Journal of International Technology and Information Management

Volume 28 | Issue 3 | Article 3

2019

Assessment of Adaptability of a Supply Chain Trading Agent's Strategy: Evolutionary Game Theory Approach

Yoon Sang Lee
Columbus State University, yslee211@gmail.com

Riyaz T. Sikora
University of Texas at Arlington, rsikora@uta.edu

Assessment of Adaptability of a Supply Chain Trading Agent's Strategy: Evolutionary Game Theory Approach

Yoon Sang Lee
(*Columbus State University*)

Riyaz T. Sikora
(*University of Texas at Arlington*)

ABSTRACT

With the increase in the complexity of supply chain management, the use of intelligent agents for automated trading has gained popularity (Collins, Arunachalam, B, et al. 2006). The performance of supply-chain agents depends on not just the market environment (supply and demand patterns) but also on what types of other agents they are competing with. For designers of such agents it is important to ascertain that their agents are robust and can adapt to changing market and competitive environments. However, to date there has not been any work done that assesses the adaptability of a trading agent's strategy in the presence of various demand and supply distributions when competing with a changing composition of agents using different strategies.

In this paper we use the concept of replicator dynamics to study the evolution of a population of strategies used by supply chain agents when the different agents are competing against each other. We also study the evolution of the population of agents' strategies in the presence of six types of adverse market conditions. In particular we test three strategies that have been presented in the literature and our results indicate that over time supply chain agents gravitate towards using the SCMaster strategy in most scenarios.

Keywords: word; multi-agent systems, replicator dynamics, reinforcement learning, artificial intelligence, e-commerce, simulation, supply chain management, evolutionary game theory

INTRODUCTION

Introduction of the Research

According to the report from Acquity Group (2014), more than 68 percent of buyers purchase goods online, and the number of respondents who spent 90 percent or more of their budgets online has doubled from 2013. As seen in the report, IT already became a force and key driver of globalization (Ejiaku 2014), and enterprises come to have a variety of options on their supply chain. However, the management of a supply chain has been a challenging problem since it is recognized as a complex adaptive system requiring dynamic and flexible reactions against the changes of environment (Choi, Dooley, and Rungtusanatham 2001). You et al. (2017) also indicated that the dynamics of supplier selection is important to improve the performance of a supply chain. To deal with these complexities in supply chain management, an intelligent trading agent can be a good alternative for the automation of supply chain management (Collins et al. 2006).

The supply chain trading agent automatically operates all aspects of supply chain management such as inventory management, procurement, bidding to auction for sales, and the scheduling for shipment and production. Since the purpose of supply chain trading agent is to maximize the profit, the performance of a supply chain trading agent is measured by its profit. One important requirement of trading agents for achieving higher profit is the adaptability of their strategy to deal with the environmental changes. For example, if the actual customer demand is lower than a trading agent's demand prediction and the agent uses a fixed strategy, then the supply chain agent would end up with an excessive inventory regardless of the actual demand amount. However, if the agent has an adaptive strategy, it would update the inventory level threshold based on the actual demand level.

In a real-world supply chain, agents using various strategies compete against each other for the resources, such as customer orders and the supply of raw materials, to increase their profits. However, testing the performance of a supply chain trading agent in real-world situations is difficult due to the lack of control of market environment (e.g., demand and supply level and competitors in the same market) and the difficulties in accessing agents' performance (profit). For such limitations, Saar-Tsechansky (2015) suggested simulation as a good alternative for this category of studies to evaluate the performance of a model. Since a supply chain is composed of multiple actors and their interactions, agent-based modeling and simulation (ABMS) can be considered a viable modeling method for a supply chain. ABMS is rooted in complex adaptive system (CAS), so it provides a modeling method for agents' dynamic interactions simulated repeatedly over time. Due to this advantage for modeling of CAS, it has gained attention in various areas such as sociology, biology, operations research, and economics (Macal and North 2009). One of the problems in managing a supply chain is the complex and stochastic demand patterns causing time-varying state. For these problems, mathematical or optimization techniques are not able to provide solutions, but artificial intelligence algorithms could provide learning ability to self-adapting agent (Mortazavi et al. 2015). Therefore, some studies on supply chain trading agent focus on the application of the adaptive learning techniques on the basis of ABMS while others use optimization-based solutions.

The performance of supply-chain agents depends on not just the market environment (supply and demand patterns) but also on what types of other agents they are competing with. Most of the studies in the literature evaluate the performance of such agents in static market conditions (i.e., supply and demand patterns) and competition environments (i.e., the composition of the competing agents does not change). In real world situations the composition of competition is dynamic; weaker agents do not survive whereas stronger agents become more established. Moreover, the performance of the agents itself depends on such composition (stronger agents do better when competing against weaker agents than when competing against other stronger agents). It is this constant feedback loop (where performance depends on composition of agents, but the composition of agents in turn depends on their performance) that we have tried to simulate in our experiments. For designers of such agents it is important to ascertain that their agents are robust and can adapt to changing market and competitive environments.

In this study we assess the adaptability of our trading agent SCMaster (Lee and Sikora (2016)) by observing it's growth in a population composed of supply chain trading agents drawn from the literature. Through this study we not only evaluate the adaptability of our agent's strategy, but also propose a simulation framework to study the survivability of a trading agent in a complex adaptive system (in a trading market). In addition, we expect that this approach could be used as a viable method in trading agent industry for the prediction of the success and evaluation of trading agents.

For our simulation, we adopted and modified the simulation framework of a supply chain trading agent competition, Trading Agent Competition - Supply Chain Management (TAC-SCM, Collins et al. 2006). In TAC-SCM literature, various studies provided models of automated supply chain management agents. Among the supply chain trading agent models, we select the main strategies of two best-performed models, DeepMaize (Kiekintveld et al. 2006) and TacTex (Pardoe and Stone, 2009), and implemented two types of agents Agent-D and Agent-T respectively to use them as competing strategies in the market. In the simulation, we used 30 agents incorporating 3 different strategies (10 agents for each strategy: 10 Agent-D instances, 10 Agent-T instances, and 10 SCMaster instances) to implement a more realistic competition scenario in supply chain markets. Each simulation run is for a period of 250 days at the end of which the agents are evaluated based on their net profit accumulated. We ran simulations in six different situations modeling unfavorable market conditions to SCMaster: ±10% of errors on the demand estimation, ±10% of errors on the profit estimation, and ±10% of distribution changes in actual demand and supply at the same time from the baseline distributions.

To test the adaptability of agents' strategy under adverse market conditions, we adopted concepts from the Evolutionary Game Theory (EGT) (J. Maynard Smith and Price 1973). The EGT has been widely used in many fields, such as economics, sociology, biology, computer science, etc., since it provides the theoretical framework to observe the competition among different strategies in a population over time using natural selection. We observe the growth rate of the agents in a population based on replicator dynamics (RD). Since the competition and the growth rate of agents in our simulation closely reflect the competition and the market's selection of strategies in real-world, our experiment can also be considered a simulated software market. In each experiment we run the simulations until we reach an equilibrium state in which the proportion of agents of each type in a population does not change over time.

Organization of the Study

The rest of this study is composed as follows. In section 2, the related studies regarding the supply chain trading agents' performance measurement and the evaluation of the adaptability of agents' strategy are discussed. In section 3, the simulation environment consisting of customers, suppliers, three type of supply chain agents, and the performance measurements are described. In section 4, the results of simulations are discussed. In the last section, we conclude and discuss future research improvements.

RELATED WORK

In the supply chain trading agent simulation, measuring the performance of agents is an important issue. In the literature for supply chain trading agents, simulations are frequently used to emulate the interaction among agents because a simulation provides a controlled experiment environment. One such supply chain trading agent simulation is provided by TAC-SCM. The TAC-SCM is a tournament in which intelligent agents compete against each other to maximize profit in the personal computer (PC) market, and its specification is given in Collins et al. (2006). In this simulation, agents win customers' sales orders, purchase components from multiple suppliers, store the components at its storage, assemble components into PCs at their factory, and ship the PCs to fulfill customer orders. The winner of simulation is the agent that yields the highest profit among six agents after multiple days of trading in the supply chain environment. In the literature of TAC-SCM, most of the agents use strategies based on optimization techniques such as greedy algorithms and a family of linear programming methods.

While the performance of agents in a game in TAC-SCM is measured by the profits of agents, it is difficult to compare the results created from different executions of simulations because each execution of simulation provides different demand and supply level. The simulation environment from Swedish Institution of Computer Science (SICS) was used by many agent development teams, but it is hard to compare the profit of a simulation to those from other simulations since the distribution of supply and demand are created randomly. To overcome the limitation of these simulation environments, the concept of demand-adjusted profit (DAP) was proposed by Wellman et al. (2006) to directly compare a result of a simulation to those from other simulations. The team from University of Minnesota suggested a model of Controlling Server, which provides the same market conditions in multiple executions of simulation in Borghetti and Sodomka (2006). In this server, their focus is to manipulate a market condition and to secure the randomness in each execution of simulation while the server provides the repeatability during multiple executions of simulation. Focusing more on the manipulation of the market condition, they suggested Market Relief and Market Pressure Agents to control the market environment. Although TAC-SCM provided a valuable framework for supply chain trading agent competition, it still has several limitations. In each round of competition the number of agents is fixed to six and no two strategies are identical. Due to these restrictions the effect of interactions from other instances of agents are smaller than that in a larger-population game. Also, we cannot observe the result yielded when agents with the same strategy compete against each other. In a simulation setting with multiple agents of same type, agents with the same strategy can compete against each other or they might form a group against the agents having different strategies. It is not possible to observe such a scenario in TAC-SCM games while such a situation is frequently found in various markets.

Another group of supply chain trading agents adopted Reinforcement Learning (RL) techniques since it is generally known that artificial intelligence algorithms could provide better performance with self-adaptive agents in a supply chain with complex and stochastic demand patterns than mathematical or optimization techniques (Mortazavi, Arshadi Khamseh, and Azimi 2015). Although these studies showed that RL techniques provide a better performance to supply chain trading agents, their simulation is conducted in a linear supply chain setting, in which there exist very few suppliers or customers in up or down streams of the supply chain. Therefore, it is also hard to predict the performance of such agents when the agents are used in a multi-agent competition environment.

When considering the characteristics of complex adaptive system, it is more reasonable to use the changes in the composition of population over time to decide which strategy provides better adaptability rather than considering a snapshot of the composition of population. Evolutionary Game Theory (EGT) (Weibull 1997) and the notion of Replicator Dynamics (RD) in EGT provides a way to study the dynamics of a population of strategies based on its payoffs (Sikora and You 2014). In addition, EGT provides the notion of Evolutionary Stable Strategy (ESS) to measure the robustness of a strategy. In EGT, a strategy is evolutionarily stable if a population of individuals using that strategy cannot be invaded by a group of alternative strategy (Liu et al. 2016).

Due to these advantages of EGT, studies using multi-agent modeling in various fields have applied the EGT concepts and observed the evolution of agents to determine which strategy is the most appropriate strategy in a given environment. Amir, Evstigneev, and Schenk-Hoppé (2013) proposed a model of financial market, in which actors invest according to their individual decision (i.e., the strategy profile in the model) to identify investment strategies that guarantee the "long-run survival" of any investor using them. In their model, investors are able to change their strategy depending on the inputs from the environment. According to Amir, Evstigneev, and Schenk-Hoppé (2013), the notion of a survival portfolio rule is similar to the notion of ESS introduced by Smith and Price (1973) and Schaffer (1988, 1989), and the concepts of survival and extinction are the ones proposed in EGT (Hofbauer and Sigmund 1998; Samuelson 1998; Sandholm 2010; Vega-Redondo 1996; Weibull 1997).

You and Sikora (2014) and Sikora and You (2014) applied the concept of RD to study reputation mechanisms and rating bias. They used agent-based modeling approach to simulate realistic market conditions. They created seller agents and buyer agents using different strategies (reputation model and strategic behaviour) and conducted simulations. Xiao et al. (2015) applied the dynamic placement of virtual machines for the power saving of data center. For a virtualized data center, the most effective issue is to allocate virtual machines dynamically. They proposed that an algorithm based on EGT can be a better solution. In their study, they used multiple decision makers in a game and allowed some of them to give up their best strategy depending on the situation. At the end of a game, virtual machines are mapped into multiplayers in the evolutionary game. Similar approach is made in Arora, Singh, and Gupta (2016). They used EGT on the cryptographic protocol selection problem for wireless sensor networks since wireless sensor networks have many aspects in common with biological behaviors in decision making and cooperation (Arora, Singh, and Gupta 2016), and the algorithm chooses a “better” strategy in various situations.

Adami, Schossau, and Hintze (2014) provided an insight on the validity of the application of agent-based method where purely mathematical treatment cannot be applied. In their study, they suggested the agent-based method as a feasible modeling method to evaluate the realistic models based on the EGT while a mathematical method still plays an essential role to describe the agent-based methods. Liu et al. (2016) proposed a general robustness measure of incentive mechanism against bounded rational players. In their study, they suggested a general robustness measure and a simulation framework for a quantitative evaluation at first, and, then, implemented and validated the framework by applying it on a simulation in which four agents have four different incentive reputation mechanisms in e-market place. The result of the simulation shows that agent-based simulation also provided the same result expected in the analytical prediction based on a modified simples analysis approach (Liu et al. 2016).

THE SIMULATION ENVIRONMENT

Our simulation model is a personal computer (PC) market where multiple agents compete against each other for higher profits. In the simulation, supply chain trading agents purchase components from a supplier, assemble them into PCs, and sell them to customers through auction. The performance of each agent is measured by their profits. The simulation is composed largely of three parts: supplier, customers, and supply chain agents.

Suppliers and Procurement Process

Suppliers in the simulation provide components of PCs, so agents need to acquire components from suppliers to assemble PCs. Suppliers and agents interact by exchanging RFQs and related offers. To purchase the components of PCs, agents send RFQs to suppliers every day. Then, suppliers return offers to agents. When issuing offers, suppliers apply a priority mechanism for offer price and supply amount allocation, so that more loyal customers could be offered lower prices and larger supply amounts. The priority mechanism named reputation, is adopted from the specification of TAC-SCM in Collins et al. (2006). The reputation is defined as shown in Eq. (1) and (2).

$$\varsigma_a = \frac{quantityPurchase_a}{quantityOffered_a} \tag{1}$$

where $quantityPurchase_a$ is the sum of the quantities in all the orders by agent a, and $quantityOffered_a$ is the sum of quantities in all the offers issued by the supplier to the agent a. By considering ς_a , suppliers may measure which agent is more likely to purchase their components for a given procurement offer. With the value of ς_a, suppliers will calculate the final reputation value as follows:

$$rep_a = \frac{\min(apr, \varsigma_a)}{apr} \tag{2}$$

where apr is the acceptable purchase ratio. This concept is used to prevent excessive punishment for not purchasing the component because agents should not be punished for purchasing from only one supplier.

When a supplier has enough capacity, it may provide a full amount to all RFQs. However, if the supplier does not have enough available capacity for all RFQs, suppliers issue partial offers to distribute the supply amount to RFQs by applying the concept of conflict in Eq. (3).

$$\forall_{\mathrm{r}} \in R_{conflict}, \;\; q_{\mathrm{r}}^{p} = q_r' - C^{conflict} \frac{q_r'(1/reputation_r^m)}{\sum_{r' \in R_{conflict}} q_{r'}' \, (1/reputation_{r'}^m)} \tag{3}$$

where $R_{conflict}$, is the group of RFQ of which sum quantity is not met by the supplier, q_{r}^{p} is new price of RFQ r for product (=component) p, q_r' is the RFQ quantity, $C^{conflict}$ is conflict amount, and $reputation_r$ is the reputation of the agent that issued RFQ r. Once partial offers are created, suppliers also create earliest-possible-offers to fulfill the amounts not filled by partial offers and send all offers to agents.

Customer and Sales Process

In the simulation customers purchase PCs from agents through a sealed first-price auction. Since the issue of the customer demand is to maintain the repeatability and randomness at the same time as indicated in Borghetti and Sodomka (2006), we combined a pre-defined demand distribution (stored in a distribution table), a Poisson distribution, and a random function to achieve the goal.

Customers use a pre-defined daily demand amount for each PC type (stored in a distribution table). To issue customer RFQs, we assumed a Poisson distribution as in the specification of Collins et al. (2006). Hence, customers create RFQs by using Poisson distribution as follows:

$$N = Poisson(Q) \tag{4}$$

where N is the number of sales RFQs of the day, and Q is the average of Poisson distribution. For the demand amount for each RFQ, customers select the due date of each RFQ randomly in between three and twelve days after the current day. After selecting the due date, customers choose the number of products randomly between three and twelve as defined in Eq. (5) to decide the demand amount of the day:

$$Qty_d^p = MIN(\, random(3,12), D_d^p) \tag{5}$$

where Qty_d^p is the order quantity on day *d*, and D_d^p is the demand amount of product p on day *d.* Once RFQs are created, customers send RFQs to trading agents, and the agents return offers to win the orders. From among the sales offers from agents, customers choose an offer having the lowest bidding price. If multiple agents bid the same lowest price, customer randomly chooses one offer.

Agents in the Simulation

The key players of this simulation are agents, and agents compete against each other to acquire suppliers' component and to win customers' sales. The revenue of each agent comes from sales of PCs to customers, and the cost of the agent occurs from the component purchase cost, storage holding cost, and penalty for late delivery. Individual agents have their own factory that can assemble any type of PC, and production is limited based on daily capacity. After purchase of components or after the production, components and completed PCs are stored in the storage. For the storage usage, the component holding and product holding costs are charged daily. For the delivery of products to customers, agents establish a shipping schedule every day, but if the production inventory does not satisfy the entire quantity in an order, agents are not able to ship the product for the order.

We created two types of competing agents by adopting two best-preforming strategies from the TAC-SCM literature. First agent was created from the main strategy of DeepMaize (Kiekintveld et al. 2006) and it is named Agent-D. Agent-D uses the concept of reference inventory trajectory and the baseline buffer level for the estimation of procurement amount. To find the best bidding price, Agent-D maximizes the sum of the increase of expected profits of sales offers. Agent-D uses supply and demand amount prediction tables and fills the tables with prediction data. In our simulation, we provided the exact prediction to Agent-D by filling the tables with the actual values used for supplier and customer's distribution tables.

Second agent, Agent-T, was created to simulate the TacTex agent in Pardoe and Stone (2009). For creation of procurement RFQ, Agent-T uses the two inventory threshold for non-CPU components and CPU components. For each sales offer, Agent-T searches a bidding price by using a greedy algorithm that maximizes the sum of the increase of expected profits of sales offers. Agent-T establishes shipment schedules depending on the availability of PC inventory at first, and it schedules additional production for unfilled orders by using a greedy algorithm. Agent-T also uses supply and demand amount prediction tables for the same purpose as that of Agent-D, so the tables are also filled with the actual values.

The third agent used was SCMaster presented in Lee and Sikora (2016). SCMaster uses several machine-learning techniques to enhance the adaptibility of its strategy to deal with the changes in environment. SCMaster uses the inventory threshold mechanism like other agents in TAC-SCM literature, but it controls the threshold dynamically to reduce the inventory cost and to maintain the appropriate level of inventory for production at the same time by combining Q-Learning, Softmax, and ε-greedy reinforcement learning algorithms. Production module is enhanced from the one used in Tac-Tex (Pardoe and Stone 2009) by applying a modified greedy algorithm. When scheduling the production, SCMaster allocates production resources to the production of the orders that are past due or that do not have enough product inventory for the completion of the orders. SCMaster's sales strategy is focused on the prediction of exact winning prices because the key part of winning orders in the auction is a better prediction of bidding price. To achieve this, SCMaster uses a sliding window protocol (R. T. Sikora and Sachdev 2008). This algorithm enables SCMaster to improve its bidding price by using finer price bands for more competitive intervals of prices and coarser price bands for less competitive intervals of prices. It does that by learning and dynamically adjusting the price bands as simulation continues. For the prediction of demand and supply amount, SCMaster also uses supply and demand amount prediction tables. In the case of SCMaster, however, we filled the tables with the exact values when running the no-estimation-errors simulation, and the tables are filled with incorrect values (e.g., 10% of demand prediction error values or 10% of supply prediction error values, etc.) when running simulations for prediction error situations. In this way, we want to test the robustness and adaptability of SCMaster in the presence of adverse market conditions that lead to estimation errors on both the supply and demand distributions.

Modification of Simulation

Even though the simulation in this study is motivated by TAC-SCM in Collins et al. (2006), several of its features were modified to implement a more realistic business environment. First, we assumed a continuum of business environment with no specific start or end of the simulation. Therefore, none of the agents in this study use any algorithm for the initial status of the simulation or the end of the simulation. The reason for not considering the concept of start, end, and data from previous simulations is that this is not possible in the real world. Secondly, we pre-define a demand distribution rather than create them by a random function. As defined in the section 3.2, we combined a pre-defined daily demand distribution and a random demand due date for the creation of an order to secure the randomness and repeatability. In addition, the distributions of the customer demand and the supply capacity are shared among competitor agents to allow them to have high prediction accuracy.

Replicator Dynamics

One of the primary goals of this paper is to assess the adaptability of our model's strategy in a trading market. One way of documenting that adaptability is to empirically study the growth of the agents' strategies in a population containing competing strategies. To do so, we borrow the concept of replicator dynamics (RD) from evolutionary game theory (Weibull, J. W., 1995). RD studies the dynamics of a population of strategies that replicates based on its payoff. In our case, the strategies represent the different SCM models used by the agents. At each time period, a population of agents takes part in the market simulation. Each agent uses a different model and performs the tasks of procurement, sales, and scheduling in accordance with its model. The agent's gain is treated as its fitness, and agents breed from one generation to the next in proportion to their average fitness. The changes of the relative composition of the population of agents using different types of models can be an indicator of the effect on change in agent behavior.

One can model RD as a set of difference equations. Let N be the total number of agents, x_{it} be the proportion of agents who are using model *i* at time t, and u_{it} be the average gain of the agents using model *i* at time *t*. Then it follows that the proportion of agents using model *i* at time *t*+1 is given by

$$x_i^{t+1} = \frac{u_i^t x_i^t N}{\sum_k u_k^t x_k^t N} = \frac{u_i^t x_i^t}{\sum_k u_k^t x_k^t} \tag{7}$$

We can study the evolution of popularity of different SCM models within a population of agents by studying the evolution of x_{it}.

After adjustment of the composition of population, a series of simulations, which are called generations, are continued until the composition of population reaches an equilibrium, a stable state where the composition of population does not change over time. In all simulations, we assumed a uniform customer demand distribution with a mean of 2000 and a standard deviation of 10 as a baseline demand distribution. In the first four simulation setups, we modified SCMaster's prediction module to cause four types of prediction errors while other agents have exact prediction of customer demand distribution. In the fifth and sixth simulation, we modified actual demand and supply distributions to cause two types of prediction errors to all agents in the simulation. The six types of modification include: ±10% of errors on the agent's internal demand estimation only for SCMaster, ±10% of errors on the agent's internal profit estimation only for SCMaster, ±10% of distribution change in actual demand and supply.

RESULTS

No Estimation Errors

This simulation is a baseline setting to compare its results to those from other simulations. In this experiment, the baseline performance of three types of agent are measured. The three types of agents in this simulation have perfect predictions on customer demand and suppliers' capacity. We used a population of 30 agents, ten for each type of SCM agent (10 Agent-D instances, 10 Agent-T instances, and 10 SCMaster instances).

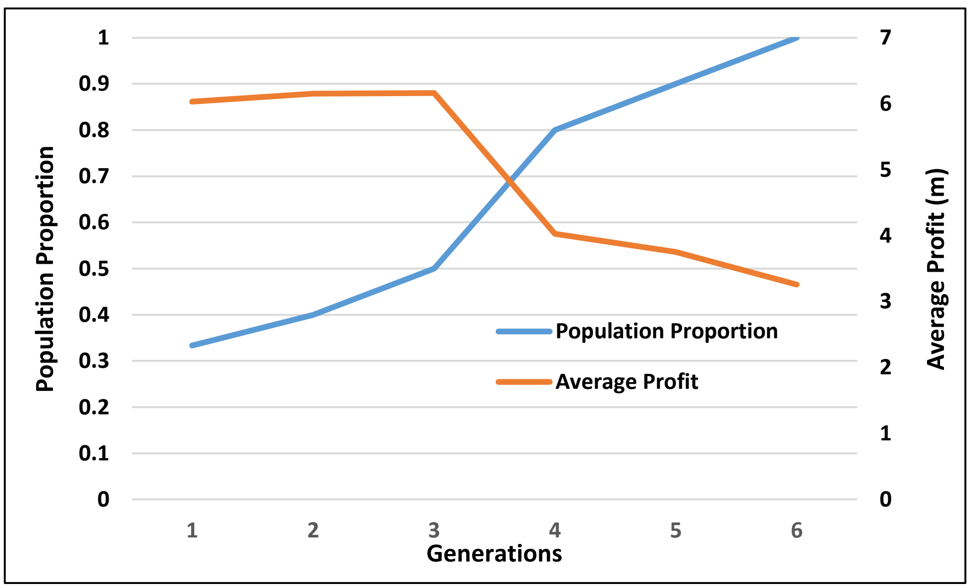
The experiments of this setting were conducted under a uniform demand distribution with a mean of 2000 and a standard deviation of 10. Table 1 represents the changes of the average profit of each agent type with their count in the population, and Figure 1 depicts the changes in the composition of population. As seen in Figure 1, Agent-D is eliminated from the population at the fourth generation, and the proportion of Agent-T diminishes from the fourth generation. Finally, at the sixth generation, SCMaster takes over the entire population, and the population reaches an equilibrium. As mentioned earlier, in real world situations the composition of competition is dynamic; weaker agents do not survive whereas stronger agents become more established. Moreover, the performance of the agents itself depends on such composition (stronger agents do better when competing against weaker agents than when competing against other stronger agents). It is this constant feedback loop (where performance depends on composition of agents, but the composition of agents in turn depends on their performance) that we have tried to simulate in our experiments and the results demonstrate this. As shown in Figure 2, although the SCMaster agents takes over the population their average profit keeps falling as their numbers grow in the population since they are now competing with stronger agents (i.e., competing amongst themselves). They earn an average profit of 6.03m when competing against different types of agents. But when they overtake the entire population they are competing against their own strategy and their average profit drops to 3.26m.

Table 1. No estimation errors

Generation	SCMaster		Agent-D		Agent-T	
	Avg. Profit (m)	Count	Avg. Profit (m)	Count	Avg. Profit (m)	Count
1	6.03	10	-2.41	10	3.56	10
2	6.15	12	-2.17	7	3.29	11
3	6.16	15	0.38	4	3.27	11
4	4.03	24	0	0	0.6	6
5	3.75	27	0	0	-7.37	3
6	3.26	30	0	0	0	0

Figure 1. The changes in the composition of population with no estimation errors

Figure 2. Performance of SCMaster with no estimation errors



Overestimation and Underestimation of Demand by 10 %

To test the robustness of the strategy of SCMaster in the presence of estimation errors, we ran an experiment with overestimation and underestimation of demand by 10% only for SCMaster while other agents have a perfect estimation on customer demand.

Table 2. Overestimation of demand by 10%

Generation	SCMaster		Agent-D		Agent-T	
	Avg. Profit (m)	Count	Avg. Profit (m)	Count	Avg. Profit (m)	Count
1	6.06	10	-3.63	10	3.44	10
2	6.02	14	-1.41	5	3	11
3	6.26	16	-12.54	3	3	11
4	5.69	19	0	0	3.24	11
5	4.72	23	0	0	2.69	7
6	3.77	26	0	0	-4.75	4
7	3.11	29	0	0	-28.88	1
8	3.16	30	0	0	0	0

Table 2 provides the results from 10% overestimation of demand. The composition of population has the same pattern as that in the baseline experiment. Agent-D is eliminated at the fourth generation. After the extinction of Agent-D, the population of Agent-T begins to decrease, and finally the population reaches the equilibrium at the eighth generation.

Table 3. Underestimation of demand by 10%

Generation	SCMaster		Agent-D		Agent-T	
	Avg. Profit (m)	Count	Avg. Profit (m)	Count	Avg. Profit (m)	Count
1	5.94	10	-4.74	10	3.57	10
2	5.94	13	-0.54	5	3.16	12
3	6.24	16	-8.38	3	3.09	11
4	5.53	18	-33.49	1	3.33	11
5	5.49	19	0	0	3.2	11
6	4.76	22	0	0	3.15	8
7	4.23	24	0	0	0.35	6
8	3.56	27	0	0	-7.22	3
9	3.21	30	0	0	0	0

Table 3 provides the results from 10% underestimation of demand. The composition of the population showed slightly different pattern from that in the baseline experiment. Agent-D is eliminated from the population from the fifth generation, and the population reaches an equilibrium at the ninth generation while Agent-T begins to decrease from the sixth generation. Based on these results we can conclude that the SCMaster's strategy is robust and adaptive enough to handle the errors in demand estimation.

Overestimation and Underestimation of Profit by 10%

Since one of the key aspects of SCMaster is its preference for producing high profit products, we test the robustness of its strategy in the presence of errors in estimation of its profits. To do that, we test its performance with overestimation and underestimation of profit by 10% only for SCMaster.

The result of simulation for the overestimation of profit is provided in Table 4. Again, the patterns of the changes in the composition of population are fundamentally the same to the one in the baseline experiment. Agent-D is eliminated from the population at the fifth generation. During that time, Agent-T maintains the same proportion of the population and SCMaster captures Agent-D's portion in the population. However, the population of Agent-T also begins to decrease from the sixth generation, and the entire population is taken up by SCMaster at the ninth generation and the population reaches an equilibrium.

Table 4. Overestimation of profit by 10%

Generation	SCMaster		Agent-D		Agent-T	
	Avg. Profit (m)	Count	Avg. Profit (m)	Count	Avg. Profit (m)	Count
1	5.91	10	-3.46	10	3.53	10
2	6.13	13	-1.17	6	3.31	11
3	6.10	16	-12.05	3	3.24	11
4	5.67	18	-33.89	1	3.31	11
5	5.40	19	0	0	3.29	11
6	4.70	22	0	0	3.04	8
7	4.25	24	0	0	0.28	6
8	3.44	28	0	0	-12.10	2
9	3.30	30	0	0	0	0

The result of simulation for the underestimation of profit is provided in Table 5. The pattern of the changes in the composition of population is similar to the one in the baseline experiment. Agent-D disappears from the population at the fifth generation and the simulation continues to the ninth generation. Agent-T maintains the same population until the extinction of Agent-D, but Agent-T declines after extinction of Agent-D. Hence, we can conclude that SCMaster's strategy is adaptive enough to overcome the over or under estimation of profit.

Table 5. Underestimation of profit by 10%

Generation	SCMaster		Agent-D		Agent-T	
	Avg. Profit (m)	Count	Avg. Profit (m)	Count	Avg. Profit (m)	Count
1	5.94	10	-2.47	10	3.44	10
2	6.00	13	-2.85	6	3.26	11
3	4.56	16	-12.83	3	1.17	11
4	5.85	18	-34.51	1	3.31	11
5	5.52	19	0	0	3.31	11
6	4.82	22	0	0	3	8
7	4.26	24	0	0	0.97	6
8	3.24	28	0	0	-12.33	2
9	3.28	30	0	0	0	0

Over/Under Demand and Supply

In this simulation we tested the adaptability of our agent's strategy when all the agents make an error of 10% (both over and under) in estimating the demand and supply distributions.

Table 6 presents the results of over estimation of the demand and supply distributions. We got similar results for under estimation of both demand and supply distributions. The changes in the composition of population reaches an equilibrium at the ninth generation. Similar to the evolution pattern in previous simulations, SCMaster occupies Agent-D's positions during the decline of Agent-D, and, SCMaster takes over the population by the ninth generation after Agent-T's elimination from the population. Based on these results, we can conclude that even in the presence of estimation errors in both supply and demand distributions, SCMaster is better able to handle the adverse market conditions.

Table 6. Overestimation of demand and supply by 10%

Generation	SCMaster		Agent-D		Agent-T	
	Avg. Profit (m)	Count	Avg. Profit (m)	Count	Avg. Profit (m)	Count
1	5.95	10	-3.5	10	3.57	10
2	6.26	13	-6.08	6	3.45	11
3	6.19	14	-3.16	5	3.42	11
4	6.43	17	-22.67	2	3.39	11
5	6.01	19	0	0	3.29	11
6	5.09	23	0	0	2.13	7
7	3.89	27	0	0	-7.16	3
8	3.84	30	0	0	0	0

Sensitivity Analysis

In previous sections, we conducted simulations in three different types of error situations under a uniform demand distribution with a mean of 2000 and a standard deviation of 10, and the results demonstrated the adaptability of the SCMaster's strategy. To test whether the results hold under more unstable environment, we performed a sensitivity analysis with a higher variance of the demand distributions. We repeated the same set of simulations in previous sections by applying a uniform demand distribution with a mean of 2000 and a standard deviation of 100.

In all these simulations the patterns of the changes in the composition of population were similar to those under the baseline simulation. Based on these results, we can conclude that the SCMaster's algorithms for the procurement and bidding to customer are adaptive enough to overcome the estimation errors created by the internal or external causes.

SCMaster as an Evolutionary Stable Strategy

In previous sections, we tested the adaptability of SCMaster's strategy in simulations in which all types of agents start with the same number of agents in the population. Another concept from evolutionary game theory that is related to the stability of an equilibrium strategy is an evolutionary stable strategy (ESS), first proposed by (John Maynard Smith 1982). In the context of evolutionary game theory, a strategy is ESS if a small proportion of that strategy can invade a population of other strategies and completely overtake it. It has been shown that a strict Nash equilibrium strategy (i.e., a stable Nash equilibrium where a player making a small change away from the equilibrium becomes worse off) also implies an ESS (Weibull 1997).

In the next set of experiments, we wanted to study whether SCMaster was such an ESS and what was the minimum proportion of SCMaster strategies needed to takeover a population of Agent-D and Agent-T strategies. We used a uniform demand distribution with a mean of 2000 and a standard deviation of 100 and used a population of 30 instances. In the first experiment with Agent–D, the minimum number of initial SCMaster strategy needed to take over the entire population was 4. Table 7 gives the evolution of the population and by the ninth generation SCMaster takes over the entire population to reach an equilibrium. In the next experiment with Agent–T, the minimum number of initial SCMaster strategies needed to take over the entire population was 8. Table 8 shows the changes in the composition of population and by seventh generation SCMaster takes over the entire population to reach an equilibrium. In the last experiment with both Agent–T and Agent-D, the minimum number of initial SCMaster strategies needed to take over the entire population was 5. Table 9 gives the changes in the composition of population and by 11th generation SCMaster takes over the entire population to reach an equilibrium. Although a single SCMaster strategy is not able to invade and take over an entire population consisting of competing strategies, a very small proportion of SCMaster strategies is able to do so in all situations. The above results imply that SCMaster is a type of evolutionary stable strategy.

Table 7. Result with Agent-D when the initial number of SCMaster = 4

Generation	SCMaster		Agent-D	
	Avg. Profit (m)	Count	Avg. Profit (m)	Count
1	3.26	4	0.2	26
2	4.02	5	0.8	25
3	5.50	7	-0.14	23
4	6.31	11	1.46	19
5	6.39	14	-0.27	16
6	5.69	19	-2.81	11
7	4.27	25	-4.84	5
8	3.30	29	-868.88	1
9	3.30	30	0	0

Table 8. Result with Agent-T when the initial number of SCMaster = 8

Generation	SCMaster		Agent-T	
	Avg. Profit (m)	Count	Avg. Profit (m)	Count
1	3.15	8	1.83	22
2	4.98	12	2.51	18
3	6.02	17	3.05	13
4	4.69	22	3.12	8
5	4.51	24	1.01	6
6	3.78	27	-7.24	3
7	3.3	30	0	0

Table 9. Result with both Agent-D and Agent-T when the initial number of SCMaster = 5

Generation	SCMaster		Agent-D		Agent-T	
	Avg. Profit (m)	Count	Avg. Profit (m)	Count	Avg. Profit (m)	Count
1	4.04	5	-5.3	13	3.21	13
2	3.86	7	-1.01	6	3.34	18
3	4.1	8	-6.39	4	3.08	19
4	4.16	10	-28.65	1	2.05	20
5	4.53	11	0	0	2.6	20
6	5.3	15	0	0	2.57	16
7	5.39	20	0	0	3.25	11
8	4.19	23	0	0	2,93	8
9	3.86	25	0	0	0.32	6
10	3.51	28	0	0	-7.27	3
11	3.15	31	0	0	0	0

CONCLUSION

In this paper we studied the evolution of strategies used by supply chain agents when the different agents are competing against each other to manage a supply chain most profitably in a single marketplace. We used multi-agent modelling to simulate a realistic marketplace consisting of different types of supply chain agents. Using the concept of replicator dynamics we studied the evolution of a population of supply chain agents consisting of three different types of strategies. Our results indicate that given a choice of strategies to use, agents prefer to switch to using our adaptive SCMaster model over time. We also simulated scenarios of adverse market conditions by inserting errors in the estimation of demand, supply and profit. In all scenarios we found the SCMaster model to be the most adaptive. This has important implications for the designers of automated and intelligent supply chain agents. As a methodology, it provides a tool to test whether a given strategy will be able to compete and survive in a population of other strategies. As a strategy to be used by a supply chain agent, it shows that SCMaster is a better strategy.

SCMaster's strategy is most adaptive among the three strategies we tested as it uses dynamic inventory control and various reinforcement learning techniques like Q-learning, Softmax, ε-greedy, and sliding window protocol to make the agent adapt dynamically to the changing environment created by competing agents. Using concepts from evolutionary game theory, we also conducted experiments that showed that SCMaster was a type of evolutionary stable strategy akin to a stable Nash equilibrium.

There are several limitations of the study presented in this paper. The foremost is that we only tested the evolution of a population of fixed strategies. In future we plan on designing experiments where the individual strategies themselves evolve over time by combining bits and pieces from the existing strategies to see if any newer and stronger strategies are discovered. The other limitation of this study is that although we tested scenarios of different demand and supply distributions these distributions or environment stayed static within each simulation. In future we want to test simulations where the environment in terms of the supply and demand distributions changes over time to reflect seasonal changes or economic cycle changes.

REFERENCES

Acquity Group. 2014. "Uncovering the Shifiting Landscape in B2B Commerce." https://www.accenture.com/t20150624T211502__w__/us-en/_acnmedia/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Industries_15/Accenture-B2B-Procurement-Study.pdf.

Adami, Christoph, Jory Schossau, and Arend Hintze. 2014. "Evolutionary Game Theory Using Agent-Based Methods." *ArXiv Preprint ArXiv:1404.0994*. http://arxiv.org/abs/1404.0994.

Amir, Rabah, Igor V. Evstigneev, and Klaus Reiner Schenk-Hoppé. 2013. "Asset Market Games of Survival: A Synthesis of Evolutionary and Dynamic Games." *Annals of Finance* 9 (2): 121–44. doi:10.1007/s10436-012-0210-5.

Arora, Srishti, Prabhjot Singh, and Ashok Ji Gupta. 2016. "Adaptive Selection of Cryptographic Protocols in Wireless Sensor Networks Using Evolutionary Game Theory." *Procedia Computer Science* 78: 358–66. doi:10.1016/j.procs.2016.02.075.

Borghetti, Brett, and Eric Sodomka. 2006. "Performance Evaluation Methods for the Trading Agent Competition." In *AAAI*, 1857–1858. http://www.aaai.org/Papers/AAAI/2006/AAAI06-307.pdf.

Choi, Thomas Y., Kevin J. Dooley, and Manus Rungtusanatham. 2001. "Supply Networks and Complex Adaptive Systems: Control versus Emergence." *Journal of Operations Management* 19 (3): 351–366.

Collins, John, Raghu Arunachalam, Norman Sadeh B, Joakim Eriksson, Niclas Finne, and Sverker Janson C. 2006. *The Supply Chain Management Game for the 2007 Trading Agent Competition*.

Collins, John, Raghu Arunachalam, Norman M. Sadeh, Joakim Eriksson, Niclas Finne, and Sverker Janson. 2006. "The Supply Chain Management Game for the 2007 Trading Agent Competition." *Http://Reports-Archive.Adm.Cs.Cmu.Edu/Anon/Isri2007/CMU-ISRI-07-100.Pdf*.

Ejiaku, Samuel A. 2014. "Technology Adoption: Issues and Challenges in Information Technology Adoption in Emerging Economies." *Journal of International Technology and Information Management* 23 (2): 5.

Hofbauer, Josef, and Karl Sigmund. 1998. *Evolutionary Games and Population Dynamics*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139173179.

Kiekintveld, Christopher, Jason Miller, Patrick R. Jordan, and Michael P. Wellman. 2006. "Controlling a Supply Chain Agent Using Value-Based Decomposition." In , 208–217. http://dl.acm.org/citation.cfm?id=1134730.

Lee, YoonSang, and Riyaz Sikora. 2016. "Design of Intelligent Agents for Supply Chain Management." In *E-Life: Web-Enabled Convergence of Commerce, Work, and Social Life*, edited by Vijayan Sugumaran, Victoria Yoon, and Michael J. Shaw, 258:27–39. Cham: Springer International Publishing. doi:10.1007/978-3-319-45408-5_3.

Liu, Yuan, Jie Zhang, Bo An, and Sandip Sen. 2016. "A Simulation Framework for Measuring Robustness of Incentive Mechanisms and Its Implementation in Reputation Systems." *Autonomous Agents and Multi-Agent Systems* 30 (4): 581–600. doi:10.1007/s10458-015-9296-2.

Macal, Charles M., and Michael J. North. 2009. "Agent-Based Modeling and Simulation." In *Winter Simulation Conference*, 86–98. Winter simulation conference. http://dl.acm.org/citation.cfm?id=1995474.

Mortazavi, Ahmad, Alireza Arshadi Khamseh, and Parham Azimi. 2015. "Designing of an Intelligent Self-Adaptive Model for Supply Chain Ordering Management System." *Engineering Applications of Artificial Intelligence* 37 (January): 207–20. doi:10.1016/j.engappai.2014.09.004.

Pardoe, David, and Peter Stone. 2009. "An Autonomous Agent for Supply Chain Management." *Handbooks in Information Systems Series: Business Computing* 3: 141–172.

Saar-Tsechansky, Maytal. 2015. "The Business of Business Data Science in IS Journals." *MIS Quarterly* 39 (4).

Samuelson, Larry. 1998. *Evolutionary Games and Equilibrium Selection*. Vol. 1. MIT press.

Sandholm, William H. 2010. *Population Games and Evolutionary Dynamics*. MIT press.

Schaffer, Mark E. 1988. "Evolutionarily Stable Strategies for a Finite Population and a Variable Contest Size." *Journal of Theoretical Biology* 132 (4): 469–78. doi:10.1016/S0022-5193(88)80085-7.

Schaffer, Mark E. 1989. "Are Profit-Maximisers the Best Survivors?: A Darwinian Model of Economic Natural Selection." *Journal of Economic Behavior & Organization* 12 (1): 29–45.

Sikora, Riyaz T., and Vishal Sachdev. 2008. "Learning Bidding Strategies with Autonomous Agents in Environments with Unstable Equilibrium." *Decision Support Systems* 46 (1): 101–14. doi:10.1016/j.dss.2008.05.005.

Sikora, Riyaz, and Liangjun You. 2014. "Effect of Reputation Mechanisms and Ratings Biases on Traders' Behavior in Online Marketplaces." *Journal of Organizational Computing and Electronic Commerce* 24 (1): 58–73. doi:10.1080/10919392.2014.866503.

Smith, J. Maynard, and G. R. Price. 1973. "The Logic of Animal Conflict." *Nature* 246 (5427): 15–18. doi:10.1038/246015a0.

Smith, John Maynard. 1982. *Evolution and the Theory of Games*. Cambridge university press.

Vega-Redondo, Fernando, ed. 1996. *Evolution, Games, and Economic Behaviour*. Oxford University Press. doi:10.1093/0198774729.001.0001.

Weibull, Jörgen. 1997. *Evolutionary Game Theory*. MIT press.

Wellman, Michael P., Patrick R. Jordan, Christopher Kiekintveld, Jason Miller, and Daniel M. Reeves. 2006. "Empirical Game-Theoretic Analysis of the TAC Market Games." In *AAMAS-06 Workshop on Game-Theoretic and Decision-Theoretic Agents*. http://wwweb.eecs.umich.edu/deepmaize/papers/WellmanGTDT-final.pdf.

Xiao, Zhijiao, Jianmin Jiang, Yingying Zhu, Zhong Ming, Shenghua Zhong, and Shubin Cai. 2015. "A Solution of Dynamic VMs Placement Problem for Energy Consumption Optimization Based on Evolutionary Game Theory." *Journal of Systems and Software* 101 (March): 260–72. doi:10.1016/j.jss.2014.12.030.

You, Liangjun, and Riyaz Sikora. 2014. "Performance of Online Reputation Mechanisms under the Influence of Different Types of Biases." *Information Systems and E-Business Management* 12 (3): 417–42. doi:10.1007/s10257-013-0229-9.

You, Liangjun, Dong-Qing Yao, Riyaz T. Sikora, and Barin Nag. 2017. "An Adaptive Supplier Selection Mechanism in E-Procurement Marketplace." *Journal of International Technology and Information Management* 26 (2): 94–116.