Learning to Locate Trading Partners in Agent Networks

John Porter
Department of Computer Science
University of Tulsa
john-porter@utulsa.edu

Kuheli Chakraborty
Department of Computer Science
University of Tulsa
kuheli@utulsa.edu

Sandip Sen
Department of Computer Science
University of Tulsa
sandip@utulsa.edu

ABSTRACT

This paper is motivated by some recent, intriguing research results involving agent-organized networks (AONs). In AONs, nodes represent agents, and collaboration between nodes are represented by corresponding edges. Agents rewire edges, i.e., change partners, to improve performance. The challenge in these domains is the search and location of desirable interaction or collaboration partners in a relatively large population. It was found that random selection of partners in each time period produced better performance but incurred larger search costs in a production and exchange economy compared to gradual rewiring of edges in the network. We propose an exponentially decaying exploration scheme that produces similar utilities to random rewiring but with much less rewiring costs. We evaluate the effects of the number of trading partners on connections on the utilities obtained by the agents. We hypothesize on the cause for the observed performance differences and verify that by showing that the observed performance differences with more realistic model of the economy.

1. INTRODUCTION

As social networks and peer-to-peer (P2P) networks have received widespread use, various forms of network topologies and their associated properties have been studied in the literature [1]. In this paper, we focus on relationships between producer and consumer agents in a distributed environment. Agents in such an Agent Oriented Network (AON) are connected, at any point in time, with a limited number of other agents but can change their connections over time [4]. To obtain utility, agents need to trade with other agents producing complementary goods. A critical decision problem affecting the viability and success of agents in such an economy is their ability to identify beneficial trading partners. Gaston and desJardins observed that randomly connecting to other agents produced more profitable trades than using more stable wiring patterns [2]. This is a counter-intuitive result, as in real economies we observe more stable and healthy partnerships between organizations in supply chains [3]. We wanted to explain this intriguing phenomena by a careful analysis of the experimental results. More importantly, we wanted to study the properties of such an AON under varying network characteristics. In this paper, we evaluate the relative effectiveness of an exponentially decaying rewiring strategy in an enhanced production and exchange model that incorporate more realistic constraints of stock limits and minimum trade volumes.

Cite as: Learning to Locate Trading Partners in Agent Networks, John Porter, Kuheli Chakraborty and Sandip Sen, *Proc. of AAMAS-2009 Adaptive and Learning Agents Workshop*, May 12, 2009, Budapest, Hungary.

2. DOMAIN MODELS

2.1 Production and Exchange Model

Gaston and desJardins have studied a simple production and exchange model to study strategies in AON exchange economies [2]. An AON is a network of agents in which the agents self-organize and can rewire their own connections to other agents [4]. The connections are unidirectional (whether one agent can request to trade with another is separate from whether the second agent can request to trade with the first) and determined unilaterally.

In the Production and Exchange Model used by Gaston and desJardins [2] every agent starts with some supply of two goods and a capacity to produce a fixed amount of only one of them. At each iteration agents choose whether to produce or exchange goods. Agents are greedy and attempt to maximize the utility they gain at each time step. They are also truthful and always provide correct information when proposing a trade.

In this model there are n agents and two goods g_1 and g_2. g_1 is only traded in whole units while g_2 is infinitely divisible. g_k^i is the amount of good k that agent currently possesses. The utility of agent i is given by the product of its stock of the two goods:

$$U^i = g_1^i g_2^i.$$

In each round the agents are chosen in random order and allowed to trade or produce. First, they have to calculate how much utility they would gain by trading. Each agent is linked to m other agents with whom it can trade. The chosen agent checks its *marginal rate of substitution* (mrs) against the mrs of each of the agents it is linked with. This value is calculated as follows and truthfully revealed:

$$mrs^i = \frac{g_2^i}{g_1^i}.$$

The agents may be able to gain by trading if their mrs's differ. The next step is to decide on the exchange price p_{ij}, which is computed as

$$p_{ij} = \frac{g_2^i + g_2^j}{g_1^i + g_1^j},$$

when agent i is negotiating with agent j. A tax τ is applied to every transaction. At this point a trade is simulated. No actual goods are exchanged until agent i chooses one trading partner. If agent i is trading one unit of g_1 for p_{ij} units of g_2 with agent j and δg_k^i is the amount of good k traded by agent i

$$\delta g_1^i = -\delta g_1^j = -(1+\tau)$$

$$\delta g_2^i = -\delta g_2^j = (1+\tau)p_{ij}.$$

This trade is repeated until the utility of neither agent will not increase from further trading. The corresponding utility gain is recorded. Once this simulation has been repeated for every agent that agent i can trade with the most profitable partner, i.e., the trade that would gain i the greatest utility is chosen. The agent then checks if producing could provide more gain in the current time period than gain from its best trade.

Every agent has a production capacity Δg_i uniformly distributed in the range $[1, q]$ for one of the goods g_1 or g_2. If i produces g_1 its change in utility after production is

$$\Delta U^i = \Delta g_1^i g_2^i.$$

Once an agent knows how much utility it can gain by producing, it can choose whether to produce or trade with its best partner. Once it has made this decision and carried out the corresponding action, the agent can choose to rewire its trading connections for the next iteration.

2.2 Enhanced Production and Exchange Model

In many real world examples, goods are consumed or agents gain utility through consuming goods rather than by just possessing them. These agents also often have a limited space available for storage which can be expensive. Therefore we propose a system of clearing. Whenever an agent has both types of goods, it combines them to create a product. Thus, no excess goods are stored. However, some agents are more efficient at this than others. Every agent must use some multiple, G, units of the good they do not produce for every unit of good that they can produce. Thus, if an agent is a producer of good 1 its new utility gain function would be:

$$\Delta U^i = \zeta \min(g_1^i, \frac{g_2^i}{G^i}).$$

The agent loses the corresponding amounts of goods 1 and 2 and gains utility times the parameter ζ.

The agents have a limited $\bar{S}$, and the maximum amount of the produced good that can be stored is $\bar{S}$ times the production rate. Agents also have a lower bound on the amount of good they have to have before they can try trading. If, at the beginning of their turn, they have less than $\underline{S}$ times their production rate then they do not look for a trading partner. However, an agent may still end up trading even if it has less than $\underline{S}$ goods. Another agent may still initiate a trade with it. Trading can be an expensive operation. So we do not want the agents to make a huge number of small trades. $\underline{S}$ and $\bar{S}$ comprise an optimal trading window for the agents. If the agent cannot find a trading partner before it reaches $\bar{S}$ then it will start losing production as the produced goods in excess of the capacity must be disposed.

As a final change to the production and exchange model we allowed continuous production. Agents could produce every turn, even if they had traded. Because of the $\underline{S}$ lower limit for trading, the agents will not attempt to trade every iteration.

3. REWIRING STRATEGIES

In this paper we evaluate three rewiring strategies **random mixture**, **random selection**, and **exploration**. The first two were used by Gaston and desJardins [2]. **Random mixture** (RM) is the simplest strategy. At each iteration agents randomly reinitialize every connection.

In **random selection** (RS) the agent first decides whether it should adapt. It keeps an exponential weighted moving average, V, of the utility gained in each iteration. The utility agent i expects to gain in the next iteration, t, is

$$V_t^i = V_{t-1}^i + \alpha(\Delta U_{t-1}^i - V_{t-1}^i).$$

If $V_t^i < \Theta$ then the agent chooses to adapt. $\alpha \in [0, 1]$ is a learning parameter and Θ a threshold.

If it chooses to adapt, it still must choose which connections to adapt. This decision is also based on an exponentially weighted moving average of connection strengths represented by connection weights. Agent i then updates its connection weight W_t^{ij} for the connection to agent j:

$$W_t^{ij} = W_{t-1}^{ij} + \beta(\Delta U_{t-1}^{ij} - W_{t-1}^{ij}),$$

where ΔU_{t-1}^{ij} is the change in utility that agent i could have received by trading with agent j on iteration t. $\beta \in [0, 1]$ is a learning parameter. The agent rewires every connection where $W_t^{ij} < \Phi$, where Φ is a threshold parameter. New connection weights are initialized to the average of the current connection weights.

We introduce a third rewiring strategy to reduce search and exploration over time. When using the **exploration** (RE) strategy, each agent has an initial exploration rate $x_0 \in (0, 1]$. This rate exponentially decays at a rate η such that $x_t = \eta x_{t-1}$. The rewiring rate is based on this x_t as well as V_t^i as described above and the base expected utility, V_0^i. In the **exploration** strategy the probability of an agent rewiring a connection is

$$p_t^i = x_t * (1 - \frac{V_t^i}{V_0^i}).$$

The base expected utility is initialized as the average expected utilities for other agents this agent connects to.

Just as in the RS strategy, the agents keep track of a weight, W_t^{ij}, for each connection. However, an agent only rewires the connection with the lowest weight, and only if the connection satisfies the constraint $W_t^{ij} < \Phi$.

4. EXPERIMENTAL RESULTS

We now discuss our experimental results. Due to space constraints, we only present results from the Enhanced Production and Exchange Model. The parameters used in the model are as follows: n=300, q=30, τ=0.05, and m was varied from 2 to 10 in steps of 2. The agent's learning parameters were set at $\alpha = \beta = \theta = \phi = 0.1$ and both the initial expected utility, V_0^i, and at the beginning of each run, the initial valuation of every connection, W_0^{ij}, were set to 1 following Gaston and desJardins[2]. The exploration strategy began with an exploration rate of $x_0 = 0.3$. The decay rate was $\eta = 0.996$. For this model we use Minimum Trade

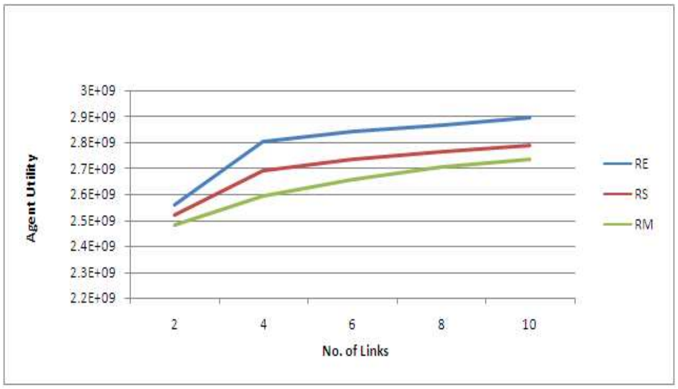

Figure 1: Effect of m in homogeneous populations using the enhanced production and exchange model for non-continuous production.

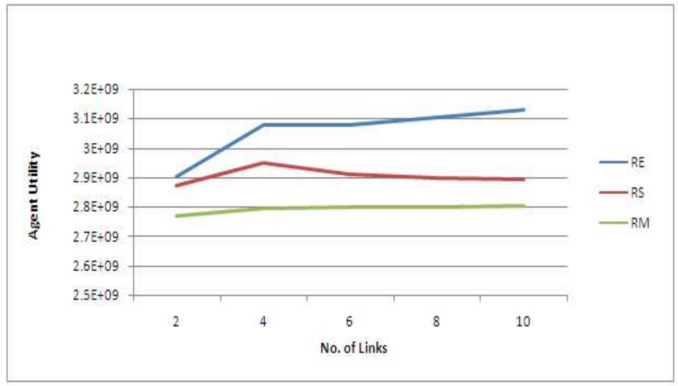

Figure 2: Effect of m in homogeneous populations using the enhanced production and exchange model with continuous production.

Volume($\underline{S}$)=3 and Storage Capacity($\bar{S}$)=4. All results are based on random generated initial network structures.

4.1 Homogeneous Populations

The first set of experiments was run with homogeneous agent populations. In a homogeneous population, there is significant and interesting effect on the performance of RE, RS and RM strategies when varying other domain characteristics like continuous and non continuous production, $\underline{S}$, $\bar{S}$, number of links, etc.

Non-Continuous Production.

The effect of number of connections on the agent utilities for the non-continuous production environment (see Figure 1) shows the advantage of the judicious exploration scheme. In contrast to the basic production and exchange model, the order of performance is RE followed by RS followed by RM. In this model all agents have to accumulate sufficient stock and maintain minimum trade volume before trading. Hence the agents are making less trades, which lower their overall utility somewhat, but this decline is more pronounced for RS and particularly RM agents compared to RE agents. RM suffers more because in contrast to the basic model randomly selected agents are less likely to be available for trading at each time instant. Since RE identifies better trading partner and repeatedly uses the same trading partner unless required to change, RE out performs RS and RM. Similarly RS also out performs RM because it identifies some good partners but not to the extent RE is able to do. When m was increased from 2 to 10 in steps of 2 the performance of RE, RM and RS strategies improve but their performance difference is maintained throughout.

Continuous Production.

When agents are allowed to produce and trade in the same time period, there is a significant increase in the performance of RE, RS and RM strategies over the non-continuous production situation. This is because the agents could produce every turn, even if they had traded, and hence gain higher utility from these additional stocks. The relative performance of the three strategies follow trends similar to the non-continuous production case. With increase in the value of m the difference between RE and RS increases. On the other hand the difference between RS and RM reduces and stabilizes for 6 or more connections.

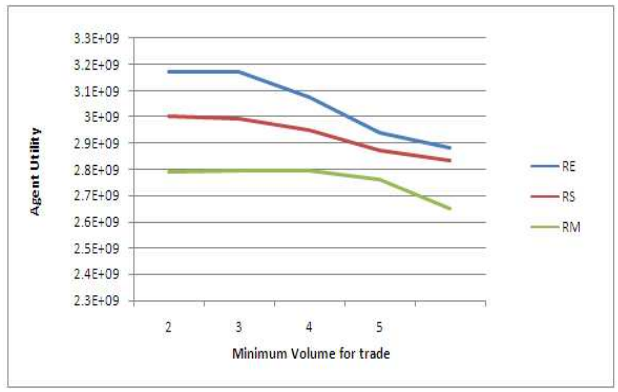

Figure 3: Effect of increasing $\underline{S}$ while holding $\bar{S}$ constant in homogeneous populations using the enhanced production and exchange model with continuous production.

Minimum Trade Volume and Storage Capacity Effects.

We next observe the effect of change in value of the Storage Capacity ($\bar{S}$) and Minimum Trade Volume ($\underline{S}$) on the performance of the rewiring strategies.

We hold $\bar{S}$ constant at 6 and increase the value of $\underline{S}$ from 2 to 6 in steps of 1. This variation significantly affects agents utility (see Figure 3). With increase in value of $\underline{S}$, the overall performance of agents gradually decreases. When $\underline{S}$=2 and $\bar{S}$=6, agents can produce till they find good trading partner to trade. With increase in value of $\underline{S}$, the trading window computed as the difference between $\bar{S}$ and $\underline{S}$ reduces. If an agent cannot find good trading partners within the trading window, it loses production opportunity as maximum storage limit is reached.

We have observed similar effects when we keep $\underline{S}$ constant and vary the value of $\bar{S}$.

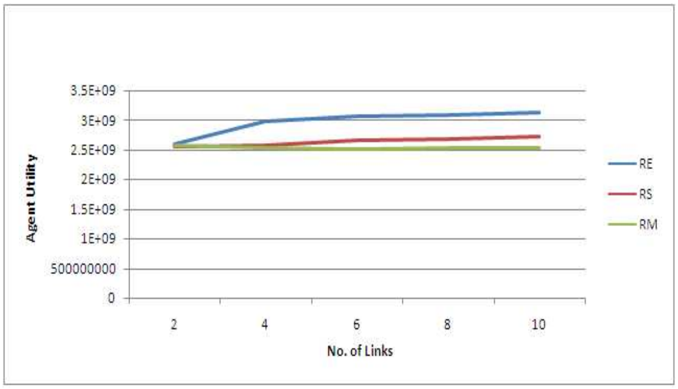

Figure 4: Effect of m in heterogeneous populations using the enhanced production and exchange model for non-continuous production..

We performed an additional experiment to compare the effects of different $\underline{S}$ and $\bar{S}$ while keeping the trading window, i.e., $\underline{S}-\bar{S}$, the same. We used two configurations: C1 with $\underline{S}$= 3 and $\bar{S}$=5, and C2 with $\underline{S}$= 4 and $\bar{S}$=6. In both cases the trading window is 2. We found that going from C1 to C2 increases the performance advantage of RE over RS and that of RS over RM. Also, with the increase in the number of connections, m, the performance of RE improves further compared to that of the performance of RS and RM.

4.2 Heterogeneous Populations

In the next set of experiments we experimented with heterogeneous agent populations. We included equal proportions of RM, RS, and RE strategies in a population of 300 agents.

Non-Continuous Production.

In this configuration, when $m = 2$, the utilities produced by all the strategies RE, RS and RM are almost equal (see Figure 4). But with increase in the value of m, e.g., when $m = 4$, agents have more trading partners per iteration when compared with $m = 2$, and able to locate desirable partners with less exploration. RE strategy produces higher utility than RS and RM strategies. For lower values of m, more exploration is necessary to locate compatible trading partners. For sufficiently high m values, therefore, the utilities of RE agents increase significantly over RS and RM.

When we compare the results of the heterogeneous population to the corresponding number of trading partners in the homogeneous population results (see Figure 1), we find that the RE strategy actually benefits at the expense of RM and RS strategy.

Continuous Production.

We also performed experiments with continuous production for heterogeneous populations. The trends are similar to the case of non-continuous production. The primary difference is that the agent utilities are higher as they have more stock to trade with.

5. DISCUSSIONS

We investigated the effects of introducing exploration into a rewiring strategy for locating effective trading partners within networks in production and exchange economies. Though random rewirings in each round can produce more utilities, it incurs significant cost for changing connections. The proposed decaying exploration rewiring strategy and a more patient random selection strategy incurs significantly lesser rewiring costs. Additionally, the exploration strategy provides certain benefits over random selection: it smooths out the rewirings over time and decreases the number of rewirings required. The performance advantage of the random rewiring strategy diminishes with higher number of connections per agent and when agents are allowed to produce while trading. Interestingly, however, the performance advantage is regained by the random rewiring strategy when all agent types are present in a heterogeneous society.

We believe that the basic production and exchange economy model is oversimplified and does not adequately represent real-life scenarios. We therefore evaluate the performance of the three rewiring strategies in an enhanced production and trade model that includes constraints on minimum trade volumes and storage capacities. In contrast to the basic model, the decaying exploration mechanism outperforms the more random rewiring strategies in this more realistic environments. This performance advantage also suggests the need for investigating smarter learning mechanisms for identifying preferred trading partners.

Acknowledgment.

This work is supported in part by a DOD-Army Research Office Grant #W911NF-05-1-0285.

6. REFERENCES

[1] L. Amaral, A.Scala, M. Barthelemy, and H. Stanley. Classes of small-world networks. *Proceedings of the National Academy of Sciences*, 97(21):11149–11152, 2000.

[2] M. E. Gaston and M. desJardins. Agent-organized networks for multi-agent production and exchange. In M. M. Veloso and S. Kambhampati, editors, *AAAI*, pages 77–82. AAAI Press; AAAI Press / The MIT Press, 2005.

[3] Y. L. Lederer-Antonucci, P. S. Greenberg, M. zur Muehlen, and G. Ralph. Establishing trust in a business-to-business collaboration: Results from an international simulation. In *Proc. of the IRMA 2003 Conference*, pages 922–924, 2003.

[4] A. Wilhite. Self-organizing production and exchange. *Computational Economics*, 21(1–2):107–123, 2003.