A POMDP Based Approach to Optimally Select Sellers in E-Marketplaces

Paper XXX

ABSTRACT

In e-marketplaces, buying agents choose among multiple sellers, finding the best one to perform a transaction. A seller's quality is mainly determined using buyers' assessments on the seller based on previous transactions. How a buyer utilizes existing information and chooses the right seller, while maximizing its utility derived from the decision process is crucial, since the value derived from a successful transaction with the seller, may be negligible relative to the cost of choosing the right seller. This paper presents SALE POMDP, a framework based on the *Partially Observable Markov Decision Process*, which helps a buyer to optimally select the right seller to perform a transaction. The framework is robust in the sense that it also allows to reason about the quality of other buying agents (advisors) in the market, who provide opinions about the sellers. Experimental results on the ART Testbed demonstrate that SALE POMDP can optimally select the right sellers and advisors (to ask opinions), effectively balancing the trade-off between the cost of obtaining and benefit of more information, than traditional trust models.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Intelligent Agents; Multiagent Systems

General Terms

Performance; Design

Keywords

Seller Selection, E-Marketplace, POMDPs.

1. INTRODUCTION

In multi-agent based e-marketplaces, self-interested selling agents can act maliciously by not delivering products with the same quality as promised. It is thus important for buying agents to analyze the quality of sellers and determine which sellers to do business with. Buyers maintain beliefs over the quality levels of the sellers, based on their previous transactions with the seller, using which they can select the best seller to interact. However, realistically, in most e-marketplaces, buyers often encounter sellers with which they have no previous experience. In such cases, they can query other buyers (called advisors) about their beliefs on the sellers' quality.

A number of trust models (*e.g.* BRS [16], TRAVOS [15], Personalized [17], BLADE [11], etc.), have been proposed by researchers in the multi-agent community to help buyers assess seller quality and choose business partners. These approaches work by combining the buyer's own belief and those of the advisors, to estimate the true quality of the seller. However, the above approaches mainly focus on accurately estimating the quality of the seller rather than optimally choosing a right seller to perform transaction with. They do not consider the underlying trade-off between the cost of estimating seller's quality and the value derived from a successful transaction with the seller. This is because the above trust models fail to reason *when* it is necessary to query advisors, though they may determine *whom* to query by analyzing the quality levels (trustworthiness) of the advisors who provide accurate opinions[1] about the sellers. Most approaches simply query all advisors whom they consider honest. This will in fact result in negligible utility for the buyer, especially when the querying advisors are large in numbers, since the cost of querying advisors may be greater than the value derived from a successful transaction with the seller. Rather than querying all potential advisors, the buying agent may query advisors until it is sufficiently sure that it has identified a seller with the expected quality. This is a strong motivation behind our work where we consider the seller selection problem as a decision problem, resolving the trade off between the cost of selecting the right seller and the value derived from the transaction with the seller.

POMDPs provide such a framework for making optimal decisions in partially observable environments, where an agent cannot directly observe the underlying state. A POMDP agent works towards the main goal of maximizing the expected discounted future reward [4] (buyers' utility in our case). Modeling the buying agent as a POMDP will help to optimally select the right seller (rather than accurately determining the quality of the sellers). Regan *et al.* [10] proposed the *Advisor* POMDP, which also considered the seller selection problem as a *Partially Observable Markov Decision Process*. However, their approach did not reason about advisors' quality levels and considered all advisors to be trustworthy. Thus, *Advisor* POMDP cannot cope with unfair rating attacks [3], where advisors may provide incorrect opinions about the sellers. Also, in *Advisor* POMDP, each advisor provides opinions about its beliefs on all sellers, when queried, resulting in a lot of unnecessary information. However, rather than estimating the quality of all sellers, the only goal should be to select the seller with high quality.

---

[1]opinion refers to the advisor's belief on the seller

In this paper, we present the detailed analysis of *(S)eller & (A)dvisor se(LE)ction POMDP (SALE POMDP)* [7]. SALE POMDP models the buying agent as a POMDP, with the main goal of optimally selecting the right seller to perform a transaction. In this process, it maintains the beliefs over the quality levels of the sellers, as well as that of the advisors. To query advisors about the quality of the sellers, it firstly determines the quality of the advisors, thereby limiting its interactions to the honest ones. It also makes provision to query advisors about their beliefs on other advisors, sufficiently using it to improve its own beliefs and make optimal decisions. We also analyze how SALE POMDP can be extended to be more usable in e-market scenarios, effectively dealing with the seller selection problem in a holistic manner than the traditional trust models. Specifically, we have: 1) modeled SALE POMDP as a *factored* POMDP, to resolve the scalability issues, which is a limiting factor in solving most POMDPs [9]; 2) modeled the advisor behavior (in providing unfair ratings) to be more realistic. Specifically two types of attacking behavior are considered, dishonest advisors who exhibit *random* behavior in providing opinions about other agents and those who behave in an *adversarial* manner, providing complementary opinions; 3) allowed to consider the influence of previous responses (history) in the POMDP formulation and 4) performed a thorough evaluation of the SALE POMDP model using the ART Testbed, comparing its effectiveness against other trust models.

The rest of the paper is organized as follows. Section 2 gives a brief discussion of related work and continues to argue the importance and uniqueness of the SALE POMDP framework. The detailed description of the SALE POMDP model is present in Section 3. In Section 4, we demonstrate how SALE POMDP is can be modeled as a *factored* POMDP. In Section 5, through extensive experimentation on the ART testbed, we demonstrate the effectiveness of SALE POMDP in optimally selecting the right sellers to perform transaction. The scalability of the *factored* approach and methods to improve the robustness of the SALE POMDP model are also discussed. Finally, Section 5 concludes the current work and proposes future work.

2. RELATED WORK

Many trust models have been proposed in literature to address the seller selection problem. They model sellers' quality (reputation), using opinions from other buyers and thereby select the best seller to interact with. The Beta Reputation System (BRS) [16] models the seller's reputation as the expected value of the beta probability distribution of the trusted advisors' beliefs about the seller. If the calculated reputation of a seller based on the opinions from the set of honest buyers falls within the rejection area (q quantile or $1-q$ quantile) of the beta distribution of an advisor's beliefs on the seller, then the advisor is considered untrustworthy. BRS is a centralized approach and is heavily based on the *majority rule*. Teacy *et al.* [15] propose TRAVOS which is also based on beta probability distribution. To calculate the reputation of a seller, it integrates buyer's own beliefs about sellers as well as the beliefs of advisors, after assessing their credibility. The personalized approach [18], models seller's quality as a weighted score of its public and private reputation. The private reputation is calculated as the expected value of the beta probability distribution of the buyer's beliefs on the seller and the public reputation is calculated us-

ing the beliefs of the advisors, after discounting them based on their trustworthiness. However, it assumes the behavior of the advisors to be consistent across all sellers.

The BLADE approach of Regan et al. [11] applies Bayesian learning to reinterpret advisors' beliefs instead of filtering them. The BLADE model allows a buyer to learn other advisors' evaluation functions on different features of the services delivered by sellers, by analyzing their beliefs. This makes it possible to adjust the advisor's beliefs, thereby coping with subjectivity and deception, while determining the quality of the seller. However, the model assumes consistent advisor behavior and also requires multiple transactions with the same seller to accurately interpret its quality. In contrast, in SALE POMDP we learn about advisors by asking other advisors, thereby avoiding the need to engage in costly transactions. Also, all the above trust models do not provide optimal decision making for the buyer on whether and from which seller to perform transaction, which is exactly what our approach tries to offer.

Few works based on POMDPs exist to address the seller selection problem using the notion of trust. The *Advisor* POMDP [10], described by $\langle \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \Omega, \mathcal{O} \rangle$, where $\mathcal{S}$ represents the possible states consisting of the seller's quality and the outcome of the transaction with seller, $\mathcal{A}$ represents the actions $\in ask, buy$, $\mathcal{T}$ represents the transition functions, $\mathcal{R}$ represents the rewards based on the outcomes of the transaction, Ω represents the observations from advisors, consisting of the advisors' beliefs about the sellers along with the certainty and $\mathcal{O}$ denotes the observation function, updates its belief about the seller via Bayes' rule, on interacting with other advisors in the environment, thereby selecting the right seller to perform transaction. However, *Advisor* POMDP does not consider advisor's quality to be a part of the state space, making it vulnerable to unfair rating attacks. Also the observations in *Advisor* POMDP contain beliefs about all sellers, resulting in an unnecessary exchange of information during the *ask* actions. Seymour *et al.* [12] proposed TI-POMDP, which adds trust modeling to the I(interactive)-POMDP framework [2]. I-POMDP is a multi-agent extension of the POMDP framework, in which each agent maintains beliefs about both the physical states of the world and the decision process models of the other agents. The TI-POMDP maintains the basic components of the I-POMDP, and adds trust modeling as a primary decision factor for the agents. In addition to the state belief model, an agent maintains and updates a trust model (a rating of the trustworthiness of the other agents) for the environment. This trust model contains an agent's level of trust in the other agents, which helps to decide whether or not to cooperate with another agent on a given task. However, this model depends heavily on communication between the agents for decidability [6]. Whereas, in our work, we aim to optimally select the sellers, under conditions of uncertainty, by securing additional information on the seller's (and advisor's) quality levels from other agents in the market, eliminating the need to explicitly engage in personal interactions with the sellers.

3. THE SALE POMDP MODEL

In this section, we introduce the SALE POMDP framework for dealing with the seller section problem by optimally selecting the right seller to perform transaction. Unlike *Advisor* POMDP, SALE POMDP assumes that advisors also

have quality levels (*trustworthiness*) and considers it as part of the state space. It also allows agents to query about the quality of other advisors in the system. Each SALE POMDP agent is described using the following elements.

- $\mathcal{S}$—a set of possible states of the environment. A state contains the quality levels[2] of each seller, advisor and the status of the current transaction. Let $\mathcal{Q}$ be the discrete set of seller quality levels and $\mathcal{U}$ be the set of advisor quality levels. Then, a state is a tuple $s = \langle \vec{q}, \vec{u}, sat \rangle$, where $\vec{q} \in \mathcal{Q}^J$ is a vector indicating the quality of each seller, $\vec{u} \in \mathcal{U}^I$ a vector indicating the quality of each advisor, and $sat \in \{-1, 0, +1, +2\}$ indicates whether the outcome of a transaction is satisfactory $(+1)$, unsatisfactory (-1), given up $(+2)$ due to the *donotbuy* action or whether no transaction took place yet (0). We also write q_j for the j-th element of $\vec{q}$ and u_i for the i-th element of $\vec{u}$. The end of the decision process (after a buy/donotbuy action) is modeled using sets of *terminal states*. That is, a terminal state is a state where $sat = +1$, $sat = +2$ or $sat = -1$. We will consider these sets of states as single states called *satisfied*, *unsatisfied* and *gaveup*.

- $\mathcal{A}$—a set of actions which can be:
  - *seller_query$_{ij}$*, ask advisor i about seller j,
  - *advisor_query$_{ii'}$*, ask advisor i about advisor i',
  - *buy$_j$*, buy from seller j.
  - *do_not_buy*, decide not to buy from any seller.

- T—a transition function that specifies $\Pr(s'|s, a)$, the probability of transferring to a state s' given that action a was taken in state s. We assume that when taking a query action, the state does not change:

$$\forall_{i,j} \quad \Pr(s'|s, seller\_query_{ij}) = \delta_{ss'} \quad (1)$$

$$\forall_{i,i'} \quad \Pr(s'|s, advisor\_query_{ii'}) = \delta_{ss'} \quad (2)$$

$\delta_{ss'}$ is the Kronecker delta *i.e.* 1 if and only if $s = s'$. When taking a *buy$_j$* action, the state will always transition to a terminal state. The transition probabilities to terminal states give a definition of the quality levels. In general, chances of transition to *satisfied* should be higher when buying from higher quality sellers. Together, the specifications of these transitions imply the assumption that quality and trust-levels are stationary for the duration of the decision process.

- R—a reward function specifying $R(s, a, s')$: a small cost is associated with ask action $R(s, seller\_query_{ij}) = R(s, advisor\_query_{ii'}) = R_{ask}$; a reward is associated with a good purchase $R(s, buy_j, s' = \langle \vec{q}, \vec{u}, sat = +1 \rangle) = R_{sat}$; a penalty is levied when the outcome of a transaction is *unsatisfied* $R(s, buy_j, s' = \langle \vec{q}, \vec{u}, sat = -1 \rangle) = R_{unsat}$ and a penalty for taking the *do_not_buy* action when in fact there is a seller of high enough quality (we use $-R_{sat}$), otherwise the reward for this action is 0. Once the state changed to *satisfied*, *unsatisfied* or *gaveup*, no further rewards are given.

- Ω—a set of observations. When a *query* action is performed the agent will receive an observation from the

[2] we assume discrete quality levels, in order to use standard POMDP solvers.

set of discriminated quality levels. That is, after a *seller_query$_{ij}$* action, the agent receives an observation $o \in \mathcal{Q}$ corresponding to the quality of seller j, while after an *advisor_query$_{ii'}$* it will get an observation $o \in \mathcal{U}$ corresponding to the quality of advisor i'. When the agent transitions to a terminal state, it receives the observation *ended*. As such $\mathcal{O} = \mathcal{Q} \cup \mathcal{U} \cup \{ended\}$.

- O—the observation function that specifies $\Pr(o|a, s)$. As in the Advisor POMDP, there is no a priori correct way to specify the observation probabilities. The probabilities picked for the observation function define the meaning of different trust levels. In general, the idea is that trustworthy advisors will give more accurate and consistent answers than untrustworthy ones.

- b^0—the initial state distribution. It is dependent on the subjective beliefs of the agent, when the need for purchasing an item arises. For simplicity, one may start with a uniform belief over the quality levels, but a different initial belief can also be obtained as a result of previous interactions. That is, once the *buy$_j$* or *do_not_buy* action is performed, the resulting belief can be used as the basis for an initial belief for a new seller selection instantiation. In a bit more detail, There are two sources of previous experience: 1) previous seller selection tasks: the modified belief state resulting from advice in a previous problem can be retained and 2) actual experiences with sellers: even though in the decision making task we model a transition to a terminal state with a deterministic *ended* observation, the actual order will result in the owner of the agent being satisfied or not and this information can be used to update the final belief of the agent's previous seller selection task giving a new initial belief for a new task[3]

- h—the *horizon* of the problem. That is the number of time steps, or *stages*, for which we want to plan. We will assume that h is infinite in this paper.

When the SALE POMDP agent interacts with the environment, it maintains a *belief* b, i.e., a probability distribution over states via Bayes' rule. That is, when $b(s)$ specifies the probability of s (for all s), we can derive b' an updated belief after taking some action a and receiving an observation o. The belief updates are performed such that they correlate the state factors in meaningful ways. For instance, observing good after *seller_query$_{ij}$* should give more weights to states where the seller is high quality $q_j = H$ and the advisor is trustworthy $u_i = T$, and less weights to states where the seller is low quality $q_j = L$ and the advisor is untrustworthy $u_i = U/A$. Similarly, observing T after *advisor_query$_{ii'}$* should put more weight on states where $u'_i = T$ and $u_i = T$, and decrease weight on states where $u'_i = U/A$ and $u_i = T$. Assuming discrete sets of states and observations, this update can be written as follows:

$$b'(s') = \frac{\Pr(s', o|b, a)}{\Pr(o|b, a)} = \frac{1}{\Pr(o|b, a)} \Pr(o|a, s') \sum_s \Pr(s'|s, a)b(s) \quad (3)$$

Here, $\Pr(o|b, a)$ is a normalization factor.

[3] In fact this can be an important mechanism to deal with advisors that are consistent but deceptive and settings in which the majority of advisors is untrustworthy.

These beliefs are the basis for decision making: a policy π maps beliefs to actions $\pi(b) = a$. The goal of solving the POMDP is to find an optimal policy that maximizes the expected discounted cumulative reward, also called *value*:

$$V(\pi) = \mathbb{E}\Big[\sum_{t=0}^{h-1} \gamma^t R(s, a, s') \mid \pi, b^0\Big], \qquad (4)$$

with $0 \leq \gamma < 1$ the discount factor.

Finding an optimal policy π^* is intractable in general (PSPACE complete [8]), however, in recent years substantial advances have been made in approximate solutions (e.g., [5, 14]). Since the SALE POMDP really is a *factored* POMDP, we try and use solvers that exploit this property such as symbolic Perseus [9] as described in the next section.

4. FACTORED REPRESENTATION

POMDPs with very large state spaces are impractical to be solved by the classic solution algorithms. However, such state spaces can be described using a set of state variables, and the effects of actions in terms of their effects on these variables. Dynamic Bayesian Networks (DBNs) with conditional probability tables (CPTs) in the form of decision-trees or graphs are often used to represent these effects compactly referred to as a *factored* representation [13]. Solvers such as Symbolic Perseus [9] exploit this *factored* structure to resolve scalability issues in POMDPs.

To illustrate the *factored* nature of the SALE POMDP, we will consider a simple case of the seller selection problem with 1 seller and 2 advisors such that $\mathcal{Q} = \{L, H\}$ denoting L(ow) and H(igh) seller quality and $\mathcal{U} = \{T, U, A\}$ denoting T(rustworty), U(ntrustworty) or A(dversarial) advisor quality levels. Here each state $s \in \mathcal{S}$ is a compilation of the quality of each seller, quality of each advisor and transaction outcome, *i.e.* $s = \langle\langle q_1\rangle, \langle u_1, u_2\rangle, sat\rangle$. To factor the SALE POMDP model, we decompose the transition and observation functions as follows.

Buy Action The DBN and CPT for the buy_1 action (buy from s_1) are shown in Fig 1. The probability of transferring to a state s' given that action buy_1 was taken in state s can be factored into a product of smaller conditional distributions with respect to its parent variables as shown in Eqn 5. $\Pr(sat'|q_1', sat, buy_1)$ signifies that the outcome of a transaction sat' at a time step depends on the previous status of the transaction sat, the true quality of the seller s_1' and the buy_1 action. The CPT for sat' variable is framed assuming that buying from a high quality seller will lead to *satisfied* with 80% probability: $\Pr(sat = -1 | \langle sat = 0, q_j = L\rangle, buy_j) = 0.8$. Similarly, we will assume that a low quality seller will lead to *unsatisfied* with probability 0.8. For the *do_not_buy* action, the outcome sat' is not dependant on the seller quality q_1' and $\Pr(sat'|sat, do\_not\_buy) = 1$.

$$\begin{aligned}
\Pr(s'|s, buy_1) &= \Pr(u_1', u_2', q_1', sat'|u_1, u_2, q_1, sat, buy_1) \\
&= \Pr(u_1'|u_1, buy_1) \times \Pr(u_2'|u_2, buy_1) \times \quad (5) \\
&\quad \Pr(q_1'|q_1, buy_1) \times \Pr(sat'|sat, q_1', buy_1)
\end{aligned}$$

Seller Query Fig 2 shows the DBN and CPT for the action $seller\_query_{11}(sq_{11})$ (query a_1 about s_1). Eqn 6 shows that observing the quality of a seller s_1 through o' is conditionally dependant on the current quality of the seller q_1' (after the $seller\_query_{11}$ action) and the quality of the advisor u_1', who provides opinion. 2 possible observations are

considered: *good* (i.e., the advisor says that the seller is high quality) denoted by g and *bad* (the seller is said to be low quality) denoted by b. The observation probabilities for the $seller\_query_{ij}$ action are such that asking a trustworthy (T) advisor gives more accurate observations, while untrustworthy ones may provide a random (U) opinion or they can be adversarial (A) in nature. We will obtain a similar figure for $seller\_query_{21}$ action.
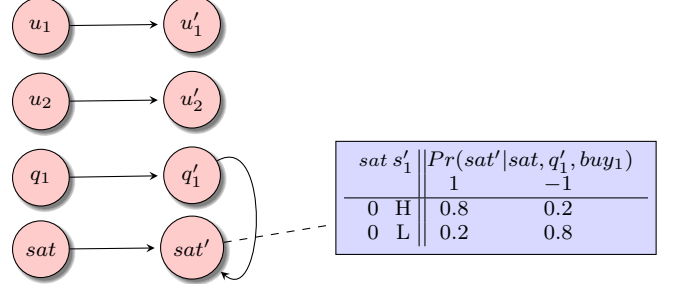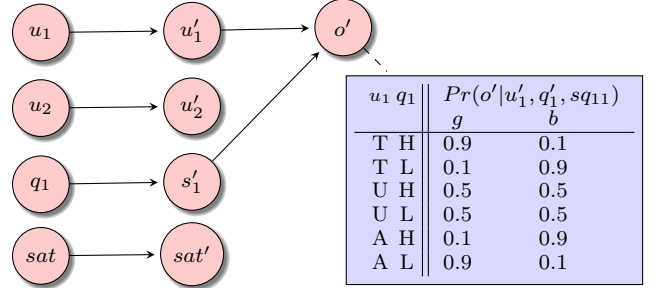

sat	s_1'	$Pr(sat'\|sat, q_1', buy_1)$	
		1	-1
0	H	0.8	0.2
0	L	0.2	0.8

Figure 1: DBN for buy_1

u_1	q_1	$Pr(o'\|u_1', q_1', sq_{11})$	
		g	b
T	H	0.9	0.1
T	L	0.1	0.9
U	H	0.5	0.5
U	L	0.5	0.5
A	H	0.1	0.9
A	L	0.9	0.1

Figure 2: DBN for $seller\_query_{11}$

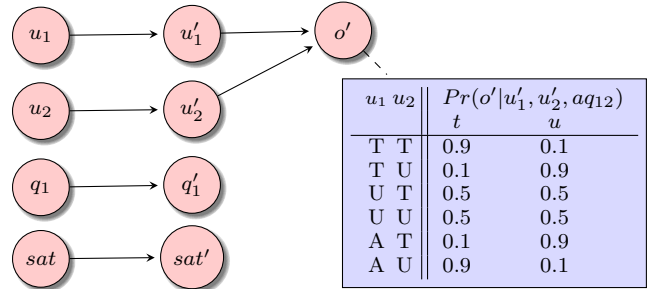| u_1	u_2	$Pr(o'\|u_1', u_2', aq_{12})$	
		t	u
T	T	0.9	0.1
T	U	0.1	0.9
U	T	0.5	0.5
U	U	0.5	0.5
A	T	0.1	0.9
A	U	0.9	0.1

Figure 3: DBN for $advisor\_query_{12}$

$$\begin{aligned}
\Pr(o'|s, sq_{11}) &= \Pr(o'|u_1', u_2', q_1', sat', sq_{11}) \\
&= \Pr(o'|u_1', q_1', sq_{11})
\end{aligned} \qquad (6)$$

Advisor Query Fig 3 shows the DBN and CPT for the action $advisor\_query_{12}(aq_{12})$ (query a_1 about a_2). Eqn 7 show that observing the quality of advisor a_2 using the variable o' (we use the same observation variable for simplicity) is conditionally dependant on the current quality of the advisor u_2' (after the $advisor\_query_{12}$ action) and the quality of the advisor a_1 (u_1'), who provides opinion. We will obtain a similar figure for $advisor\_query_{21}$ action.

$$\begin{aligned}
\Pr(o'|s, aq_{12}) &= \Pr(o'|u_1', u_2', s_1', sat', aq_{12}) \\
&= \Pr(o'|u_1', u_2', aq_{12})
\end{aligned} \qquad (7)$$

The conditional probability tables in the above figures can be further compressed by exploiting context-specific independence which refers to the fact that some random variables are independent of each other in some contexts (i.e., for some specific values of other random variables). In general, context-specific independence allows one to compactly represent probability and utility tables using Horn rules [107] decision trees (DTs) or algebraic decision diagrams (ADDs) [1]. We will show one such ADD for the CPT in Fig 2.
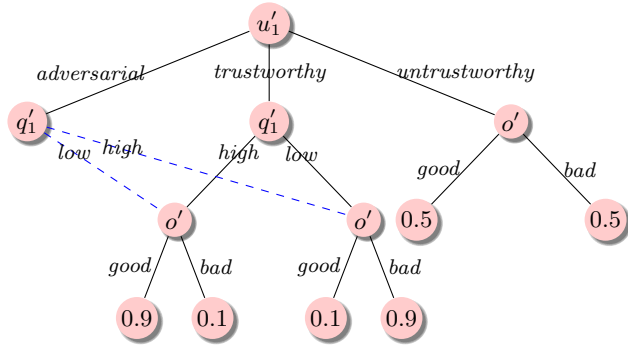

Figure 4: ADD for $seller\_query_{11}$

Fig. 4 illustrates the ADD representations of the CPT for the observation variable o' in Fig 2. The value of o' for some trust assignment for variables u'_1, q'_1 can be found by following the corresponding branch in the ADD. Note that when u'_1 is *adversarial* in nature, it will result in the exact compliment of the values obtained when traversed through the path when u'_1 is *trustworthy*. This can be compactly represented as shown in Fig. 4. Also, when the advisor behaves in a random manner $u'_1 = untrustworthy$, o' is independent of q'_1 and results in a value of 0.5 whether q'_1 is of *high* or *low* quality. Hence, by using the notion of context-specific independence, the ADD in Fig. 4 can encode with only 6 leaves the 12 values of the CPT in Fig 2. Thus we show that by exploiting the features of conditional-independence and context-specific independence in SALE POMDP, we can use solvers such as Symbolic Perseus to find the optimal policy, especially for POMDPs with a large number of states.

5. EVALUATION

6. CONCLUSION AND FUTURE WORK

7. REFERENCES

[1] C. Boutilier, N. Friedman, M. Goldszmidt, and D. Koller. Context-specific independence in bayesian networks. In *Proceedings of the Twelfth international conference on Uncertainty in artificial intelligence*, pages 115–123. Morgan Kaufmann Publishers Inc., 1996.

[2] P. J. Gmytrasiewicz and P. Doshi. A framework for sequential planning in multi-agent settings. *J. Artif. Intell. Res.(JAIR)*, 24:49–79, 2005.

[3] A. A. Irissappane, S. Jiang, and J. Zhang. Towards a comprehensive testbed to evaluate the robustness of reputation systems against unfair rating attack. In *UMAP Workshops*, volume 12, 2012.

[4] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial intelligence*, 101(1):99–134, 1998.

[5] H. Kurniawati, D. Hsu, and W. S. Lee. Sarsop: Efficient point-based pomdp planning by approximating optimally reachable belief spaces. In *Robotics: Science and Systems*, pages 65–72, 2008.

[6] K. W. Mayo. Multi-objective constraint satisfaction for mobile robot area defense. Technical report, DTIC Document, 2010.

[7] F. A. Oliehoek, A. A. Gokhale, and J. Zhang. Reasoning about advisors for seller selection in e-marketplaces via pomdps. In *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS) Workshop on Trust in Agent Societies*, pages 67–78, 2012.

[8] C. H. Papadimitriou and J. N. Tsitsiklis. The complexity of markov decision processes. *Mathematics of operations research*, 12(3):441–450, 1987.

[9] P. Poupart. *Exploiting structure to efficiently solve large scale partially observable Markov decision processes*. PhD thesis, University of Toronto, 2005.

[10] K. Regan, R. Cohen, and P. Poupart. The advisor-pomdp: A principled approach to trust through reputation in electronic markets. In *Proceedings of the Third Annual Conference on Privacy, Security and Trust (PST)*, pages 121–130. Citeseer, 2001.

[11] K. Regan, P. Poupart, and R. Cohen. Bayesian reputation modeling in e-marketplaces sensitive to subjectivity, deception and change. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 2006.

[12] R. Seymour and G. L. Peterson. A trust-based multiagent system. In *Proceedings of IEEE International Conference on Computational Science and Engineering (CSE)*, volume 3, pages 109–116. IEEE, 2009.

[13] G. Shani. Task-based decomposition of factored pomdps. 2013.

[14] D. Silver and J. Veness. Monte-carlo planning in large pomdps. In *Advances in Neural Information Processing Systems*, pages 2164–2172, 2010.

[15] W. Teacy, J. Patel, N. Jennings, and M. Luck. TRAVOS: Trust and reputation in the context of inaccurate information sources. *Autonomous Agents and Multi-Agent Systems*, 12(2):183–198, 2006.

[16] A. Whitby, A. Jøsang, and J. Indulska. Filtering out unfair ratings in bayesian reputation systems. In *Proceedings of the International Joint Conference on Autonomous Agenst Systems (AAMAS) Workshop on Trust in Agent Societies (TRUST)*, 2004.

[17] J. Zhang. *Promoting Honesty in E-Marketplaces: Combining Trust Modeling and Incentive Mechanism Design*. PhD thesis, University of Waterloo, 2009.

[18] J. Zhang and R. Cohen. Evaluating the trustworthiness of advice about seller agents in e-marketplaces: A personalized approach. *Electronic Commerce Research and Applications*, 7(3):330–340, 2008.