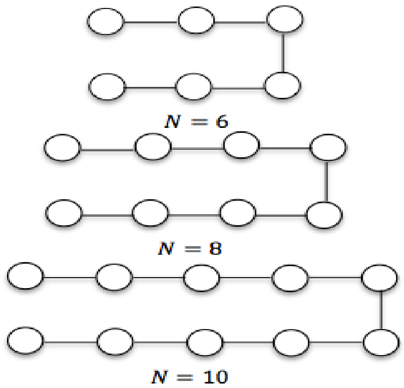

Fig. 4 Scheme of interaction for weak connectivity

Each player in a group has a learning model, i.e., REM stands for Roth-Erev Model, BSM stands for Borgers-Sarin Model, and SM stands for Satisfying Model. Table 2 shows the parameters of the learning models.

TABLE 2 PARAMETERS OF THE LEARNING MODELS

REM	$\nu = 10^{-9}$ $\phi = 10^{-5}$ $\epsilon = 10^{-5}$ $\omega^{-} = 0.02$ $\omega^{+} = 0.01$
BSM	$\lambda = 0.5$ $\beta = h = 0.01$
SM	$\delta = h = 0.01$

At each time t, each player takes an action depending on the probability of the action. They will receive a reward from this game. After that, they update the probabilities of the actions based on their learning model. Finally, they interact to share the aspiration level (only for interact group) and update their aspiration level (for interact and non-interact group).

A. *First Scenario*

In this simulation we compare interact group with non-interact group. Both groups consist of $N = 5$ players. At each time t, every player plays the public goods game. We use endowment $e = 2$, so that each player can take an action from $A = \{0,1,2\}$ and the public good $g = 9.9$. This condition satisfies $0 < g < Ne = 10$. We set initial value of $q_i(a, t = 0) = 1$ so that $pr_i(a, t = 0) = \frac{1}{|A|} = \frac{1}{3} \forall a \in A$ for each player using learning model REM. For each player that uses learning model BSM, we also use the initial value of probability $pr_i(a, t = 0) = \frac{1}{|A|} = \frac{1}{3} \forall a \in A$, and for SM players we set the initial action randomly.

No information will be shared in non-interact group and the aspiration level will be updated by using Eq. (5) and Eq. (6) for REM, Eq. (11) for BSM, and Eq. (12) for SM. The information will be shared in interact group by using the mechanism of sharing aspiration. The aspiration level will be updated by using Eq. (14) and Eq. (15) for REM and Eq. (16) and Eq. (17) for BSM and SM. We use the scheme of interaction as in Figure 2 in this scenario. The initial value of aspiration level and the learning model of each member in each group can be seen in Table 3.

TABLE 3 PARAMETERS OF THE FIRST SCENARIO

Number of players (N)	**Public good (g)**	**Initial aspiration level ($\rho_i(t = 0)$)**	**Learning model**
5 Interact Group	$g = 9.9$	5.5; 6.5; 7.5; 8.5; 9.5	$REM; BSM; SM;$ $REM; BSM$
5 Non-Interact Group	$g = 9.9$	5.5; 6.5; 7.5; 8.5; 9.5	$REM; BSM; SM;$ $REM; BSM$

We set the same initial aspiration level and the same learning models for each member in each group (for comparison purpose). The parameters of learning models are shown in Table 2. We run the simulation for 50 trials in 10000 iterations.

Figure 5 shows the average reward for each group for 50 trials and 10000 iterations. The average value of reward in interact group can reach the public goods value, i.e., g =9.9. It means that all players in the group contribute all of their endowment, i.e., $e = 2$. While in non-interact group, the average value of reward is only about 8, lower than the public goods value.

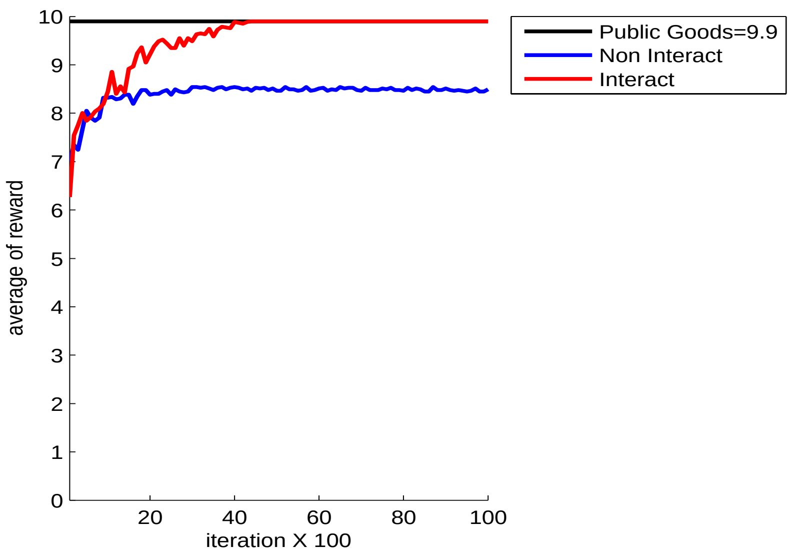

Fig. 5 Comparison of average reward between interact group and non-interact group

Figure 6 and Figure 7 also confirm this result. Figure 6 shows the average value of aspiration level for each member in interact group, while Figure 7 shows the average value of aspiration level for each member in non-interact group. As we can see in Figure 6, the average of aspiration level of all members almost has the same value from the beginning (see the small figure inside Figure 6) and converges to the public goods value, i.e., $g = 9.9$. This condition affects the behavior of each player in interact group to the cooperative behavior. All members in interact group contribute all of their endowment, i.e., $e = 2$ so that they get the reward of 9.9 which is equal to the public goods and his/her aspiration level as well.

On the other hand, the average of aspiration level in non-interact group is not converging to the same value (Figure 7). Each player still has his/her own aspiration level so that they still have their own value to be contributed to the public goods. No interaction and communication in this group, so that no information can be shared.

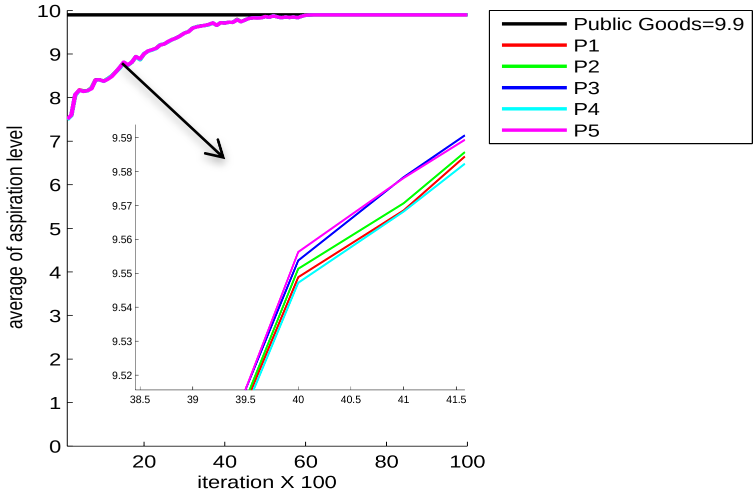

Fig. 6 The average of aspiration level for interact group

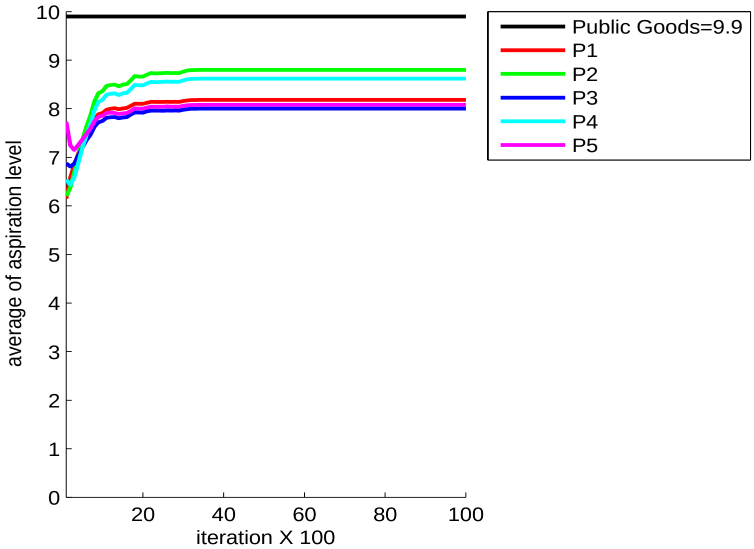

Fig. 7 The average of aspiration level for non-interact group

B. Second Scenario

We use the scheme of interaction in Figure 3 for strong connectivity and in Figure 4 for weak connectivity. The initial value of aspiration level and the learning model of each member in each group can be seen in Table 4.

TABLE 4 PARAMETERS OF THE SECOND SCENARIO

Number of players (N)	**Public good (g)**	**Initial aspiration level ($\rho_{i(t=0)}$)**	**Learning model**
6	$g = 11.9$	0.5; 1.5; 2.5; 3.5; 4.5; 5.5	$REM; BSM; SM; REM; BSM; SM$
8	$g = 15.9$	0.5; 1.5; 2.5; 3.5; 4.5; 5.5; 6.5; 7.5	$REM; BSM; SM; REM; BSM; SM; REM; BSM$
10	$g = 19.9$	0.5; 1.5; 2.5; 3.5; 4.5; 5.5; 6.5; 7.5; 8.5; 9.5	$REM; BSM; SM; REM; BSM; SM; REM; BSM; SM; REM$

We use endowment $e = 2$, so that each player can take an action from $A = \{0,1,2\}$. We set initial value of $q_i(a, t = 0) = 1$ so that $pr_i(a, t = 0) = \frac{1}{|A|} = \frac{1}{3} \forall a \in A$ for each player using learning model REM. For each player that uses learning model BSM we also use the initial value of probability $pr_i(a, t = 0) = \frac{1}{|A|} = \frac{1}{3} \forall a \in A$, and for SM players we set the initial action randomly. All parameters of the models are shown in Table 2. We run the simulation for 50 trials in 100000 iterations to see the long-term dynamics of the model. The output of this scenario is the normalized average reward. Values of the normalized average reward are in the range 0 and 1. The values close to 1 indicate full cooperative behavior. This normalization is needed because we have different values of g.

As we can see in Figure 8, the strong connectivity is more quickly to converge to full cooperative behavior than the weak connectivity. As long as there is a connectivity that involves all players, the players will get the information that coordinates their strategy. As we can see in the weak connectivity with $N = 6$ (Figure 4), the first player can get the information from the second player, the second player can get the information from the third player, and so on. Because the connection is in two directions, there is a chain of information in the long run so that the players can coordinate their strategy. However, as the number of players increases, i.e., $N = 10$, full cooperative behavior is hard to achieve.

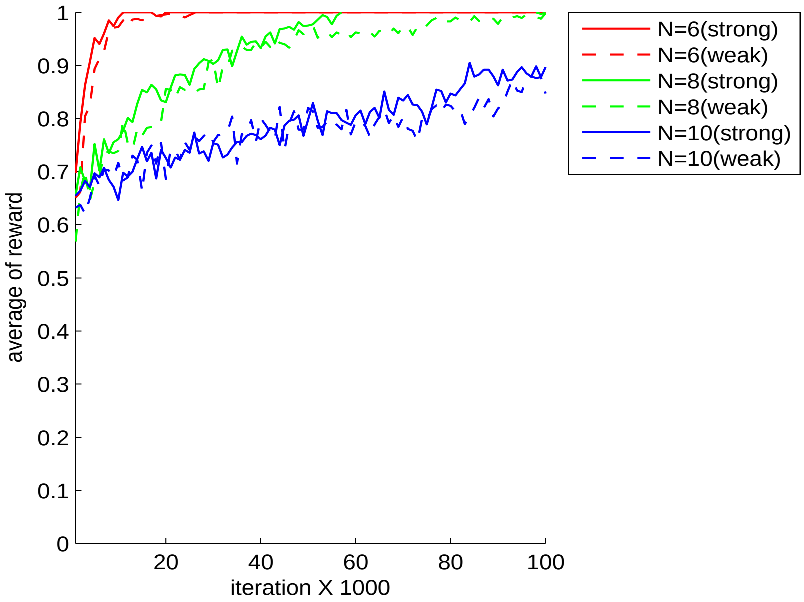

Fig. 8 The average of reward for strong and weak connectivity with various numbers of players

VI. CONCLUSION

Based on the simulation, we found two interesting results. The first one is that with interaction and sharing aspiration all players can improve their cooperative behaviors in a group. Players with different initial aspirations and learning models can adjust their aspirations through interaction and sharing aspiration process. In this situation, a player compares his/her aspiration with the group's aspiration level via sharing aspiration mechanism as long as his/her aspiration is below the group's aspiration level that he/she receives. Within this process, a player can share his/her aspiration with whom he/she interacts. By increasing aspiration of one player, then the aspirations of other players are also increased. This process makes the aspiration level of all members of the group converge to the cooperative results, i.e., all players contribute all of their endowment to the public goods. This situation is also confirmed by social comparison theory, i.e., pressure towards uniformity [17].

The second result is about the complexity of sharing aspiration in a group. A group with few numbers of players can improve their cooperative behavior more quickly in strong connectivity. A player in a group that consists of many players and in the weak connectivity needs more time to coordinate their action towards a goal. As long as there is communication (interaction/connectivity) in the group, cooperative behavior can be maintained. This result is in line with the previous experimental studies [7, 8]. However, as the number of players increases, the level of cooperative behavior decreases.

The proposed mechanism of sharing aspiration uses some clues, i.e., individuals' expectations, information, and communication obtained from experimental studies to investigate people's behavior in social dilemmas. These three clues are important to spur cooperative behavior in a group [2-8]. However, in this research we also combine this mechanism with aspiration-based reinforcement-learning models that can make players choose his/her action based on experience, i.e., aspiration level. As long as the members of the group are allowed to share their aspiration level, the cooperative behavior can be promoted.

One of the main assumptions of our model is that all players are willing to share their aspiration through interaction. This assumption is difficult to be fulfilled in reality. People might not want to share information about their aspiration. Our future research will investigate how to design a mechanism that can make people willing to share their aspiration.

The proposed model gives us insight into the aspiration level, which determines the behavior of the players in a group playing public goods game. However, in future research we would like to confirm our findings with experimental tests.

REFERENCES

[1] Robyn M. Dawes and David M. Messick, “Social Dilemmas”. International Journal of Psychology, vol. 35, iss. 2, pp. 111-116, 2000.

[2] David De Cremer and Mark Van Vugt, “Intergroup and Intragroup Aspects of leadership in Social Dilemmas: A Relational Model of Cooperation”. Journal of Experimental Social Psychology, vol. 38, pp. 126-136, 2002.

[3] Mark Van Vugt, Sarah F. Jepson, Claire M. Hart, and David De Cremer, “Autocratic Leadership in Social Dilemmas: A Threat to Group Stability”. Journal of Experimental Social Psychology, vol. 40, pp. 1-13, 2004.

[4] Robert Kurzban and Peter Descioli, “Reciprocity in Groups: Information-seeking in a Public Goods Game”. European Journal of Social Psychology, vol. 38, pp. 138-158, 2008.

[5] Robert Kurzban and Daniel Houser, “Individual Differences in Cooperation in Circular Public Goods Game”. European Journal of Social Psychology, vol. 15, pp. S37-S38, 2001.

[6] Timothy Hopthrow and Lorne G. Hulbert, “The Effect of Group Decision Making on Cooperation in Social Dilemmas”. Group Processes & Intergroup Relations, vol. 8, iss. 1, pp. 89-100, 2005.

[7] Cristina Bicchieri and Azi Lev-On, “Computer-Mediated Communication and Cooperation in Social Dilemmas: An Experimental Analysis”. Politics, Philosophy & Economics, 97, pp. 170-176, 2007.

[8] Kerr N and C. Kaufmann-Gilliland, “Communication, Commitment, and Cooperation in Social Dilemmas”. Journal of Personality and Social Psychology, vol. 66, pp. 513-529, 1994.

[9] Alvin E. Roth and Ido Erev, “Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term”. Games and Economic Behavior, vol. 8, pp. 164-212, 1995.

[10] R. Bush and F. Mosteller, “Stochastic Model of Learning”. New York: John Wiley and Sons, 1955.

[11] Thorndike Edward L. “Animal Intelligence: Experimental Studies”. New York: Mac Millan, 1911.

[12] Erev I and Roth A.E, “Predicting How People Play Games: Reinforcement Learning and Experimental Games with Unique Mixed Strategy Equilibrium”. American Economic Review, vol. 88, pp. 848-881, 1998.

[13] Erev I and A. Rapoport, “Coordination, Magic, and Reinforcement Learning in a Market Entry Game”, Games and Economic Behavior, vol. 23, pp. 146-175, 1998.

[14] Palomino F and F. Vega-Redondo, “Convergence of Aspirations and (Partial) Cooperation in the Prisoner’s Dilemma”. International Journal of Game Theory, vol. 28, iss. 4, pp. 465-488, 1999.

[15] Jonathan Bendor, Dilip Mookherjee, and Debraj Ray, “Aspiration-Based Reinforcement Learning in Repeated Interaction Games: An Overview”. International Game Theory Review, vol. 3, pp. 159-174, 2001.

[16] Flache A and M.W. Macy, “The Power Law of Learning”, Journal of Conflict Resolution, vol. 46, iss. 5, pp. 629-653, 2002.

[17] Festinger L, “A Theory of Social Comparison Processes”. Human Relation, vol. 7, pp. 117-140, 1954a.

[18] Anna G and A. Rapoport, “Embedding Social Dilemmas in Intergroup Competition Reduces Free-riding”, Organizational Behavior and Human Decision Making, vol. 101, pp. 184-199, 2006.

[19] Rapoport A and Amaldoss W, “Social Dilemmas Embedded in Between-Group Competition: Effects of Contest and Distribution Rules”, in M. Foddy, M. Smithson, S. Schneider, and M. Hogg (Eds.), Resolving Social Dilemmas: Dynamic, structural, and intergroup aspects, pp. 66-99, 1999.

[20] Borgers T and R. Sarin, “Dynamic Consistency and Non-expected Utility Models of Choice Under Uncertainty”. Journal of Economic Theory, vol. 77, pp. 1-14, 1997.

[21] H. A. Simon, “A Behavioral Model of Rational Choice”. Quarterly Journal of Economics, vol. 69, pp. 99-118, 1955.

[22] Jeffrey R. Stimpson and Michael A. Goodrich, “Learning to Cooperate in a Social Dilemma: A Satisficing Approach to Bargaining”, In Proceedings of the 20th International Conference on Machine Learning, pp. 728-735, 2003.

[23] Bendor J, Diermeir D, and Ting M, “The Empirical Content of Behavioral Models of Adaption”. Mimeo, Graduate School of Business, Stanford University, 2002.

[24] Lin Xiao and Stephen Boyd, “Fast Linear Iterations for Distributed Averaging”, Systems and Control Letters, vol. 53, pp. 65-78, 2004.