JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 30, 313-332 (2014)

A Proposal for Intelligent Agents with Adjusting Self-Reputation Capability for Preventing Fraud in Multi-Agent Societies*

SHENG-WEN WANG[1], CHUNG-YUAN HUANG[2,*] AND CHUEN-TSAI SUN[1]
[1]*Department of Computer Science*
National Chiao Tung University
Hsinchu, 300 Taiwan
E-mail: sww0915@gmail.com; ctsun@cs.nctu.edu.tw
[2,*]*Department of Computer Science and Information Engineering*
Chang Gung University
Taoyuan, 333 Taiwan
E-mail: gscott@mail.cgu.edu.tw

Intelligent agents are increasingly being used for tasks such as e-commerce auto-bidding. Cooperation and defection challenges commonly found in multi-agent systems make the design and construction of a reliable reputation mechanism an important research and engineering goal. Here we describe our proposal for an intelligent agent model that addresses learning issues in terms of self-reputation. The agents in our proposed model are capable of evaluating their behaviors based on a mix of public and private interest considerations, and of testing various solutions aimed at mitigating self-discrepancy and meeting social standards. Our results indicate promise for integrating a self-reputation mechanism into a self-learning framework in a manner that encourages the early emergence of social cooperation among multiple agents and specific cooperation with individual agents.

Keywords: intelligent agent, e-commerce, self-reputation, prisoners' dilemma, multi-agent society, cooperation and defection

1. INTRODUCTION

Web 3.0 applications are increasingly using intelligent agents for tasks such as auto-bidding for products on eBay [1], auto-ordering equity stocks [2], and shopping for cheaper products [3]. Web 3.0 developers are experimenting with artificial intelligence (AI) techniques to create ways through which intelligent agents can "live" in Internet communities in ways that resemble how humans live in real-world communities [4-7]. However, due to finite resource and anonymity issues, intelligent agents within the Web 3.0 framework face two human-like problems: (a) cooperation and defection, and (b) conflicts between public and private interests [8-10]. The issues are more serious in Internet networks because Internet-based agents have greater mobility and anonymity [11, 12], increasing the potential for fraudulent behavior, betrayal, irresponsible comments, rumors, lack of cooperation, and a focus on private interests that stands in the way of maximizing benefits [13].

A large number of researchers are looking at ways of adding reputation mechanisms

Received December 19, 2012; revised June 17, 2013; accepted July 24, 2013.
Communicated by Jonathan Lee.

to agent models as a solution to these problems [10, 14-20]. Social scientists and game theory researchers are using reputation as a means for understanding how an individual's past behaviors affect current strategies and interactions involving co-workers [21, 22]. A shared finding in both disciplines is that most individuals, especially game players, are more willing to work and/or cooperate with opponents who have higher reputation values [9, 23-25]. This has led to recent efforts in the fields of computer science and AI to embed mechanisms in agent models for purposes of measuring reputation value and determining trustworthiness. In contrast, past agent models have emphasized learning from external environments only (*i.e.*, world model) − for instance, limiting efforts to looking for reliable co-workers with strong positive reputations [10, 18, 26-32]. However, this emphasis on external factors neglects an important point: the identification of ideal co-workers does not mean that those co-workers will automatically cooperate with someone who has a lower reputation value [33-35]. This can lead to situations where low-repute-tion agents spend most of their time searching for and being rejected by high-reputation agents, or where high-reputation agents take advantage of their status to cover up mistakes or to bully their co-workers − actions that can lead to a declining reputation value [33]. These scenarios underscore the need for agent introspection.

In this paper we will describe a Self-Reputation Calculation (SRC) intelligent agent model that emphasizes internal learning and self-reputation adjustment. According to our proposed model, low-reputation agents use their self-reputation calculations to build a strong understanding of the need to continuously cooperate with co-workers and to consider public interest as a primary goal − in other words, to consider self-reputation enhancement through cooperation as a primary learning target. The proposed model functions in a similar manner to the eBay system of having users complete a number of successful transactions in order to build up their self-reputations for future transactions. The model assumes that high-reputation buyers and sellers who are capable of adjusting their self-reputations are likely to protect the interests of others and to seek optimum balances between private and public interests in order to achieve maximum benefits. Thus, our SRC model increases the potential for cooperation among agents and decreases the potential for fraudulent behavior. Note that in the current e-commerce environment, agent relationships (both competitive and cooperative) are increasingly long-term, with decisions on both sides influenced by past interaction results. Accordingly, each agent is not only required to consider the other's past behaviors when selecting response strategies, but also to consider its own reputation based on other agents' feedback.

We used two agent-based social networks − cellular automata and Watts-Strogatz (WS) small-world [36] − to construct a model that supports agent efforts to calculate self-reputation values and to make strategy adjustments to enhance their self-reputations. SRC agents were added for purposes of playing Iterated Prisoner's Dilemma (IPD) games, with two agents sharing a link within a social network. Our study goals were: (a) to construct agents capable of learning internally, focusing on self-reputation, and reducing reliance on betrayal strategies so as to enhance self-reputation; (b) to have SRC agents consider public interests when pursuing their private interests, thereby increasing their willingness to cooperate with others so as to achieve an optimum balance between the two; and (c) to add SRC agents to artificial social networks to encourage all agents to consider self-reputation so that social cooperation emerges more quickly.

2. RELATED WORK

2.1 Self-Reputation and Artificial Multi-Agent Societies

Whereas researchers and designers working with reputation mechanisms in virtual societies are concerned with improving e-community reliability, capacity, and performance, AI researchers are increasingly considering environmental, community, and cooperation issues in their studies [37]. Another trend in intelligent and autonomous agent research comes from the social sciences, where there is considerable interest in adding reputation and trust mechanisms to multi-agent systems [18, 38, 39]. Researchers in these disciplines have identified three reputation characteristics:

1. Reputation is context-dependent. An individual's identity consists of multiple roles and/or characteristics, and each one may have its own reputation – for example, professional, personal, and family life reputations.
2. Reputation can be classified as public or private. Examples of public reputation include professional licenses awarded by public institutions. Private reputation evaluations are dependent on personal experiences.
3. Reputation knowledge is disseminated via information exchanges within social networks [40].

According to Mui *et al.*'s [41] reputation classification structure shown in Fig. 1, direct (also known as "encounter-derived") reputations are constructed from personal negotiations or interactions, while indirect (or "observed") reputations are based on word-of-mouth information from other agents. The eBay rating system is an example of indirect/observed reputation [28]. Indirect reputation can be further classified as prior-derived (*e.g.*, racial discrimination or gender bias; also known as preconceived), group-derived (*i.e.*, a group that an agent belongs to), or propagated (*i.e.*, obtained by asking other agents for information obtained by word-of-mouth). According to these characteristics and classifications, reputation cannot be separated from social networks. Social network analysts study individual social relations within networks in order to establish overall pictures of social network structure [42]; this explains our motivation to use social network analysis rather than the attributes of single individuals to establish a reputation model [38]. Obviously, the more relationship data that are obtained, the more complete the social network analysis.

Some ethologists and sociologists describe self-reputation as the sum of identity self-evaluation – that is, the sum of all evaluations of an individual's various community roles [43, 44]. Others assert that self-reputation affects all aspects of development and change in human self-discipline and self-respect, and that all individuals have the power to control, reflect on, and correct their behaviors and ideas [45, 46]. Humans are also capable of acknowledging their social images in ways that affect their self-cognition; once established, human self-reputation serves as a guide for making self-inferences through decisions driven by internal motives [47-49]. However, Smith and Mackie [50, 51] are among researchers who disagree with claims that individuals can achieve self-understanding by observing their own behaviors – in other words, they resist a self-centered view of social behavior that perceives similarities in the ways that self-images and self-reputations are formed. Instead, they believe that individual social behavior is strongly affected by judgments expressed by significant members of one's social network.

If true, this means it is easier to objectively observe one's neighbors than oneself – a scenario that supports a cycle in which self-reputation supports an understanding of others' reputations and vice versa [33]. It also infers that one's self-reputation is strongly influenced by the perceptions of others in terms of the acknowledgment, interpretation, and correction of erroneous self-reputations. This "looking-glass self" is important for establishing correct attitudes and concepts because it supports an understanding of how others' perspectives differ from our own. Without this ability and willingness to use the eyes of others for self-reflection, self-reputation can easily become blurred [52, 53].

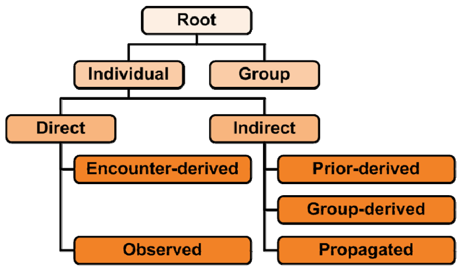


Fig. 1. Mui *et al.*'s [41] reputation classifications.

2.2 Intelligent Agent Models

In most agent models, architectures, rules, and knowledge are key factors affecting performance in terms of problem-solving speed and solution quality. Intelligent agent-based models can be divided into two categories based on knowledge structure and adaptation method: rule-based and learning-based [54]. For rule-based agents, domain experts are required to prepare and embed in advance all knowledge, information, and rules required for agent-solving problems. Rule-based agents use a combination of this embedded knowledge and external information to develop appropriate behaviors [54]. Their main drawbacks are complex updating requirements and lack of customization capability. Learning-based agents solve problems, collect experience, and compute statistical information toward the goal of establishing and adjusting optimal problem solving strategies [54]. Learning-based agents must convert this information into a form that they can store according to their knowledge bases and rule sets.

The "adaptive agent-based" structure of today's mainstream intelligent agents (Fig. 2) represents a mix of rule-based and learning-based models [54]. To explain this model (which is best understood as a sensing, planning, and acting model), we will use the extended classifier system (XCS), a problem-independent and adaptive learning-based agent model that has four components: performance, reinforcement, rule discovery, and a finite classifier population (Fig. 3) [55]. Stored classifiers use a horizontal competition mechanism to control the system and vertical cooperation to perform tasks. The performance component governs interactions with the target problem. The input interface transmits current target problem states to the performance component, and determines dominant classifiers according to exploration/exploitation criteria. Following execution, any action advocated by a dominant classifier receives feedback. Similar to Q-learning, the reinforcement (credit assignment) component uses an algorithm to update the parameters of classifiers that advocate output action. The rule discovery component uses a genetic algorithm to search for more efficient or more general classifiers, and to discard incorrect classifiers.

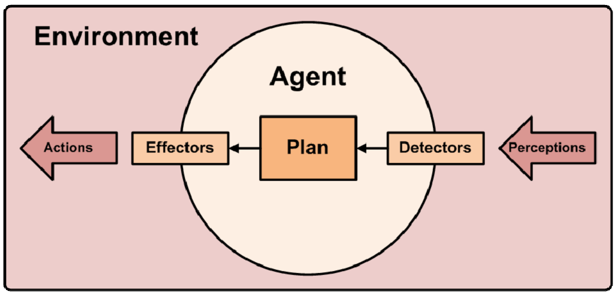


Fig. 2. Standard version of sense-plan-act agents that interact with their environments via input-output (detector-effector) interfaces.

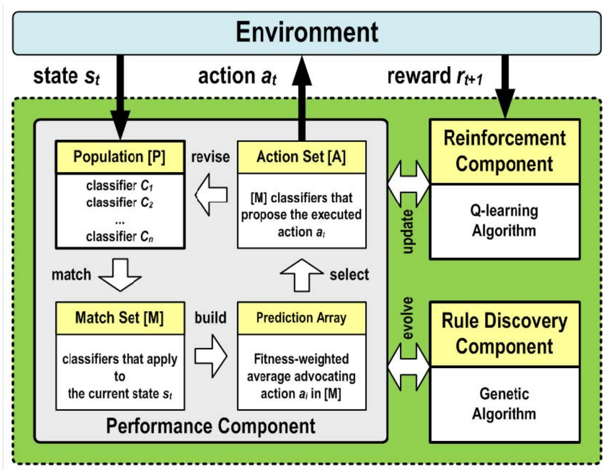


Fig. 3. Extended classifier system (XCS) architecture.

An XCS agent senses the current state *St* during discrete time *t*, and compares it with all rules in the learning classifier component *P* to find matching rules for forming a rule set to be placed in buffer *M*. The most appropriate *M* rule is selected based on a combination of fitness value, cost estimate, and experience accumulation. The rule is accepted or rejected according to external data (trial and error), thereby encouraging constant improvement in learning effectiveness to achieve optimum performance. While researchers in various disciplines have proposed intelligent agent learning algorithms (*e.g.*, neural networks, genetic algorithms, Q-Learning), a structure based on "sensing-planning-action" is still the most widely accepted and used design type [56].

2.3 Iterated Prisoner's Dilemma

The first Prisoner's Dilemma (PD) puzzles were devised by Merrill Flood and Melvin Dresher in 1950 as part of a Rand Corporation game theory project tied to global nuclear strategies; the term "prisoner's dilemma" was created by Albert Tucker [57]. PD game payoff matrices have been used to simulate and explain many types of human and non-human interactions in such fields as economics [58], politics and sociology [59], and the biological sciences, especially ethology and evolutionary biology [60]. Many models of natural processes have been established based on extended PD game series.

Our proposed model makes use of the Iterated Prisoner's Dilemma (IPD), an extension of the original in which the same two opponents play multiple non-single sets of PD games. The IPD, considered fundamental to certain theories of human cooperation and trust, is also referred to as the "Peace-War Game" [61]. Assuming that the basic IPD game accurately represents transactions requiring trust between two agents, a multi-player version can be used to create a cooperative behavior model for any population. This explains its popularity among researchers: already by 1975, Grofman and Pool estimated that more than 2,000 scholarly articles had been published on IPD-related topics [62]. In 1984, Axelrod suggested that IPD games could be used to study cooperative behaviors in societies primarily consisting of selfish individuals, and to emphasize the importance of players remembering their opponents' previous behaviors [61].

According to game theory, the PD explains why two individuals may not cooperate even though it is in their best interest to do so. A typical example involves two suspects arrested by the police. Since the police do not have enough evidence for a clear conviction, they separate the two prisoners and offer freedom to a defector if he testifies against the other suspect [63]. A payoff matrix for a typical PD scenario is presented in Table 1. As shown, each prisoner gets a higher pay-off if he betrays the other. According to the Table 1 payoff schedule, prisoner A is much better off defecting regardless of what prisoner B does, therefore prisoner A logically should defect. Since the game is symmetric, prisoner B should act the same way, but when both defect, each receives a lower reward than if both cooperate. The result of rational decision making is that both players will be worse off if one chooses to lessen the sentence of his opponent, even at the self-inflicted cost of spending more time in jail.

A long list of researchers in the natural and social sciences have used a combination of an agent-based non-zero PD model with simulation-based computer models to explore conflicts between the public good and private interests. In most cases the research goal has been to determine relationships between rational strategies and evolutionary results from cooperation or selfish betrayal [9, 63]. The payoff matrix used in the present study (Table 2) serves as an illustrative example. $R = 3$ represents the reward for dual cooperation, $T = 5$ the temptation of unilateral betrayal, $S = 0$ the punishment for unilateral cooperation, and $P = 1$ the punishment for dual betrayal. According to Dresher and Flood's original definition, two key PD conditions are $T > R > P > S$ and $2 \times R > T + S$. The first condition states that the two participants will betray each other once they understand the $T > R$ and $P > S$ conditions, and therefore choose the second best option (P, P). The second condition is that prisoners cannot avoid the dilemma by alternately betraying each other. In other words, the alternate betrayal payoff is inferior to the dual cooperation payoff.

Table 1. Prisoner's Dilemma Scenarios.

	Prisoner B cooperates	*Prisoner B defects*
Prisoner A cooperates	Each serves 1 month sentences	Prisoner A: 1 year sentence Prisoner B: goes free
Prisoner A defects	Prisoner A: goes free Prisoner B: 1 year sentence	Each serves 3 month sentences

Table 2. Payoff Matrix.

		Player B	
		Cooperation (C)	**Defection (D)**
Player A	**Cooperation (C)**	Reward for mutual cooperation (R = 3 , $R = 3$)	"Sucker's payoff" and temptation to defect ($T = 5$, $S = 0$)
	Defection (D)	Temptation to defect and "sucker's payoff" ($S = 0$, $T = 5$)	Punishment for mutual defection ($P = 1$, $P = 1$)

Note: Numbers given in each cell are payoffs for both players, with player A's payoff listed first.

3. MODEL AND SIMULATION ENVIRONMENT

3.1 SRC Model

According to the agent learning pattern shown in Fig. 4, detectors and effectors are responsible for communicating with the external environment, and the learning element is responsible for improving agent learning capability [54]. The performance element is responsible for sensing external environmental states and selecting appropriate responses, the critic component for evaluating current learning effects, and the problem generator for determining the best learning direction for an agent in a new situation ("new experience").

We used the agent model described in [54] to develop a SRC agent model consisting of performance, learning, and SRC elements (Fig. 5). The third element helps agents use past experiences as information to be added to knowledge bases, thus supporting self-reputation understanding, explanations, and predictions. The learning and self-reputation calculation elements act in coordination to support agent efforts to judge external environments (opponent strategies) and internal cognition, which moves SRC agents closer to a human intelligence model. We believe the ability of an agent to reflect on its self-reputation will result in increased social benefits and faster collaborative behaviors [18, 28, 37, 39]. Our proposed model combines social expectation strategies with a reputation evaluation algorithm. The following pseudo code describes the process that an agent uses to give his opponent a reputation score:

```
function  compute the reputation values of agent's opponents  (agent A_i) return a List is
begin
   Step 1. Assume that agent A_i has n opponents so that O_i = (o_{i,0}, o_{i,1}, o_{i,2}, ··· , o_{i,n-1})^g during generation g;
```

$o_{i,j}$ represents agent A_i's jth opponent ($j \in \{0, 1, 2, \ldots, n-1\}$).

Step 2. At the end of generation g we arrive at the numerical sequence $C_i = (c_{0,i}, c_{1,i}, c_{2,i}, \ldots, c_{n-1,i})^g$, with $c_{j,i}$ representing the number of times that opponent $o_{i,j}$ makes a cooperative move during a PD interaction with agent A_i in generation g.

Step 3. Calculate the average value (avg_i) and standard deviation (std_i) of the numberical sequence C_i.

Step 4. Compute the reputation value $r_{i,j}$ of agent A_i's opponent $o_{i,j}$ as $\lceil (c_{j,i} - avg_i)/std_i \rceil$.

Step 5. Return the reputation values $R_i = (r_{i,0}, r_{i,1}, r_{i,2}, \ldots, r_{i,n-1})^g$ of agent A_i's opponents O_i in generation g.

end.

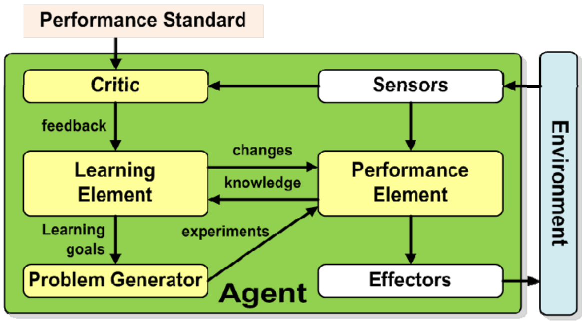


Fig. 4. A general learning agent model.

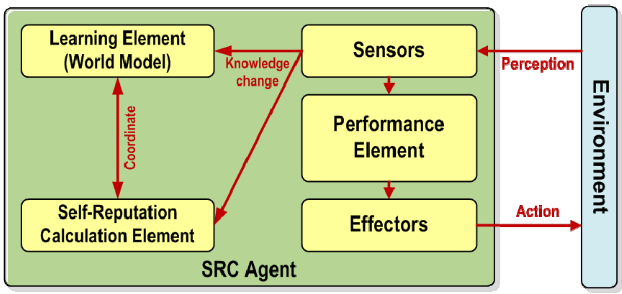


Fig. 5. Self-reputation calculation (SRC) agent model.

The self-reputation mechanism gives agents access to self-knowledge of their fitness degrees and reputation values within their respective groups. As shown in Fig. 6, fitness and reputation are categorized as high, mid, or low, resulting in nine possible agent-opponent interaction types. For example, an agent with a high degree of fitness and low degree of reputation usually follows an always-betray strategy (also known as ALL-D, or "villain strategy") to achieve better performance. However, an ALL-D strategy does not produce a higher public good value because acts of betrayal decrease the ability of other individuals to promote their own interests. Thus, an intelligent agent with an SRC model must be taught that an ALL-D strategy will negatively affect its self-reputation. By learning from its opponent's strategy, a self-adjusting agent may achieve higher fitness and reputation values.

The following steps refer to an algorithm that can be used to calculate self-fitness and self-reputation values for an agent A_i. Fitness value is defined in terms of the IPD payoff matrix mentioned in section 2.3 (Table 2). Agent fitness will be described in greater detail in section 3.3, step 4. The following pseudo-code refers to the self-reputation calculation algorithm:

```
procedure self reputation calculation (agent A_i) is
begin
   Step 1. Assume that agent A_i has n opponents so that O_i = (o_{i,0}, o_{i,1}, o_{i,2}, ..., o_{i,n-1})^g during generation g;
           o_{i,j} represents agent A_i's jth opponent (j ∈ {0, 1, 2, ..., n − 1}).
   Step 2. Assume that af_i is agent A_i's fitness value.
   Step 3. Agent A_i asks all o_{i,j} opponents about their evaluations of her reputation r_{j,i}.
           In the numerical sequence AR_i = (r_{0,i}, r_{1,i}, r_{2,i}, ..., r_{n-1,i})^g, r_{j,i} represents opponent o_{i,j}'s
           evaluation of agent A_i's reputation. The information is used to calculate
           the average value ar_i of numerical sequence AR_i; the resulting ar_i represents
           agent A_i's average reputation in group O_i.
   Step 4. Agent A_i collects reputation information for opponent O_i from all other opponents,
           and uses it to create the numerical sequence OR_i = (r_{0,1}, ..., r_{0,n-1}, r_{1,0}, r_{1,2}, ..., r_{1,n-1}, ..., r_{n-1,0}, ..., r_{n-1,n-2})^g
           with length n × (n − 1). r_{j,k} represents opponent o_{i,j}'s evaluation of opponent o_{i,k}'s reputation.
           Also as part of this step, the average value (OR_avg_i) of numerical sequence OR_i and
           standard deviation (OR_std_i) are calculated.
   Step 5. Agent A_i collects fitness value data for all opponents O_i and creates the numerical
           sequence OF_i = (f_0, f_1, f_2, ..., f_{n-1})^g with length n. f_j represents the fitness value of opponent o_{i,j}.
           Also as part of this step, the average value (OF_avg_i) of numerical sequence OF_i and
           standard deviation (OF_std_i) are calculated.
   Step 6. if Agent A_i's fitness value af_i is less than (OF_avg_i - OF_std_i) then
              Fitness_i = LOW
           else if Agent A_i's fitness value af_i is greater than (OF_avg_i + OF_std_i) then
              Fitness_i = HIGH
           else
              Fitness_i = MIDDLE
           end if
   Step 7. if Agent A_i's reputation ar_i is less than (OR_avg_i - OR_std_i) then
              Reputation_i = LOW
           else if Agent A_i's reputation ar_i is greater than (OR_avg_i + OR_std_i) then
              Reputation_i = HIGH
           else
              Reputation_i = MIDDLE
           end if
   Step 8. if Fitness_i == LOW or Reputation_i == LOW then
              comment Agent A_i's strategy is not suitable for surviva within its group.
              Agent A_i uses crossover and mutation operations to adjust its strategy.
              comment refer to the pseudo-code of IPD simulation experiment.
           end if
end.
```

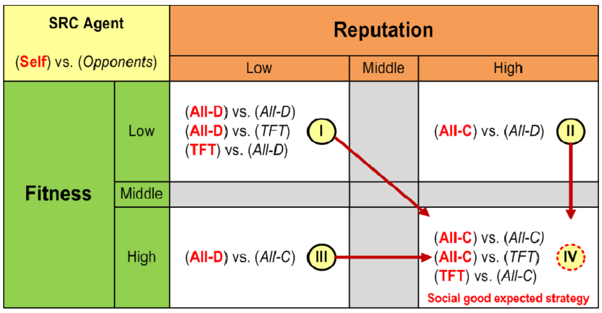


Fig. 6. Agent fitness score and reputation index matrix.

3.2 Strategies

Within each PD game set, agents are free to choose between cooperation and betrayal, with betrayal always resulting in greater benefits. However, if both players choose betrayal, total rewards between them will be less than those from mutual cooperation. In e-commerce, the possibility exists of two agents interacting more than once, thus transforming the game into an IPD. After a period of play, survival pressure is likely to result in mutual long-term cooperation, which has the potential to support both the public interest and individual fitness [61].

In addition to always-betray (ALL-D) and always-cooperate (ALL-C), two other strategies have attracted research interest. The first is the "win-stay, lose-shift" strategy – also known as the PAVLOV strategy. This strategy applies Pavlovian psychological theory by proposing that an agent will adhere to one strategy until its income goes below a threshold, causing it to switch to the opposite strategy [64-66]. The other is "tit-for-tat," (TFT), in which an agent always chooses cooperation during the first round of a game, and then imitates its opponent's strategy in subsequent rounds [61]. TFT is based on four principles: (a) it must be friendly – that is, the user must never be the initial traitor; (b) once an opponent chooses betrayal, a TFT player must immediately retaliate; (c) tolerance is required – that is, as soon as an opponent stops its betrayal behavior, an agent must immediately stop its own betrayal behavior; and (d) the strategy must be transparent.

As shown in Table 3, IPD game agents generally follow a memory-1 deterministic strategy (*i.e.*, they memorize the previous set's results): CC (all cooperate), CD (I cooperate but my opponent defects), DC (I defect but my opponent cooperates), or DD (both defect). These strategies are expressed as (S_{cc}, S_{cd}, S_{dc}, S_{dd}). Since an agent has two responses to choose from, there are 2^4 = 16 possible memory-1 deterministic strategies. Note that S_0 = (C, C, C, C), S_1 = (C, C, C, D), …, S_{15} = (D, D, D, D) where S_0 = (C, C, C, C) denotes the ALL-C strategy, S_5 = (C, D, C, D) the TFT strategy, S_6 = (C, D, D, C) the PAVLOV strategy, and S_{15} = (D, D, D, D) the ALL-D strategy.

Table 3. Memory-1 deterministic strategies.

(S_{cc}, S_{cd}, S_{dc}, S_{dd})		Note
S_0	(C, C, C, C)	all-cooperation strategy (ALL-C)
S_1	(C, C, C, D)	
S_2	(C, C, D, C)	
S_3	(C, C, D, D)	
S_4	(C, D, C, C)	
S_5	(C, D, C, D)	tit-for-tat strategy (TFT)
S_6	(C, D, D, C)	win-stay, lose-shift strategy (PAVLOV)
S_7	(C, D, D, D)	
S_8	(D, C, C, C)	
S_9	(D, C, C, D)	
S_{10}	(D, C, D, C))	
S_{11}	(D, C, D, D)	
S_{12}	(D, D, C, C)	
S_{13}	(D, D, C, D)	
S_{14}	(D, D, D, C)	
S_{15}	(D, D, D, D)	all-defection strategy (ALL-D)

3.3 Simulation Social Networks

Our proposed simulation environment[1] consists of two layers: a multi-agent top layer and an agent contact/social network bottom layer. The bottom layer is either a cellular automata or a Watts and Strogatz (WS) small-world network. As shown in Fig. 7, the first is a two-dimensional $n \times n$ regular network with high degrees of local clustering and separation [67], and the second has the properties of a high degree of local clustering and a low degree of separation [68]. To compare IPD simulation results for these network types, we stipulated that the numbers of nodes ($N = n \times n$, representing the number of agents) and edges ($M = 4N$ the number of agent contacts and interactions) in each must be equal. Each node has an average of eight adjacent nodes. As shown in Fig. 8 (a), we also stipulated that one cell is equivalent to one node, and that each node is connected to eight adjacent nodes. After all nodes are established and connected, a rewiring probability ρ is used to determine whether or not individual edges must be rewired. If rewiring is necessary, one of the two original nodes (one on each side of an edge) is discarded and replaced with a randomly selected new node (Fig. 8 (b)).

Our simulation experiment consisted of eight steps:

```
main procedure IPD simulation experiment is
begin
   Step 1. Set environmental parameters and initial values for evolutionary computations (Table IV).
           comment Environmental parameters include total number of time steps for each simulation
                   experiment, strategy and color mapping table, agent memory capacity, and human
                   contact and interactive social network values (i.e., numbers of nodes and edges,
                   neighbor patterns, and small-world network rewiring probability).
           comment Evolutionary computation parameters include total number of agents/nodes, crossover
                   rate, mutation rate, and total number of generations/time steps.
   Step 2. Use experimental requirements to choose and construct the most appropriate network:
           (1) cellular automata or (2) two-dimensional WS small-world.
   Step 3. Set time step t to 0.
   Step 4. Based on network connection patterns, have the nodes at the ends of links A_i and A_j execute q
           IPD rounds. Using the payoff matrix shown in Table 2, calculate scores for A_i (as_i) and A_j (as_j),
           both ranging from q × S to q × T (i.e., from 0 to 5q). Use these scores as fitness values for
           agents A_i (af_i) and A_j (af_j), with af_i ← as_i and af_j ← as_j.
   Step 5. Agent calculate their relative fitness (Fitness_i) and reputation (Reputation_i) values,
           which represent the evaluations of all A_i opponents.
   Step 6. Each agent determines whether or not to make strategy adjustments.
           The current strategy is considered inappropriate when Fitness_i == LOW or Reputation_i == LOW.
           comment Agents that need to adjust their strategies use evolutionary computation crossover
                   operation to combine their original strategies with strategies used by opponents with
                   high fitness values. They also use evolutionary computation mutation operation to
                   randomly change their strategies.
   Step 7. t ← t + 1
   Step 8. if t < Time_Step_Limit then
              goto Step 4
           end if
   Step 9. Terminate the experiment.
end.
```

[1] The simulation environment used in this project was built using JAVA programming language. For source code, please contact the corresponding author.

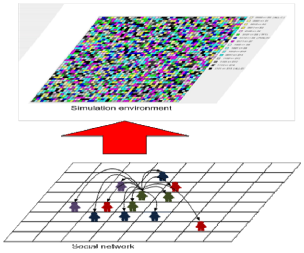
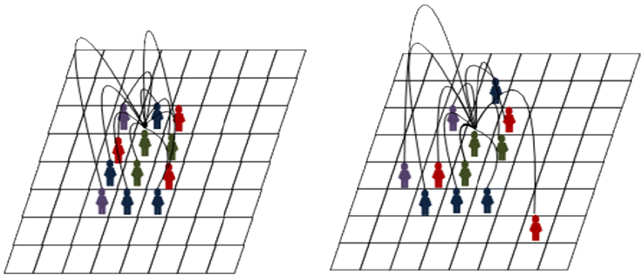


Fig. 7. A simulation model.



(a) Cellular automata. (b) Two-dimensional WS small-world network.
Fig. 8. Human contact and interactive social networks at the lower layers of our proposed simulation model.

Table 4. Experimental parameters.

Parameter	Default value	Description
TIME_STEP_LIMIT	100	Total number of generations for each simulation experiment.
q	100	Total number of interactions between an agent and its opponent.
Network Type		If NetworkType = CA, a 2D cellular automata is built; if NetworkType = WS-SWN, a 2DWatts and Strogatz small-world network is built.
W	50	Width of 2D cellular automata.
H	50	Height of 2D cellular automata.
N	2,500	Total number of nodes (agents). Default value= W×H.
E	10,000	Total number of edges. Default value = (50 × 50 × 8)⁄2.
Neighborhood	Moore	Von Neumann/Moore neighborhood pattern with periodic boundary condition.
ρ	0.01	Generating such a network begins with 2D cellular automata with periodic boundary conditions. Each link is randomly rewired to new node according to rewiring probability.
C	1	Agent memory capacity.
P_c	0.7	Crossover rate of the genetic algorithm used in this work.
P_m	0.01	Mutation rate of the genetic algorithm used in this work.

4. RESULTS AND DISCUSSION

Results from experiments using 0% (control), 10%, 30%, 50% and 100% SRC agents are shown in Fig. 9 (cellular automata) and 10 (two-dimensional WS small-world network). Initial parameter settings were identical. Blue curves indicate significant improvement in societal interest when all agents were capable of adjusting their self-reputations; note that a dynamic equilibrium was achieved within very few generations. Since this is an unrealistic scenario, we focused on experiments involving small numbers of SRC agents with less emphasis on agent architectures and/or learning mechanisms. As indicated by the red (10%) and green (30%) curves in both figures, small numbers of SRC agents were capable of exerting significant influence in either cellular automata or WS networks, thus counteracting the effects of ALL-D agents and promoting coopera-

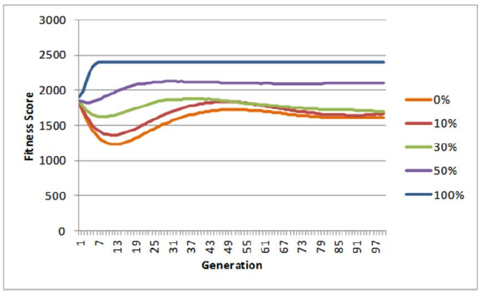


Fig. 9. Effects of adding different percentages of various SRC agent to cellular automata.

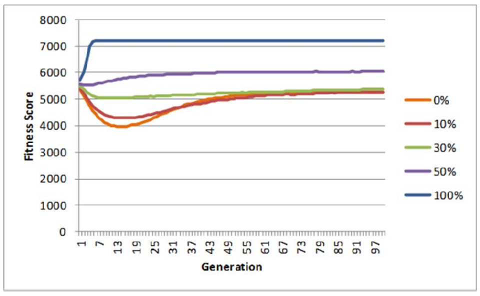


Fig. 10. Effects of adding different percentages of various SRC agents to a WS small-world network.

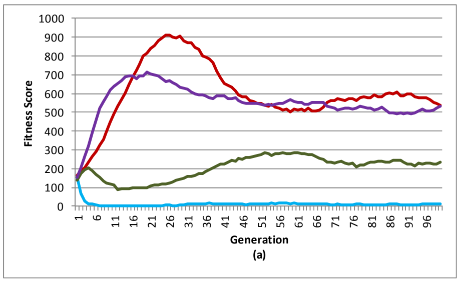
tion as standard behavior. The results indicate that the proposed SRC model successfully encourages agents to consider self-reputation for purposes of reducing betrayal behavior in the overall community and promoting the early emergence of social cooperation.

Agents who follow the TFT strategy cooperate easily with ALL-C agents and generally defeat ALL-D agents, yet they suffer from asymmetric memory of preceding sets, sometimes resulting in two TFT agents engaging in long-term repetitive breaches of contracts. In contrast, the focus of a PAVLOV strategy is on profit, with outcomes clearly categorized as success/failure and win/lose. When one round results in a win, that agent uses the same strategy for the next round, and when a loss occurs, the agent switches to another strategy. To determine how SRC agents influence individual phases of an experimental simulation, we analyzed evolutionary dynamics and equilibria among the four IPD game strategies by adding (a) 0% SRC agents (Figs. 11 (a) and (b), control group), (b) 100% SRC agents (Figs. 11 (c) and (d), extreme special condition), or (c) 10% SRC agents (Figs. 11 (e) and (f)). The data in Figs. 11 (a)-(e) are considered part of Category 1 (cellular automata); all other figures are part of Category 2 (two-dimensional WS small world social networks). Our primary focus was on Figs. 11 (e) and (f) data.

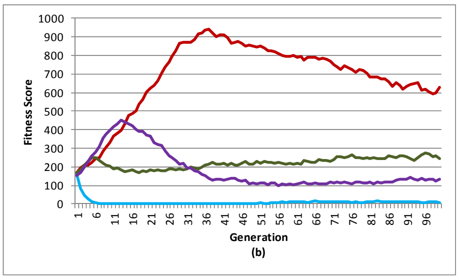
The Fig. 11 (a) data reflect two conditions: cellular automata as the underlying social network, and zero agents capable of adjusting their self-reputations. Numbers of the four agent types were similar (within 5%) during the first two generations. After the third generation, the number of ALL-D agents increased sharply and the number of ALL-C agents decreased at a comparable rate. Larger numbers of TFT agents emerged when the number of ALL-D agents reached a certain value; after 20 generations their number exceeded the number of ALL-D agents, resulting in a rapid decline in ALL-D agents. The number of PAVLOV agents started to increase around generation 30; by generation 60 the number of TFT agents fell below those of ALL-D agents, triggering a rapid increase in the number of ALL-D and PAVLOV agents. At generation 80, TFT agents once again exceeded ALL-D agents, and the numbers of PAVLOV and ALL-C agents stabilized.

Fig. 11 (b) represents WS small-world network conditions and zero SRC agents capable of adjusting their self-reputations. During the early stages of evolution, growth and decline trends for the four strategies were similar to those shown in Fig. 11 (a). After 30 generations, the number of ALL-D agents stabilized at a fixed number (lower than the cellular automata equivalent) and did not change from that point forward. Due to the low degree of separation that is characteristic of two-dimensional WS small-world networks, a dynamic equilibrium was achieved between generations 50 and 60 – faster than in the cellular automata scenario. Note that when all agents in a cellular automata simulation were capable of self-reputation adjustment, ALL-D agents quickly acknowledged the lack of social acceptance for their strategy and quickly used their self-adjustment mechanisms to match the expectations of adjacent agents, resulting in the complete disappearance of ALL-D agents between generations 3 and 4, and rapid stability among agents using the other three strategies. Again, all initial parameter settings for Figs. 11 (c) (cellular automata) and (d) (WS small-world network) were identical, meaning that the only significant difference between them was the increased sensitivity of an agent-based society due to random long-range shortcuts associated with WS networks. The slightest change in the strategy of a single agent could affect the entire community by causing a large number of other agents to change their strategies. However, due to the low separation characteristic of WS networks, these societies quickly adjusted and achieved new equilibrium states.

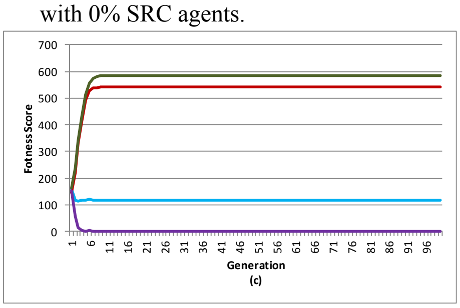
Last, the data shown in Figs. 11 (a), (c) and (e) indicate that when agents were embedded in the SRC model, the number of ALL-D agents decreased rapidly while the number of TFT agents remained high, indicating that SRC agents clearly acknowledged the potential damage to their reputations if they followed an ALL-D strategy. However, always following an ALL-C strategy increased the potential for a SRC agent to be taken advantage of, resulting in a lower fitness value. Accordingly, they learned the importance of cooperating with other agents who were committed to an ALL-C strategy, and used a defection strategy with all others. These results support the three goals as stated in our introduction section.



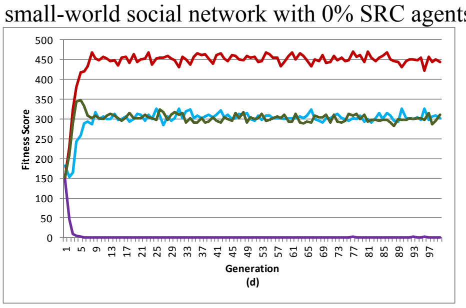
(a) Evolution dynamics in a cellar automata with 0% SRC agents.



(b) Evolution dynamics in a two-dimensional WS small-world social network with 0% SRC agents.

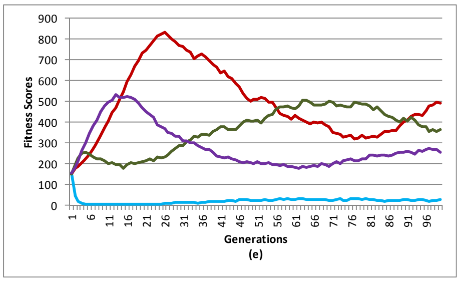


(c) Evolution dynamics in a cellular automata with 100% SRC agents.

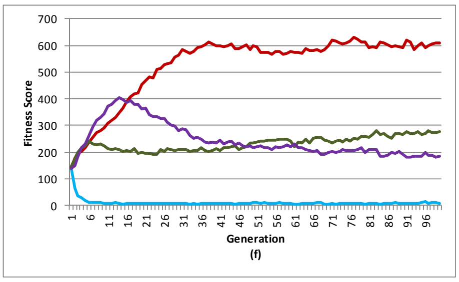


(d) Evolution dynamics in a two-dimensional WS small-world social network with 100% SRC agents.

Fig. 11. Evolution dynamics for four well-known Prisoner's Dilemma strategies.



(e) Evolution dynamics in a cellular automata with 10% SRC agents.



(f) Evolution dynamics in a two-dimensional WS small-world social network with 10% SRC agents.

Fig. 11. (Cont'd) Evolution dynamics for four well-known Prisoner's Dilemma strategies.

From our comparison of the effects of various percentages of SRC agents in cellular automata and small-world networks, we found that even a small number could trigger positive change in the form of increased society-wide benefits, suggesting that our proposed model has potential utility for solving private/public interest conflicts in social science research. We observed four evolutionary stages:

1. **Generations 1-10:** agents built an understanding of the effects of an ALL-D strategy. In the interest of self-survival, a large number of agents adopted this strategy, resulting in societal chaos and lack of trust among individuals. Apparently agents who observed their partners following an ALL-D strategy felt compelled to do the same to ensure their survival. This ALL-D clustering phenomenon resulted in many negative effects.
2. **Generations 11-20:** the effects of self-adjusting reputation capability among SRC agents emerged. A small number of SRC agents who originally used the ALL-D strategy started to use the TFT strategy in reaction to social pressure, even though a cluster of SRC agents gave the appearance of waiting for outside assistance in order to continue using their preferred ALL-D strategy. Over generations, the growing number of TFT agents served as a balance for ALL-D agents, thus preventing further chaos. Further, as the number of TFT agents grew, they started to surround the remaining ALL-D agents, cutting them off from the influences of all non-TFT strategies.
3. **Generations 21-40:** the number of TFT agents declined. Due to the asymmetric memory problem associated with this strategy, there was a disintegration of TFT agent clustering, even though the numbers of TFT agents surrounding ALL-D agents did not change. PAVLOV agents also started to emerge during this stage. Lacking an asymmetric memory problem, they were more capable of cooperating with other friendly agents (*i.e.*, ALL-D, TFT and PAVLOV strategy agents), which explains their increasing numbers.
4. **Generations 41-100:** evolutionary equilibrium was achieved. ALL-C agents started to appear in regions dominated by PAVLOV agents. During this stage, ALL-C and PAVLOV agents interacted in ways that rewarded both sides. Although this scenario is often found in human societies, such societies are at-risk of damage from internal

mutations and external invasions. Repetition of this particular process can lead to stable agent equilibrium.

7. CONCLUSION

In this paper we described our proposal for a Self-Reputation Capability (SRC) model in which agents are given the power to calculate and interpret their self-reputation values and to adjust their IPD strategies accordingly. Our primary conclusions are (a) the proposed model is successful in encouraging agents to adjust their strategies to achieve the best balance between self-reputation and private interests, thus making it likely that an agent will suppress its betrayal behavior in the form of a defection strategy in order to increase cooperation with co-workers; and (b) overall social cooperative behavior is likely to emerge much earlier compared to other models. It is our hope that the SRC model will support the efforts of smart object researchers interested in improving intelligent agent internal cognition and external learning capability.

REFERENCES

1. M.-P. Wellman, A. Greenwald, and P. Stone, *Autonomous Bidding Agents: Strategies and Lessons from the Trading Agent Competition* (*Intelligent Robotics and Autonomous Agents*), 2007, The MIT Press.
2. S.-H. Chen and C.-H. Yeh, "Evolving traders and the business school with genetic programming: A new architecture of the agent-based artificial stock market," *Journal of Economic Dynamics and Control*, Vol. 25, 2001, pp. 363-393.
3. P. Maes, R.-H. Guttman, and A.-G. Moukas, "Agents that buy and sell," *Communications of the ACM*, Vol. 42, 1999, pp. 81-91.
4. J. Hendler, "Web 3.0 emerging," *Computer*, Vol. 42, 2009, pp. 111-113.
5. A. Gutowska and A. Sloane, "Modelling the B2C marketplace: Evaluation of a reputation metric for e-commerce," in *Proceeding of the 5th International Conference on Web Information Systems and Technologies*, 2010, pp. 212-226.
6. R.-D. Morris, "Web 3.0: Implications for online learning," *TechTrends*, Vol. 55, 2011, pp. 42-46.
7. M. Tasner, *Marketing in the Moment: The Practical Guide to Using Web 3.0 Marketing to Reach Your Customers First*, 2010, Ft Press.
8. F. McGillivray and A. Smith, "Trust and cooperation through agent-specific punishments," *International Organization*, Vol. 54, 2000, pp. 809-824.
9. R. Axelrod, *The Complexity of Cooperation: Agent-based Models of Competition and Collaboration*, Princeton University Press, 1997.
10. S.-D. Ramchurn, D. Huynh, and N.-R. Jennings, "Trust in multi-agent systems," *The Knowledge Engineering Review*, Vol. 19, 2005, pp. 1-25.
11. S. Garfinkel and G. Spafford, *Web Security, Privacy & Commerce*, 2011, O'Reilly Media.
12. J. Hogg, *Web Service Security: Scenarios*, *Patterns*, *and Implementation Guidance for Web Services Enhancements (WSE) 3.0*, 2006, Microsoft Press.
13. A.-P. Mintz, *Web of Deception: Misinformation on the Internet*, 2002, Information

Today, Inc.
14. A. Jøsang, "Trust-based decision making for electronic transactions," in *Proceedings of the 4th Nordic Workshop on Secure Computer Systems*, 1999, pp. 1-21.
15. D.-H. Mcknight and N.-L. Chervany, "What trust means in e-commerce customer relationships: an interdisciplinary conceptual typology," *International Journal of Electronic Commerce*, 2001, Vol. 6, pp. 35-59.
16. J. Carter, E. Bitting, and A.-A. Ghorbani, "Reputation formalization for an information-sharing multi-agent system," *Computational Intelligence*, Vol. 18, 2002, pp. 515-534.
17. P. Resnick and R. Zeckhauser, "Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system," *Advances in Applied Microeconomics: A Research Annual*, 2002, Vol. 11, pp. 127-157.
18. W. Wang, G. Zeng, and L. Yuan, "A reputation multi-agent system in semantic web," *Agent Computing and Multi-Agent Systems*, Vol. 4088, 2006. pp. 211-219.
19. Y. Wang and K.-J. Lin, "Reputation-oriented trustworthy computing in e-commerce environments," *IEEE Internet Computing*, 2008, Vol. 12, pp. 55-59.
20. W. Liang, G. Yajun, and Q. Mei, "A reputation-based trust evaluation model for P2P e-commerce," *International Journal of Distributed Sensor Networks*, Vol. 5, 2009, pp. 39-39.
21. D. Kreps and R. Wilson, "Reputation and imperfect information," *Journal of Economic Theory*, Vol. 27, 1982, pp. 253-279.
22. D. Bromley, *Reputation*, *Image and Impression Management*, John Wiley & Sons Ltd., NY, 1993.
23. R. Arnold, "The evolution of cooperation," *The Review of Austrian Economics*, Vol. 1, 1987, pp. 227-229.
24. E. Adler, "The emergence of cooperation: national epistemic communities and the international evolution of the idea of nuclear arms control," *International Organization*, Vol. 46, 2009, pp. 101-145.
25. E. Maskin, "Evolution, cooperation, and repeated games," *Games*, *Groups*, *and the Global Good*, 2009, pp. 79-84.
26. E. Chang, F. Hussain, and T. Dillon, *Trust and Reputation for Service-Oriented Environments: Technologies for Building Business Intelligence and Consumer Confidence*, Johns Wiley & Sons, Hoboken, NJ, 2007.
27. S. Jiang and J. Li, "A reputation-based trust mechanism for P2P e-commerce systems," *Journal of Software*, Vol. 18, 2007, pp. 2551-2563.
28. A. Jøsang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision," *Decision Support Systems*, Vol. 43, 2007, pp. 618-644.
29. J. Shouxu and L. Jianzhong, "A reputation-based trust mechanism for P2P e-commerce systems, *Journal of Software*, Vol. 18, 2007, pp. 2551-2563.
30. T. Teo and J. Liu, 'Consumer trust in e-commerce in the United States, Singapore and China," *Omega*, Vol. 35, 2007, pp. 22-38.
31. A. Caballero, *et al.*, "A trust and reputation model as adaptive mechanism for multi-agent systems," *Inteligencia Artificial*, Vol. 13, 2009, pp. 42.
32. S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed., Prentice Hall, Upper Saddle River, NJ, 2009.
33. E. Aronson, T. D. Wilson, and R. M. Akert, *Social Psychology*, 7th ed., Pearson,

Boston, MA, 2010.
34. O. Leimar and P. Hammerstein, "Evolution of cooperation through indirect reciprocity," in *Proceedings of the Royal Society of London*, *Series B: Biological Sciences*, Vol. 268, 2001, pp. 745.
35. J. Sachs, *et al.*, "The evolution of cooperation," *The Quarterly Review of Biology*, Vol. 79, 2004, pp. 327-348.
36. M. E. Newman and D. J. Watts, "Scaling and percolation in the small-world network model," *Physical Review E*, Vol. 60, 1999, pp. 7332.
37. J. Sabater and C. Sierra, "Review on computational trust and reputation models," *Artificial Intelligence Review*, Vol. 24, 2005, pp. 33-60.
38. J. Sabater and C. Sierra, "Reputation and social network analysis in multi-agent systems," in *Proceedings of the 1st International Joint Conference on Autonomous Agents and Multi-agent Systems*, 2002, pp. 475-482.
39. T. Huynh, N. Jennings, and N. Shadbolt, "An integrated trust and reputation model for open multi-agent systems," *Autonomous Agents and Multi-Agent Systems*, Vol. 13, 2006, pp. 119-154.
40. B. Yu and M. Singh, "Distributed reputation management for electronic commerce," *Computational Intelligence*, Vol. 18, 2002. pp. 535-549.
41. L. Mui, M. Mohtashemi, and A. Halberstadt, "Notions of reputation in multi-agents systems: a review," in *Proceedings of the 1st International Joint Conference on Autonomous Agents and Multi-Agent Systems*, 2002, pp. 280-287.
42. J. Scott, "Social network analysis," *Sociology*, Vol. 22, 1988, pp. 109.
43. D. Morris, *The Naked Ape: A Zoologist's Study of the Human Animal*, Vintage Books, 1994.
44. R. Luhtanen and J. Crocker, "A collective self-esteem scale: Self-evaluation of one's social identity," *Personality and Social Psychology Bulletin*, Vol. 18, 1992, pp. 302-318.
45. J. Sabater and C. Sierra, "Social regret, a reputation model based on social relations," *ACM SIGecom Exchanges*, Vol. 3, 2001, pp. 44-56.
46. J. Sabater and C.-S. Regret, "A reputation model for gregarious societies," in *Proceedings of the 4th Workshop on Deception*, *Fraud and Trust in Agent Societies*, 2001, pp. 194-195.
47. D.-J. Bem, "Self-perception theory," *Advances in Experimental Social Psychology*, Vol. 6, 1972, pp. 1-62.
48. D. Cicchetti and D. Tucker, "Development and self-regulatory structures of the mind," *Development and Psychopathology*, Vol. 6, 1994, pp. 533-533.
49. H. S. Gross, "The analysis of the self: A systematic approach to the psychoanalytic treatment of narcissistic personality disorders," *The Journal of Nervous and Mental Disease*, Vol. 157, 1973, pp. 470.
50. E. R. Smith and D. M. Mackie, *Social Psychology*, Worth Publishers, NY, 2000.
51. W. A. Mason, F. R. Conrey, and E. R. Smith, "Situating social influence processes: Dynamic, multidirectional flows of influence within social networks," *Personality and Social Psychology Review*, Vol. 11, 2007, pp. 279-300.
52. S. M. Andersen and S. Chen, "The relational self: an interpersonal social-cognitive theory," *Psychological Review*, Vol. 109, 2002, pp. 619.
53. D. M. Tice and H. M. Wallace, "The reflected self: Creating yourself as (you think)

others see you," *Handbook of Self and Identity*, 2003, pp. 91.
54. S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed., 2010, Prentice Hall.
55. L. Bull, *Applications of Learning Classifier Systems*, Vol. 150, Springer Verlag, NY, 2004.
56. N. Gilbert and P. Terna, "How to build and use agent-based models in social science," *Mind & Society*, Vol. 1, 2000, pp. 57-72.
57. W. Poundstone, *Prisoner's Dilemma*, Doubleday, NY, 1993.
58. R. H. Frank, T. Gilovich, and D. T. Regan, "Does studying economics inhibit cooperation?" *The Journal of Economic Perspectives*, Vol. 7, 1993, pp. 159-171.
59. M. W. Macy and R. Willer, "From factors to actors: Computational sociology and agent-based modeling," *Annual Review of Sociology*, 2002, pp. 143-166.
60. M. A. Nowak and K. Sigmund, "Evolutionary dynamics of biological games," *Science's STKE*, Vol. 303, 2004, pp. 793.
61. R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *Science*, Vol. 211, 1981, pp. 1390-1396.
62. B. Grofman and J. Pool, "How to make cooperation the optimizing strategy in a two-person game," *Journal of Mathematical Sociology*, Vol. 5, 1977, pp. 173-186.
63. R. Axelrod, "The evolution of strategies in the iterated prisoner's dilemma," *The Dynamics of Norms*, 1997, pp. 1-16.
64. D. Kraines and V. Kraines, "Pavlov and the prisoner's dilemma," *Theory and Decision*, Vol. 26, 1989, pp. 47-79.
65. D. Kraines and V. Kraines, "Evolution of learning among Pavlov strategies in a competitive environment with noise," *Journal of Conflict Resolution*, Vol. 39, 1995, pp. 439-466.
66. M. Nowak and K. Sigmund, "A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game," *Nature*, Vol. 364, 1993, pp. 56-58.
67. K. Preston and M. J. B. Duff, *Modern Cellular Automata: Theory and Applications*, Springer, NY, 1984.
68. D. Watts and S. Strogatz, "Collective dynamics of small-world networks," *Nature*, Vol. 393, 1998, pp. 440-442.

Sheng-Wen Wang (王聖文) received his BS and MS degrees in Computer Science from the Tamkang and the National Tsing Hua University, Taiwan, in 1999 and 2001, respectively. He is currently a Ph.D. candidate in the Department of Computer Science, National Chiao Tung University. His current research interests include social simulations and multi-agent systems.

Chung-Yuan Huang (黃崇源) received his MS in Computer Information and Science (2000) and his Ph.D. in Computer Science (2005), both from the National Chiao Tung University, Taiwan. He is currently an Associate Professor in the Department of Computer Science and Information Engineering and a member of the Research Center for Emerging Viral Infections at Chang Gung University, Taiwan. His research interests include complex adaptive networks and systems, agent-based modeling and simulation for social science research, and computational epidemiology.

Chuen-Tsai Sun (孫春在) received his BS degree in Electrical Engineering (1979) and his MA in History (1984) from National Taiwan University. He earned his Ph.D. in Computer Science (1992) from the University of California at Berkeley. He joined the faculty at National Chiao Tung University in 1992. He is currently a joint Professor of Department of Computer Science and Graduate Institute of Education, National Chiao Tung University, Taiwan. His research interests include digital games, digital learning, social networking, and simulation modeling in the social sciences.