Journal of Universal Computer Science, vol. 11, no. 1 (2005), 193-209
submitted: 15/7/04, accepted: 6/12/04, appeared: 28/1/05 © J.UCS

Information Quality Assurance by Lazy Exploration of Information Source Combinations Space in Open Multi-Agent Systems

Jisun Park
(The Laboratory for Intelligent Processes and Systems
The University of Texas at Austin, USA
jisun@lips.utexas.edu)

K. Suzanne Barber
(The Laboratory for Intelligent Processes and Systems
The University of Texas at Austin, USA
barber@lips.utexas.edu)

Abstract: Information quality assurance under the existence of uncertainty can be investigated in the context of soft security, where an agent maintains trustworthiness evaluations of its information sources to assist in the evaluation of incoming information quality from those sources. Since dependency inherently exists in a system where agents do not have self-sufficient sensing or data collection capabilities, finding an appropriate set of information sources is important for assuring the quality of information and for increasing the agent's goal achievement. This research proposes an approach for selecting information sources as partners. In order to increase the efficiency and the accuracy, we use trustworthiness, information cost and goal coverage as the metrics for information valuation while adopting a lazy exploration of information sources combination space. Experimental results show that the proposed approach increases the efficiency and results in quality information acquisition.

Keywords: multi-agent systems, agent, information quality
Categories: I.2.8, I.2.11, H.3.3

1 Introduction

Information quality assurance under the existence of uncertainty can be investigated in the context of soft security [Rasmusson and Janson 1996], where an agent maintains trustworthiness evaluations of its information sources to assist in the evaluation of incoming information quality from those sources. Various trustworthiness evaluation mechanisms have been proposed to handle the uncertainty of both information and information sources [Dragoni and Giorgini 1997; Schillo, Funk et al. 2000; Barber and Kim 2003; Barber and Park 2003; Falcone, Pezzulo et al. 2003]. Trustworthiness evaluation can be used to find the best partners from whom to gather information. This research investigates the problems occurred by deploying the existing technologies of trustworthiness evaluation for information quality assurance, and proposes a heuristic to address the problems.

In Multi-Agent Systems (MAS), an agent's goals impose information requirements on each agent. This means that a set of information are required for an agent to achieve its goals. For example, the agents which track the locations of targets need at least a piece of information which can be transformed to the locations of the targets. Therefore, satisfying information requirements is a necessary condition for goal achievement. Dependencies exist between agents (and information sources) when each agent cannot be complete in its information acquisition capability and/or in tasks it can fulfill to achieve goals. If an agent is dependent on other agents, it is dependent on them with respect not only to capability, but also to the reliability meaning that depending on an unreliable entity can cause an agent itself to become unreliable. Therefore, in order for an agent to achieve its goals it is necessary to find and interact with the partners who can provide the required and reliable information. Finding the most appropriate information providing partners increases the quality of information as well as the achievability of the goals which require the corresponding information.

Deploying the notion of trustworthiness is a useful for handling the uncertainty and selecting the best partners. However, since most of the proposed trustworthiness evaluation mechanisms consider a single information source when evaluating the trustworthiness, it is necessary to repeat the evaluation process multiple times when there are many sources providing various types of information. Moreover, the trustworthiness evaluation of multiple information sources often requires the exhaustive search of possible source combination. If the exhaustive search of information combination space is not performed, it cannot be said that the best combination of information sources is selected given a trustworthiness evaluation mechanism. In other word, if an agent wants to find the best combination of sources given trustworthiness evaluation about the sources, the agent needs to find an information source combination which the agent is the most confident with, and the information source combination with which the agent is the most confident can be found by looking at every possible combination of information sources which is often very expensive. Therefore, we proposed a scheme to select the best information source combination efficiently by adaptively exploring the search space which consists of information source combination while keeping the quality of the resulting information. We interchangeably call the selected information sources as partners.

The partner selection for information quality assurance not only helps the goal achievement of the agents, but also contributes to the robustness of a system. According to [Schillo, Burckert et al. 2001], robustness is the ability to maintain “safety-responsibilities” [Wooldridge, Jennings et al. 1999] even with the occurrence of disturbing events. In other words, robustness needs to be related to faults in systems. In MAS, openness and uncertainty necessitate different kinds of faults handlings. Various replication schemes [Goodmand, Skeen et al. 1983; Davcec and Burkhard 1985; Budhiraja, Marzullo et al. 1993; Schneider 1993] have offered significant advances for fault-tolerance in traditional distributed systems, but they do not work in the face of maliciousness, innocuous quality degradation which are typical in open systems. Robustness can be enhanced by forming an information exchange with reliable partners, not in the way that agents confront the faults but by reducing the possibility of fault occurrence.

The goal of this research is to enhance reliability and robustness of an open MAS by deploying the notion of trustworthiness as well as information dependency, thus to assure the quality of information despite the level of uncertainty surrounding both information and sources. In particular, two main issues addressed by allowing an agent to adapt its dependencies are: *1) efficient partner selection* and *2) information quality assurance.*

The paper is organized as follows. In the next section, the overview of the problem and the approach is described. Section 3 provides the proposed algorithm, and in section 4, the experimental results are presented. Section 5 concludes the paper by summarizing the contributions and ongoing works.

2 Overview

When an agent needs information, the agent collects the necessary information either by sending requests for the information to an appropriate set of sources (in pull-based information acquisition system) or by taking the necessary information from the available sources of the information (in push-based information acquisition system). In either case, it is necessary for agents to distinguish the appropriate providers from an arbitrary set of sources. Thus, we do not limit the scope of selecting the partners to any one of the above cases.

In this section, the detailed description of the problem and the overview of the proposed approach are described.

2.1 Notation

When an agent ***a*** requires information from external sources to achieve its goals *G(a)*, the agent is dependent on both the information and the sources. The reliability (referred to as trustworthiness in the research) of the information provider, *s* from agent ***a***'s perspective is abstractly represented as $\phi_a(s)$. $R_a = \{r_1, r_2, ..., r_n\}$ is the set of information agent ***a*** requires. The *Information Pool (IP)* of agent ***a*** is a set of all tuples $< r_k, s, \Phi_a(s) >$, where *s* is a provider of r_k. The *Information Source Combination Pool (ICP)* of agent ***a*** is a set of tuple sets, where each tuple set $X_j = \{< r_1, s_1, \Phi_a(s_1) >, < r_m, s_k, \Phi_a(s_k) >, ..., < r_n, s_n, \Phi_a(s_n) >\}$ is a set of tuple combinations which satisfies the information requirements.

When an agent needs information, the agent is dependent on information sources providing the information it needs, constructing the relationship (i.e. dependency) with the sources of the information. Unilateral relationships (dependencies) are dominated by one end of the provider-consumer pair, the consumer. If a consumer filters out bad providers and the providers are not concerned about it or are unable to influence the filtering process, the relationship is unilateral. In this case, what the consumer evaluates about the providers dominates the relationship determination. Mutual relationships can be constituted by agreement among the stakeholders (i.e. both providers and consumers). For most cases, mutual relationship can be realized when the existence of the dependency increases the benefits for involved participants. We investigate the unilateral relationship so that the agents can determine their dependencies on their own intentions.

2.2 Valuation of Information Sources

The valuation of information sources is defined in terms of three factors – goal coverage, information cost, and trustworthiness. The combined measure of the information source's value is a weighted sum of the three factors and the weight can be decided depending on the application necessity.

2.2.1 Information Cost

When an agent has unlimited resources for the information acquisition process, the agent should pursue the highest quality information possible. However, an agent is often limited by the information costs it can afford; therefore, the agent must address the quality and efficiency tradeoffs between acquiring quality information at reasonable cost [Park and Barber 2004].

In this research, information cost is derived from the message-passing and computational burden required to communicate information. In many information networks, such as ad-hoc networks, agents are not connected to every source. As a result, information must pass through other sources (sources willing to relay information) to arrive at the requesting agent, and consequently information cost is increased. In this research, we assume that information cost is directly proportional to the length of the path from the requesting agent to the original information source (we denote the cost of passing information along one link as a single cost unit). However, an agent cannot be assured of the identity of the original information provider; it only knows the resulting information cost it is asked to pay by its neighbor (its direct information provider not requiring relaying sources). Neighboring information sources have an incentive to report truthful information costs; requesting artificially high costs may discourage the agent from utilizing that source (see trustworthiness valuation in Section 2.2.3).

Figure 1 demonstrates a network of agents; each agent also serves as an information source to other agents. The network is drawn as an incomplete graph, in which connections between agents signify communication links. For example, in this graph, a_1 can only communicates directly with a_2, a_3, a_5, and a_6. If a_1 requires information from a_2 which must be obtained from a_8, who, in turn, must obtain the information from a_9, a_2 will convey to a_1 at a cost of three units. If, however, a_3 is an original provider of the same information, a_3 may convey to a_1 at a cost of one unit.

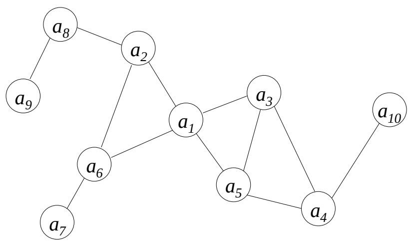

Figure 1: Agents are distributed information sources with limited information acquisition and communication.

To determine the cost incurred by an agent in fulfilling a set of its goals, we define the following properties:

N(a) = {s | s is a neighbor of a}
$R_a(g)$ = {r | r is an information element required to achieve agent a's goal, g}
PROV(s) = {r | r is an information provided by source s }.

N(a) is the set of neighbors of a, $R_a(g)$ is the set of information which must be held by agent a in order for a to achieve its goal g, where $\bigcup_g R_a(g) = R_a$, and *PROV(s)* is the set of information which s provides, either originally or after obtaining from other sources. The total cost, *TotalCost(a, g)*, incurred by agent a in achieving a single goal g is calculated as:

$$TotalCost(a,g) = \sum_{s,r} Cost(s,r)$$

, where *Cost(s,r)* denotes the cost of information r from source s, and

1. $s \in N(a)$ (s is a neighbor of a),
2. $r \in PROV(s)$ (s provides information r), and
3. $r \in R_a(g)$ (r is required for a's goal g).

2.2.2 Goal Coverage

An agent may not possess all the information it requires. Thus, an agent needs to obtain information from the information sources which are capable of satisfying its information requirements. The relevance of a set of information from a set of information sources is decided by the degree to which the agent's information requirements are met. The notion of requirement priority is introduced to describe the importance of each requirement. The priority on each information requirement is represented by *PRIO(a, r)*, and the assignment of priority value is decided by the information requesting agent with the assumption that the information requirements are mutually exclusive. The priority assignment scheme used in this research is to give a higher priority to the information requirements which contribute to more goals. In that case, the priority of an information requirement is defined as follows:

$$PRIO(a,r) = \sum_{all\ n,\ s.t.\ g_n \text{ requires } r_k} \frac{k_{g_n}}{\text{number of information elements required by } g_n},$$

$$\text{,where } k_{g_n} \text{ is a weight on } g_n \in G(a)$$

This definition of information requirement priority enables the expression of information relevance based on the importance of requirements and contribution of the information to each goal. Figure 2 offers a priority assignment example.

$$
\begin{array}{|l|}
\hline
(1)\ A = \{a_1, a_2, a_3, a_4\} \\
(2)\ G(a_1) = \{g_1, g_2\},\ k_1 = 2, k_2 = 1 \\
\quad g_1 \text{ requires } \{r_1, r_2\} \\
\quad g_2 \text{ requires } \{r_2, r_3\} \\
(3)\ R(a_1) = \{r_1, r_2\} \cup \{r_2, r_3\} = \{r_1, r_2, r_3\} \\
\hline
\end{array}
$$

Figure 2: Priority Assignment Example

This example is derived from a_1's point of view where a_1 has two goals g_1, g_2. g_1 is assigned a weight of 2 ($k_1=2$) and requires two information elements r_1, r_2. g_2 is assigned a weight of 1 ($k_2=1$) and requires two information elements r_2, r_3. Priorities on the required information are determined based on the priority calculation given above (*PRIO*(*a*, *r*)) as follows.

$$
\begin{aligned}
PRIO(a_1, r_1) &= \tfrac{2}{2} = 1 \\
PRIO(a_1, r_2) &= \tfrac{2}{2} + \tfrac{1}{2} = 1.5 \\
PRIO(a_1, r_3) &= \tfrac{1}{2} = 0.5
\end{aligned}
$$

Relevance is measured in terms of priorities on information requirements. Relevance measures include source coverage, goal coverage, and total coverage. Source coverage represents the relevance of information provided by a specific information source for a goal. Goal coverage is defined to be the relevance of information provided by a set of information sources for a goal. Total coverage represents the relevance of information provided by a set of information sources for all goals.

Source coverage is represented by *SourceCoverage(a, s, g)* where *s* is a information source and *g* is a goal of agent *a*. It represents each source's contribution to the requirements of a single goal. Goal coverage is represented by *GoalCoverage(a, S, g)* where *S* is a set of information sources and *g* is a goal of agent *a*. Goal coverage can be calculated by performing a set union operation on each source's contribution to the requirements. *TotalCoverage(a, S, G)* is a total coverage of a set of information sources for goal set *G*. Goal coverage can be calculated by performing a set union operation on the requirements satisfaction by all sources and is normalized to be between 0 and 1. Formal representations of the coverages are presented as follows:

$$
SourceCoverage(a, s, g) = \frac{\displaystyle\sum_{\substack{all\ r_i, s.t. g \text{ requires } r_i \\ \wedge r_i \in PROV(s,a)}} PRIO(a, r_i)}{\displaystyle\sum_{all\ r_j,\ s.t.\ g \text{ requires } r_j} PRIO(a, r_j)}
$$

$$
GoalCoverage(a, S, g) = \frac{\displaystyle\sum_{\substack{all\ r_i, s.t. g \text{ requires } r_i \\ \wedge r_i \in \bigcup_{s \in S} PROV(s,a)}} PRIO(a, r_i)}{\displaystyle\sum_{all\ r_j,\ s.t.\ g \text{ requires } r_j} PRIO(a, r_j)}
$$

$$TotalCoverage(a,S,G) = \frac{\sum_{all\ g \in G} GoalCoverage(a,S,g)}{|G|}$$

If an agent is concerned only about the relevance of information when selecting partners, the objective is to maximize *TotalCoverage*. However, there can be multiple different source combinations which maximize *TotalCoverage*. The choice of information source combination can affect the robustness of the goal achievement. Assume there are 2 instances of source combinations which maximize *TotalCoverage* where one instance consists of the information sources with low *SourceCoverage* values and another instance of source combination consists of the information sources with high *SourceCoverage* values. If an agent prefers the first case - sources with low *SourceCoverage*, the agent is less dependent on the undesirable (as well as desirable) behavior of those information sources. On the other hand, if an agent prefers the second case – sources with high *SourceCoverage*, the agent is more dependent on the undesirable (as well as desirable) behavior of those information sources. The decision of which information source combination to choose in those cases is a design consideration of agent designers.

2.2.3 Trustworthiness

Trustworthiness of an information source can be represented by the probability that the information provided by the source is true or the probability distribution of the error, from the true value, of the provided information [Fullam and Barber 2004]. Since the agents do not know the true values of information provided to them, the confidence or certainty the agents convey on the beliefs depends on the quality of the provided information, where the quality of information is defined as the accuracy of the information based on the estimated true values. Although we do not limit our trustworthiness evaluation to a specific evaluation mechanism for generality, the trustworthiness evaluation algorithm proposed in [Fullam and Barber 2004] is used for experiments and implementations because the algorithm effectively represents the quality of information provided in terms of accuracy and consistency. In [Fullam and Barber 2004], a belief revision algorithm based on a set of policies for information valuation is proposed. Belief is an agent's perspective model of truth, or something believed as true, on some subject at some time. The policies include the preference to the information from the reliable sources with high certainty in information quality as well as the preference to agreed upon information from as many sources as possible. We use the trustworthiness model used for the policy of preference to source reliability. In this model, belief distribution mean μ_B and belief distribution standard deviation σ_B are used to represent a belief. Source distribution is a distribution of source reports represented by source distribution mean μ_{s_i} and source distribution standard deviation σ_{s_i} . The trustworthiness of an information source (called reputation or reliability in [Fullam and Barber 2004] is modeled as a distribution of source report errors. Since the reliability of information source concerns the errors of the reported information and the agent does not know the truth value of the information, the distribution ρ_{s_i} of the source report errors use the mean difference α between the reported values and the belief of the agent. Therefore, the distribution ρ_{s_i} after N timestep is represented by its mean μ_{ρ,s_i} and standard deviation σ_{ρ,s_i} .

$$\mu_{\rho,s_i} = \frac{\sum_{t=1}^{N}\left(\frac{1}{\sigma_B(t)\sigma_{s_i}(t)}\alpha(t)\right)}{\sum_{t=1}^{N}\left(\frac{1}{\sigma_B(t)\sigma_{s_i}(t)}\right)}$$

$$\sigma_{\rho,s_i} = \sqrt{\frac{\sum_{t=1}^{N}\left(\frac{1}{\sigma_B(t)\sigma_{s_i}(t)}(\mu_{\rho,s_i}-\alpha(t))^2\right)}{\sum_{t=1}^{N}\left(\frac{1}{\sigma_B(t)\sigma_{s_i}(t)}\right)}}$$

, where $\alpha(t) = \mu_{s_i}(t) - \mu_B(t)$. Since the mean value describes the accuracy and the standard deviation describes the consistency of the information source, we define the trustworthiness of an information source as a linear combination of those two values as:

$$\phi_a(s) = \frac{1}{1+e^{B(\xi|\mu_{\rho,s}|+(1-\xi)\sigma_{\rho,s}-z)}}, \text{ where } 0 \le \xi \le 1$$

The weight factor ξ is decided depending on whether an agent ***a*** values accuracy or consistency. B is the growth rate and ***z*** is a domain-specific bias parameter. Depending on the domain-specific bias parameter, an agent can take an optimistic or pessimistic trustworthiness evaluation approach. When an agent evaluate the trustworthiness of multiple sources, it is reasonable to take an average of the selected sources.

Since we have defined the metrics for valuating the information sources, the partner selection process will be presented from the next section in more detail.

2.3 Information Sharing Networks

Information Sharing Networks (ISN) refer to a system where the agents share their information either by providing necessary information or by using (consuming) the information provided by other agents. The partner selection process aims to filter out bad information sources; thus, the objective is to consume only good information from good sources. Good sources are the information sources which provide information close to the true value so that the agents' confidence on its resulting beliefs based on that information is high. In the same context, the best information sources are the information providers which provide information which is closest to the truth value among the potential providers, so that the confidence of the agents about the resulting beliefs is the highest possible. The naïve approach to finding the best information sources is to investigate every possible combination of information sources which satisfies an agent's information requirement for its goals and pick the one which yields the highest confidence. Figure 3 depicts this naïve approach. There are two main concerns in the process. The first is to build the *Information Source Combination Pool* (ICP) from the *Information Pool* (IP) Recall that the *Information Source Combination Pool* (ICP) is a set of all information sources combination satisfying the information requirement and *Information Pool* (IP) is a set of mappings from an information requirement to an information source with trustworthiness value (see Section 2.1 for formal definitions). The size of ICP increases exponentially as the amount of information $(|R|)$ and corresponding sources $(|S|)$ increase. The second is to find the best (or near-best) set of information and sources from ICP.

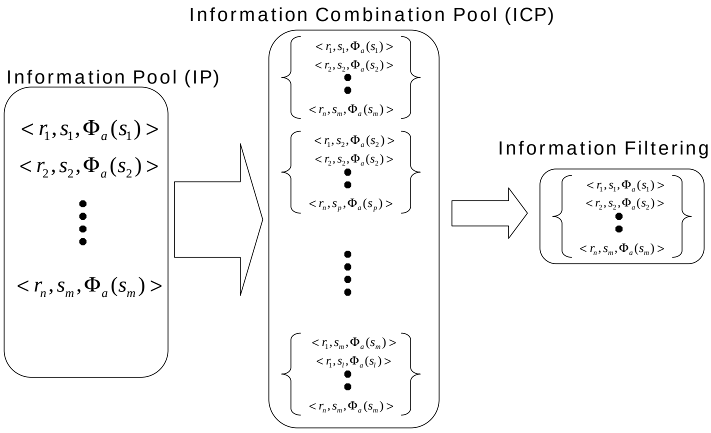

Figure 3: Naïve Approach for Partner Selection

The size of ICP depends on the agent's information requirements as well as the number of sources satisfying those information requirements. Let ***N*** be the number of information sources $(|S|)$, $n(r_i)$ be the number of potential sources for information r_i. If an agent requires ***M*** information $(M = |R_a|)$, the possible amount of information and information source combinations (|ICP|) is $\prod_{i=1}^{M}\left(\sum_{j=1}^{n(r_i)} {}_{n(r_i)}C_j\right)$, which increases exponentially as the number of information required increases. As an example, suppose agent ***x*** requires information $\{r_1, r_2, r_3\}$, information source 1 provides $\{r_1, r_3\}$, source 2 provides $\{r_1, r_2\}$, and source 3 provides $\{r_2, r_3\}$ (Figure 4). In this case, the total number of possible combinations is 27. Comparing all the elements in ICP is a simple way to find the best ISN, but it is significantly complex in computation and memory. Table 1 shows the size of ICP for different situations.

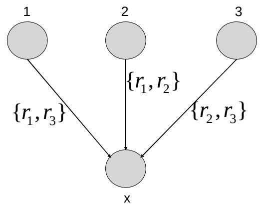

Figure 4: Potential information sources

M	$n(r_i)$	$\|ICP\|$
1	1	1
5	3	16807
10	5	$8.1963\text{x}10^{14}$
50	10	$3.1173\text{x}10^{150}$

Table 1: Size of ICP for different M and $n(r_i)$

2.4 Approach

Because of the large size of ICP, it is expensive to find the best combination of information sources by exploring the whole ICP. Instead of the exhaustive search in ICP, we propose a lazy exploration of the search space. In order to structure the search space and make this approach efficient, we make use of the notion of dependency. It is reasonable to assume that information dependency is inherent since all the agents can not always be self-sufficient, having all the information they need for all their goals. Information dependency can be quantified by the following equation, where $N_p = |S_p|$, S_p is a set of sources which are selected as information sources $(N_p \leq N)$, also $n_p(r_j)$ is the number of selected information sources for information r_j

$$Dependency(a) = \frac{N_p \sum_{all\ j, \text{s.t. } r_j \in R_a} n_p(r_j)}{\sum_{all\ k, \text{s.t. } r_k \in R_a} n(r_k)}$$

For example, assuming agent ***x*** filters out information from source 2, the dependency from Figure 4 is 4x2/6 = 1.33. Since the dependency is represented in terms of information itself and the information sources, dependency increases if an agent receives more information from a fixed number of sources as well as if an agent receives information from more sources. Figure 5 shows an ICP instance from Figure 4, which is represented by a lattice derived using the notion of dependency. The ICP lattice is populated by nodes which are the possible combinations of information sources satisfying the information requirement except for the null bottom node. Bottom node is not a reachable node, and there exist edges between nodes when they have minimal dependency difference. Minimal dependency difference exists between two combinations of sources when they are different only in one element. Therefore, the minimal dependency difference is gained by addition or reduction of one source to/from a node. The parents of a node have higher dependency and the children of a node have lower dependency. Each node is denoted by $(S_1, S_2, ..., S_N)$, where S_k is a set of information sources providing r_k. In Figure 5, we can see that the node $(\{2\},\{3\},\{3,1\})$ has two parents which are $(\{2\},\{2,3\},\{3,1\})$ and $(\{1,2\},\{2\},\{3,1\})$, and has two children which are $(\{2\},\{3\},\{3\})$ and $(\{2\},\{3\},\{1\})$.

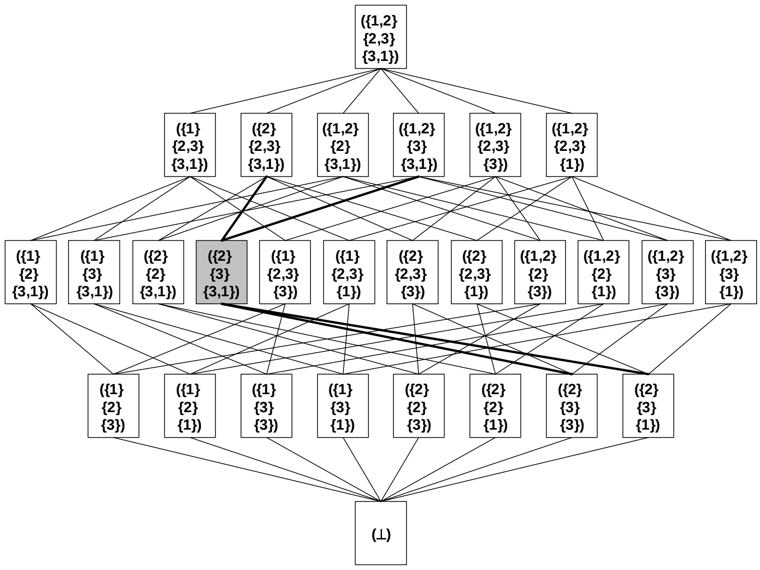

Figure 5: Information Source Combination Pool built from Figure 4

3 Algorithm

At a given state, the selection of the best combination of information sources is performed by using the most recent reliability evaluations of the sources. In each node in ICP, we have an aggregate reliability of the corresponding source combination. The aggregate reliability $\Phi_a(C)$ of S is calculated by the following equation.

$$\Phi_a(C) = \sum_{r_i} avg(\phi_a(s)), \text{ where } s \in C \text{ and } r_i \in PROV(s)$$

Given a previous information source combination which is a node in ICP and the most recent evaluation of the sources, an agent looks for the parents and children of the node. The node which has the highest reliability $\Phi_a(S)$ is selected as a new information source combination. Since the agent does not investigate the whole ICP lattice, the approach is called a lazy exploration. When the information sources are in a relatively static (reliability of sources is constant and openness does not result in any information source additions or deletions), the best combination is reached quickly.

When the information sources are relatively dynamic with respect to the reliability and openness, the selected source combination may not be the best combination, but since the selection is refined at each timestep, an agent gradually reaches a better or "good-enough" source combination.

Figure 6 summarize the proposed algorithm.

```
Given:
  R_a: information requirements
  S: potential sources
  φ_a(s) : most recent evaluations of potential sources (trustworthiness)


PartnerSelection (a_i):
    build parents nodes and children nodes
    for current node, parents, children
      calculate Φ_a(C)
   select a information source combination C such that
      C = argmax_{C'} Φ_a(C') , C'∈ {current node, parents, children}
```

Figure 6: Partner Selection by Lazy Exploration

4 Experiments

Experiments were performed in a UAV target tracking domain where UAVs are agents tracking targets and UAVs are sources sharing detected target locations. In this domain, there are moving targets, and information sources (UAVs) track the location of the moving targets. The location of the target is represented by a Cartesian coordinate of floating numbers. Information sources are assumed to be imperfect meaning there exist errors in the locations provided to the agent (UAV). There can be bad sources which have relatively higher errors in the target location information they provide.

In order to simulate the partner selection process, a Topology UI was implemented as in

Figure ***7*** and Figure 8. In the Topology UI, it is shown which information sources provide the necessary information and which information sources are selected as partners.

Figure ***7*** represents all the communications connections between partners; thus the density of connectivity. Each node represents the potential information sources and the edges represent the selection of information sources. Therefore, if there is an edge between two nodes, the two agents (information sources) exchange the information either in a unidirectional way or in a bidirectional way. Figure 8 shows agent 1's connections to its selected partners. Each agent node is colored based on agent 1's evaluation of the respective agent's trustworthiness. The legends for categories of trustworthiness are given in the left lower panel in Figure 8.

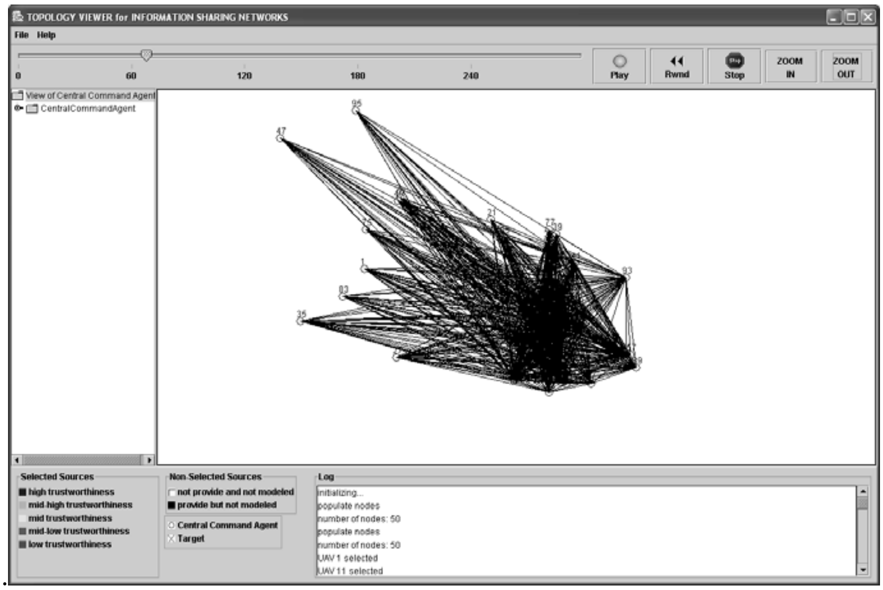

Figure 7: Topology UI (Global View)

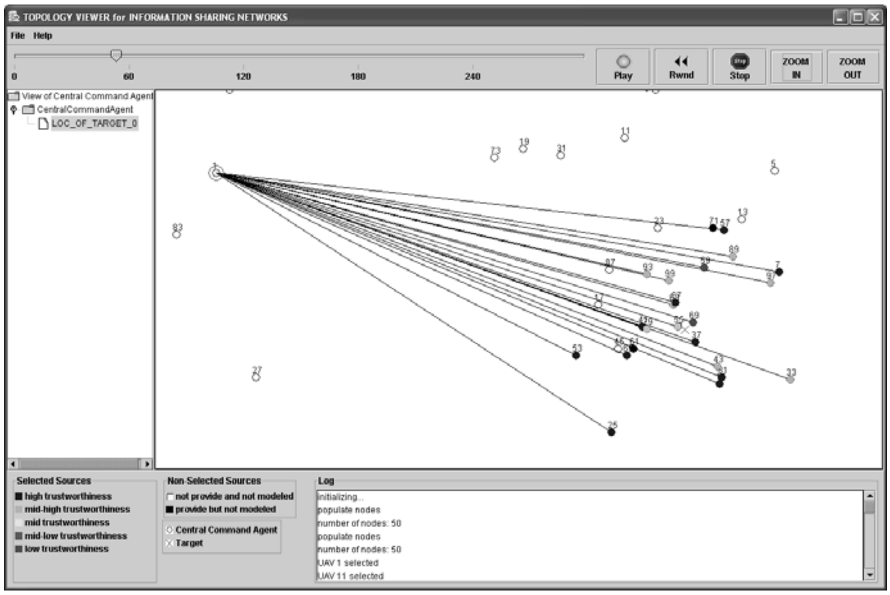

Figure 8: Topology UI (Local View)

In order to compare the efficiency of the proposed scheme, we compare the actual running time of the naïve exhaustive exploration of the search space and the proposed lazy exploration by measuring the CPU time of building the beliefs (locations).

Accuracy of the lazy exploration is also compared with that of the naïve approach. Mean Square Error (MSE) is used for accuracy measure and the error is a weighted sum of deviations from the true location and standard deviation of the error distribution.

Figure 9 shows the efficiency between the naïve exhaustive approach and the lazy exploration for different number of UAVs. In this experiment, CPU time for iterating 100 timesteps is measure for each of the approaches with varying the number of UAVs. In the case of exhaustive search, the CPU time increases exponentially as the number of UAVs increases. However, the CPU time for lazy exploration increases almost linearly.

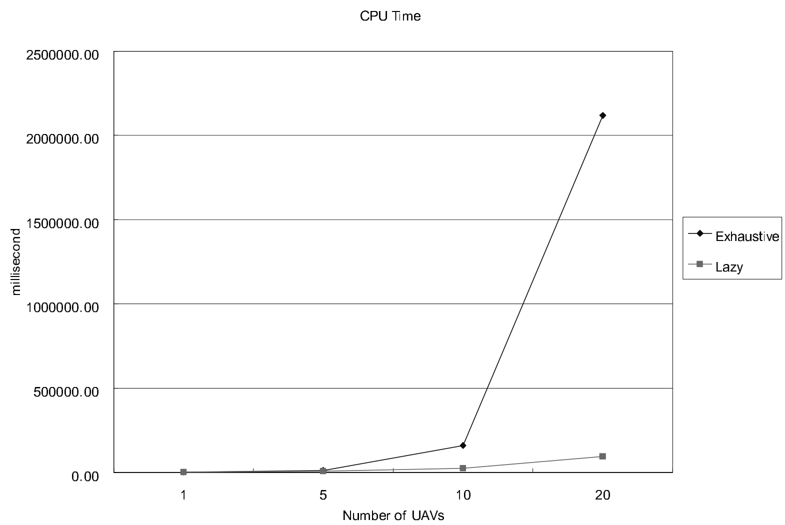

Figure 9: Comparison of running time

Figure 10 shows the accuracy of the naïve approach and the lazy exploration for different number of UAVs. In this experiment, 100 timesteps are iterated to calculate the Mean Square Error of each approach. The decrease in the quality of acquired information is very small considering the enhancement in efficiency.

In the above experiments, we can verify that the lazy exploration of the information sources combination space results in “good-enough” information very efficiently.

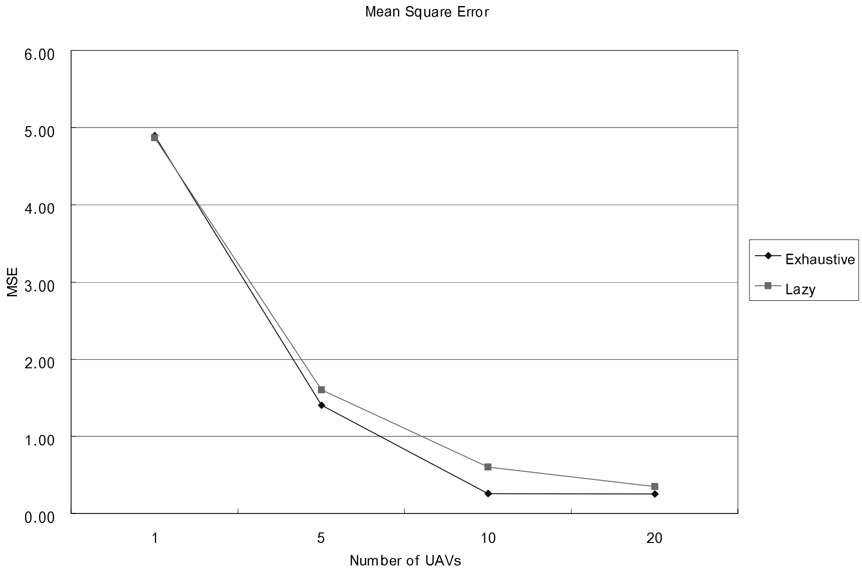

Figure 10: Comparison of Accuracy

5 Conclusions

Information quality assurance under the existence of uncertainty can be investigated in the context of soft security, where an agent maintains trustworthiness evaluations of its information sources to assist in the evaluation of incoming information quality from those sources. Since dependency inherently exists in a system where agents do not have self-sufficient sensing or data collection capabilities, finding an appropriate set of information sources is important for assuring the quality of information and for increasing the agent's goal achievement. In order to valuate the information sources, we proposed three metrics – information cost, goal coverage and trustworthiness. Information cost of acquiring necessary information is derived from the message-passing and computational burden required to communicate information. Goal coverage is the amount of contribution information sources make. Regarding the trustworthiness evaluation, we adopt an approach which takes into account various policies for evaluating the information. The policies include the priority to maximum, corroborated information and to the reliable sources with high certainty. With the metrics for information source valuation, lazy exploration of information source combination space is adopted for efficiently find the best partners to result in the best or "good-enough" information. Lazy exploration is efficient since the search space expands exponentially as the number of information sources increases. Compared to a naïve exhaustive search, the lazy exploration shows almost a linear increase of time complexity as the number of information sources increases. While the proposed approach is efficient, experiment results also shows the accuracy of the approach.

Acknowledgement

This research was supported in part under a subcontract with ScenPro, Inc. funded by U.S. Department of Defense, Small Business Technology Transfer (STTR) Program, Contract Number F30602-99-C-0193.

References

K. S. Barber and J. Kim (2003). Soft Security: Isolating Unreliable Agents from Society. *Trust, Reputation, and Security: Theories and Practice*. Lecture Notes in Artificial Intelligence M. Singh, Springer**:** 224-234.

K. S. Barber and J. Park (2003). Autonomy Affected by Beliefs: Building Information Sharing Networks with Trustworthy Providers. In Proc. of workshop on Autonomy, Delegation, and Control at the 2nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS2003), 11-16 Melbourne, Australia.

N. Budhiraja, K. Marzullo, et al. (1993). The Primary-Backup Approach. *Distributed Systems*. S. Mullender, Addison-Wesley**:** 199-216.

D. Davcec and A. Burkhard (1985). Consistency and Recovery Control for Replicated Files. In Proc. of the 10th Symposium on Operating Systems Principles, 87-96.

A. F. Dragoni and P. Giorgini (1997). Learning Agents' Reliability through Bayesian Conditioning: a simulation study. In Proc. of Learning in DAI Systems, Springer-Verlag.

R. Falcone, G. Pezzulo, et al. (2003). A fuzzy approach to a belief-based trust computation. *Lecture Notes on Artificial Intelligence*. **2631:** 73-86.

K. Fullam and K. S. Barber (2004). Using Policies for Information Valuation to Justify Beliefs. In Proc. of the Third International Conference on Autonomous and Multi Agent Systems (AAMAS 2004), 404-411 New York City, New York.

N. Goodmand, D. Skeen, et al. (1983). A Recovery Algorithm for a Distributed Database System. In Proc. of the 2nd ACM SIGATC-SIGMOD Symposium on Principles of Database Systems.

J. Park and K. S. Barber (2004). Finding Information Sources by Model Sharing in Open Multi-Agent Systems. In Proc. of Workshop on Agents for Ubiquitous Computing (UbiAgents) at AAMAS 2004, New York City, NY.

L. Rasmusson and S. Janson (1996). Simulated social control for secure Internet commerce. In Proc. of New Security Paradigms '96, ACM Press.

M. Schillo, H. J. Burckert, et al. (2001). Towards a Definition of Robusness for Market-Style Open Multi-Agent Systems. In Proc. of The Fifth International Conference on Autonomous Agents (Agents-01), 75-76.

M. Schillo, P. Funk, et al. (2000). "Using Trust for Detecting Deceitful Agents in Artificial Societies." *the Applied Artificial Intelligence Journal, Special Issue on Deception, Fraud and Trust in Agent Societies*: 825-848.

F. B. Schneider (1993). Replication Management using the State-Machine Approach. *Distributed Systems.* S. Mullender, ACM Press - Addison Wesley**:** 169-197.

M. J. Wooldridge, N. R. Jennings, et al. (1999). A Methodology for Agent-Oriented Analysis and Design. In Proc. of Third International Conference on Autonomous Agents (Agents-99), 69 - 83 Seattle, WA, ACM Press.