Collaboration among Competitive Agents in Information Sharing Networks [1]

Jisun Park and K. Suzanne Barber [2]

The Laboratory for Intelligent Processes and Systems
The University of Texas at Austin
Austin, TX, USA

Abstract

Agents often need information from other information providers or agents to build a situational picture for goal achievement. However, in competitive environments, such as an open system populated by self-interested agents, agents who can provide information may not want to provide information unless the information supply aids in their own goal achievement. This research assumes that agents can act as information providers and consumers at the same time, and proposes a method, by introducing *Competitive Collaborating Agents (CoCoAgents)*, for deciding the multi-dimensional strategies regarding requesting information and providing information. For the strategy as an information consumer, stochastic games are deployed so that an agent can explicitly take into account its own request strategy and other agents' strategy for providing information. The adaptive degree of collaboration constitutes an agent's strategy as an information provider. Following the best strategy for information providers and consumers lead the agents to construct collaborative relationships among competitive but complementary agents.

Key words: Information Sharing Networks, Partner Selection, Collaboration, Degree of Collaboration.

1 Introduction

Agents are goal-driven and goal achievement is often dependent on efficient acquisition of necessary information [6,10] since the information can be used

[1] This research is sponsored in part by the Defense Advanced Research Project Agency (DARPA) Taskable Agent Software Kit (TASK) program, F30602-00-2-0588. This research is also sponsored in part by a subcontract from The University of Texas Applied Research Labs Office of Naval Research Grant N00014-05-1-0857, entitled Improvised Explosive Device. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon.
[2] Email: `{jisun, barber}@lips.utexas.edu`

for building up the situational picture of the environment, or the acquisition of a certain information element in itself can be an agent's goal. When an agent does not have self-sufficient information acquisition capability, meaning that an agent cannot completely satisfy its own information requirements (i.e., a set of essential information elements required to achieve a goal) by itself, the agent needs to meet the information requirements by requesting deficient information from other agents or external information sources. However, when agents are self-interested, they are competitive since each agent is interested in achieving its own goal, and the requested agents may not want to waste their own resources by providing requested information to requesters unless the information supply contributes to the provider's own goal achievement in the future. In other words, agents need to be motivated to provide requested information in the way that the information supply is beneficial to those providers. The potential benefit information providers can expect is the reciprocal supply of deficient information from the requesting agents. A prerequisite condition for establishing this reciprocally beneficial relationship is information interdependence, meaning the agents at both ends are partially or completely complementary about the deficient information requirements and available information for others. However, even with the existence of information interdependence it is difficult for agents to figure out the best way to request information and respond to others' requests when there are multiple agents with multiple information elements which need to be provided or are available for supply. In addition, the information requirements can dynamically change over time depending on the ratio of the information requirements satisfaction from the previous interaction, or agents may need a completely different set of information to achieve different goals or sub-goals. Therefore, it is necessary for an agent to model the other agents' action selection about information exchange as well as to figure out the best actions based both on the model of other agents and its dynamic information requirements. In this paper, an approach for enabling agents to efficiently make multidimensional decisions about how to request information and how to respond to requests, based on past interaction history and expected rewards, is proposed by adopting a stochastic game theory. Also, the emergence of collaborative relationships leading to prevailing information exchange among competitive but complementary agents is empirically demonstrated with the proposed approach. The remainder of this paper is organized as follows. Section 2 provides a brief survey on related work. In section 3, competitive collaborating agents (*CoCoAgents*) are introduced with an overview of stochastic games and the explanation about each dimension of strategy for the agents in detail. Section 4 shows experimental setup and results to demonstrate the emergence of collaboration in information sharing networks. Section 5 summarizes this paper and suggests future work.

2 Related Work

Modeling other agents' behavior has been an important research topic in Multi-Agent Systems because the model of other agents can provide the clues for the most appropriate actions for an agent. There has been a significant amount of literature related to the modeling of other agents, with trust research (e.g., [2,3]) representing the most relevant area. Research addressing trust in an agent society provides various perspectives for evaluating other agents as well as the methods for identifying the most potentially trustworthy partners [11]. However, the trust research is mostly scoped within modeling and evaluating the trustworthiness of the agents, and research about how the trust models can affect goal achievement in a dynamic environment has not been fully explored. This paper demonstrates how agents can proactively pursue their goals using an estimated model of other agents in a dynamic environment as well as how agents can establish collaborative relationships.

Strategies for building collaborative relationships with other agents have also been investigated as a key capability of agents in Multi-Agent Systems. Maximilien and Singh [9] proposed a method for selecting appropriate service providers based on Quality of Service (QoS). They provide thorough metrics for evaluating service providers from the consumer's point of view. However, the selection is assumed to be unidirectional and does not take into account the dynamics of service requirements. The goal or required services can change over time and in multi-agent systems each agent often can be a service provider and a service consumer at the same time. In order to overcome this limitation, we propose a method for deciding multidimensional strategies about selecting appropriate information (or service) providers for each required information element, as well as an appropriate way of responding to other agents' requests considering the expected rewards for goal achievement. Sichman et al. [18] proposed a social reasoning mechanism using dependence networks. Dependence networks are used for defining a taxonomy of dependence situation and can be used for making up the capabilities of agents. A data structure called *external description*, which is composed of goals, actions, resources, plans, is used to store the information about others, and the *external description* is used to build the dependence networks of each agent. While this approach provides a useful and descriptive mechanism for building collaborative relationships with others, the adaptation to dynamics of goal achievement (e.g. goal change, partial achievement) is not explicitly taken care of. Sen et al. [12,13,14] proposed a decision mechanism for constructing collaborative relationships among self-interested agents based on expertise reciprocity. The decision mechanism aids in agents' decisions about whether to accept other agents' help requests for tasks or not. The decision is based on cost metric from past interaction and expected future savings by reciprocal collaborative relationships. This work differs from our approach in that they assume system-wide goals and focus on giving information provider's collaboration strategies, while this paper provides strategies for both providers and requesters. Also, this paper uses a goal-oriented metric for designing the reward structure, while they consider reduction of cost.

3 CoCoAgents: Competitive Collaborating Agents

The Competitive Collaborating Agents (CoCoAgents) are self-interested, but need to exchange information with appropriate counterparts. *CoCoAgents*' goals impose a set of information requirements which need to be satisfied along with a set of information available for other agents. Therefore, *CoCoAgents* can act as an information consumer and as a provider at the same time, and two dimensional decisions have to be made.

- Strategy for requesting information: an agent's strategy for requesting determines how effectively the agent receives information elements from other information providers (which is to say, other agents). The acquisition of information elements in turn contributes to an agent's overall rewards.
- Strategy for responding to other agents' requests: an agent's strategy for responding to information requests from other agents determines how many requests the agent responds to. The reciprocal sharing of information between agents may ultimately allow a given agent to better acquire needed information elements, which in turn contributes to an agent's overall rewards.

Each dimension of strategy decisions is not independent of the other dimension. An agent's strategy about how to provide information may affect other agents' strategy for how to request information as well as how to respond to the agent's requests, which can eventually affect the agent's strategy for how to request information. If the decisions about requests and responses are made simultaneously, the space size for strategies can easily become intractable, and an unrealistic assumption about system-wide synchronous message delivery of requests and replies needs to be made for simultaneous decisions. Since each dimension of the strategy contributes to the utility (rewards) in different ways, separating each dimension of strategies and determining how the dimensions are related aids in agents' decision process with regard to designing rewards and reducing search space. In the following subsections, an overview of stochastic game theory and its adoption for information sharing networks are provided, followed by details about each dimension of strategy and the relationship between them.

3.1 Overview of Stochastic Games

A stochastic game [15,5,8] is a tuple $(S, A_1, \ldots, A_n, r_1, \ldots, r_n, p)$, where S is the state space, A_i is the action space for player i, n is the number of players, $r_i : S \times A_1 \times \ldots \times A_n \mapsto R$ is the reward function of player i, and

$p : S \times A_1 \times \ldots \times A_n \mapsto \Delta(S)$ is the transition probability where $\Delta(S)$ represents the set of probability distribution over S and p satisfies the condition.

$$(1) \qquad \sum_{s' \in S} p(s'|s, a_1, \ldots, a_n) = 1, a_i \in A_i$$

At a state $s \in S$, each agent selects actions $a_1, \ldots, a_n$, and receives rewards $r_i(s, a_1, \ldots, a_n)$, and makes a transition to $s' \in S$ based on the transition probability. The objective of each agent is to find an action strategy which maximizes the discounted sum of rewards represented by v_i (Equation 2), where π_i is the action strategy of player i, $\beta \in [0, 1)$ is a discount factor, r_i^t is a reward for player i at time t, and s_0 is an initial state.

$$(2) \qquad v_i(s, \pi_1, \ldots, \pi_n) = \sum_{t=0}^{\infty} \beta^t E(r_i^t | \pi_1, \ldots, \pi_n, s_0 = s)$$

In information sharing networks, each state is represented by the information requirements, and the rewards are defined in terms of the coverage of information requirements [4] attained by taking an action for information request and corresponding reception of the necessary information. Each agent's action space consists of two separate spaces: actions for information request as an information consumer and actions for information supply as an information provider. Accordingly, the action space for each agent can be represented by $A_i = A_i^{prov} \times A_i^{cons}$, where A_i^{prov} is the action space consisting of the available actions as an information provider and A_i^{cons} is the action space with available actions as an information consumer. The joint action space $A = A_1^{prov} \times A_1^{cons} \times \ldots \times A_n^{prov} \times A_n^{cons}$ can be decomposed into 2 terms based on agents' roles. From information consumer's perspective, the actions to be taken into account by agent i for calculating rewards are its actions for requesting information and the corresponding actions of other agents as providers (Equation 3). The action space as an information provider (Equation 4) consists of agent's actions for responding to other agents' requests and corresponding request actions by other agents.

$$(3) \qquad A_i^{cons} \times \prod_{j=1, j \neq i}^{n} A_j^{prov}$$

$$(4) \qquad A_i^{prov} \times \prod_{j=1, j \neq i}^{n} A_j^{cons}$$

The rationale for decomposing the action space into multiple dimensions based on roles is that each dimension contributes to the expected reward in different ways. However, since they are not completely independent, the effect of one action dimension on the other also needs to be considered. The following subsections present detail about modeling action spaces and strategy selection for an information consumer and an information provider.

3.2 Strategy for Information Consumers

An agent i has a set of information requirements $RQ(i) = \{info_k\}$ and a set of available information $PR(i) = \{info_p\}$ for other agents and itself. A state of an agent is represented by the information requirements. The action space at a given state is dependent on the current information requirements and the available information providers. The reward from actions at a given state is calculated by the coverage measure which is the percentage of information requirements satisfied by the actions. An available action for an information consumer is a set of requests to a subset of available information providers. When interacting agents commit actions, the information requirements for an information requester can therefore change, and the requester's state transits to the next state. In the next state, due to the change in information requirements, available action sets are different from the previous state unless the agent stays in the same state as the previous one. Therefore, at each state, agents play different games, which make the deployment of a stochastic game the most appropriate for information sharing networks. The strategy for this stochastic game is stationary and in a stochastic game with stationary strategies, there exists at least one Nash Equilibrium point [7] where the strategies in Nash Equilibrium is defined as a tuple of strategies $(\pi_1^*, \ldots, \pi_n^*)$ such that for all $s \in S$ and $i = 1, \ldots, n$,

$$v_i(s, \pi_1^*, \ldots, \pi_n^*) \geq v_i(s, \pi_1^*, \ldots, \pi_{i-1}^*, \pi_i, \pi_{i+1}^*, \ldots, \pi_n^*), \quad \forall \pi_i \tag{5}$$

This paper experimentally shows how agents reach the equilibrium by finding the best response given the action spaces, reward structure, and observation about other agents.

An action $a_i^{cons} = \{REQUEST_i(j, info_k)\}$ denotes that agent i requests information element k from agent j. The requests have to satisfy the constraints that all the required information elements are requested and each element can be requested from a single information provider. The strategy of agent i for requesting information, π-in$_i$, is a path of actions which maximize the discounted sum of rewards (in Equation 6) when the other agents have the strategy, π-out, for providing information. π-out$_{-i}$ in Equation 6 represents strategies of agents but agent i for providing information.

$$v_i(s, \pi\text{-in}_i, \pi\text{-out}_{-i}) = \sum_{t=0}^{\infty} \beta^t E(r_i^t | \pi\text{-in}_i, \pi\text{-out}_{-i}, s_0 = s) \tag{6}$$

From the information consuming agents' point of view, the information providers' actions for their own requests are recognized as being comprised of either to provide $(PROVIDE_j(info_k, i))$ or not to provide the requested information $(\neg PROVIDE_j(info_k, i))$. Since it is assumed to be impossible to know the exact model of other agents' actions and the selection of the best-response in a competitive environment, an agent maintains an estimated probabilistic model of other agents' actions by observation. *The reception rate* $RR(i, j)$ $(\in [0, 1])$ is the probability of j's replying to i's requests and can be

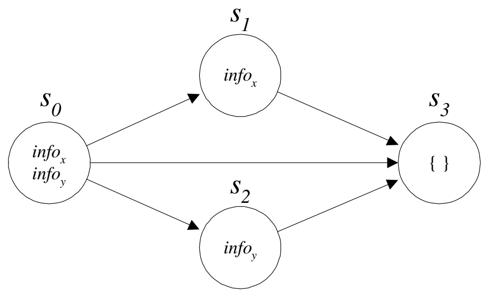

Fig. 1. State transition for agent 0 with $RQ(0) = \{info_x, info_y\}$

calculated by counting the number of requests replied by agent j. The set of actions by both providers and a consumer lead a static state transition to a certain state with a probability of 1.0, but since the probability of actions being committed by providers can vary depending on the providers' strategy, rewards are scaled by $RR(i, j)$. The opponent modeling Q-learning method [16] takes a similar approach for building a model of other agents' action and calculating the expected rewards. [16] constructs a table of probabilities that other agents take actions at a given state and the expected rewards are scaled by the probability so that an agent can calculate the best action.

Figure 1 shows a simple example of agent 0's state transition graph with $RQ(0) = \{info_x, info_y\}$, meaning agent 0 requires $info_x$ and $info_y$. Supposing that there are two other agents - agent 1, agent 2 - which can provide the required information, $PR(1) = \{info_x, info_y\}$, $PR(2) = \{info_x, info_y\}$, the available actions for agent 0 ($a^{cons}_{i,j}$, where j is an index of an action) at s_0 as an information consumer are

$$
\begin{aligned}
a^{cons}_{0,0} &= \{REQUEST_0(1, info_x), REQUEST_0(2, info_y)\} \\
a^{cons}_{0,1} &= \{REQUEST_0(2, info_x), REQUEST_0(1, info_y)\} \\
a^{cons}_{0,2} &= \{REQUEST_0(1, info_x), REQUEST_0(1, info_y)\} \\
a^{cons}_{0,3} &= \{REQUEST_0(2, info_x), REQUEST_0(2, info_y)\}
\end{aligned}
$$

For $a^{cons}_{0,0}$ at state s_0, the other agents' actions $\{PROVIDE_1(info_x, 0)$, $PROVIDE_2(info_y, 0)\}$ lead agent 0 to state s_3. The reward for the action set is 1.0 because all the requirements are satisfied. The probability of the actions $\{PROVIDE_1(info_x, 0)$, $PROVIDE_2(info_y, 0)\}$ being taken is $RR(0, 1) \times RR(0, 2)$, therefore, the expected reward is scaled to $(1.0 \times RR(0, 1) \times RR(0, 2))$. In another situation, with the same action set for agent 0 at state s_0, the other agents' actions can lead agent 0 to state s_2 ($\{PROVIDE_1(info_x, 0)$, $\neg PROVIDE_2(info_y, 0)\}$). In this case, the reward of the actions is 0.5 because 50% of the information requirements can be satisfied. The reward is scaled by $RR(0, 1) \times (1 - RR(0, 2))$ so the expected reward becomes $(0.5 \times RR(0, 1) \times (1 - RR(0, 2)))$. In the same way, Table 1 lists action space for agents from the initial state and corresponding rewards for

Table 1
Actions and Expected Rewards for agent 0 at s_0

A_0^{cons}	A_1^{prov}	A_2^{prov}	Expected Reward
$(1, info_x)\ (2, info_y)$	$(info_x, 0)$	$(info_y, 0)$	$1.0 \cdot RR(0,1)RR(0,2)$
	$(info_x, 0)$	$\neg(info_y, 0)$	$0.5 \cdot RR(0,1)(1 - RR(0,2))$
	$\neg(info_x, 0)$	$(info_y, 0)$	$0.5 \cdot (1 - RR(0,1))RR(0,2)$
	$\neg(info_x, 0)$	$\neg(info_y, 0)$	0
$(2, info_x)\ (1, info_y)$	$(info_y, 0)$	$(info_x, 0)$	$1.0 \cdot RR(0,1)RR(0,2)$
	$\neg(info_y, 0)$	$(info_x, 0)$	$0.5 \cdot (1 - RR(0,1))RR(0,2)$
	$(info_y, 0)$	$\neg(info_x, 0)$	$0.5 \cdot RR(0,1)(1 - RR(0,2))$
	$\neg(info_y, 0)$	$\neg(info_x, 0)$	0
$(1, info_x)\ (1, info_y)$	$(info_x, 0)\ (info_y, 0)$	.	$1.0 \cdot RR(0,1)RR(0,1)$
	$\neg(info_x, 0)\ (info_y, 0)$	.	$0.5 \cdot (1 - RR(0,1))RR(0,1)$
	$(info_x, 0)\ \neg(info_y, 0)$	.	$0.5 \cdot RR(0,1)(1 - RR(0,1))$
	$\neg(info_x, 0)\ \neg(info_y, 0)$	.	0
$(2, info_x)\ (2, info_y)$	.	$(info_x, 0)\ (info_y, 0)$	$1.0 \cdot RR(0,2)RR(0,2)$
	.	$(info_x, 0)\ (info_y, 0)$	$0.5 \cdot (1 - RR(0,2))RR(0,1)$
	.	$(info_x, 0)\ (info_y, 0)$	$0.5 \cdot RR(0,2)(1 - RR(0,2))$
	.	$(info_x, 0)\ (info_y, 0)$	0

agent 0. In the table, actions for agent 0 are denoted by $(k, info)$ which means that agent 0 requests $info$ from agent k. The actions for providers (agent 1 and 2) are denoted by either $(info, j)$ or $\neg(info, j)$ which means the agent provides $info$ to agent j or does not provide info to agent j respectively. The expected rewards by providing information are not awarded directly but indirectly expected through the potential increase in the probability of receiving necessary information.

3.3 Strategy for Information Providers

Although the strategy for the information provider denoted by π-out indirectly affects the expected rewards by potentially affecting the other agents' responses thus the expected rewards, the direct objective of the information provider's strategy is to adjust the degree of collaboration with other agents so that collaborative relationships can be built among the complementary agents. Adapting the strategy as an information provider can help *CoCoAgents* avoid

Algorithm 1. Adaptive DoC
$if \Delta RR(i,j) \geq 0 \wedge DoC(i,j) \geq 0.5$
$\quad DoC(i,j) \leftarrow DoC(i,j) + \alpha(MAX\_DOC - DoC(i,j))$
$if \Delta RR(i,j) \geq 0 \wedge DoC(i,j) < 0.5$
$\quad DoC(i,j) \leftarrow (1+\alpha)DoC(i,j)$
$if \Delta RR(i,j) < 0 \wedge DoC(i,j) \geq 0.5$
$\quad DoC(i,j) \leftarrow DoC(i,j) - \alpha(MAX\_DOC - DoC(i,j))$
$if \Delta RR(i,j) < 0 \wedge DoC(i,j) < 0.5$
$\quad DoC(i,j) \leftarrow (1-\alpha)DoC(i,j)$

both being exploited by other agents and being isolated from other agents when the agents are not self-sufficient in their information acquisition capability. The action space for information providers is defined by the *degree of collaboration (DoC)* which carries the agent's amount of intention for collaboration with each agent. Agent i's degree of collaboration with agent j is denoted by $DoC(i,j)$ and it is modeled as a probability distribution of responses to requests from other agents. Therefore, agents as information providers stochastically make decisions about whether to reply or not, depending on the degree of collaboration. The degree of collaboration is triggered by the change of the reception rate $RR(i,j)$ and the previous degree of collaboration as in Algorithm 1.

MAX_DOC is the maximum degree of collaboration, 1.0 and MIN_DOC is the minimum degree of collaboration, 0.0. A constant α determines the amount of increase or decrease (i.e., higher α leads to a greater change for $DoC(i,j)$). When the reception rate $RR(i,j)$ increases from the previous reception rate, the degree of collaboration $DoC(i,j)$ increases. The amount of increase or decrease depends on the previous degree of collaboration. For example, as the previous degree of collaboration becomes closer to 0.0 or 1.0, the degree of collaboration is less affected by the change of the reception rate, while the amount of change becomes relatively high as the previous degree of collaboration is closer to 0.5. These rules leverage the construction of collaborative relations or non-collaborative relationships by accelerating the change of DoC near unstable relationships (i.e., DoC is near 0.5) and decelerating the change of DoC in relatively stable relationships (i.e., DoC is near 1.0 or 0.0).

Updating the degree of collaboration only when the reception rate is updated can make an agent vulnerable to other agents' exploitation when the agent does not make requests to those exploiting agents. Therefore, the degree of collaboration ($DoC(i,j)$) decays over time when an agent i does not request information from agent j as follows.

$$DoC(i,j) \leftarrow \beta DoC(i,j), \quad 0 \ll \beta < 1 \tag{7}$$

In order to minimize the effect of temporal decay and let the degree of collaboration more dependent on the direct interaction, the decay rate β is assumed to be close to 1.0.

3.4 Emergence of Collaboration

In game theory, the prisoner's dilemma (PD) is a general-sum game where two players try to receive the higher rewards by choosing an action between $cooperation$ and $defect$. In the non-iterated PD, $defect$ by both players is a Nash Equilibrium point though it is not Pareto-Optimal solution. In the iterated prisoner's dilemma (IPD), Tit-for-Tat [1] has been the most effective strategy to achieve the maximum rewards. Tit-for-Tat is to cooperate until the opponent defects. When the opponent defects, the player punishes the opponent by selecting $defect$. Tit-for-Tat strategy is cooperative approach and the success of the strategy has provided the foundation for the explanation of the emergence of cooperation in human society, in a group of animal, or in politics [1].

The similarity between Tit-for-Tat and the adaptive degree of collaboration proposed in this paper can be found by investigating a matrix game which can be observed in a state between two agents. Once we assume that two agents are partially or completely complementary and both agents request information from each other, the reward matrix for agent 1 (player 1) can be modeled as in Table 2. Note that since the requirements and requests change depending on the previous actions, the agents play different games at each state. However,the reward structure observed in each game provides the intuition on how the collboration can emerge.

As can be seen in Table 2, the game is basically a Prisoner's Dilemma when considering the four extreme cases ($DoC = 1$ and $DoC = 0$ for each agent). ΔE is the expected gain represented by the product of attained coverage and the probability of achiving the coverage ($\Delta coverage \prod_{-i} Pr(A_{-i}^{prov})$). $cost_c$ is the cost of requesting and receiving information, and $cost_p$ is the cost for providing requested information. The difference is that the action space is continuous in this game. The adaptation of the degree of collaboration in the proposed method is performed in a similar way to what Tit-for-Tat does. When the opponent has been a collaborative provider (increasing RR), the degree of collaboration is kept higher while the degree of collaboration for the opponent decreases if the opponent tends to be a non-cooperative provider. The emergence of collaboration such that the total rewards are maximized is shown in the experiments.

4 Experiments

Experiments were performed to show how self-interested, competitive agents adopt their strategies and how collaboration emerges between complementary agents in information sharing networks. In order to find the best response strategies, agents use the extended policy iteration method for a stochastic

Table 2
Player 1's reward

	$DoC = 1$	...	$DoC = 0$
$DoC = 1$	$\frac{\Delta E}{cost_c + cost_p}$	...	$\frac{\Delta E}{cost_c + cost_p} - \frac{\Delta E}{cost_c}$
...	...	...	...
$DoC = 0$	$\frac{\Delta E}{cost_c}$	...	0

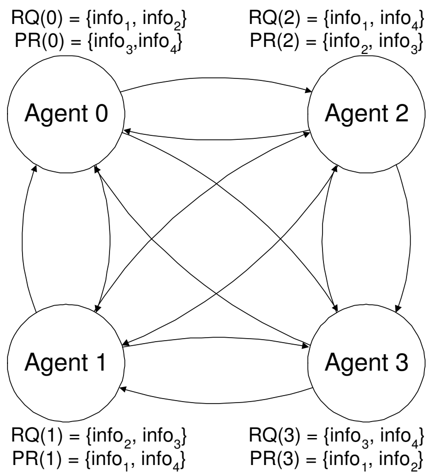

Fig. 2. Experimental Setup

game proposed by Pollatschek and Avi-Itzhak [17].

4.1 Experimental Setup

In the experiment, 4 agents (agent 0, agent 1, agent 2, and agent 3) are deployed with pre-assigned information requirements and available information elements for others (Figure 2). It is assumed that each agent knows which agents are potential information providers for its information requirements. The communication links are bidirectional and the agents are fully connected. Message delivery is not synchronized with simulation timesteps so agents do not necessarily request or reply at every timestep. When the initial information requirements are satisfied, the agent reaches a state with an empty set of requirements. The information requirements are then reset to be the initial information requirements.

The requirements RQs and available information PRs in Figure 2 were designed so that an agent is completely complementary with another agent while it is still partially complementary with the remaining agents. For example, agent 0 is completely complementary with agent 3 while partially complementary with agents 1 and 2.

Two sets of experiments were conducted to show how *CoCoAgents* per-

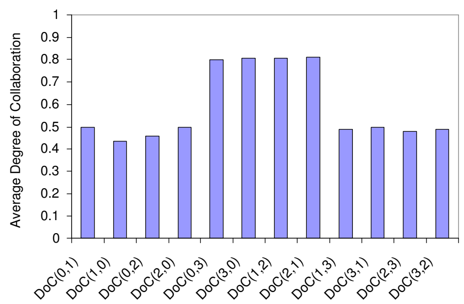

Fig. 3. Average Degree of Collaboration from Exp1

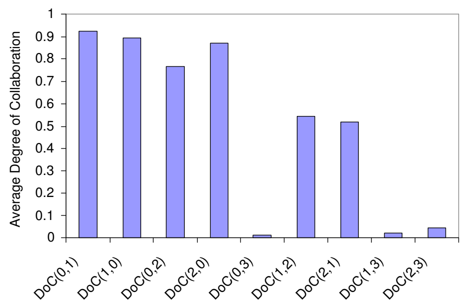

Fig. 4. Average Degree of Collaboration from Exp2

form in different situations. In the first experiment (Exp1), all agents are *CoCoAgents*, meaning that all agents are able to select the best response to other agents' strategies using the proposed approach. In the second experiment (Exp2), three agents are *CoCoAgents*, but agent 3 is a completely selfish agent which does not respond to other agents' request. Each experiment was conducted for 5 runs of 3000 timesteps, and α was set to 0.25 and β was set to 0.99. Initial *DoC* for each experiment was 0.5.

4.2 Experimental Results

In the first set of experiments (Exp1), all agents are *Competitive Collaborating Agents (CoCoAgents)*. Figure 3 shows the average degree of collaboration between agents. Initially, $DoC(i, j)$ is set to 0.5 and is updated whenever the reception rate $RR(i, j)$ is updated and decays over time when agent i does not send requests to agent j. From the figure, agent 0 has the highest degree of collaboration with agent 3 ($DoC(0, 3)$) while it maintains values near 0.5 for agent 1 and 2 ($DoC(0, 1)$, $DoC(0, 2)$), which is because agent 0 does not request information from agent 1 and 2 once it starts to request from agent 3. Agent 3 also yields a higher average degree of collaboration with agent 0 ($DoC(3, 0)$) than with other agents ($DoC(3, 1)$, $DoC(3, 2)$). Similarly, agent 1 and agent 2 show a higher average degree of collaboration with each other ($DoC(1, 2)$, $DoC(2, 1)$) than with other agents.

In the second set of experiments (Exp2), agent 0, 1, 2 are *CoCoAgents*, but agent 3 is a selfish agent which requests necessary information from all available information providers but does not respond to other agents' requests. Figure 4 plots the average degree of collaboration between agents in Exp2. In this figure, the average degrees of collaboration with agent 3 ($DoC(0, 3)$, $DoC(1, 3)$, $DoC(2, 3)$) are low resulting in isolation of the selfish agent from other agents.

One observation to note from Figure 4 is that the reciprocal degrees of collaboration between agents 1 and agent 0 ($DoC(1, 0)$, $DoC(0, 1)$) are higher than the degrees of collaboration between agent 1 and 2 ($DoC(1, 2)$, $DoC(2, 1)$), despite the completely complementary relationship between agent 1 and agent

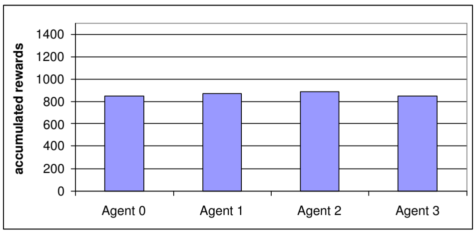

Fig. 5. Sum of Rewards (Exp1)

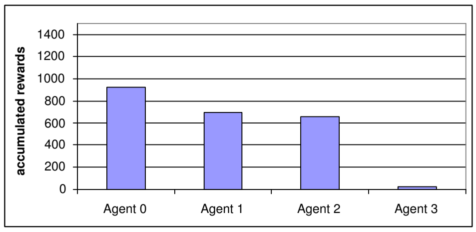

Fig. 6. Sum of Rewards (Exp2)

2. Also, the reciprocal degrees of collaboration between agents 2 and agent 0 ($DoC(2,0)$, $DoC(0,2)$) are higher than the degrees of collaboration between agent 1 and 2 ($DoC(1,2)$, $DoC(2,1)$). In the initial stage of interaction, agent 0 may request $info_1$ either from agent 1 or agent 3. However, since agent 3 never returns requested information, agent 0's requests for $info_1$ are likely to be directed mostly toward agent 1. In the same way, agent 0's requests for $info_2$ are likely to be directed mostly toward agent 2. Since agent 0's only information provider for $info_1$ and $info_2$ become agent 1 and agent 2 respectively while agent 1 and agent 2 still have 2 information providers per information (agent 0 and 2, agent 0 and 1 respectively) among which each agent has to choose a single provider per information, it is likely that agent 0 can have a relatively higher chance of receiving necessary information resulting in the increase of the degrees of collaboration with agent 1 and agent 2, which in turn derives the higher chance of building bilateral collaborative relationships with agent 1 and agent 2. Therefore, with the existence of asymmetric information requirements and available information providers as in this case, the highest collaborative relationships do not necessarily occur between completely complementary agents, but can be constructed between partially complementary agents.

Figure 5 plots the average sum of rewards from the first experiment (Exp1). In the experiment, each agent receives a relatively even amount of rewards. Experiment 2 (Figure 6) shows that agent 0 receives the highest rewards since agent 0 establishes a high degree of collaboration with agent 1 and 2 while agents 1 and 2 also establish high degree of collaboration. Agent 3 cannot accumulate enough rewards since agent 3 is isolated from the other agents.

5 Conclusion

Agents are goal-driven and goal achievement is often dependent on acquiring required information. When agents do not possess a self-sufficient capability for information acquisition, the agents need to request the necessary information from external information sources (e.g., other agents). Among competitive agents, information exchange may not happen unless the information exchange aids in the providers' accomplishment of a goal. In order to motivate the information exchange, the existence of reciprocally beneficial relationships between agents is required, and the agents need to select the most appropriate actions both for requesting information and providing information under these conditions. Because of environmental dynamics and the action selection strategy of other interacting agents, the available action sets and the current information requirements can change dynamically. In order to accommodate this dynamism, a method for decomposing multidimensional strategies is proposed in information sharing networks.

From an information consuming agent's point of view, the rewards depend both on its own strategy for requesting information and the information providers' strategy for providing information. Therefore, when an agent is in an information consumer mode, the agent needs to model the other agents' actions for providing information, and explore the search space established using a stochastic game. A stochastic game provides the most feasible model for information sharing networks because it takes into account the actions not only for a given agent but also for other agents which affect the agent's payoffs.

The degree of collaboration is used to create an action space of information providers. An agent maintains the degree of collaboration with respect to other agents to adjust the probability of responding to other agent's requests. The reciprocal sharing of information between agents may ultimately allow a given agent to better acquire needed information elements, which in turn contributes to an agent's overall rewards.

Experimental results show that the agents equipped with the proposed method converge to collaborative relationships. Also, it is shown that selfish agents are isolated from the society by reducing the degree of collaboration and the number of requests to the selfish agents. As an ongoing effort for extending the current results, the extension of the proposed strategy for information providers will be investigated. Also, the reward structure will be extended to accommodate the quality of information in addition to the information requirement coverage to examine various aspects of impact to goals by information acquisition. Comparing the effects of various reinforcement learning algorithms for information consumers' strategy can also be an interesting issue for future work.

References

[1] Axelrod, R., *The Evolution of Cooperation*, New York, Basic Books, 1985.

[2] Ashri, R., Ramchurn, S., Sabater, J., Luck, M., Jennings, N., *Trust Evaluation Through Relationship Analysis*, In proceedings of International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 05), pp. 1005-1011, The Netherlands, 2005.

[3] Barber, K. Suzanne., Kim, Joonoo., *Soft Security: Isolating Unreliable Agents from Society. Trust, Reputation, and Security*, LNCS, p. 224-233, 2002.

[4] Barber, K. S., Park, J., *Agent Belief Autonomy in Open Multi-Agent Systems*, Agents and Computational Autonomy, LNCS, Springer-Verlag p. 7- 16, 2004.

[5] Bowling, Michael., Veloso, Manuela., *An Analysis of Stochastic Game Theory for Multiagent Reinforcement Learning*, Technical Report, CMU-CS-00-165, Carnegie Mellon University, 2000.

[6] Castelfranchi, Cristiano., *Guarantees for Autonomy in Cognitive Agent Architecture*, In proceedings of ECAI Workshop on Agent Theories, Architectures, and Languages, pp. 56-70, 1994.

[7] Fink, A. M., *Equilibrium in a stochastic n-person game*, Journal of Science in Hiroshima University, Series A-I, 28:89-93, 1964.

[8] Hu, Junling., Wellman, Michael., *Nash Q-Learning for General-Sum Stochastic Games*, Journal of Machine Learning Research, vol 4, pp. 1039-1069, 2003.

[9] Maximilien, E. M., Singh, M. P., *Agent-Based Trust Model Involving Multiple Qualities*, In proceedings of International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 05), pp. 519-526, The Netherlands, 2005.

[10] Park, Jisun., Barber, K. Suzanne., *Information Quality Assurance by Lazy Exploration of Information Sources Combination Space in Open Multi-Agent Systems*, Journal of Universal Computer Science, vol. 11, p. 193-209, Feb, 2005.

[11] Rasmusson, L., Jansson, S., *Simulated Social Control for Secure Internet Commerce*, In proceedings of the New Security Paradigms '96, ACM Press, p. 18-26, 1996.

[12] Saha, Sabyasachi., *Reciprocity based Cooperation in a limited resource environment*, In proceedings of International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 04), pp. 1402-1403, Melbourne, Austiralia, 2004.

[13] Saha, Sabyasachi., Sen, Sandip., Dutta, Partha Sarathi. *Helping Based on Future Expectations*, In proceedings of International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 03), pp. 289-295, Melbourne, Australia, 2003.

[14] Sen, Sandip., Dutta, Partha Sarathi., Saha, Sabyassachi., *Emergence and Stability of Collaborations Among Rational Agents*, Cooperative Information Agents VII, LNAI 2782, Springer-Verlag, pp. 192-205, 2003.

[15] Shapley, L. S.,*Stochastic Games*, In proceedings of National Academy of Science USA, 39, pp. 1095-1100, 1953.

[16] Uther, William., Veloso, Manuela., *Adversarial Reinforcement Learning*, Technical Report, CMU-CS-03-107, Carnegie Mellon University, Jan, 2003.

[17] Vrieze, O. J., *Stochastic Games with Finite Sate and Action Spaces*, Number 33. CWI Tracts, 1987.

[18] Sichman, J. S., Conte, R., Demazeau, Y., Castelfranchi, C., *A Social Reasoning Mechanism Based on Dependence Networks*, In Proceedings of the Eleventh European Conference on Artificial Intelligence (ECAI 94), A. Cohn. pp. 188-192, John Wiley & Sons, Ltd. 1994.